Journal of Statistical Physics, Vol. 105, Nos. 3/4, November 2001 (© 2001)

Clustering in a Self-Gravitating One-Dimensional Gas at Zero Temperature

Christophe Giraud[1]

Received April 12, 2001; revised June 15, 2001

We study a system of gravitationally interacting sticky particles. At the initial time, we have n particles, each with mass $1/n$ and momentum 0, independently spread on $[0, 1]$ according to the uniform law. Due to the confining of the system, all particles merge into a single cluster after a finite time. We give the asymptotic laws of the time of the last collision and of the time of the kth collision, when $n \to \infty$. We prove also that clusters of size k appear at time $\sim n^{-1/2(k-1)}$. We then investigate the system at a fixed time $t < 1$. We show that the biggest cluster has size of order $\log n$, whereas a typical cluster is of finite size.

KEY WORDS: Sticky particles; gravitational interacting; uniform law; Brownian bridge.

1. INTRODUCTION

The dynamics of gravitationally interacting sticky particles are a model that has been suggested by Zeldovich[(11)] and other authors to investigate the formation of large scale structure in the universe, see ref. 10 for a survey article. We focus in this paper on the one-dimensional case. Sticky particles are particles which collide in a completely inelastic way. When particles meet, they form a new massive particle with conservation of mass and momentum. More precisely, when particles with mass m_i and m_j collide, they merge into a single particle with mass $m_i + m_j$, which follows the trajectory of their center of mass. It must be noticed that in the one dimensional case, only the nearest neighboring particles can collide. Following,[(1, 4)] we consider the case when these sticky particles are attracting

[1] Laboratoire de Probabilités et Modèles Aléatoires, Université Pierre et Marie Curie, et C.N.R.S. UMR 7599, 175, rue du Chevaleret, F-75013 Paris, France; e-mail: christophe.giraud@ens.fr

0022-4715/01/1100-0585$19.50/0

each other with forces proportional to the product of their masses, independently of the interparticle distance. Rigorously, the dynamic between collisions is governed by the Hamiltonian

$$H=\sum_i \frac{p_i^2}{2m_i}+\gamma \sum_{i\neq j} m_i m_j\,|x_i-x_j|,$$

where x_i, m_i, p_i denote the location, mass and momentum of the particle i and γ is the gravitational constant. The acceleration of a particle is then proportional to the difference between the total masses at its right and at its left. It is to be mentioned that these dynamics give rise to global weak solutions to the system of conservation laws

$$\begin{cases} \partial_t \rho+\partial_x(\rho u)=0 \\ \partial_t(\rho u)+\partial_x(\rho u^2)=-\gamma\rho\,\partial_x \Phi \\ \partial_{xx}\Phi=\rho \end{cases}$$

where $\rho(x,t)$, $u(x,t)$, $\Phi(x,t)$ are meant to represent the velocity, density and gravitational potential at x at time t. This connection occurs when the initial density $\rho(.,0)$ is a purely singular measure supported on a finite or countable set, see ref. 3 for further explanations.

The subject that is of interest in the study of gravitationally interacting sticky particles is the mass distributions induced by small perturbations of an initial homogeneous state. When a finite number of identical particles are initially spread on the regular lattice sites $x_j=ja$, $j=1\cdots n$, with mass m and momentum 0, they all merge simultaneously into a single cluster at the characteristic time $t^*=1/\sqrt{\gamma\rho}$, where $\rho=m/a$, see ref. 4, Section 3. Martin, Piasecki, Bonvin and Zotos[1,4] and also Suidan[9] have focused on the evolution of the system, when the initial homogeneous state is perturbed by introducing random uncorrelated velocities to the initial particles. More specifically, they mainly work with a system of n initial particles located on the lattice sites $x_j=ja$, $j=1\ldots n$, with mass m and uncorrelated velocities distributed according to the Gaussian law. They study the statistics of the continuum limit $a\to 0$, while keeping $\rho=m/a$ constant. In the concluding remarks of ref. 4, they raise the question of the evolution of a system, in which the perturbation of the initial homogeneous state should result from the randomization of the initial locations of the particles. This question mainly motivates the present work.

We focus thenceforth on a system of initially n gravitationally interacting sticky particles with mass $1/n$, momentum 0 and which are independently spread on $[0,1]$ according to the uniform law. Up to the change of time $t' = \sqrt{\gamma}\, t$, we can fix $\gamma = 1$. It should be mentioned that the evolution of the system is isomorphic to the evolution of n gravitationally interacting sticky particles with unit mass and momentum 0, independently spread at the initial time on $[0, n]$ according to the uniform law. We shall mainly investigate here the asymptotics of the statistics when n tends to infinity.

We start with giving in section 2 some material needed to study the system. In section 3 and 4, we specify the asymptotic laws of the first and last collisions, as well as the time scale of appearance of a cluster of finite size k. In section 5, we determine the size of scale of the biggest cluster at a fixed time t. Some elements on the evolution of a marked particle are presented in the last section.

2. PRELIMINARIES

2.1. Analyzing the System

We shall give in this section some key results for our analysis. We first properly define the system. For any $n \geqslant 1$, $(X_{n,i}, i = 1,..., n)$ shall denote n independent random variables with uniform law on $[0, 1]$. We write $0 \leqslant X_{n:1} \leqslant \cdots \leqslant X_{n:n} \leqslant 1$ for the ordered statistics. We consider henceforth a system of n particles of mass $1/n$ spread at the initial time on the sites $X_{n,i}$ with momentum 0.The particle initially located at $X_{n:i}$ should be called the ith particle. These particles are assumed to evolve as time runs according to the dynamics of gravitationally interacting sticky particles described previously.

Our investigations are mainly based on an analysis made independently by Martin and Piasecki[4] (see also ref. 1) and E *et al.*[3] Let us consider the k particles $(i+1,..., i+k)$. Recall that masses and momenta are conserved during collisions and that the acceleration acting on these k particles is equal to the difference of the masses at their right and at their left. As a consequence, if these k particles have not collide with surrounding particles before time t, their center of mass $G(i+1,..., i+k)$ follows the trajectory (see for close formulaes (11) in ref. 1 and (1–19) in ref. 3)

$$G(i+1,..., i+k)(s) = G(i+1,..., i+k)(0) + \frac{(n-(k+2i))\, s^2}{2n} \quad \text{for any } s \leqslant t.$$

We shall state now the key of the analysis of the system: a necessary and sufficient condition for the k particles $(i+1,..., i+k)$ to merge into a single cluster of size k at time t, is that these particles did not collide with surrounding particles before time t and that for any partition into two subclusters $(i+1,\ldots, i+r)$ and $(i+r+1,\ldots, i+k)$, the centers of mass of these subclusters cross before time t (see formula (6) in ref. 4 and also formula (1–12) in ref. 3). In particular, the condition

$$G(i+1,\ldots, i+r)(t) \geqslant G(i+r+1,\ldots, i+k)(t), \qquad \text{for} \quad r=1,\ldots, k-1. \tag{1}$$

is a necessary condition for the merging of $(i+1,\ldots, i+k)$ into a single cluster of size k before time t. It is also a sufficient condition for the merging of $(i+1,\ldots, i+k)$ into a cluster of size at least k. Indeed, this results from the fact that when the cluster $(i-l,\ldots, i)$ collide with the cluster $(i+1,\ldots, 1+r)$ at time s, their trajectories cross and for $t \geqslant s$

$$G(i-l,\ldots, i)(t) \geqslant G(i-l,\ldots, i+r)(t) \geqslant G(i+1,\ldots, i+r)(t),$$

see Lemma 6 in ref. 3 (and also Lemmas 2 and 3 there) for very close arguments. Expressing condition (1) in terms of the initial locations of the particles gives

$$\frac{1}{r}\sum_{j=1}^{r} X_{n:i+j} - \frac{1}{k-r}\sum_{j=r+1}^{k} X_{n:i+j} + \frac{kt^2}{2n} \geqslant 0, \qquad \text{for} \quad r=1,\ldots, k-1,$$

which is not easily amenable to mathematical analysis. We thus look for (weaker) necessary and sufficient conditions which only involve $X_{n:i+k} - X_{n:i+1}$. First, it follows from the inequality

$$X_{n:i+1} - X_{n:i+k} \leqslant \frac{1}{r}\sum_{j=1}^{r} X_{n:i+j} - \frac{1}{k-r}\sum_{j=r+1}^{k} X_{n:i+j},$$

that a sufficient condition for the merging of $(i+1,\ldots, i+k)$ into a cluster of size at least k before time t is

$$X_{n:i+1} - X_{n:i+k} + \frac{kt^2}{2n} \geqslant 0. \tag{2}$$

Second, when $(i+1,\ldots, i+k)$ merge into a cluster of size k before time t, the trajectory $s \to G(i+1)(s)$ of the particle $i+1$ crosses the trajectory $s \to G(i+1,\ldots, i+k)(s)$ before time t, which implies that $G(i+1)(t) \geqslant G(i+1,\ldots, i+k)(t)$. For the same reasons, the inequality $G(i+k)(t) \leqslant G(i+1,\ldots, i+k)(t)$ holds, which enables us to formulate a simplier necessary condition for the merging of $(i+1,\ldots, i+k)$ into a single cluster of size k before time t as

$$X_{n:i+1} - X_{n:i+k} + \frac{kt^2}{n} \geqslant 0. \tag{3}$$

We next want to give necessary and sufficient conditions for the existence of a cluster of size at least k at time t. It follows from the previous analysis that a sufficient condition is

$$\exists i \in \{0,\dots, n-k\} \qquad \text{such that for} \quad r=1,\dots, k-1,$$
$$\frac{1}{r}\sum_{j=1}^{r} X_{n:i+j} - \frac{1}{k-r}\sum_{j=r+1}^{k} X_{n:i+j} + \frac{kt^2}{2n} \geqslant 0, \tag{4}$$

and thus it suffices that

$$\exists i \in \{0,\dots, n-k\} \qquad \text{such that} \quad X_{n:i+1} - X_{n:i+k} + \frac{kt^2}{2n} \geqslant 0. \tag{5}$$

Let us give now a necessary condition. There exists a cluster of size at least k at time t if and only if, for some time $s \leqslant t$ there exists a cluster of size between k and $2k$, which only occurs if there exists $s \leqslant t$, $k \leqslant p \leqslant 2k$, and $i \in \{0,\dots, n-p\}$ such that for any $r \in \{1,\dots, n-p\}$, $G(i+1,\dots, i+r)(s) \geqslant G(i+r+1,\dots, i+p)(s)$. A necessary condition is thus

$$\exists k \leqslant p \leqslant 2k \;\text{ and }\; i \in \{0,\dots, n-p\} \;\text{ such that for }\; r=1,\dots, p-1,$$
$$\frac{1}{r}\sum_{j=1}^{r} X_{n:i+j} - \frac{1}{p-r}\sum_{j=r+1}^{p} X_{n:i+j} + \frac{pt^2}{2n} \geqslant 0, \tag{6}$$

and a fortiori

$$\exists k \leqslant p \leqslant 2k \;\text{ and }\; i \in \{0,\dots, n-p\} \;\text{ such that }\; X_{n:i+1} - X_{n:i+p} + \frac{pt^2}{n} \geqslant 0. \tag{7}$$

Armed of this collection of necessary or sufficient conditions, we are ready to start our investigations, after recalling some basic features on empirical and quantile processes.

2.2. Uniform Empirical and Quantile Processes

Our study involves some basic and some more advanced results on the uniform quantile process. For the convenience of the reader, we recall a few basic facts in this field. We refer to ref. 8 for a classical text-book. We associate to the n i.i.d. random variables $(X_{n,i}, i=1, \dots, n)$ with uniform law on $[0, 1]$, the uniform empirical distribution function

$$G_n(t) := \frac{1}{n}\sum_{i=1}^{n} \mathbf{1}_{\{X_{n,i} \leqslant t\}}$$

and its left-continuous inverse $G_n^{-1}(t) := \inf\{x\colon G_n(x) \geqslant t\}$. It is easily seen that

$$G_n^{-1}(t) = X_{n:i} \qquad \text{for} \quad \frac{i-1}{n} < t \leqslant \frac{i}{n},$$

and $G_n^{-1}(0) = 0$, where $X_{n:i}$ denotes the ordered statistics. The law of large numbers implies that $G_n(t) \xrightarrow{n\to\infty} t$ a.s. and as a consequence $G_n^{-1}(t) \xrightarrow{n\to\infty} t$ a.s. Actually, if I is the identity function on $[0, 1]$, it follows from the central limit theorem, that the so-called uniform empirical process $u_n := \sqrt{n}(G_n^{-1} - I)$ converges in law towards a Brownian bridge, when n tends to infinity. In analogy with the uniform empirical process, we define the uniform quantile process as $v_n = \sqrt{n}(G_n^{-1} - I)$, i.e.,

$$v_n(t) = \sqrt{n}(X_{n:i} - t), \qquad \text{for} \quad \frac{i-1}{n} < t \leqslant \frac{i}{n},$$

and $v_n(0) = 0$. In light of the formula $v_n = -u_n(G_n^{-1}) + \sqrt{n}(G_n \circ G_n^{-1} - I)$, it is easily seen that the uniform quantile process converges in law towards a Brownian bridge, when n tends to infinity. Furthermore, the present work also relies on some more advanced properties on the ordered statistics related to the modulus of continuity of the uniform quantile process, cf. Section 14-7 in ref. 8.

3. LAST COLLISION

A consequence of the confining of the system is that every particles have merged into a single cluster after a finite time. We have already recall that when particles are initially located on the lattice sites i/n, they all collide simultaneously at time 1. We study in this section the effect on the last collision of a randomization of the initial location. We show that in this case the last collision occurs between two macroscopic clusters at time of order $1 + O(1/\sqrt{n})$.

Theorem 3.1. In the n-particles system, the last collision occurs a.s. between two macroscopic clusters at time $T_n^{\text{l.c.}}$ which follows the convergence in law when $n \to \infty$

$$\sqrt{n}\,(T_n^{\text{l.c.}} - 1) \xrightarrow{\text{law}} \sup_{x\in[0,1]} \left(\frac{1}{1-x} \int_x^1 b(t)\,dt - \frac{1}{x} \int_0^x b(t)\,dt \right),$$

where b denotes a Brownian bridge.

It should be noticed that the law of

$$\sup_{x \in [0,1]} \left(\frac{1}{1-x} \int_x^1 b(t)\,dt - \frac{1}{x} \int_0^x b(t)\,dt \right)$$

is not degenerated since we have the inequalities

$$\left| \int_0^1 b(t)\,dt \right| \leqslant \sup_{x \in [0,1]} \left(\frac{1}{1-x} \int_x^1 b(t)\,dt - \frac{1}{x} \int_0^x b(t)\,dt \right) \leqslant 2 \sup_{x \in [0,1]} |b(t)|.$$

Proof of Theorem 3.1. We first focus on the time $T_n^{\text{l.c.}}$ of last collision. According to condition (1), the last collision occurs before time t if and only if

$$\text{for} \quad r = 1,..., n-1, \qquad \frac{1}{n-r} \sum_{i=r+1}^{n} X_{n:i} - \frac{1}{r} \sum_{i=1}^{r} X_{n:i} \leqslant \frac{t^2}{2},$$

which implies

$$(T_n^{\text{l.c.}})^2 = 2 \sup_{r=1,..., n-1} \left(\frac{1}{n-r} \sum_{i=r+1}^{n} X_{n:i} - \frac{1}{r} \sum_{i=1}^{r} X_{n:i} \right).$$

We want to express the time of last collision in terms of the uniform quantile process that has been introduced in Section 2.2. It follows from the equalities

$$\int_0^{r/n} v_n(t)\,dt = \sqrt{n} \sum_{i=1}^{r} \frac{1}{n} X_{n:i} - \sqrt{n}\, \frac{r^2}{2n^2}$$

and

$$\int_{r/n}^1 v_n(t)\,dt = \sqrt{n} \sum_{i=r+1}^{n} \frac{1}{n} X_{n:i} - \sqrt{n} \left(\frac{1}{2} - \frac{r^2}{2n^2} \right)$$

that

$$(T_n^{\text{l.c.}})^2 = 1 + \frac{2}{\sqrt{n}} \sup_{r=1,..., n-1} \left(\frac{1}{1-r/n} \int_{r/n}^1 v_n(t)\,dt - \frac{n}{r} \int_0^{r/n} v_n(t)\,dt \right).$$

Recall that v_n converges in law to a Brownian bridge b when $n \to \infty$ (see, e.g., ref. 8, Chap. 3). We shall prove now the convergence

$$\sqrt{n}\,((T_n^{\text{l.c.}})^2 - 1) \xrightarrow{\text{law}} 2 \sup_{x \in [0,1]} \left(\frac{1}{1-x} \int_x^1 b(t)\,dt - \frac{1}{x} \int_0^x b(t)\,dt \right) \text{ when } n \to \infty, \tag{8}$$

from which follows the convergence given in Theorem 3.1, since $(T_n^{\text{l.c.}})^2-1 = (T_n^{\text{l.c.}}+1)(T_n^{\text{l.c.}}-1)$ and $T_n^{\text{l.c.}} \to 1$ in probability.

In order to prove (8) we define for any left continuous with right limits functions v

$$f_n(v) := \sup_{r=1,\dots,n-1} \frac{1}{1-r/n}\int_{r/n}^{1} v(t)\,dt - \frac{n}{r}\int_0^{r/n} v(t)\,dt$$

$$\text{and} \quad f_\infty(v) := \sup_{x\in[0,1]} \frac{1}{1-x}\int_x^1 v(t)\,dt - \frac{1}{x}\int_0^x v(t)\,dt.$$

We shall prove that for any Lipschitz bounded function g,

$$\mathbb{E}(g(f_n(v_n))) \xrightarrow{n\to\infty} \mathbb{E}(g(f_\infty(b))).$$

By the triangle inequality

$$\begin{aligned}
&|\mathbb{E}(g(f_n(v_n))) - \mathbb{E}(g(f_\infty(b)))| \\
&\quad \leqslant |\mathbb{E}(g(f_\infty(v_n)) - g(f_n(v_n)))| + |\mathbb{E}(g(f_\infty(v_n)) - g(f_\infty(b)))|.
\end{aligned} \tag{9}$$

The second term tends to 0 when $n\to\infty$, since $g\circ f_\infty$ is a continuous bounded functional and $v_n \xrightarrow{\text{law}} b$. We focus now on the first term. For $\frac{r-1}{n} < x \leqslant \frac{r}{n}$ we have the inequalities

$$\begin{aligned}
&\left|\frac{1}{x}\int_0^x v_n(t)\,dt - \frac{n}{r}\int_0^{r/n} v_n(t)\,dt\right| \\
&\leqslant \left|\frac{1}{x}-\frac{n}{r}\right|\int_0^x |v_n(t)|\,dt + \frac{n}{r}\int_x^{r/n} |v_n(t)|\,dt \\
&\leqslant \left(\frac{n}{r-1}-\frac{n}{r}\right)\int_0^{r/n} |v_n(t)|\,dt + \frac{n}{r}\int_{(r-1)/n}^{r/n} |v_n(t)|\,dt \\
&\leqslant \frac{1}{r-1}\sup_{t\in[0,r/n]} |v_n(t)| + \frac{1}{r}\sup_{t\in[(r-1)/n,\,r/n]} |v_n(t)| \\
&\leqslant \begin{cases} 2\sup\limits_{t\in[0,\sqrt{n}/n]} |v_n(t)| & \text{for } 2\leqslant r\leqslant\sqrt{n} \\ 2\sup\limits_{t\in[0,1]} |v_n(t)|/\sqrt{n} & \text{for } r\geqslant\sqrt{n}+1 \end{cases} \\
&\leqslant \frac{2}{\sqrt{n}}\sup_{t\in[0,1]} |v_n(t)| + 2\sup_{t\in[0,\sqrt{n}/n]} |v_n(t)| =: K_n,
\end{aligned}$$

and this upper bound remains true for $r=1$. For any $\varepsilon>0$, we can write

$$\begin{aligned}
&|\mathbb{E}(g(f_n(v_n))-g(f_\infty(v_n)))|\\
&\quad\leqslant |\mathbb{E}(g(f_n(v_n))-g(f_\infty(v_n));\, K_n\leqslant\varepsilon)|+2\,\|g\|_\infty\,\mathbb{P}(K_n>\varepsilon)\\
&\quad\leqslant M\varepsilon+2\,\|g\|_\infty\,\mathbb{P}(K_n>\varepsilon),
\end{aligned}$$

where $\|g\|_\infty=\sup_{t\in\mathbb{R}}|g(t)|$ and M is the Lipschitz modulus of g. We take first the limit $n\to\infty$, and then $\varepsilon\to 0$ to obtain

$$\lim_{n\to\infty}\mathbb{E}(g(f_n(v_n))-g(f_\infty(v_n)))=0.$$

Substituting this result in the inequality (9) gives formula (8). The proof of the convergence in law of $T_n^{\mathrm{l.c.}}$ is complete.

Let us check now the first assertion of Theorem 3.1. The asymptotic masses of the two last clusters are given by the abscissa x_0 for which

$$\alpha(x)=\frac{1}{1-x}\int_x^1 b(t)\,dt-\frac{1}{x}\int_0^x b(t)\,dt$$

reaches its maximum. Indeed, the asymptotic masses of the two last clusters are given by x_0 and $1-x_0$. All that we need is to show that x_0 is different from 0 and 1 with probability 1. We can write

$$\alpha(x)=\frac{1}{1-x}\left(\int_0^1 b(t)\,dt-\frac{1}{x}\int_0^x b(t)\,dt\right).$$

Suppose for example that $\int_0^1 b(t)\,dt$ is positive. Then $\alpha(1)\leqslant 0\leqslant\alpha(0)$. We shall prove that with probability 1 there exists $X\in\,]0,1[$ such that $\int_0^X b(t)\,dt<0$, which implies in particular that $\alpha(X)>\alpha(0)$ and finally that x_0 is different from 0 and 1.

The process $I(x)=\int_0^x b(t)\,dt$ is symmetric and adapted to the filtration $\mathscr{F}_x=\sigma(b(t);\,t\leqslant x)$. Let A denotes the event

$$A:=\bigcap_{p\geqslant 1}\bigcup_{n\geqslant p}\{I(1/n)<0\},$$

that belongs to the σ-field

$$\bigcap_{x>0}\mathscr{F}_x$$

which is trivial. The probability of the event A is therefore 0 or 1. Since A is the decreasing limit when $p \to \infty$ of the events

$$\bigcup_{n \geqslant p} \{I(1/n) < 0\}$$

and since

$$\mathbb{P}\left(\bigcup_{n \geqslant p} \{I(1/n) < 0\}\right) \geqslant \mathbb{P}(I(1/p) < 0) = \frac{1}{2}$$

the probability of the event A is larger than $1/2$ and hence is 1. As a consequence I takes negative value after 0 with probability 1.

4. FIRST AGGREGATIONS

We now turn our attention on the first aggregations. We start by determining the scale of size of the time of appearance of a cluster of size k in a n-particles system.

Theorem 4.1. Let k be an integer, and t_n a sequence of positive time. When $n \to \infty$, the probability that there exists a cluster of size at least k at time t_n among a n-particles system tends to 0 if $n^{1/2(k-1)} t_n \to 0$ and tends to 1 if $n^{1/2(k-1)} t_n \to \infty$.

A notable consequence of Theorem 4.1 is the asymptotic laws of the times $T_{n:j}$ of jth collision.

Corollary 4.1. For any integer k, the k-uplet $(\sqrt{n}\, T_{n:1},..., \sqrt{n}\, T_{n:k})$ converges in law to $(\sqrt{\mathsf{e}_1},..., \sqrt{\mathsf{e}_1 + \cdots + \mathsf{e}_k})$, where $\mathsf{e}_1,..., \mathsf{e}_k$ are independent random variables with exponential law of parameter 1.

The evolution of the system at small times may be thus describe for large n as follows. Particles start to aggregate pairwise at time $\asymp n^{-1/2}$. At time $\asymp n^{-1/4}$ clusters of size 3 appear, whereas we shall wait time of order $n^{-1/6}$ to see clusters of size 4, and so on. The fact that clusters of size 3 appear before clusters of size 4 may be physically explained by the few number ($\asymp n^{1/4}$) of clusters at time $\asymp n^{-1/4}$, so that the probability they meet together is infinitesimal.

Proof of Theorem 4.1. We first give an upper bound to the probability of existence of a cluster of size at least k at time t_n. Recall by condition (7), that this probability is less than the probability of the event

$$\exists k \leqslant p \leqslant 2k \text{ and } i \in \{0,..., n-p\} \text{ such that } X_{n:i+1} - X_{n:i+p} + \frac{pt_n^2}{n} \geqslant 0.$$

It is known (see, e.g., Proposition 8-2-1 in ref. 8) that

$$(X_{n:i};\, i=1,\ldots,n) \overset{\text{law}}{\sim} \left(\frac{\mathsf{e}_1+\cdots+\mathsf{e}_i}{\Gamma_{n+1}};\, i=1,\ldots,n\right),$$

where $(\mathsf{e}_i;\, i=1,\ldots,n+1)$ are independent exponential variables of parameter 1 and $\Gamma_{n+1}=\mathsf{e}_1+\cdots+\mathsf{e}_{n+1}$. We rexpress the previous condition in terms of e_i:

$$\exists k\leqslant p\leqslant 2k \text{ and } i\in\{0,\ldots,n-p\} \text{ such that } \mathsf{e}_{i+2}+\cdots+\mathsf{e}_{i+p}\leqslant\frac{\Gamma_{n+1}}{n}\,pt_n^2.$$

We thus have for upper bound to the probability of existence of a cluster of size at least k at time t_n

$$\begin{aligned}\mathbb{P}\left(\frac{\Gamma_{n+1}}{n}>2\right)&+\sum_{p=k}^{2k}\sum_{i=0}^{n-p}\mathbb{P}(\mathsf{e}_{i+2}+\cdots+\mathsf{e}_{i+p}\leqslant 2pt_n^2)\\ &\leqslant n\sum_{p=k}^{2k}\mathbb{P}(\mathsf{e}_1+\cdots+\mathsf{e}_{p-1}\leqslant 2pt_n^2)+\mathbb{P}\left(\frac{\Gamma_{n+1}}{n}>2\right).\end{aligned}$$

Since $\Gamma_{n+1}/n\xrightarrow{\text{a.s.}}1$ according to the law of large number and

$$\mathbb{P}(\mathsf{e}_1+\cdots+\mathsf{e}_p\leqslant\lambda)=e^{-\lambda}\sum_{j=p}^{\infty}\frac{\lambda^j}{j!}\overset{\lambda\to 0}{\sim}\frac{\lambda^p}{p!},\tag{10}$$

the probability of existence of a cluster of size at least k at time t_n tends to 0 when $nt_n^{2(k-1)}\to 0$ (or in other words $n^{1/2(k-1)}t_n\to 0$). The first part of Theorem 4.1 follows.

We give now a lower bound to the probability of existence of a cluster of size at least k at time t_n. Using the condition (5), one notices that the latter is larger than the probability of the event

$$\exists i\leqslant\frac{n-k}{k}\qquad\text{such that}\quad X_{n:ki+1}-X_{n:ki+k}+\frac{kt_n^2}{2n}\geqslant 0,$$

which itself is at least

$$\mathbb{P}\left(\exists i\leqslant\frac{n-k}{k}\bigg/\sum_{j=2}^{k}\mathsf{e}_{ki+j}\leqslant\frac{kt_n^2}{4}\right)-\mathbb{P}\left(\frac{\Gamma_{n+1}}{n}\leqslant\frac{1}{2}\right).$$

Since for $i \leqslant (n-k)/k$ the events

$$\left\{\sum_{j=2}^{k} \mathsf{e}_{ki+j} \leqslant \frac{kt_n^2}{4}\right\}$$

are independent and identically distributed, we have

$$\begin{aligned}\mathbb{P}\left(\exists i \leqslant \frac{n-k}{k} \Big/ \sum_{j=2}^{k} \mathsf{e}_{ki+j} \leqslant \frac{kt_n^2}{4}\right)\\ = 1-(1-\mathbb{P}(\mathsf{e}_1+\cdots+\mathsf{e}_{k-1} \leqslant kt_n^2/4))^{[(n-k)/k]},\end{aligned}$$

where $[x]$ denotes the integer part of x. If t_n tends to 0 and $nt_n^{2(k-1)} \to \infty$ we obtain with formula (10)

$$\begin{aligned}&\mathbb{P}\left(\exists i \leqslant \frac{n-k}{k}, \sum_{j=2}^{k} \mathsf{e}_{ki+j} \leqslant \frac{kt_n^2}{4}\right)\\ &\overset{n\to\infty}{=} 1-\exp\left(-\left[\frac{n-k}{k}\right]\frac{(kt_n^2/4)^{k-1}}{(k-1)!}(1+o(1))\right) \xrightarrow{n\to\infty} 1.\end{aligned}$$

Now the law of large numbers ensures that

$$\mathbb{P}\left(\frac{\Gamma_{n+1}}{n} \leqslant \frac{1}{2}\right) \to 0, \qquad \text{when} \quad n \to \infty,$$

so putting pieces together we obtain the second part of Theorem 4.1.

Proof of Corollary 4.1. Let us call $T_n^{(3)}$ the time of appearance of the first cluster of size at least 3. The particles j and $j+1$ merge into a single cluster at time $t < T_n^{(3)}$ if and only if

$$X_{n:j+1} - X_{n:j} \leqslant \frac{t^2}{n}.$$

If we write $\delta_{n:i}$ for the ith smallest spacing between the $X_{n:j+1}$ and $X_{n:j}$, we have the equalities $T_{n:1}^2 = n\delta_{n:1}, \ldots, T_{n:k}^2 = n\delta_{n:k}$ on the event $\{T_{n:k} < T_n^{(3)}\}$ or equivalently on the event $\{n\delta_{n:k} < (T_n^{(3)})^2\}$. Recall the identity in law

$$(X_{n:i};\, i=1,\ldots,n) \overset{\text{law}}{\sim} \left(\frac{\mathsf{e}_1+\cdots+\mathsf{e}_i}{\Gamma_{n+1}};\, i=1,\ldots,n\right),$$

where $(\mathsf{e}_i;\, i=1,\ldots,n+1)$ are independent exponential variables of parameter 1 and $\Gamma_{n+1} = \mathsf{e}_1+\cdots+\mathsf{e}_{n+1}$. The spacing $\delta_{n:i}$ has the same law as

$m_n^{[i]}/\Gamma_{n+1}$, where $m_n^{[i]}$ denotes the ith smallest variable e_j. In order to evaluate the asymptotic law of $(n^2\delta_{n:1},..., n^2\delta_{n:k})$, we shall study the asymptotics of $(nm_n^{[1]},..., nm_n^{[k]})$ since $\Gamma_{n+1}/n \to 1$ a.s. The law of $(m_n^{[1]},..., m_n^{[k]})$ is given for $0 < s_1 < \cdots < s_k$ by

$$\begin{aligned}
&\mathbb{P}(m_n^{[1]} \in ds_1,..., m_n^{[k]} \in ds_k)\\
&\quad = \mathbb{P}\Bigg(\bigcup_{i_1 \neq \cdots \neq i_k} \{\mathsf{e}_{i_1} \in ds_1,..., \mathsf{e}_{i_k} \in ds_k\} \bigcap_{i \notin \{i_1,..., i_k\}} \{\mathsf{e}_i > s_k\}\Bigg)\\
&\quad = \sum_{i_1 \neq \cdots \neq i_k} \mathbb{P}(\mathsf{e}_{i_1} \in ds_1) \cdots \mathbb{P}(\mathsf{e}_{i_k} \in ds_k) \prod_{i \notin \{i_1,..., i_k\}} \mathbb{P}(\mathsf{e}_i > s_k)\\
&\quad = \frac{n!}{(n-k)!} e^{-s_1} \cdots e^{-s_k} e^{-(n-k)\, s_k}\, ds_1 \cdots ds_k,
\end{aligned}$$

where the second equality stems from the independence of the exponential variables. We obtain the convergence

$$\begin{aligned}
&\mathbb{P}(nm_n^{[1]} \in ds_1,..., nm_n^{[k]} \in ds_k)\\
&\quad = \frac{n!}{n^k(n-k)!} e^{-s_1/n} \cdots e^{-s_{k-1}/n} e^{-(n-k+1)\, s_k/n}\, ds_1 \cdots ds_k \xrightarrow{n \to \infty} e^{-s_k}\, ds_1 \cdots ds_k,
\end{aligned}$$

which means that

$$(nm_n^{[1]},..., nm_n^{[k]}) \xrightarrow{\text{law}} (\mathsf{e}'_1,..., \mathsf{e}'_1 + \cdots + \mathsf{e}'_k),$$

where $\mathsf{e}'_1,..., \mathsf{e}'_k$ are independent exponential variables of parameter 1. Since Theorem 4.1 ensures that $n(T_n^{(3)})^2 \to \infty$ a.s., and thus that $n\delta_{n:k} \ll (T_n^{(3)})^2$ a.s., we finally obtain the convergence

$$(nT^2_{n:1},..., nT^2_{n:k}) \xrightarrow{\text{law}} (\mathsf{e}'_1,..., \mathsf{e}'_1 + \cdots + \mathsf{e}'_k).$$

Corollary 4.1 follows.

5. SIZE OF THE LARGEST CLUSTER AT A FIXED TIME *t*

We focus now on the system at a given time $t \in \,]0, 1[$. It is of interest to estimate the size $L_n(t)$ of the largest cluster at time t. The following theorem ensures that $L_n(t) \asymp \log n$.

Theorem 5.1. For any $t \in]0, 1[$, there exists two constants $0 < C_t \leqslant C'_t < \infty$ such that

$$\lim_{n \to \infty} \mathbb{P}(C_t \log n \leqslant L_n(t) \leqslant C'_t \log n) = 1.$$

This result implies in particular, that the first macroscopic cluster (i.e., cluster of size $\asymp n$) appears at a time which tends to the critical time 1, when n tends to infinity (see Theorem 3.1). A referee raised the question of describing the evolution of the gas at times closed to the critical time. This interesting problem remains open.

Proof of Theorem 5.1. We give in a first time an upper bound to the size $L_n(t)$ of the largest cluster. According to condition (6) a necessary condition for the existence of a cluster of size k at time t is

$$\exists k \leqslant p \leqslant 2k \text{ and } i \in \{0,..., n-p\} \text{ such that for } r = 1,..., p-1,$$
$$\frac{1}{r} \sum_{j=1}^{r} X_{n:i+j} - \frac{1}{p-r} \sum_{j=r+1}^{p} X_{n:i+j} + \frac{pt^2}{2n} \geqslant 0,$$

which may be written

$$\exists k \leqslant p \leqslant 2k \text{ and } i \in \{0,..., n-p\} \text{ such that for } r = 1,..., p-1,$$
$$\frac{1}{p-r} \sum_{j=r+1}^{p} \left(X_{n:i+j} - X_{n:i} - \frac{j}{n}\right) - \frac{1}{r} \sum_{j=1}^{r} \left(X_{n:i+j} - X_{n:i} - \frac{j}{n}\right) \leqslant p\, \frac{t^2-1}{2n}.$$

We shall deal with the quantity

$$\delta_n(k) = \max_{1 \leqslant j \leqslant k} \max_{0 \leqslant i \leqslant n+1-j} \left| X_{n:i+j} - X_{n:i} - \frac{j}{n} \right|$$

which has been introduced by Mason[5] in order to describe the oscillation modulus of the uniform quantile process v_n. It follows from the inequality

$$\frac{1}{p-r} \sum_{j=r+1}^{p} \left(X_{n:i+j} - X_{n:i} - \frac{j}{n}\right) - \frac{1}{r} \sum_{j=1}^{r} \left(X_{n:i+j} - X_{n:i} - \frac{j}{n}\right) \leqslant 2\delta_n(p),$$

that when the condition

$$2\delta_n(2k) < \frac{k}{n} \left(\frac{1-t^2}{2}\right)$$

holds, no cluster of size larger than k can exist. When $k_n \sim c \log n$ with $c > 0$, Mason has shown (see Theorem 2(II') in ref. 5) that

$$\frac{n\delta_n(k_n)}{k_n} \xrightarrow{n\to\infty} \alpha_c^+ - 1 \quad \text{a.s.},$$

where α_c^+ is the unique solution larger than 1 of $\alpha_c^+ - \log \alpha_c^+ - 1 = 1/c$. Since $\alpha_c^+ \to 1$ when $c \to \infty$, there exists $C'_t < \infty$ such that

$$4(\alpha_{2C'_t}^+ - 1) < \frac{1-t^2}{2},$$

so for n large enough

$$2\delta_n(2k_n) < \frac{k_n}{n}\left(\frac{1-t^2}{2}\right) \quad \text{a.s.},$$

and there exists a.s. no cluster of size larger than $C'_t \log n$.

We now give a lower bound to the size of the largest cluster at time t. Recall from (5) that a sufficient condition for the existence of a cluster of size larger than k at time t is

$$\exists i \in \{0,\ldots, n-k\} \qquad \text{such that} \quad X_{n:i+1} - X_{n:i+k} + \frac{kt^2}{2n} \geqslant 0,$$

which leads us to consider

$$D_n^-(k) := \min_{0\leqslant i\leqslant n-k} (X_{n:i+k} - X_{n:i+1}).$$

A sufficient condition for the existence of a cluster of size at least k in terms of $D_n^-(k)$ is thus

$$D_n^-(k) \leqslant \frac{kt^2}{2n}. \tag{11}$$

The analog to formula (15) in ref. 5 (see also Theorem 2(12) in ref. 8) for $D_n^-(k)$ is when $k_n \sim c \log n$

$$\frac{nD_n^-(k_n)}{k_n} \overset{n\to\infty}{\sim} \alpha_c^- \quad \text{a.s.},$$

where α_c^- is the unique solution less than 1 of $\alpha_c^- - \log \alpha_c^- - 1 = 1/c$. Since $\alpha_c^- \to 0$ when $c \to 0$, there exists a constant C_t such that $\alpha_{C_t}^- < t^2/2$, and for n large enough

$$D_n^-(k_n) \leqslant \frac{k_n t^2}{2n} \quad \text{a.s.}$$

Combining this with formula (11) one obtains for n large enough the a.s. existence of a cluster of size larger than $C_t \log n$, which completes the proof of Theorem 5.1.

6. EVOLUTION OF A MARKED PARTICLE

We should like to estimate the typical size of a cluster. In this direction, we study the size of the cluster which contains a marked particle, say for example the particle number $[n/2]$ ($[\cdot]$ denotes the integer part function). The reason we choose a particle "in the middle" is that we want to avoid the side effects. The following theorem claims that particle $i_n = [n/2]$ does not collide with the others at infinitesimal times and belongs to a finite cluster at any time $t \in]0, 1[$.

Theorem 6.1. Let $S_n(t)$ denotes the size of the cluster which contains particle $i_n = [n/2]$ at time t.

(i) If t_n is a sequence of times decreasing to 0, then

$$\mathbb{P}(S_n(t_n) = 1) \xrightarrow{n \to \infty} 1.$$

(ii) If k_n is sequence of integer increasing to ∞ and $t \in]0, 1[$, then

$$\mathbb{P}(S_n(t) \geqslant k_n) \xrightarrow{n \to \infty} 0.$$

Proof of Theorem 6.1. According to condition (3), a necessary condition for the merging of $(i+1,..., i+k)$ into a cluster of size k before time t is

$$X_{n:i+1} - X_{n:i+k} + \frac{kt^2}{n} \geqslant 0.$$

In particular, a necessary condition for $S_n(t)$ to be larger than k is

$$\exists k \leqslant p \leqslant n, \text{ and } 0 \leqslant j \leqslant p-1, \text{ such that } X_{n:i_n-j} - X_{n:i_n-j+p-1} + \frac{pt^2}{n} \geqslant 0.$$

Recall that

$$(X_{n:i};\, i=1,..., n) \overset{\text{law}}{\sim} \left(\frac{\mathsf{e}_1+\cdots+\mathsf{e}_i}{\Gamma_{n+1}};\, i=1,..., n\right),$$

where $(\mathsf{e}_i;\, i=1,..., n+1)$ are independent random variables with exponential law of parameter 1 and $\Gamma_{n+1}=\mathsf{e}_1+\cdots+\mathsf{e}_{n+1}$. We can give an upper bound to $\mathbb{P}(S_n(t)\geqslant k)$ in terms of e_i: for any $\mu>1$

$$\begin{aligned}\mathbb{P}(S_n(t)\geqslant k) &\leqslant \mathbb{P}\left(\bigcup_{p=k}^{n}\bigcup_{j=0}^{p-1}\left\{X_{n:i_n-j}-X_{n:i_n-j+p-1}+\frac{pt^2}{n}\geqslant 0\right\}\right)\\ &\leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n}\geqslant\mu\right)+\sum_{p=k}^{n} p\,\mathbb{P}(\mathsf{e}_1+\cdots+\mathsf{e}_{p-1}\leqslant \mu p t^2).\end{aligned}$$

The Cramer's large deviation inequality (see ref. 2) yields

$$\mathbb{P}\left(\frac{\mathsf{e}_1+\cdots+\mathsf{e}_p}{p}\leqslant\lambda\right)\leqslant \exp(-p\Lambda^*(\lambda)),$$

with

$$\Lambda^*(\lambda)=\sup_{s\leqslant 0}\,(\lambda s-\Lambda(s))=\lambda-1-\log\lambda.$$

We thus obtain the upper bound for μ such that $\frac{k}{k-1}\,\mu t^2<1$:

$$\mathbb{P}(S_n(t)\geqslant k)\leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n}\geqslant\mu\right)+\sum_{p=k}^{n} p\exp\left(-(p-1)\,\Lambda^*\left(\frac{p}{p-1}\,\mu t^2\right)\right). \tag{12}$$

We first focus on the case $t_n \xrightarrow{n\to\infty} 0$. Under the assumption that $t_n^2<1/4$ formula (12) gives for $k=2$ and $\mu=2$

$$\begin{aligned}\mathbb{P}(S_n(t_n)\geqslant 2) &\leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n}\geqslant 2\right)+\sum_{p=2}^{n} p\exp\left(-(p-1)\,\Lambda^*\left(\frac{p}{p-1}\,2t_n^2\right)\right)\\ &\leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n}\geqslant 2\right)+\sum_{p=2}^{n} \exp\,(\log p-(p-1)\,\Lambda^*(4t_n^2)).\end{aligned}$$

Let us consider the exponential term. Expanding Λ^*, we obtain

$$\begin{aligned}\log p-(p-1)\,\Lambda^*(4t_n^2) &= \log p-(p-1)(4t_n^2-1-2\log(2t_n))\\ &\leqslant \log p+(p-1)\log(2t_n)+(p-1)(1+\log(2t_n))\,.\end{aligned}$$

As soon as $t_n \leqslant 1/4$, the term $\log p + (p-1)\log(2t_n)$ is negative for any $p \geqslant 2$, so we have the upper bound

$$\mathbb{P}(S_n(t_n) \geqslant 2) \leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n} \geqslant 2\right) + \sum_{p=2}^{n} \exp((p-1)(1+\log(2t_n)))$$
$$\leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n} \geqslant 2\right) + \frac{\exp(1+\log(2t_n))}{1-\exp(1+\log(2t_n))}.$$

A consequence of the previous inequality is that $\mathbb{P}(S_n(t_n) \geqslant 2)$ tends to 0 when n tends to infinity. The first part of Theorem 6.1 is proved.

We now focus on the case $t \in \,]0, 1[$ and k_n is an increasing sequence of integer. Formula (12) may be written in this case as

$$\mathbb{P}(S_n(t) \geqslant k_n) \leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n} \geqslant \mu\right) + \sum_{p=k_n}^{n} p \exp\left(-(p-1)\, \Lambda^*\left(\frac{k_n}{k_n-1}\, \mu t^2\right)\right).$$

Since $t < 1$, we can choose $\mu > 1$ such that $\mu t^2 < 1$ and then for n large enough $\frac{k_n}{k_n-1}\, \mu t^2 \leqslant \delta < 1$. Under these assumptions we obtain

$$\mathbb{P}(S_n(t) \geqslant k_n) \leqslant \mathbb{P}\left(\frac{\Gamma_{n+1}}{n} \geqslant \mu\right) + \sum_{p=k_n}^{n} p \exp\left(-(p-1)\, \Lambda^*(\delta)\right),$$

with $\Lambda^*(\delta) > 0$. It follows that $\mathbb{P}(S_n(t) \geqslant k_n)$ tends to 0 when n tends to infinity. This conclude the proof of Theorem 6.1.

7. CONCLUDING REMARKS

We would like to emphasize the difference of behaviour between the system considered in the present paper, and the Gaussian one studied by Martin *et al.*[1,4] The major difference is the scarcity of collisions in our case compared to the Gaussian case, due to the static initial condition of the gas. Computer numerical simulations lead Bonvin and al. to conjecture the existence in the Gaussian case of $\asymp \sqrt{n}$ aggregates of size $\asymp \sqrt{n}$ at fixed time $t < t^*$, whereas we have seen that in our case a typical cluster is of finite size at time $t < t^* = 1$, and has in any case a size bounded by $C'_t \log n$. The proliferation of collisions in the Gaussian case is a consequence of the existence of particles with high initial kinetic energy, which collect quickly many neighboring particles. We must underline at this point that in the case we consider, the scarcity of collisions is a characteristic phenomenon that explains as well the appearance of clusters of size k before those of size

$k+1$, as the somewhat small size of the largest cluster at time $t<1$. It is also to be noticed that the last collision occurs in different ways in the two cases. In the Gaussian one the time of last collision do not converge to the characteristic time t^* (see formula (37) in ref. 4) and the last collision involves a macroscopic together with a microscopic cluster (see ref. 4, Section 6 and also ref. 1, Section 3(ii)). This phenomenon results again from the existence of particles with high kinetic energy. Some of them are near the sides and they flee far away from the system at small times. We shall conclude with an interesting comment made by a referee. There is a qualitative gap between the evolution of the system starting at zero temperature[2] and the evolution starting at low temperature. Indeed, as soon as the temperature of the initial particles is not strictly zero, the gas follows the behaviour described in refs. 1 and 4.

ACKNOWLEDGMENTS

I should like to thank Jean Bertoin for drawing my attention to the subject and especially for his guidance, mathematical or not. I do also thank a referee for his stimulating comments.

REFERENCES

1. J. C. Bonvin, Ph. A. Martin, J. Piasecki, and X. Zotos, Statistics of mass aggregation in a self-gravitating one dimensional gas, *J. Statist. Phys.* **91**:177–197 (1998).
2. J. D. Deuschel and D. W. Stroock, *Large Deviations*, Pure Appl. Math., Vol. 137 (Academic Press, 1989).
3. W. E, Y. G. Rykov, and Ya. G. Sinai, Generalized variational principles, global weak solutions and behaviour with random initial data for systems of conservation laws arising in adhesion particle dynamics, *Com. Math. Phys.* **177**:349–380 (1996).
4. P. A. Martin and J. Piasecki, Aggregation dynamics in a self-gravitating one-dimensional gas, *J. Statist. Phys.* **84**:837–857 (1996).
5. D. M. Mason, A strong limit theorem for the oscillation modulus of the empirical quantile process, *Stoch. Process. Appl.* **17**:127–136 (1984).
6. D. M. Mason, G. R. Shorack, and J. A. Wellner, Strong limit theorems for oscillation moduli of the uniform empirical process, *Z. Wahrschein. verw. Gebiete* **65**:127–136 (1983).
7. J. S. Rao and M. Sobel, Incomplete dirichlet integrals with applications to ordered uniform spacing, *J. Multivariate Anal.* **10**:603–610 (1980).
8. G. R. Shorack and J. A. Wellner, *Empirical Processes with Applications to Statistics* (Wiley, 1986).

[2] The zero temperature relies to the static initial state of the gas, which gets of course heated as time runs by the conversion of potential energy into kinetic energy. It will then cool down again by dissipative collisions.

9. T. Suidan, Adhesion dynamics on the line: The mass aggregation process, *J. Statist. Phys.* **101**:893–903 (2000).
10. M. Vergassola, B. Dubrulle, U. Frisch, and A. Noullez, Burgers' equation, devil's staircases and the mass distribution function for large scale structures, *Astron. Astrophys.* **289**:325–356 (1994).
11. Ya. B. Zeldovich, Gravitational instability: An approximate theory for large density pertubations, *Astron. Astrophys.* **5**:84–89 (1970).