HAL
open science

One-sided convergence in the Boltzmann-Grad limit

Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, Sergio Simonella

To cite this version:

Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, Sergio Simonella. One-sided convergence in the Boltzmann-Grad limit. Annales de la Faculté des Sciences de Toulouse. Mathématiques., 2018, 27 (5), 10.5802/afst.1589 . hal-01413299

HAL Id: hal-01413299

https://hal.science/hal-01413299v1

Submitted on 9 Dec 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ONE-SIDED CONVERGENCE IN THE BOLTZMANN-GRAD LIMIT

THIERRY BODINEAU, ISABELLE GALLAGHER,
LAURE SAINT-RAYMOND AND SERGIO SIMONELLA

Abstract. We review various contributions on the fundamental work of Lanford [20] deriving the Boltzmann equation from hard-sphere dynamics in the low density limit.

We focus especially on the assumptions made on the initial data and on how they encode irreversibility. The impossibility to reverse time in the Boltzmann equation (expressed for instance by Boltzmann's H-theorem) is related to the lack of convergence of higher order marginals on some singular sets. Explicit counterexamples single out the microscopic sets where the initial data should converge in order to produce the Boltzmann dynamics.

1. Introduction

1.1. **Goals.** The Boltzmann equation was introduced at the end of the nineteenth century to predict the statistical behavior of a perfect gas out of thermodynamic equilibrium. This equation expresses the transport and collisions of microscopic particles (atoms) which are supposed to interact typically as elastic hard spheres.

However the resulting dynamics exhibits very different features compared to the reversible deterministic system of hard spheres, which is a Hamiltonian system. The Boltzmann equation generates indeed a semi-group with a Lyapunov functional (the entropy increases along the evolution), and an attractor as time goes to infinity (the density converges to thermodynamic equilibrium). These discrepancies between the microscopic and the macroscopic descriptions were the starting point of some very violent controversy opposing for instance Boltzmann to Loschmidt [11, 8, 9, 24]. There is still an important challenge in understanding the origin of the **non-reversible** Boltzmann equation and the conditions under which it can provide a good approximation of the microscopic dynamics. We refer to [22, 23] for a review on the irreversibility and on the key role played by entropy and to [33] for a modern perspective on Loschmidt's argument. In this paper, we will focus on a more quantitative analysis of the mathematical aspects leading to the emergence of irreversibility.

The convergence result describing at best up to now this transition is due to Lanford [20]. It states that the Boltzmann equation can be obtained as the limit of the deterministic dynamics in a box of size 1

- in the low density regime, i.e. as the number of particles $N \to \infty$, their size $\varepsilon \to 0$, with the additional condition that the inverse mean free path $N\varepsilon^{d-1}$ remains of order 1 (where d is the space dimension);
- up to excluding some pathological situations which occur with vanishing probability in this limit;
- provided that initially the particles are distributed independently.

One important restriction is that this convergence result holds only for short times, which is not enough for observing any relaxation towards equilibrium. Despite many efforts, this restriction has not been removed to this day. There is no attempt in the present paper to improve the convergence time. Our goal here is to study the appearance of irreversibility which already occurs for short times.

More precisely, we intend to discuss in detail the assumptions on the initial data in Lanford's theorem, as they encode all the information on the future evolution. The statement is the following.

Theorem 1.1 ([20]). *Consider a system of N hard-spheres of diameter ε on the d-dimensional periodic box $\mathbb{T}^d = [0,1]^d$ (with $d \geq 2$), initially "independent" and identically distributed with continuous density f_0 such that*

$$\left\| f_0 \exp(\mu + \frac{\beta}{2}|v|^2) \right\|_{L^\infty(\mathbb{T}^d_x \times \mathbb{R}^d_v)} \leq 1 \,, \tag{1.1}$$

for some $\beta > 0, \mu \in \mathbb{R}$. For instance, we can choose the initial distribution of N particles with minimal correlations, due only to the non overlapping conditions :

$$f_{N,0}(x_1, v_1, \ldots, x_N, v_N) = \frac{1}{\mathcal{Z}_N} \prod_{i=1}^N f_0(x_i, v_i) \prod_{i \neq j} \mathbf{1}_{|x_i - x_j| > \varepsilon} \,, \tag{1.2}$$

denoting by $\mathcal{Z}_N$ the partition function, that is the normalizing constant for $f_{N,0}$ to be a probability.

In the Boltzmann-Grad limit $N \to \infty$ with $N\varepsilon^{d-1} = 1$, the one particle distribution $f_N^{(1)} = f_N^{(1)}(t,x,v)$ converges almost everywhere to the solution of the Boltzmann equation

$$\begin{cases} \partial_t f + v \cdot \nabla_x f = Q(f,f) \,, \\ Q(f,f)(v) := \displaystyle\iint_{\mathbb{S}^{d-1} \times \mathbb{R}^d} [f(v')f(v_1') - f(v)f(v_1)] \left((v_1 - v) \cdot \nu\right)_+ dv_1 d\nu \,, \\ v' = v + \nu \cdot (v_1 - v)\, \nu \,, \quad v_1' = v_1 - \nu \cdot (v_1 - v)\, \nu \,, \end{cases} \tag{1.3}$$

with initial data f_0, on a time interval $[0, t^*]$ where t^* depends only on the parameters β, μ of (1.1).

Extensions of this result to different interaction potentials have been recently achieved in [15, 26].

As asserted by Boltzmann himself, the absence of contradiction between reversible microscopic (Newton) equations and the non-reversible Boltzmann equation is due to the fact that only particular, "typical" solutions to the former equation are well approximated by f. The way to give a precise meaning to this typicality is to introduce a statistical description of the initial state, which is in fact the point of view of Theorem 1.1 [21, 32].

The goal of the present paper is to analyze in detail the proof of Lanford's theorem in order to point out where irreversibility shows up. We shall see that part of the information is lost in the convergence process as some pathological sets of configurations with vanishing measure are neglected. These sets turn out to be not time-reversal invariant and the possibility to retrace one's steps fades in the limit.

Furthermore note that, in Theorem 1.1, the weak notion of convergence at time t prevents us from iterating the result as written. Describing more precisely the geometry of the microscopic sets, we shall use a notion of **one-sided convergence** holding at positive times as well as at time zero. Thus we will obtain a refined statement of the theorem (Theorem 2.1) compatible both with the **irreversibility** and the **time-concatenation** (semigroup) properties of the limiting equation (Sect. 3). A similar notion of one-sided convergence has been introduced by Denlinger in [14], see also [18] for a first, non quantitative version.

In order to characterize precisely the (small) sets where the convergence of the initial data is essential, we shall finally construct explicit examples of measures which are badly behaved exclusively in those regions, leading to a violation of Theorem 1.1 (Sect. 4).

1.2. **Microscopic dynamics.** In the following we denote, for $1 \leq i \leq N$, $z_i := (x_i, v_i)$ and $Z_N := (z_1, \dots, z_N)$. With a slight abuse we say that Z_N belongs to $\mathbb{T}^{dN} \times \mathbb{R}^{dN}$ if $X_N := (x_1, \dots, x_N)$ belongs to $\mathbb{T}^{dN}$ and $V_N := (v_1, \dots, v_N)$ to $\mathbb{R}^{dN}$. The phase space is denoted by

$$\mathcal{D}_\varepsilon^N := \left\{ Z_N \in \mathbb{T}^{dN} \times \mathbb{R}^{dN} \,/\, \forall i \neq j \,, \quad |x_i - x_j| > \varepsilon \right\},$$

where $|\cdot|$ stands for the distance on the torus. We now distinguish pre-collisional configurations from post-collisional ones by defining for indexes $1 \leq i \neq j \leq N$

$$\partial\mathcal{D}_\varepsilon^{N\pm}(i,j) := \Big\{ Z_N \in \mathbb{T}^{dN} \times \mathbb{R}^{dN} \,/\, |x_i - x_j| = \varepsilon \,, \quad \pm(v_i - v_j)\cdot(x_i - x_j) > 0 \\ \text{and}\, \forall (k,\ell) \in [1,N]^2 \setminus \{(i,j)\} \text{ with } k \neq \ell \,, |x_k - x_\ell| > \varepsilon \Big\}.$$

Given a post-collisional configuration Z_N on $\partial\mathcal{D}_\varepsilon^{N+}(i,j)$, we define $Z'_N \in \partial\mathcal{D}_\varepsilon^{N-}(i,j)$ as the (pre-collisional) configuration having the same positions $(x_k)_{1\leq k\leq N}$, the same velocities $(v_k)_{k\neq i,j}$ for non interacting particles, and the following pre-collisional velocities for particles i and j

$$v'_i := v_i - \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j)$$
$$v'_j := v_j + \frac{1}{\varepsilon^2}(v_i - v_j)\cdot(x_i - x_j)(x_i - x_j)\,.$$

Defining the Hamiltonian

$$H_N(V_N) := \frac{1}{2}\sum_{i=1}^N |v_i|^2\,, \tag{1.4}$$

we consider the Liouville equation in the $2Nd$-dimensional phase space $\mathcal{D}_\varepsilon^N$

$$\partial_t f_N + \{H_N, f_N\} = 0\,, \tag{1.5}$$

with specular reflection on the boundary, meaning that if Z_N belongs to $\partial\mathcal{D}_\varepsilon^{N+}(i,j)$ then

$$f_N(t, Z_N) = f_N(t, Z'_N)\,. \tag{1.6}$$

We have denoted $\{\cdot,\cdot\}$ the Poisson bracket defined by

$$\{f, g\} := \nabla_{V_N} f \cdot \nabla_{X_N} g - \nabla_{X_N} f \cdot \nabla_{V_N} g\,.$$

The Liouville equation (1.5) writes therefore

$$\partial_t f_N + V_N \cdot \nabla_{X_N} f_N = 0\,,$$

with initial data given by (1.2) and the condition (1.1).

Remark 1.1. *Note that although the boundary condition* (1.6) *seems to introduce a symmetry between pre-collisional and post-collisional configurations, what has to be prescribed for the system to be well-posed is the density on post-collisional configurations for positive times, and for pre-collisional configurations for negative times, which are the incoming configurations for the transport equation* (1.5)*.*

We recall, as shown in [1] for instance, that the set of initial configurations leading to ill-defined characteristics (due to grazing collisions, clustering of collision times, or collisions involving more than two particles) is of measure zero in $\mathcal{D}_\varepsilon^N$.

1.3. **Propagation of chaos.** We define the marginals on $\mathcal{D}_\varepsilon^n$ (extending by zero outside) by

$$f_N^{(n)}(t, Z_n) := \int f_N(t, Z_N)\, dz_{n+1} \dots dz_N. \tag{1.7}$$

Then one can show formally as in [16, 20] and [13, 15] that the first marginal, which describes the typical evolution of the gas, evolves according to

$$\begin{aligned} (\partial_t + v \cdot \nabla_x) f_N^{(1)}(t, x, v) &= (N-1)\varepsilon^{d-1} \\ &\times \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} \Big(f_N^{(2)}(t, x, v', x + \varepsilon\nu, v_1') - f_N^{(2)}(t, x, v, x - \varepsilon\nu, v_1) \Big) \big((v_1 - v) \cdot \nu \big)_+ d\nu dv_1 \,, \end{aligned} \tag{1.8}$$

with v', v_1' as in (1.3). This equation can be interpreted by saying that a particle at $z = (x, v)$ moves in a straight line until it collides with one of the remaining $N - 1$ particles with velocity v_1. The velocities v', v_1' after the collision are then updated and the source term is determined by the joint distribution $f_N^{(2)}$.

The notion of propagation of chaos (Stoßzahlansatz) lies at the heart of the derivation of Boltzmann's equation (1.3). Heuristically, one would like to write that when two particles at configurations $z = (x, v)$ and $z_1 = (x+\varepsilon\nu, v_1)$ collide then the marginal distribution factorizes

$$\lim_{N\to\infty} \Big| f_N^{(2)}(t, z, z_1) - f_N^{(1)}(t, z) f_N^{(1)}(t, z_1) \Big| = 0 \,. \tag{1.9}$$

This statement of the Stoßzahlansatz is far from a mathematical assertion as $f_N^{(2)}$ is only defined almost surely in $\mathbb{T}^{2d} \times \mathbb{R}^{2d}$ and not on sets of codimension 1. A more standard notion of propagation of chaos is given by the following definition.

Definition 1.2 (Chaos property)**.** *The sequence of measures f_N is said asymptotically chaotic at time t if there exists a measurable $f(t)$ on $\mathbb{T}^d \times \mathbb{R}^d$ such that, almost surely in (z, z_1) in $(\mathbb{T}^d \times \mathbb{R}^d)^2$,*

$$\begin{aligned} &\lim_{N\to\infty} f_N^{(1)}(t, z) = f(t, z) \,, \\ &\lim_{N\to\infty} \Big| f_N^{(2)}(t, z, z_1) - f(t, z) f(t, z_1) \Big| = 0 \,. \end{aligned} \tag{1.10}$$

In (1.10) the coordinates z, z_1 are fixed independently of N and ε (contrary to (1.9)). As a consequence, this notion turns out to be too weak to derive Boltzmann equation from the microscopic evolution.

We shall see in Section 2 that the proof of Theorem 1.1 is not based on proving propagation of chaos but on a more global convergence of all the marginals. One of the goals of this paper is to quantify the refined notion of convergence (see Theorem 2.1) which is strictly needed in Lanford's argument. The propagation of chaos (1.10) can be derived as a byproduct.

2. Lanford's proof

In order to understand how the assumptions on the initial data come into play, we have to look more precisely at the proof of Theorem 1.1. Theorem 1.1 is actually the corollary of a more precise result. Lanford's result indeed provides the convergence of all marginals $f_N^{(n)}$

defined in (1.7) to the solutions $f^{(n)}$ of an infinite system of coupled equations

$$
\begin{aligned}
&\partial_t f^{(n)} + \sum_{i=1}^{n} v_i \cdot \nabla_{x_i} f^{(n)} = C^0_{n,n+1} f^{(n+1)}, \\
&\Big(C^0_{n,n+1} f^{(n+1)}\Big)(x_1, v_1, \dots, x_n, v_n) \\
&\quad := \sum_{i=1}^{n} \iint_{\mathbb{S}^{d-1}\times\mathbb{R}^d} \Big(f^{(n+1)}(x_1, v_1, \dots, x_i, v_i', \dots x_n, v_n, x_i, v_{n+1}') \\
&\qquad\qquad - f^{(n+1)}(x_1, v_1, \dots, x_i, v_i, \dots x_n, v_n, x_i, v_{n+1}) \Big) \big((v_{n+1} - v_i)\cdot \nu\big)_+ \, dv_{n+1} d\nu \,,
\end{aligned}
\tag{2.1}
$$

which is the so-called Boltzmann hierarchy. Chaotic families of the form $f^{(n)} = f^{\otimes n}$ with f solution to the Boltzmann equation are specific solutions to this hierarchy, where

$$
f^{\otimes n}(Z_n) := \prod_{i=1}^{n} f(z_i) \,.
$$

The connection between Boltzmann hierarchy and Boltzmann equation is discussed in [31].

The starting point of the proof is to write an explicit representation of the n particle distribution $f_N^{(n)}$ as a superposition of different $(n+s)$-particle pseudo-dynamics, with weights depending on the initial data. More precisely, by averaging and iterating Duhamel's formula for the N-particle distribution f_N, we end up with a series expansion for $f_N^{(n)}$ in which the term of order s corresponds to pseudo-dynamics involving s collisions and is therefore expressed as an operator acting on the initial $(n+s)$-particle distribution $f_{N,0}^{(n+s)}$ (see Section 2.1).

The strategy of proof then relies on two main steps.

- First we obtain a uniform bound on the series expansion, which explains the short time restriction in Theorem 1.1. In the following, we restrict our attention to times smaller than the radius of analyticity of the series.
- The convergence to the solution of the Boltzmann hierarchy then follows from the convergence of the trajectories representing the different pseudo-dynamics (note that these trajectories are related to the representation formula and that they do not coincide in general with the physical trajectories of the particles, e.g. [27] for further discussions). The convergence of pseudo-trajectories fails to hold when there are recollisions (see page 8 for a precise definition of recollisions). A geometric argument shows however that, for any fixed n, the set of initial configurations with n particles leading to such recollisions is of vanishing measure in the $N \to \infty$ limit.

Note that all the information on these bad sets is forgotten in the limit: this is related to irreversibility, that is to the impossibility of going back to the initial state. Furthermore the convergence of the first marginal to the solution of the Boltzmann equation in the case of factorized initial data such as (1.2) is due to a uniqueness property for the Boltzmann hierarchy; this follows from the uniform bound on the hierarchy obtained in the first step of the above strategy.

2.1. **The series expansions.** A formal computation based on Green's formula (see [10, 20, 15] for instance) leads to the following BBGKY hierarchy for $n < N$

$$
(\partial_t + \sum_{i=1}^{n} v_i \cdot \nabla_{x_i}) f_N^{(n)}(t, Z_n) = \big(C_{n,n+1} f_N^{(n+1)}\big)(t, Z_n) \,,
\tag{2.2}
$$

on $\mathcal{D}_\varepsilon^n$ with the boundary condition as in (1.6)

$$
f_N^{(n)}(t, Z_n) = f_N^{(n)}(t, Z_n') \quad \text{on} \quad \partial\mathcal{D}_\varepsilon^{n+}(i,j) \,.
$$

The collision term is defined by

$$\begin{aligned}
&\big(C_{n,n+1} f_N^{(n+1)}\big)(Z_n) := (N-n)\varepsilon^{d-1} \\
&\times \Big(\sum_{i=1}^n \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_N^{(n+1)}(\dots, x_i, v_i', \dots, x_i+\varepsilon\nu, v_{n+1}')\big((v_{n+1}-v_i)\cdot\nu\big)_+ d\nu dv_{n+1} \\
&\quad - \sum_{i=1}^n \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_N^{(n+1)}(\dots, x_i, v_i, \dots, x_i+\varepsilon\nu, v_{n+1})\big((v_{n+1}-v_i)\cdot\nu\big)_- d\nu dv_{n+1}\Big), \\
&\text{with} \quad v_i' := v_i - (v_i - v_{n+1})\cdot\nu\,\nu\,, \quad v_{n+1}' := v_{n+1} + (v_i - v_{n+1})\cdot\nu\,\nu\,.
\end{aligned} \tag{2.3}$$

The closure for $n = N$ is given by the Liouville equation (1.5). Note that the collision integral is split into two terms according to the sign of $(v_i - v_{n+1})\cdot\nu$ and we used the trace condition on $\partial\mathcal{D}_\varepsilon^{N+}(i, n+1)$ to express all quantities in terms of pre-collisional configurations.

To obtain the Boltzmann hierarchy, we compute the formal limit of the transport and collision operators when ε goes to 0. Recall that for fixed n, then $(N-n)\varepsilon^{d-1} \to 1$ in the Boltzmann-Grad limit. Thus the limit hierarchy is given by

$$(\partial_t + \sum_{i=1}^n v_i\cdot\nabla_{x_i}) f^{(n)}(t, Z_n) = \big(C_{n,n+1}^0 f^{(n+1)}\big)(t, Z_n) \tag{2.4}$$

in $(\mathbb{T}^d\times\mathbb{R}^d)^n$, where $C_{n,n+1}^0$ are the limit collision operators defined by (2.1). We denote by $(f_0^n)_{n\in\mathbb{N}}$ a family of initial data for this hierarchy (which will be specified later).

Iterating Duhamel's formula for the BBGKY hierarchy (2.2), we get

$$f_N^{(n)}(t) = \sum_{s=0}^{N-n} Q_{n,n+s}(t) f_{N,0}^{(n+s)}\,, \tag{2.5}$$

where we have defined

$$\begin{aligned}
Q_{n,n+s}(t) f_{N,0}^{(n+s)} := \int_0^t \int_0^{t_{n+1}} \dots \int_0^{t_{n+s-1}} &\mathbf{S}_n(t-t_{n+1}) C_{n,n+1} \mathbf{S}_{n+1}(t_{n+1}-t_{n+2}) C_{n+1,n+2} \\
&\dots \mathbf{S}_{n+s}(t_{n+s}) f_{N,0}^{(n+s)}\, dt_{n+s}\dots dt_{n+1}
\end{aligned}$$

denoting by $\mathbf{S}_n$ the group associated with free transport in $\mathcal{D}_\varepsilon^n$ with specular reflection on the boundary.

Remark 2.1. *Note that, for fixed N, the operator $C_{n,n+1}$ is a trace on a manifold of codimension 1 and thus it is a priori not defined on L^∞ functions. What makes sense is the combination $\int dt_{n+1} C_{n,n+1}\mathbf{S}_{n+1}(t_{n+1}-t_{n+2})$ (see* [3, 30, 15] *and Figure* 1*).*

For $t \geq 0$, one has $t_{n+1} - t_{n+2} \geq 0$, it is therefore necessary to express the collision operator in terms of pre-collisional configurations. In a symmetric way, for $t \leq 0$, one has $t_{n+1} - t_{n+2} \leq 0$, and we have to express the collision operator in terms of post-collisional configurations (see Remark 1.1*).*

Similarly, for the Boltzmann hierarchy (2.4)

$$f^{(n)}(t) = \sum_{s=0}^\infty Q_{n,n+s}^0(t) f_0^{(n+s)}\,, \tag{2.6}$$

where we have defined

$$\begin{aligned}
Q_{n,n+s}^0(t) f_0^{(n+s)} = \int_0^t \int_0^{t_{n+1}} \dots \int_0^{t_{n+s-1}} &\mathbf{S}_n^0(t-t_{n+1}) C_{n,n+1}^0 \mathbf{S}_{n+1}^0(t_{n+1}-t_{n+2}) C_{n+1,n+2}^0 \\
&\dots \mathbf{S}_{n+s}^0(t_{n+s}) f_0^{(n+s)}\, dt_{n+s}\dots dt_{n+1},
\end{aligned}$$

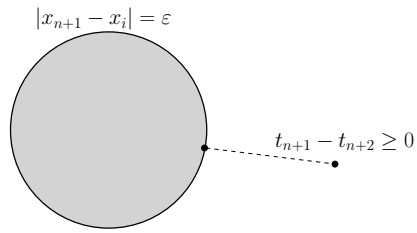

FIGURE 1. The grey domain is an excluded region and its boundary is the surface $|x_{n+1} - x_i| = \varepsilon$ corresponding to a collision between particles i and $n+1$. The admissible configurations are outside this domain and can be parametrised by a point of the surface and a non negative time $t_{n+1} - t_{n+2}$, provided that the velocities are pre-collisional.

denoting by $\mathbf{S}_n^0$ the group associated with free transport in $(\mathbb{T}^d \times \mathbb{R}^d)^n$.

Let us denote $|C_{s,s+1}|$, $|Q_{n,n+s}|$ the operators obtained by summing the absolute values of all the elementary terms. The energy $H_s = \frac{1}{2}\sum_{i=1}^s |v_i|^2$ is conserved by the transport so that

$$\mathbf{S}_s\left(\exp\left(-\beta H_s\right)\mathbf{1}_{\mathcal{D}_\varepsilon^s}\right) = \exp\left(-\beta H_s\right)\mathbf{1}_{\mathcal{D}_\varepsilon^s}\,,$$

and from the loss estimates on the collision operators (see [15] for instance)

$$|C_{s,s+1}|\left(\exp\left(-\beta H_{s+1}\right)\mathbf{1}_{\mathcal{D}_\varepsilon^{s+1}}\right) \leq C\beta^{-d/2}\Big(s\beta^{-\frac{1}{2}} + \sum_{1\leq i\leq s}|v_i|\Big)\exp\left(-\beta H_s\right)\mathbf{1}_{\mathcal{D}_\varepsilon^s}\,,$$

we get the Cauchy-Kowalevsky type iterated estimate for $\tilde\beta < \beta$

$$|Q_{n,n+s}|(t)\left(\exp\left(-\beta H_{n+s}\right)\mathbf{1}_{\mathcal{D}_\varepsilon^{n+s}}\right) \leq C^{n+s}C^s_{\beta,\tilde\beta}\exp\left(-\tilde\beta H_n\right)\,, \tag{2.7}$$

with $C_{\beta,\tilde\beta} = \beta^{-(d+1)/2}t/(\beta-\tilde\beta)$.

Using the initial data (1.2) and the condition (1.1), we deduce following [35] an upper bound on the marginals from (2.5)

$$\forall t \leq t^*, \qquad f_N^{(n)}(t) \leq \exp((\lambda t - \mu)n)\exp\Big((\lambda t - \beta)H_n\Big)\,, \tag{2.8}$$

where λ, chosen large enough, depends on β, μ, and t^* is such that $\lambda t^* = \beta/2$. The convergence time in Lanford's Theorem 1.1 is given by t^*.

Similar estimates hold for the limit operators $Q^0_{n,n+s}$ and $\mathbf{S}^0_s$, as well as for the solution of the Boltzmann hierarchy.

2.2. **Geometrical representation as a superposition of pseudo-dynamics.** The usual way to study the s-th term of the representation formula is to introduce some pseudo-dynamics describing the action of the operator $Q_{n,n+s}$. We first extract combinatorial information on the collision process: we describe the adjunction of new particles (in the backward dynamics) by ordered trees.

Definition 2.2 (Collision trees)**.** *Let $n \geq 1$, $s \geq 1$ be fixed. An (ordered) collision tree $a \in \mathcal{A}_{n,n+s}$ is defined by a family $(a(i))_{n+1\leq i\leq n+s}$ with $a(i) \in \{1,\dots,i-1\}$.*

Note that $|\mathcal{A}_{n,n+s}| \leq n(n+1)\dots(n+s-1)$.

Once we have fixed a collision tree $a \in \mathcal{A}_{n,n+s}$, we can reconstruct pseudo-dynamics starting from any point in the n-particle phase space $Z_n = (x_i, v_i)_{1\leq i\leq n}$ at time t.

Definition 2.3 (Pseudo-trajectory)**.** *Given $Z_n \in \mathcal{D}_\varepsilon^n$, consider a collection of times, angles and velocities $(T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}) = (t_i, \nu_i, v_i)_{n+1\leq i\leq n+s}$ with $0 \leq t_{n+s} \leq \dots \leq$*

$t_{n+1} \leq t$. We then define recursively the pseudo-trajectories in terms of the backward BBGKY dynamics as follows

- *in between the collision times t_i and t_{i+1} the particles follow the i-particle backward flow with specular reflection;*
- *at time t_i^+, particle i is adjoined to particle $a(i)$ at position $x_{a(i)} + \varepsilon\nu_i$ provided it remains at a distance ε from all the others, and with velocity v_i. If $(v_i - v_{a(i)}(t_i^+))\cdot\nu_i > 0$, velocities at time t_i^- are given by the scattering laws*

$$\begin{aligned} v_{a(i)}(t_i^-) &= v_{a(i)}(t_i^+) - (v_{a(i)}(t_i^+) - v_i)\cdot\nu_i\,\nu_i\,, \\ v_i(t_i^-) &= v_i + (v_{a(i)}(t_i^+) - v_i)\cdot\nu_i\,\nu_i\,. \end{aligned} \tag{2.9}$$

We denote by $z_i(a, Z_n, T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}, \tau)$ the position and velocity of the particle labeled i, at time τ (provided $\tau < t_i$). We also define $G_{n+1,n+s}(a)$ as the set of parameters $(T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s})$ such that the pseudo-trajectory $Z_{n+s}(a, Z_n, T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}, \tau)$ exists up to time 0, meaning that by adjunction of a new particle, there is no overlap. The configuration obtained at the end of the tree, i.e. at time 0, is $Z_{n+s}(a, Z_n, T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}, 0)$.

Similarly, we define the pseudo-trajectories associated with the Boltzmann hierarchy. These pseudo-trajectories evolve according to the backward Boltzmann dynamics as follows

- *in between the collision times t_i and t_{i+1} the particles follow the i-particle backward free flow;*
- *at time t_i^+, particle i is adjoined to particle $a(i)$ at exactly the same position $x_{a(i)}$. Velocities are given by the laws (2.9).*

We denote this flow by $Z^0_{n+s}(a, Z_n, T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}, \tau)$.

With these notations, the representation formulas (2.5) and (2.6) for the marginals of order n can be rewritten respectively

$$\begin{aligned} f_N^{(n)}(t, Z_n) = \sum_{s=0}^{N-n} C^{(N,n,s)} \sum_{a\in\mathcal{A}_{n,n+s}} \int_{G_{n+1,n+s}(a)} dT_{n+1,n+s} d\Omega_{n+1,n+s} dV_{n+1,n+s} \\ \Big(\prod_{i=n+1}^{n+s} \big((v_i - v_{a(i)}(t_i))\cdot\nu_i\big) f_{N,0}^{(n+s)}\big(Z_{n+s}(a, Z_n, T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}, 0)\big)\,, \end{aligned}$$

with $C^{(N,n,s)} := (N-n)\dots\big(N-(n+s-1)\big)\varepsilon^{(d-1)s} = 1 + O((n+s)^2/N)$ and

$$\begin{aligned} f^{(n)}(t, Z_n) = \sum_{s=0}^{\infty} \sum_{a\in\mathcal{A}_{n,n+s}} \int_{\mathcal{T}_{n+1,n+s}} dT_{n+1,n+s} \int_{\mathbb{S}^s} d\Omega_{n+1,n+s} \int_{\mathbb{R}^{2s}} dV_{n+1,n+s} \\ \times \Big(\prod_{i=n+1}^{n+s} \big((v_i - v^0_{a(i)}(t_i))\cdot\nu_i\big) f_0^{(n+s)}\big(Z^0_{n+s}(a, Z_n, T_{n+1,n+s}, \Omega_{n+1,n+s}, V_{n+1,n+s}, 0)\big)\,, \end{aligned}$$

denoting $\mathcal{T}_{n+1,n} = \emptyset$ and

$$\mathcal{T}_{n+1,n+s} := \big\{(t_i)_{n+1\leq i\leq n+s} \in [0,t]^s \,/\, 0 \leq t_{n+s} \leq \dots \leq t_{s+1} \leq t\big\}\,.$$

The question is then to describe the asymptotic behavior of the BBGKY pseudo-trajectories. We actually split them into two classes :

- pseudo-trajectories having no recollision, i.e. such that particles interact only at the times of adjunction of new particles, and are transported freely between two such times;
- pseudo-trajectories involving recollisions.

Note that no recollision occurs in the Boltzmann hierarchy as the particles have zero diameter.

2.3. **Bad configurations.** The transport semigroups $\mathbf{S}_n$ (with recollisions) and $\mathbf{S}_n^0$ (free transport) play a key role in the discrepancies between the BBGKY series (2.5) and the Boltzmann series (2.6). In a given time interval, both transports coincide if no recollision occurs which will be the typical case for fixed n and ε small. However, specific configurations lead to recollisions and we define below the corresponding geometric sets.

Denote by B_R the ball of $\mathbb{R}^d$ centered at zero and of radius R, and fix a time T much bigger than the radius of analyticity t^* given in (2.8) as well as a parameter $\varepsilon_0 \gg \varepsilon$. The sets of bad configurations of n particles are defined as

$$\begin{aligned} \mathcal{B}_{\varepsilon_0}^{n-} &:= \Big\{ Z_n \in (\mathbb{T}^d \times B_R)^n, \qquad \exists u \in [0,T]\,, \exists i,j, \qquad |x_i - x_j - u(v_i - v_j)| \le \varepsilon_0 \Big\}\,, \\ \mathcal{B}_{\varepsilon_0}^{n+} &:= \Big\{ Z_n \in (\mathbb{T}^d \times B_R)^n, \qquad \exists u \in [0,T]\,, \exists i,j, \qquad |x_i - x_j + u(v_i - v_j)| \le \varepsilon_0 \Big\}\,, \end{aligned} \tag{2.10}$$

where $|\cdot|$ stands for the distance on the torus. This means that, starting from $\mathcal{B}_\varepsilon^{n-}$ at time t, the backward free flow on $\mathcal{D}_\varepsilon^n$ will involve at least one recollision between t and $t-T$, and starting from $\mathcal{B}_\varepsilon^{n+}$, the forward free flow on $\mathcal{D}_\varepsilon^n$ will involve at least one recollision between t and $t+T$. In particular outside these sets, we have

$$\begin{aligned} \big(\mathbf{S}_n(t) - \mathbf{S}_n^0(t)\big)\Big(f_{N,0}^{(n)}(1 - \mathbf{1}_{\mathcal{B}_\varepsilon^{n+}})\Big) &= 0 \quad \text{for } t \in [0,T]\,, \\ \big(\mathbf{S}_n(t) - \mathbf{S}_n^0(t)\big)\Big(f_{N,0}^{(n)}(1 - \mathbf{1}_{\mathcal{B}_\varepsilon^{n-}})\Big) &= 0 \quad \text{for } t \in [-T,0]\,. \end{aligned}$$

The first term in each series (2.5) and (2.6) involves the transport, both first terms coincide when $\pm t > 0$ for configurations which are outside the bad set $\mathcal{B}_\varepsilon^{n\pm}$. We stress the fact that similar sets have been already introduced by Denlinger in [14] and previously identified in [4] as key sets (see Appendix A of [4] for a discussion on the irreversibility).

The following result is an easy calculation.

Proposition 2.4. *The bad sets are ordered*

$$\mathcal{B}_\varepsilon^{n\pm} \subset \mathcal{B}_{\varepsilon'}^{n\pm} \quad \forall \varepsilon' \ge \varepsilon\,.$$

Their measure is controlled by

$$|\mathcal{B}_\varepsilon^{n\pm}| \le (CR^d)^n n^2 RT \varepsilon^{d-1}\,,$$

and the intersection is much smaller

$$|\mathcal{B}_\varepsilon^{n+} \cap \mathcal{B}_\varepsilon^{n-}| \le (CR^d)^n \big(n^2\varepsilon^d + n^4 R^2 T^2 \varepsilon^{2(d-1)}\big)\,.$$

We now suppose that $t \ge 0$ since the situation when $t \le 0$ can be deduced by a simple symmetry in t and v. The next terms in the series expansion (2.5), (2.6) involve some averaging with respect to the parameters $(t_i, v_i, \nu_i)_{n+1 \le i \le n+s}$ describing the adjunction of new particles. What can be proved is that, provided that the n-particle backward flow $\mathbf{\Psi}_n$ on $\mathcal{D}_\varepsilon^n$ does not lead to a recollision, then the probability of having a recollision (involving at least one of the added particles) is very small.

2.4. **Convergence outside bad configurations.** Let us first prove that the solutions are close by eliminating bad trajectories. By definition, the set of good configurations with k particles will be such that the particles remain, by backward free flow, at a distance $\varepsilon_0 \gg \varepsilon|\log\varepsilon|$ for a time $T \gg t^*$, i.e. that they belong to the set

$$\mathcal{G}_k(\varepsilon_0) := (\mathbb{T}^d \times B_R)^k \setminus \mathcal{B}_{\varepsilon_0}^{k-}\,.$$

For particles in $\mathcal{G}_k(\varepsilon_0)$, the transport $\mathbf{\Psi}_k$ on $\mathcal{D}_\varepsilon^k$ coincides with the free flow $\mathbf{\Psi}_k^0$ on $(\mathbb{T}^d \times B_R)^k$. Thus, if at time t the configurations Z_k, Z_k^0 are such that

$$\forall i \le k\,, \qquad |x_i - x_i^0| \le \varepsilon|\log\varepsilon|\,, \qquad v_i = v_i^0 \tag{2.11}$$

and Z_k^0 belongs to $\mathcal{G}_k(\varepsilon_0)$, then the configurations $\mathbf{\Psi}_k(u)Z_k$, $\mathbf{\Psi}_k^0(u)Z_k^0$ will remain at distance less than $O(\varepsilon|\log\varepsilon|)$ for $u\in[0,t]$. Recall that the distance $|\cdot|$ is on the torus.

One can show that good configurations are stable by adjunction of a $(k+1)^{\text{th}}$-particle next to a particle labelled by $m_k\le k$, provided some bad sets are removed. More precisely, let $Z_k^0=(X_k^0,V_k)$ be in $\mathcal{G}_k(\varepsilon_0)$ and $Z_k=(X_k,V_k)$ with positions close to X_k^0 and same velocities (cf. (2.11)). Then, by choosing the velocity v_{k+1} and the deflection angle ν_{k+1} of the new particle $k+1$ outside a bad set $B_{m_k}(Z_k^0)$, both configurations Z_k and Z_k^0 will remain close to each other. Of course, immediately after the adjunction, the particles m_k and $k+1$ will not be at distance ε_0, but v_{k+1},ν_{k+1} can be chosen such that the particles drift rapidly far apart and after a short time $\delta>0$ the configurations Z_{k+1} and Z_{k+1}^0 are again in the good sets $\mathcal{G}_{k+1}(\varepsilon_0/2)$ and $\mathcal{G}_{k+1}(\varepsilon_0)$.

Proposition 2.5 ([15])**.** *We fix parameters $\varepsilon\ll\varepsilon_0,\delta\ll 1$ such that*

$$|\log\varepsilon|\varepsilon\ll\varepsilon_0\ll\min(\delta R,1)\,. \tag{2.12}$$

Given $Z_k^0=(X_k^0,V_k)\in\mathcal{G}_k(\varepsilon_0)$ and $m_k\le k$, there is a subset $B_{m_k}(Z_k^0)$ of $\mathbb{S}^{d-1}\times B_R$ of small measure

$$\big|B_{m_k}(Z_k^0)\big|\le CkR^d\gamma(\varepsilon,\varepsilon_0)\quad\textit{with}\quad\gamma(\varepsilon,\varepsilon_0):=|\log\varepsilon|\left(\left(\frac{\varepsilon}{\varepsilon_0}\right)^{d-1}+(RT)^d\varepsilon_0^{d-1}+\left(\frac{\varepsilon_0}{R\delta}\right)^{d-1}\right), \tag{2.13}$$

such that good configurations close to Z_k^0 are stable by adjunction of a collisional particle close to the particle $x_{m_k}^0$ in the following sense. Let $Z_k=(X_k,V_k)$ be a configuration of k particles satisfying (2.11), i.e. $|X_k-X_k^0|\le|\log\varepsilon|\varepsilon$. Given $(\nu_{k+1},v_{k+1})\in(\mathbb{S}^{d-1}\times B_R)\setminus B_{m_k}(Z_k^0)$, a new particle with velocity v_{k+1} is added at $x_{m_k}+\varepsilon\nu_{k+1}$ to Z_k and at $x_{m_k}^0$ to Z_k^0. Two possibilities may arise

- *For a pre-collisional configuration $\nu_{k+1}\cdot(v_{k+1}-v_{m_k})<0$ then*

$$\forall u\in]0,t]\,,\quad\begin{cases}\forall i\ne j\in[1,k]\,,\quad|(x_i-u\,v_i)-(x_j-u\,v_j)|>\varepsilon\,,\\ \forall j\in[1,k]\,,\quad|(x_{m_k}+\varepsilon\nu_{k+1}-u\,v_{k+1})-(x_j-u\,v_j)|>\varepsilon\,.\end{cases} \tag{2.14}$$

Moreover after the time δ, the $k+1$ particles are in a good configuration

$$\forall u\in[\delta,t]\,,\quad\begin{cases}(X_k-uV_k,V_k,x_{m_k}+\varepsilon\nu_{k+1}-u\,v_{k+1},v_{k+1})\in\mathcal{G}_{k+1}(\varepsilon_0/2),\\ (X_k^0-uV_k,V_k,x_{m_k}^0-u\,v_{k+1},v_{k+1})\in\mathcal{G}_{k+1}(\varepsilon_0)\,.\end{cases} \tag{2.15}$$

- *For a post-collisional configuration $\nu_{k+1}\cdot(v_{k+1}-v_{m_k})>0$ then the velocities are updated*

$$\forall u\in]0,t]\,,\quad\begin{cases}\forall i\ne j\in[1,k]\setminus\{m_k\}\,,\quad|(x_i-u\,v_i)-(x_j-u\,v_j)|>\varepsilon\,,\\ \forall j\in[1,k]\setminus\{m_k\}\,,\quad|(x_{m_k}+\varepsilon\nu_{k+1}-u\,v'_{k+1})-(x_j-u\,v_j)|>\varepsilon\,,\\ \forall j\in[1,k]\setminus\{m_k\}\,,\quad|(x_{m_k}-u\,v'_{m_k})-(x_j-u\,v_j)|>\varepsilon\,,\\ |(x_{m_k}-u\,v'_{m_k})-(x_{m_k}+\varepsilon\nu_{k+1}-u\,v'_{k+1})|>\varepsilon\,.\end{cases} \tag{2.16}$$

Moreover after the time δ, the $k+1$ particles are in a good configuration

$$\forall u\in[\delta,t],\quad\begin{cases}\big(\{x_j-u\,v_j,v_j\}_{j\ne m_k},x_{m_k}-u\,v'_{m_k},v'_{m_k},x_{m_k}+\varepsilon\nu_{k+1}-u\,v'_{k+1},v'_{k+1}\big)\in\mathcal{G}_{k+1}(\varepsilon_0/2),\\ \big(\{x_j^0-u\,v_j,v_j\}_{j\ne m_k},x_{m_k}^0-u\,v'_{m_k},v'_{m_k},x_{m_k}^0-u\,v'_{k+1},v'_{k+1}\big)\in\mathcal{G}_{k+1}(\varepsilon_0)\,.\end{cases} \tag{2.17}$$

We refer to [15] for a complete proof of Proposition 2.5 and simply recall that it can be obtained from the following control on free trajectories (note that compared to [15] there is an additional loss of a $|\log\varepsilon|$ which is due to the action of the scattering operator and is actually missing in [15]).

Lemma 2.6. *Given $T > 0$, $\varepsilon \ll \delta \ll 1$ and $\varepsilon|\log\varepsilon| \ll \varepsilon_0 \ll \min(\delta R, 1)$, consider two points x_1^0, x_2^0 in $\mathbb{T}^d$ such that $|x_1^0 - x_2^0| \geq \varepsilon_0$, and a velocity $v_1 \in B_R$. Then there exists a subset $K(x_1^0 - x_2^0, \varepsilon_0, \varepsilon)$ of $\mathbb{R}^d$ with measure bounded by*

$$|K(x_1^0 - x_2^0, \varepsilon_0, \varepsilon)| \leq CR^d|\log\varepsilon| \left(\left(\frac{\varepsilon}{\varepsilon_0}\right)^{d-1} + (Rt)^d \varepsilon^{d-1} \right)$$

and a subset $K_\delta(x_1^0 - x_2^0, \varepsilon_0, \varepsilon)$ of $\mathbb{R}^d$, the measure of which satisfies

$$|K_\delta(x_1^0 - x_2^0, \varepsilon_0, \varepsilon)| \leq CR^d|\log\varepsilon| \left(\left(\frac{\varepsilon_0}{R\delta}\right)^{d-1} + (Rt)^d \varepsilon_0^{d-1} \right)$$

such that for any $v_2 \in B_R$ and x_1, x_2 such that $|x_1 - x_1^0| \leq |\log\varepsilon|\varepsilon$, $|x_2 - x_2^0| \leq |\log\varepsilon|\varepsilon$, the following results hold :

- *If $v_1 - v_2 \notin K(x_1^0 - x_2^0, \varepsilon_0, \varepsilon)$, then*
$$\forall u \in [0, t]\,, \quad |(x_1 - u\, v_1) - (x_2 - u\, v_2)| > \varepsilon.$$
- *If $v_1 - v_2 \notin K_\delta(x_1^0 - x_2^0, \varepsilon_0, \varepsilon)$*
$$\forall u \in [\delta, t]\,, \quad |(x_1 - u\, v_1) - (x_2 - u\, v_2)| > \varepsilon_0\,.$$

Proposition 2.5 is the elementary step for adding a new particle. This step can be iterated in order to build inductively good pseudo-trajectories Z and Z^0. Note that after adding a new particle, velocities remain identical at each time in both configurations, but their positions differ due the exclusion condition in the BBGKY hierarchy which induces a shift of ε at each creation of a new particle.

To estimate $Q_{n,n+s}(t)f_{N,0}^{(n+s)} - Q_{n,n+s}^0(t)f_0^{(n+s)}$, we then split the integration domain in several pieces

- pseudo-trajectories with large energy $H_{n+s}(Z_{n+s}) \geq R^2 \gg 1$;
- pseudo-trajectories with collisions separated by less than a time $\delta \ll 1$;
- pseudo-trajectories (with moderate energy and collisions well separated in time) having recollisions;
- good pseudo-trajectories in the sense of Proposition 2.5.

Bad pseudo-trajectories have a small contribution to the integrals thanks to (2.13) while good pseudo-trajectories of the BBGKY and Boltzmann hierarchies can be coupled.

2.5. **Convergence of initial data.** To estimate the contribution of good pseudo-trajectories, we have then to combine the continuity of $f_0^{(n+s)}$ together with an estimate on the difference $f_{N,0}^{(n+s)} - f_0^{(n+s)}$ between initial data on the set of initial configurations which may be reached by such pseudo-dynamics: since we only consider pseudo-trajectories leading to good configurations, what we need to compute is $(f_{N,0}^{(s)} - f_0^{\otimes s})(1 - \mathbf{1}_{\mathcal{B}_\varepsilon^{s+} \cap \mathcal{B}_\varepsilon^{s-}})$.

With the specific choice of initial data (1.2) in Theorem 1.1, one can prove (see [15] for instance) that the initial data of both hierarchies are close, in the sense that for $s \geq 2$

$$\left|(f_{N,0}^{(s)} - f_0^{\otimes s})(1 - \mathbf{1}_{\mathcal{B}_\varepsilon^{s+} \cap \mathcal{B}_\varepsilon^{s-}})\right| \leq C^s \exp(-\beta H_s)\varepsilon\,. \tag{2.18}$$

This condition implies that $f_{N,0}^{(s)}$ is almost chaotic on the singular sets $\mathcal{B}_\varepsilon^{s+} \setminus \mathcal{B}_\varepsilon^{s-}$ (which are relevant for the forward equation) and $\mathcal{B}_\varepsilon^{s-} \setminus \mathcal{B}_\varepsilon^{s+}$ (which are relevant for the backward equation). Note that, compared to Definition 1.2, this is much stronger as it provides a quantitative description of the factorization on sets depending on ε.

It remains to gather all error estimates and to use the continuity property (2.7) for the operators $Q_{n,s+n}$. We define the weighted norm

$$\|f_n\|_{L^\infty_{\beta,n}} := \|f_n \exp(\beta H_n)\|_{L^\infty}\,,$$

with H_n the Hamiltonian (1.4). Fixing the parameters $\varepsilon_0, \delta, s, n$ such that

$$|\log\varepsilon|\varepsilon \ll \varepsilon_0 \ll \min(\delta R, 1)\,, \quad n+s \le |\log\varepsilon|\,,$$

and choosing $R \le C|\log\varepsilon|$, the error term $\gamma(\varepsilon,\varepsilon_0)$ from Proposition 2.5 converges to 0. The term by term convergence is then obtained from the following estimate, thanks to the previous analysis and (2.18).

Proposition 2.7. *There is $\gamma > 0$ such that*

$$\begin{aligned}
&\Big\|\Big(Q_{n,n+s}(t) f^{(n+s)}_{N,0} - Q^0_{n,n+s}(t) f^{(n+s)}_0\Big)(1-\mathbf{1}_{\mathcal{B}^{n-}_{\varepsilon_0}})\Big\|_{L^\infty_{\beta',n}} \\
&\quad \le C^{n+s}\left(\frac{t\beta^{-(d+1)/2}}{\beta-\beta'}\right)^s \Big[\Big(\exp(-(\beta-\beta')R^2/4) + (n+s)^2\frac{\delta}{t} + \gamma(\varepsilon,\varepsilon_0)\Big)\|f^{(n+s)}_{N,0}\|_{L^\infty_{\beta,n+s}} \\
&\qquad\qquad + \big\|\big(f^{(n+s)}_{N,0} - f^{(n+s)}_0\big)(1-\mathbf{1}_{\mathcal{B}^{(n+s)-}_{\varepsilon_0}})\big\|_{L^\infty_{\beta,n+s}} \\
&\qquad\qquad + \|f^{(n+s)}_0 \exp(\beta H_{n+s})\|_{W^{1,\infty}_x(L^\infty_v)}(n+s)\varepsilon\Big]\,,
\end{aligned}$$

with $\beta' < \beta$.

2.6. **A refined convergence statement.** The previous argument shows that once recollisions have been discarded, pseudo-trajectories are stable as $\varepsilon \to 0$, in the sense that their distance to the corresponding Boltzmann pseudo-trajectory converges to 0. The only assumptions used to obtain the convergence of the marginals for times $t \in [0, t^*]$ are that the initial data f_0 has some regularity in space (the Lipschitz bound appearing in Proposition 2.7 could be weakened to Hölder continuity) and the initial marginals satisfy the uniform growth condition

$$\sup_N f^{(n+s)}_{N,0} \le C^{n+s}\exp(-\beta H_{n+s}) \tag{2.19}$$

together with the convergence

$$\Big|(f^{(n+s)}_{N,0} - f^{\otimes(n+s)}_0)(1-\mathbf{1}_{\mathcal{B}^{(n+s)-}_{\varepsilon_0}})\Big| \le C^{n+s}\exp(-\beta H_{n+s})\varepsilon\,.$$

Actually any positive power of ε would do in the above estimate. Note that not all configurations in $\mathcal{D}^{n+s}_\varepsilon \setminus \mathcal{B}^{(n+s)-}_{\varepsilon_0}$ are reached by the good pseudo-trajectories. Actually a very small subset $V^+_{n+s,n} \subset \mathcal{D}^{n+s}_\varepsilon \setminus \mathcal{B}^{(n+s)-}_{\varepsilon_0}$ of these configurations can be reached since one has the condition that, looking at the forward flow, if one particle disappears at each collision, we should end up with n particles within a time T (see Figure 2). This imposes s conditions on the configuration Z_{n+s}. Note that, by definition, configurations of $V^+_{n+s,n}$ have at least one collision when evolved by the free flow $\mathbf{\Psi}_{n+s}$. Taking into account the additional constraint on the order of collisions, we can prove the following result.

Proposition 2.8. *The set of admissible initial configurations (reached by pseudo-dynamics associated with the forward BBGKY hierarchy) satisfies*

$$|V^+_{n+s,n}| \le (CR)^{s+n}(n+s-1)\dots n(RT\varepsilon^{(d-1)})^s\,.$$

Furthermore,

$$V^+_{n+s,n} \subset \mathcal{B}^{(n+s)+}_\varepsilon \setminus \mathcal{B}^{(n+s)-}_\varepsilon\,.$$

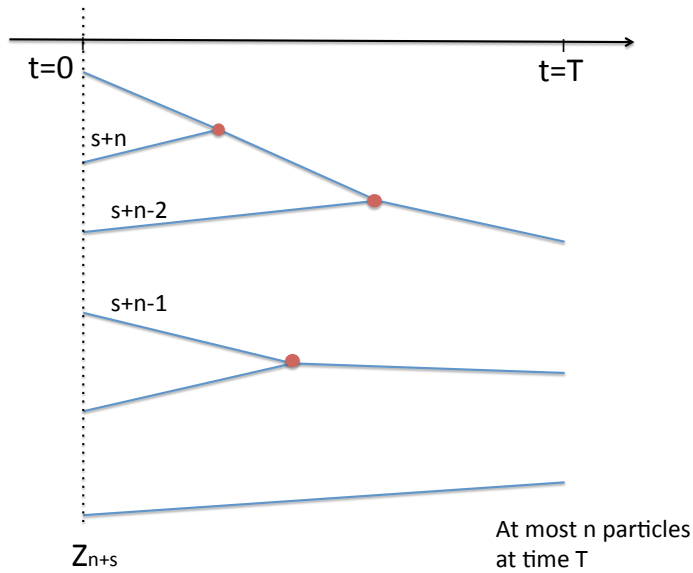

FIGURE 2. Admissible initial configurations.

We thus can state the following refined version of Lanford's theorem which provides quantitative convergence estimates outside the bad sets.

Theorem 2.1. *Consider a system of N hard-spheres of diameter ε on $\mathbb{T}^d = [0,1]^d$ (with $d \geq 2$), initially distributed according to some density $f_{N,0}$ satisfying the growth condition* (2.19) *for some $\beta > 0$, together with the convergence*

$$\forall n \in [1, N]\,, \quad \left|(f_{N,0}^{(n)} - f_0^{\otimes n})\,(1 - \mathbf{1}_{\mathcal{B}_{\varepsilon_0}^{n-}})\right| \leq C^n \exp(-\beta H_n)\varepsilon^a\,, \tag{2.20}$$

for some $a > 0$ and for $|\log \varepsilon|\varepsilon \ll \varepsilon_0 \ll 1$. Denote by f the solution of the Boltzmann equation (1.3)*. Then, in the Boltzmann-Grad limit $N \to \infty$ with $N\varepsilon^{d-1} = 1$, the marginal $f_N^{(n)}$ converges to $f^{\otimes n}$ uniformly on $(\mathcal{D}_\varepsilon^n \setminus \mathcal{B}_{\varepsilon_0}^{n-}) \times [0, t^*]$, i.e. there exists $\gamma'(\varepsilon, \varepsilon_0)$ converging to 0 such that uniformly in $t \in [0, t^*]$,*

$$\left|\big(f_N^{(n)}(t) - f^{\otimes n}(t)\big)(1 - \mathbf{1}_{\mathcal{B}_{\varepsilon_0}^{n-}})\right| \leq C^n \exp(-\beta' H_n)\gamma'(\varepsilon, \varepsilon_0)\,,$$

with $\beta' < \beta$ and t^* introduced in (2.8)*.*

Compared to [4], this theorem provides a description of the geometry of the bad sets along the evolution, and quantitative estimates of their measures. Note that a similar notion of one-sided convergence has been introduced by Denlinger in [14].

3. Irreversibility and time concatenation

Note that the very same proof shows that, in the Boltzmann-Grad limit, the marginal $f_N^{(n)}$ converges to $\tilde{f}^{\otimes n}$ where $\tilde{f}$ is the solution of the **reverse** Boltzmann equation

$$\begin{cases} \partial_t \tilde{f} + v \cdot \nabla_x \tilde{f} = -Q(\tilde{f}, \tilde{f}), \\ Q(f,f)(v) := \displaystyle\iint_{\mathbb{S}^{d-1}\times\mathbb{R}^d} [f(v')f(v_1') - f(v)f(v_1)]\,\big((v_1 - v)\cdot\nu\big)_+\, dv_1 d\nu\,, \end{cases} \tag{3.1}$$

uniformly on $(\mathcal{D}_\varepsilon^n \setminus \mathcal{B}_{\varepsilon_0}^{n+})$ and times from 0 to $-t^*$. This convergence requires only the growth condition (2.19) and the initial convergence

$$\left|(f_{N,0}^{(n)} - f_0^{(n)})(1 - 1_{\mathcal{B}_{\varepsilon_0}^{n+}})\right| \leq C^n \exp(-\beta H_n)\varepsilon^\gamma, \tag{3.2}$$

for some $\gamma > 0$.

We thus have a **symmetric situation for negative and positive times**, which indicates once more that the initial data play a very special role distinguishing between the direct and reverse Boltzmann dynamics.

3.1. **Irreversibility.**

3.1.1. *At the macroscopic level.* Recall that the Boltzmann dynamics admits a Lyapunov functional. Indeed, using the well-known facts (see [13]) that the mappings $(v, v_1) \mapsto (v_1, v)$ (microscopic exchangeability) and $(v, v_1, \nu) \mapsto (v', v_1', \nu)$ (microscopic reversibility) have unit Jacobian determinants and preserve the cross-section, one can show that formally for any test function φ

$$\int Q(f,f)\varphi dv = \frac{1}{4}\iiint [f'f_1' - ff_1](\varphi + \varphi_1 - \varphi' - \varphi_1')((v_1 - v)\cdot\nu)_+ \, dv dv_1 d\nu\,, \tag{3.3}$$

with the short notation $f' = f(v')$, $f_1' = f(v_1')$, $f_1 = f(v_1)$, and similarly for φ.

Disregarding integrability issues, we choose $\varphi = \log f$ in (3.3), and use the properties of the logarithm, to find

$$\begin{aligned} D(f) &\equiv -\int Q(f,f)\log f dv \\ &= \frac{1}{4}\int_{\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{S}^{d-1}} (f'f_1' - ff_1)\log\frac{f'f_1'}{ff_1}\,((v - v_1)\cdot\nu)_+ \, dv dv_1 d\nu \geq 0\,. \end{aligned} \tag{3.4}$$

The so-defined entropy production is therefore a nonnegative functional in agreement with the second principle of thermodynamics.

This leads to Boltzmann's H-theorem, stating that the entropy is (at least formally) a Lyapunov functional for the Boltzmann equation.

Proposition 3.1. *Let $f = f(t,x,v)$ be a smooth solution to the Boltzmann equation* (1.3) *with initial data f_0 of finite relative entropy with respect to some Gaussian (equilibrium) distribution $M = M(v)$*

$$\int f_0 \log\frac{f_0}{M} dv dx < +\infty\,.$$

Then, for all $t \geq 0$

$$\int f\log\frac{f}{M}(t,x,v)dvdx + \int_0^t\int D(f)(\tau,x)dxd\tau \leq \int f_0\log\frac{f_0}{M}dvdx\,. \tag{3.5}$$

The classical interpretation of the H-theorem is that entropy measures the quantity of microscopic information that is known on the system. **Irreversibility is related to a loss of information in our description of the dynamics**.

Note that, for negative times, the distribution is evolved according to the reverse Boltzmann dynamics, and we have

$$\int f\log\frac{f}{M}(-t,x,v)dvdx - \int_0^{-t}\int D(f)(\tau,x)dxd\tau \leq \int f_0\log\frac{f_0}{M}dvdx\,,$$

so that the global picture for the entropy should look like Figure 3.

It is important to realize that the **loss of reversibility is already present at the level of the Boltzmann hierarchy** and does not come from some averaging or projection

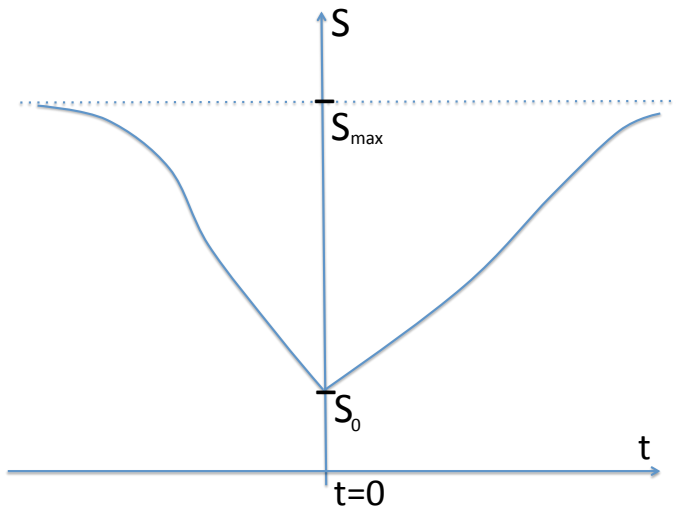

FIGURE 3. Variations of the entropy S.

in phase space. In particular, it has nothing to do with the chaos assumption. Indeed, it can be shown that the Boltzmann hierarchy is irreversible: from the Hewitt-Savage theorem (see [17]) and the symmetry assumption on the labels, we indeed know that the initial data can be decomposed as a superposition of chaotic initial data, i.e. that there exists a measure π on the space of probability densities such that

$$f_0^{(n)} = \int g_0^{\otimes n} d\pi(g_0) \text{ for any } n \in \mathbb{N}^* .$$

Then, by linearity of the Boltzmann hierarchy (2.1), we deduce that the family $(f^{(n)}(t))_{n\in\mathbb{N}^*}$ defined by

$$f^{(n)}(t) = \int (g(t))^{\otimes n} d\pi(g_0)$$

where $g(t)$ is the solution to the Boltzmann equation with initial data g_0, is a solution to the Boltzmann hierarchy. Since the entropy is nondecreasing for each solution of the Boltzmann equation, we deduce that

$$S(t) := -\int (g(t)\log g(t)) d\pi(g_0)$$

is nondecreasing, which encodes irreversibility.

This result means that microscopic information has been lost in the limiting process.

3.1.2. *At the microscopic level.* Let us now consider an intermediate time τ, positive but strictly smaller than Lanford's time t^* in Theorem 1.1. We would like to reverse time and look at the convergence of the BBGKY hierarchy on $[\tau', \tau]$ for $\tau' < \tau$, starting from time τ.

Assume that at time 0 the data $f_{N,0}$ is almost factorized, in the sense of (1.2). We have seen that, for any fixed n, as $\varepsilon \to 0$, the marginals $(f_N^{(n)})$ converge uniformly to $f^{\otimes n}$ outside from $\mathcal{B}_{\varepsilon_0}^{n-}$, where f solves the Boltzmann equation :

$$\|(f_N^{(n)}(t) - f^{\otimes n}(t))(1 - \mathbf{1}_{\mathcal{B}_{\varepsilon_0}^{n-}})\|_\infty \to 0 \text{ as } \varepsilon \to 0 .$$

By definition, starting from τ, f is a solution of the **backward** Boltzmann equation on $[0, \tau]$ (namely of the problem at final values). This would obviously not be the case if we were

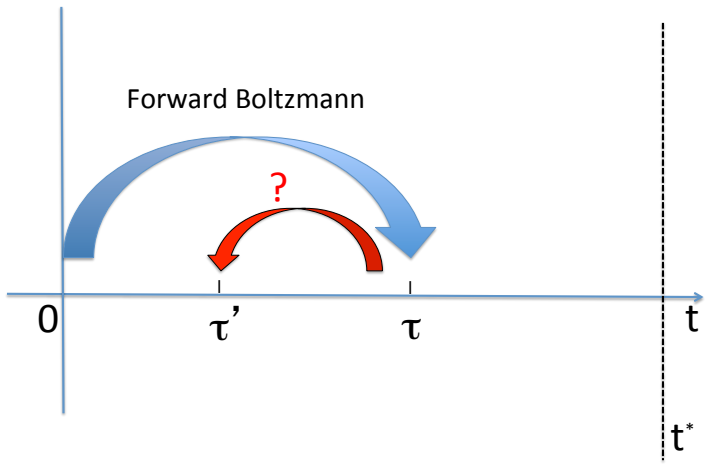

FIGURE 4. Irreversibility. The measure at time zero leads to the Boltzmann equation, but it is not possible to apply Lanford's theorem on $[\tau', \tau]$ taking the measure at time τ as initial condition.

imposing chaotic data at time τ similar to the one at time 0. Indeed, the entropy would be decreased by the reverse Boltzmann equation while it is increased along the backward Boltzmann dynamics. This means therefore that **the structure of the family** $\left(f_N^{(n)}(\tau)\right)_{n\leq N}$ **is very different from the chaotic structure of the initial data**.

Now let us turn to the question of irreversibility. We fix two times $0 < \tau' < \tau < t^*$. Consider the representation formula (2.5) for the marginals $f_N^{(n)}$

$$f_N^{(n)}(\tau') = \sum_{s=0}^{N-n} \int_0^{\tau'} \int_0^{t_{n+1}} \dots \int_0^{t_{n+s-1}} \mathbf{S}_n(\tau' - t_{n+1}) C_{n,n+1} \mathbf{S}_{n+1}(t_{n+1} - t_{n+2}) C_{n+1,n+2} \\ \dots \mathbf{S}_{n+s}(t_{n+s}) f_{N,0}^{(n+s)} \, dt_{n+s} \dots dt_{n+1} \, .$$

It can be written starting from time τ instead of 0, meaning

$$f_N^{(n)}(\tau') = \sum_{s=0}^{N-n} \int_\tau^{\tau'} \int_\tau^{t_{n+1}} \dots \int_\tau^{t_{n+s-1}} \mathbf{S}_n(\tau' - t_{n+1}) C_{n,n+1} \mathbf{S}_{n+1}(t_{n+1} - t_{n+2}) C_{n+1,n+2} \\ \dots \mathbf{S}_{n+s}(t_{n+s} - \tau) f_N^{(n+s)}(\tau) \, dt_{n+s} \dots dt_{n+1} \, ,$$

since the Liouville equation (1.5) satisfied by f_N is reversible and autonomous with respect to time (it generates a group of evolution). As usual for analytic functions, the radius of convergence of the series at τ is at least $t^* - \tau$.

What we would need to apply the refined version of Lanford's theorem (Theorem 2.1) starting from time τ and moving back to τ' is the convergence of $f_N^{(n+s)}(\tau)$ on the sets $V_{n+s,n}^-$ which consist of the configurations of $n+s$ particles at time τ reached by good pseudo-dynamics having s collisions on $[\tau', \tau]$. Note that these pseudo-trajectories are built forward as they go from time τ' to τ and that we have

$$V_{n+s,n}^- \subset \mathcal{B}_\varepsilon^{(n+s)-} \setminus \mathcal{B}_\varepsilon^{(n+s)+} \, ,$$

which is the symmetric counterpart of Proposition 2.8.

Recollisions of the backward dynamics are indeed exactly collisions of the forward pseudo-dynamics. This implies that we have no information about the convergence of $f_N^{(n)}(\tau)$ on the sets $V^-_{n+s,n}$, and that **we cannot prove the convergence to the reverse Boltzmann dynamics on** $[\tau', \tau]$ **starting from** τ (which is consistent with the fact that the reverse Boltzmann dynamics is not the backward Boltzmann dynamics!). For the same reasons the argument behind the so-called Loschmidt's paradox fails. Indeed if at time τ we invert all the velocities and consider $f_N^{(n)}(\tau, X_n, -V_n)$ as initial data, we cannot apply Theorem 2.1 so that there is no contradiction with the backtracking of marginals. The same argument was already put forward in [4].

Remark 3.2. *Evolving a chaotic data by the reverse Boltzmann dynamics gives a systematic method to construct data for which the Boltzmann-Grad limit fails to hold, even though we do have a weak chaos property in the sense of Definition* 1.2*. In Section* 4*, we show a more explicit construction leading to an almost chaotic initial data, with modifications of the second order correlations on a small set, such that the limiting dynamics is free transport (far from the Boltzmann dynamics).*

3.2. **Time-concatenation.** Another important feature of the limiting equation is that one can iterate in the sense of the following Proposition.

Proposition 3.3. *Let f be a smooth solution of the Boltzmann equation* (1.3) *on* $[0, \tau]$ *with initial data f_0, and assume there is a smooth solution f of the Boltzmann equation on* $[\tau, t^*]$ *with data $f(\tau)$ at τ. Then, f is the same solution of the Boltzmann equation on* $[0, t^*]$.

This property is a simple consequence of the fact that the Boltzmann equation is a local in time partial differential equation, with no memory effect. It is a kind of Markov property of the underlying process.

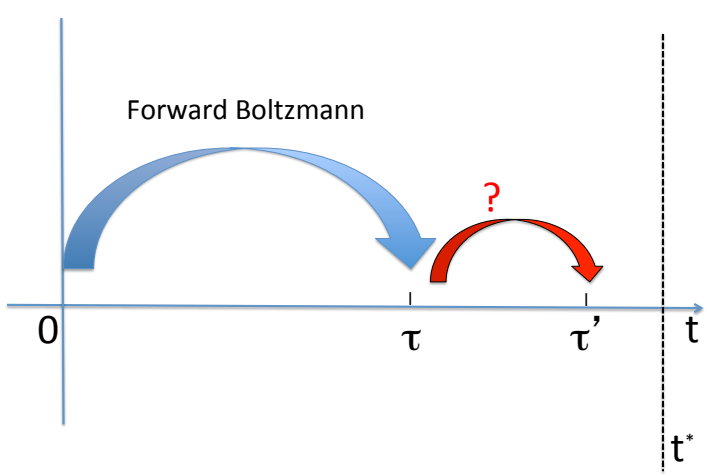

FIGURE 5. Time-concatenation. The measure at time zero leads to the Boltzmann equation on $[0, \tau]$ and it is possible to re-apply Theorem 2.1 on $[\tau, \tau']$ taking the measure at τ as initial condition.

Let $\tau < \tau'$ denote intermediate times, positive but strictly smaller than Lanford's time t^*. As previously, we denote by f_N the solution to the Liouville equation with chaotic initial data in the sense of Theorem 1.1. If we want to iterate Lanford's convergence proof on $[\tau, \tau']$, what we need (in addition to the uniform L^∞ a priori estimate) is the convergence of $f_N^{(n+s)}(\tau)$ on the sets $V^+_{n+s,n}$ reached by good (backward) pseudo-trajectories. By definition, we have $V^+_{n+s,n} \subset \mathcal{D}^{n+s}_\varepsilon \setminus \mathcal{B}^{(n+s)-}_{\varepsilon_0}$.

And from the refined version of Lanford's theorem (Theorem 2.1), we have that

$$\|\big(f_N^{(n)}(\tau) - f^{\otimes n}(\tau)\big)(1 - \mathbf{1}_{\mathcal{B}^{n-}_{\varepsilon_0}})\|_\infty \to 0 \quad \text{ as } \varepsilon \to 0\,.$$

Combining both properties, we deduce that we can iterate the convergence as long as the growth condition (2.19) is satisfied.

Remark 3.4. *Note that the main limiting factor to extend the convergence time is the loss with respect to β in the estimate* (2.7)*. The previous iteration argument fails therefore to improve the time of convergence in Lanford's theorem for initial data of the form* (1.2)*.*

For initial data close to equilibrium, it is proved in [6, 7] *that one can actually reach times of the order $O(\log\log\log N)$. The proof relies on global a priori bounds, it consists in designing a subtle pruning procedure to get rid of the contribution of super-exponential collision trees and then to express the contribution of all other dynamics in terms of the initial data.*

4. Chaotic initial data leading to different dynamics

At large scales, the propagation of chaos (1.10) holds and the measure factorizes, but the memory of the Hamiltonian dynamics remains encoded in $f_N(t)$ on very specific configuration sets of size vanishing with ε. We are not able to describe the refined structure of the correlations in the density $f_N(t)$, but we are going to introduce an example which illustrates how constraints on very small sets may change the nature of the dissipative dynamics. Unlike the one obtained by reversing velocities (see Remark 3.2), this example will be totally explicit.

Using the notation (2.10) of the bad sets, we consider the initial data

$$\hat{f}_{N,0}(Z_N) := \frac{1}{\hat{\mathcal{Z}}_N} \prod_{i=1}^N f_0(x_i, v_i)\, \mathbf{1}_{\{Z_N \notin \mathcal{B}^{N+}_\varepsilon\}}\,, \tag{4.1}$$

where $\mathcal{B}^{N+}_\varepsilon$ is the set such that some collision occurs between the N particles within a time T. Contrary to the definition (2.10), we choose T as a short time and set $T = \delta > 0$. By construction the measure $\hat{f}_{N,0}$ will evolve according to free transport on the time interval $[0, \delta]$ as there are no interactions between the particles. In particular, the evolution of the first marginal $\hat{f}^{(1)}_{N,0}$ is no longer approximated by the Boltzmann equation in the time interval $[0, \delta]$ and there is no dissipation, even at the level of the marginals.

In the following, we are even going to argue that, at a macroscopic scale, the structure of the measure (4.1) behaves essentially as the one of the initial data $f_{N,0}$ given in (1.2) for which Lanford's Theorem holds. In particular, we deduce that a chaos property (1.10) holds for the measure $\hat{f}_{N,0}$. The key point is that the two measures differ on very singular sets which are exactly the relevant sets for the microscopic evolution.

To prove this, it is convenient to rephrase the measure (4.1), which has a fixed number of particles, in a slightly different setting where N is varying. The terminology "canonical" and "grand canonical" ensemble (inherited from statistical physics) is used, respectively, for the two pictures. In the new setting one introduces "rescaled correlation functions" $f^{(j)}_{\varepsilon,0}$ describing the same macroscopic behaviour as the marginals $\hat{f}^{(j)}_{N,0}$. For our present purpose the $f^{(j)}_{\varepsilon,0}$ have some remarkable advantage, as they can be dealt with by using methods of

expansion developed in different contexts [19, 25] (for complications of the cluster expansion techniques due to a canonical formulation, see [29]).

4.1. **The grand canonical formalism.** The grand canonical phase space is

$$\mathcal{D}_\varepsilon = \cup_{n\geq 0}\mathcal{D}_\varepsilon^n$$

(actually $\mathcal{D}_\varepsilon^n = \emptyset$ for n large, due to the exclusion). Given $(f_{n,0})_{n\geq 0}$ we assign the collection of probability densities for the configuration $Z_n \in \mathcal{D}_\varepsilon^n$, $n = 0, 1, \ldots$:

$$\frac{1}{n!}W_{\varepsilon,0}^n(Z_n) := \frac{1}{\mathcal{Z}_\varepsilon}\frac{\mu_\varepsilon^n}{n!}f_{n,0}(Z_n)\,,$$

where $\mu_\varepsilon = \varepsilon^{-d+1}$. The normalization constant $\mathcal{Z}_\varepsilon$ is given by

$$\mathcal{Z}_\varepsilon := \sum_{n\geq 0}\frac{\mu_\varepsilon^n}{n!}\hat{\mathcal{Z}}_n \quad \text{with} \quad \hat{\mathcal{Z}}_n := \int dZ_n f_{n,0}\,.$$

$\{W_{\varepsilon,0}^n\}_{n\geq 0}$ defines the grand canonical state on $\mathcal{D}_\varepsilon$, normalized as

$$\sum_{n\geq 0}\frac{1}{n!}\int W_{\varepsilon,0}^n(Z_n)dZ_n = 1\,.$$

The total number of particles $\mathcal{N}$ is random and distributed according to

$$\mathbb{P}_{\mu_\varepsilon}\big(\mathcal{N} = n\big) = \frac{1}{\mathcal{Z}_\varepsilon}\frac{\mu_\varepsilon^n}{n!}\hat{\mathcal{Z}}_n\,.$$

The choice $\mu_\varepsilon = \varepsilon^{-d+1}$ ensures that the average number of particles grows as ε^{-d+1}, hence the inverse mean free path remains of order 1 (Boltzmann-Grad scaling)

$$\lim_{\varepsilon\to 0}\mathbb{E}_{\mu_\varepsilon}\big(\mathcal{N}\big)\varepsilon^{d-1} = \kappa > 0\,. \tag{4.2}$$

We postpone this check to the end of the section.

Let us define the j-particle correlation function, $j = 1, 2, \ldots$. The idea is to count how many groups of j particles fall, in average, in a given configuration $Z_j = (z_1, \ldots, z_j)$:

$$\rho_{\varepsilon,0}^{(j)}(z_1, \ldots, z_j) = \Big\langle \sum_{k_i\neq k_j}\delta_{\zeta_{k_1}}(z_1)\ldots\delta_{\zeta_{k_j}}(z_j)\Big\rangle\,,$$

where we are labelling the particles and indicating their (random) configuration by $\zeta_1, \cdots, \zeta_n$, and the brackets denote average with respect to the grand canonical state. In terms of the densities it is

$$\rho_{\varepsilon,0}^{(j)}(Z_j) = \sum_{n=0}^{\infty}\frac{1}{n!}\int dz_{j+1}\cdots dz_{j+n}W_{\varepsilon,0}^{j+n}(Z_{j+n})\,.$$

In the case with minimal correlations, i.e. when

$$f_{n,0} := \prod_{i=1}^n f_0(x_i, v_i)\prod_{i\neq j}\mathbf{1}_{|x_i-x_j|>\varepsilon}\,, \tag{4.3}$$

one has

$$
\begin{aligned}
f^{(j)}_{\varepsilon,0}(Z_j) &:= \mu_\varepsilon^{-j}\rho^{(j)}_{\varepsilon,0}(Z_j)\\
&= \Big[f_0^{\otimes j}\prod_{1\le i<k\le j}\mathbf{1}_{|x_i-x_k|>\varepsilon}\Big]\\
&\times \frac{1}{\mathcal{Z}_\varepsilon}\sum_{n\ge 0}\frac{\mu_\varepsilon^n}{n!}\int dz_{j+1}\cdots dz_{j+n} f_0^{\otimes n}\left(\prod_{i=1}^{j}\prod_{k=j+1}^{j+n}\mathbf{1}_{|x_i-x_k|>\varepsilon}\right)\left(\prod_{j+1\le i<k\le j+n}\mathbf{1}_{|x_i-x_k|>\varepsilon}\right)\\
&\le f_0^{\otimes j}\mathbf{1}_{\{Z_j\notin\mathcal{B}_\varepsilon^{j+}\}}(Z_j)\,, \qquad (4.4)
\end{aligned}
$$

where the last inequality follows by removing the constraint between the j particles and the rest of the system. Note that the rescaled correlation functions $f^{(j)}_{\varepsilon,0}$ are quantities of order 1 in ε.

The Boltzmann equation can be derived for both ensembles [4, 32, 28].

Theorem 4.1 ([4])**.** *Consider a system of hard-spheres of diameter ε on the d-dimensional periodic box $\mathbb{T}^d=[0,1]^d$ (with $d\ge 2$), initially in the grand canonical state with $f_{n,0}$ given by* (4.3) *and f_0 satisfying* (1.1)*.*

Then, as $\varepsilon\to 0$, the rescaled correlation function $f^{(1)}_\varepsilon$ converges almost everywhere to the solution of the Boltzmann equation (1.3) *with initial data f_0, on a time interval $[0,t^*]$ where t^* depends only on the parameters β,μ of* (1.1)*.*

4.2. **A counterexample.** A natural reformulation of (4.1) with varying number of particles is obtained as follows. Define

$$
\frac{1}{n!}W^n_{\varepsilon,0}(Z_n) := \frac{1}{\mathcal{Z}_\varepsilon}\frac{\mu_\varepsilon^n}{n!}f_0^{\otimes n}(Z_n)\mathbf{1}_{\{Z_n\notin\mathcal{B}_\varepsilon^{n+}\}} = \frac{1}{\mathcal{Z}_\varepsilon}\frac{\mu_\varepsilon^n}{n!}f_0^{\otimes n}(Z_n)\prod_{i<j}(1+\zeta_{ij})\,, \qquad (4.5)
$$

where $\mu_\varepsilon=\varepsilon^{-d+1}$ and $\zeta_{ij}=\zeta(z_i,z_j)=-\mathbf{1}_{\mathcal{C}}(z_i,z_j)$ with $\mathcal{C}$ the set leading to a collision

$$
\mathcal{C} := \Big\{(z_i,z_j)\in(\mathbb{T}^d\times\mathbb{R}^d)^2,\quad \exists s\in[0,\delta],\quad \big|x_i-x_j+s(v_i-v_j)\big|\le\varepsilon\Big\}.
$$

The normalization constant $\mathcal{Z}_\varepsilon$ is given as above by

$$
\mathcal{Z}_\varepsilon := \sum_{n\ge 0}\frac{\mu_\varepsilon^n}{n!}\hat{\mathcal{Z}}_n \quad\text{with}\quad \hat{\mathcal{Z}}_n := \int dZ_n f_0^{\otimes n}\prod_{i<j}(1+\zeta_{ij})\ .
$$

By construction, the grand canonical density (4.5) evolves according to the free transport dynamics in the time interval $[0,\delta]$,

$$
\forall t\le\delta,\qquad f^{(j)}_\varepsilon(t,Z_j) := \mu_\varepsilon^{-j}\rho^{(j)}_\varepsilon(t,Z_j) = \mathbf{S}^0_j(t)f^{(j)}_{\varepsilon,0}(Z_j)\,. \qquad (4.6)
$$

The rescaled correlation functions $f^{(j)}_{\varepsilon,0}$ obey some of the assumptions required to apply Lanford's theorem, in particular the key L^∞ bound holds thanks to (4.4). Moreover, we will see in Proposition 4.1 below that a chaos property holds in a sense stronger than (1.10). Nevertheless the correlation functions are irregular at the microscopic scale on the sets $\mathcal{B}_\varepsilon^{j+}$ so that Lanford's proof cannot apply and there is no contradiction with (4.6). Note that the constraints are imposed only in the forward direction, thus we expect to get the reverse Boltzmann equation for negative times.

To conclude this example, we will show that the state is chaotic.

Proposition 4.1. *The measure $\{W^n_{\varepsilon,0}\}_{n\geq 0}$ is asymptotically chaotic, uniformly outside a bad set of configurations in $\mathcal{D}_\varepsilon$. More precisely, there exists $f^{(1)}:\mathbb{T}^d\times\mathbb{R}^d\to\mathbb{R}^+$ such that*

$$\begin{aligned}&\lim_{\varepsilon\to 0}\sup_z\left|f^{(1)}_{\varepsilon,0}(z)-f^{(1)}(z)\right|=0\,,\\ &\lim_{\varepsilon\to 0}\sup_{Z_j\notin\mathcal{B}^{j+}_{\varepsilon|\log\varepsilon|}}\left|f^{(j)}_{\varepsilon,0}(Z_j)-f^{(1)}_{\varepsilon,0}(z_1)\cdots f^{(1)}_{\varepsilon,0}(z_j)\right|=0\,,\end{aligned}\tag{4.7}$$

for all $j\geq 2$.

The result for $j=2$ will follow by applying Theorem 2.3 of [25] (recalled below) where the decay of correlations has been estimated by means of cluster expansion.

Theorem 4.2 ([25])**.** *Assume that there exist non negative functions a and b such that*

$$\begin{aligned}&\forall n,\quad \forall(z_1,\dots,z_n)\,,\qquad \prod_{1\leq i<j\leq n}(1+\zeta_{ij})\leq\prod_{i=1}^n e^{b(z_i)}\,,\\ &\forall z_i\,,\quad \varepsilon^{1-d}\int f_0(z_j)|\zeta_{ij}|e^{a(z_j)+2b(z_j)}dz_j\leq a(z_i)\,.\end{aligned}\tag{4.8}$$

Then, for almost all z_1,z_2,

$$\begin{aligned}&\left|f^{(2)}_{\varepsilon,0}(z_1,z_2)-f^{(1)}_{\varepsilon,0}(z_1)f^{(1)}_{\varepsilon,0}(z_2)\right|\leq f_0(z_1)f_0(z_2)e^{a(z_1)+a(z_2)+2b(z_1)+2b(z_2)}\\ &\qquad\cdot\Big\{|\zeta_{12}|+\sum_{m\geq 1}\mu^m_\varepsilon\int dZ'_m f_0^{\otimes m}(Z'_m)|\zeta(z_1,z'_1)\zeta(z'_1,z'_2)\cdots\zeta(z'_m,z_2)|\prod e^{a(z'_i)+2b(z'_i)}\Big\}\,.\end{aligned}$$

Proof of Proposition 4.1 *when $j=2$.* Assumptions (4.8) of Theorem 4.2 hold by choosing $b=0$ and $a(v)=c\delta(1+|v|)$ with δ small enough, for some constant c (depending on β,μ,d of (1.1)). As a consequence, Theorem 4.2 leads to

$$\begin{aligned}&\left|f^{(2)}_{\varepsilon,0}(z_1,z_2)-f^{(1)}_{\varepsilon,0}(z_1)f^{(1)}_{\varepsilon,0}(z_2)\right|\leq f_0^{\otimes 2}e^{c\delta(2+|v_1|+|v_2|)}\\ &\qquad\cdot\Big\{|\zeta_{12}|+\sum_{m\geq 1}\mu^m_\varepsilon\int dZ'_m f_0^{\otimes m}(Z'_m)|\zeta(z_1,z'_1)\zeta(z'_1,z'_2)\cdots\zeta(z'_m,z_2)|e^{c\delta m+c\delta\sum_{i=1}^m|v'_i|}\Big\}\,.\end{aligned}$$

For δ small, the prefactor is bounded by $c\,e^{-\frac{\beta}{4}v^2}$ as f_0 satisfies (1.1). Moreover, for (z_1,z_2) outside $\mathcal{B}^{2+}_{\varepsilon|\log\varepsilon|}$, the first term ζ_{12} is equal to 0. Then the proof of (4.7) boils down to showing that

$$\lim_{\varepsilon\to 0}\sum_{m\geq 1}c^m\varepsilon^{m(1-d)}\int dZ'_m|\zeta(z_1,z'_1)\zeta(z'_1,z'_2)\cdots\zeta(z'_m,z_2)|e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+\sum_{i=1}^m|v'_i|^2\right)}=0\tag{4.9}$$

uniformly out of $\mathcal{B}^{2+}_{\varepsilon|\log\varepsilon|}$. Given a velocity v, we define a cylinder associated with $z_1=(x_1,v_1)$ by

$$\mathcal{R}(z_1,v):=\Big\{x\in\mathbb{T}^d,\qquad\exists s\in[0,\delta],\quad\big|x-x_1+s(v-v_1)\big|\leq\varepsilon\Big\}\,.$$

The measure of $\mathcal{R}(z_1,v)$ is of order $\varepsilon^{d-1}\delta|v-v_1|$.

We first treat the term $m=1$ and show that for some constant $C>0$

$$\varepsilon^{1-d}\int dz'_1|\zeta(z_1,z'_1)\zeta(z'_1,z_2)|e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+|v'_1|^2\right)}\leq\frac{C}{|\log\varepsilon|^{1/2}}\,.\tag{4.10}$$

Given z_1,z_2, we distinguish two cases to evaluate the measure of the overlap $\mathcal{R}(z_1,v'_1)\cap\mathcal{R}(z_2,v'_1)$. Let α be the angle between the axis of both cylinders, i.e. the angle between v'_1-v_1 and v'_1-v_2.

• *Suppose that* $|\sin(\alpha)| \geq \varepsilon|\log\varepsilon|^{1/2}$, then the angle between both axis is large enough so that the overlap $\mathcal{R}(z_1, v_1') \cap \mathcal{R}(z_2, v_1')$ has a volume of order at most $\varepsilon^{d-1}/|\log\varepsilon|^{1/2}$. We get

$$\varepsilon^{1-d} \int dz_1' \mathbf{1}_{\{|\sin(\alpha)|\geq\varepsilon|\log\varepsilon|^{1/2}\}} |\zeta(z_1, z_1')\zeta(z_1', z_2)| e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+|v_1'|^2\right)} \leq \frac{C}{|\log\varepsilon|^{1/2}} \,. \tag{4.11}$$

• *Suppose that* $|\sin(\alpha)| \leq \varepsilon|\log\varepsilon|^{1/2}$, then the cylinders $\mathcal{R}(z_1, v_1')$ and $\mathcal{R}(z_2, v_1')$ are almost parallel and they are anchored at x_1, x_2. Recall that (z_1, z_2) is outside $\mathcal{B}^{2+}_{\varepsilon|\log\varepsilon|}$ so that $|x_1 - x_2| \geq \varepsilon|\log\varepsilon|$. The length of both cylinders is less than $\delta\left(|v_1' - v_1| + |v_1 - v_2|\right)$, thus they can overlap only if θ, the angle between $x_1 - x_2$ and $v_1' - v_1$, is small enough.

- Suppose first that the lines $x_1 + \lambda(v_1' - v_1)$ and $x_2 + \mu(v_1' - v_2)$ intersect at some point u (see Figure 6). Then the length $\ell = \min\{|u - x_1|, |u - x_2|\}$ satisfies

$$\ell = \frac{|\sin\theta|}{|\sin\alpha|}|x_1 - x_2| \,.$$

For the intersection to occur one needs that $\ell \leq \delta\left(|v_1' - v_1| + |v_1 - v_2|\right)$ so that we get the condition on θ

$$|\sin\theta| \leq \delta\left(|v_1' - v_1| + |v_1 - v_2|\right) \frac{|\sin\alpha|}{|x_1 - x_2|} \leq C\,\frac{|v_1' - v_1| + |v_1 - v_2|}{|\log\varepsilon|^{1/2}} \,.$$

- If the two lines in the picture do not intersect (as will happen in general for $d > 2$), the above inequality can be proved by a similar argument. Define u, v as the points in the first and second lines where the distance 2ε between both lines is reached. Then we can project all vectors orthogonally to $u - v$, and we get exactly the same picture.

As a conclusion, we get that θ should belong to a solid angle of order $\left(\frac{|v_1'-v_1|+|v_1-v_2|}{|\log\varepsilon|^{1/2}}\right)^{d-1}$. Integrating over x_1' and $v_1' - v_1$, we deduce that

$$\varepsilon^{1-d} \int dz_1' \mathbf{1}_{\{|\sin\alpha|\leq\varepsilon|\log\varepsilon|^{1/2}\}} |\zeta(z_1, z_1')\zeta(z_1', z_2)| e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+|v_1'|^2\right)} \leq \frac{C}{|\log\varepsilon|^{(1/2)(d-1)}} \,.$$

Combined with (4.11), this completes (4.10).

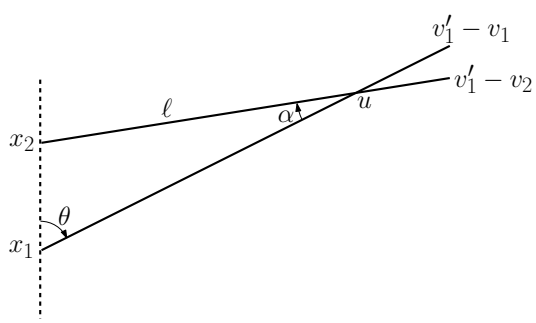

FIGURE 6.

We now show that the contribution of the term m is bounded by

$$\varepsilon^{m(1-d)} \int dZ_m' |\zeta(z_1, z_1')\zeta(z_1', z_2') \cdots \zeta(z_m', z_2)| e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+\sum_{i=1}^m |v_i'|^2\right)} \leq \frac{C^m \delta^{m-1}}{|\log\varepsilon|^{1/2}} \tag{4.12}$$

for some constant C. Summing over m this will complete the derivation of (4.9) for δ small enough.

To estimate the case $m = 1$, we simply used the fact that $|x_1 - x_2| \geq \varepsilon|\log\varepsilon|$. Suppose that x_2' is such that $|x_1 - x_2'| \geq \varepsilon|\log\varepsilon|$. Then integrating with respect to z_1' leads to

$$\varepsilon^{m(1-d)} \int dZ_m' \mathbf{1}_{\{|x_1-x_2'|\geq\varepsilon|\log\varepsilon|\}} |\zeta(z_1, z_1')\zeta(z_1', z_2')\cdots\zeta(z_m', z_2)| e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+\sum_{i=1}^m |v_i'|^2\right)}$$
$$\leq \frac{C}{|\log\varepsilon|^{1/2}} \varepsilon^{(m-1)(1-d)} \int dz_2' \ldots dz_m' |\zeta(z_2', z_3')\cdots\zeta(z_m', z_2)| e^{-\frac{\beta}{8}\left(v_2^2+\sum_{i=2}^m |v_i'|^2\right)} ,$$

where we applied an estimate as (4.10) using part of the exponential factor, and removed the constraint $\mathbf{1}$ in the upper bound. Finally, we can integrate term by term as the constraint on z_i' depends only on z_{i+1}'. This leads to a contribution of the form $C\varepsilon^{d-1}\delta(|v_i'| + |v_{i+1}'|)$ for each constraint. After integrating the velocities, we obtain an upper bound $C^{m-1}\delta^{m-1}\varepsilon^{(m-1)(d-1)}(1 + |v_2|)e^{-\frac{\beta}{8}v_2^2}$ which implies an estimate as in (4.12).

It remains to consider the set $\{|x_1 - x_2'| \leq \varepsilon|\log\varepsilon|\}$. We first integrate over z_2'

$$\varepsilon^{m(1-d)} \int dZ_m' \mathbf{1}_{\{|x_1-x_2'|\leq\varepsilon|\log\varepsilon|\}} |\zeta(z_1, z_1')\zeta(z_1', z_2')\cdots\zeta(z_m', z_2)| e^{-\frac{\beta}{4}\left(v_1^2+v_2^2+\sum_{i=1}^m |v_i'|^2\right)}$$
$$\leq \varepsilon^{m(1-d)} C\varepsilon^d |\log\varepsilon|^d e^{-\frac{\beta}{4}\left(v_1^2+v_2^2\right)} \int dz_1' |\zeta(z_1, z_1')| e^{-\frac{\beta}{4}|v_1'|^2}$$
$$\times \int dz_3' \ldots dz_m' |\zeta(z_3', z_4')\cdots\zeta(z_m', z_2)| e^{-\frac{\beta}{4}\sum_{i=3}^m |v_i'|^2}.$$

This breaks the cluster into two independent parts which can be estimated separately by the product of the volume of the cylinders, leading to a higher order contribution $\varepsilon|\log\varepsilon|^d C^m \delta^{m-1}$. This completes the derivation of (4.12) and the proof of (4.7) for $j = 2$.

□

The statement for $j = 1$ is also similar and follows from the cluster expansion of [25]. In fact Theorem 2.2 and Proposition 6.1 in [25] imply that $f_{\varepsilon,0}^{(1)}$ is uniformly bounded by a geometric series for δ small.

Note that, in particular, the scaling condition (4.2) holds, with κ uniformly bounded in δ. Indeed, since there exists a (nontrivial) measurable nonnegative $f^{(1)}$ such that $f_{\varepsilon,0}^{(1)} \to f^{(1)}$ as $\varepsilon \to 0$, it follows that

$$\varepsilon^{d-1}\mathbb{E}_{\mu_\varepsilon}(\mathcal{N}) = \varepsilon^{d-1} \int \rho_{\varepsilon,0}^{(1)}(z)dz = \mu_\varepsilon^{-1} \int \rho_{\varepsilon,0}^{(1)}(z)dz = \int f_{\varepsilon,0}^{(1)}(z)dz \to \kappa$$

where $\kappa := \int f^{(1)}(z)dz$.

The case $j > 2$ can be treated similarly, however the expressions are more lengthy and we refer to [34] for details.

5. Concluding remarks

5.1. **Some wrong ideas about irreversibility.** The previous analysis brings a more precise understanding of Loschmidt's paradox : it indicates where the irreversibility of the Boltzmann description appears in the limiting process.

We would like first to comment upon some of the possible explanations which can be found in the literature.

- The direction of time in the Boltzmann dynamics is **not related to an arbitrary choice in writing the collision operator**. Once the initial data is prescribed, one has no choice in expressing the collision operator in terms of pre-collisional configurations for positive times, and in terms of post-collisional configurations for negative times. As explained in Remark 2.1, this is the only way to define properly the traces

by using the transport operator. This is also related to the fact that only the distribution of ingoing configurations has to be prescribed for the transport equation (see Remark 1.1).

- Irreversibility is not **a direct consequence of chaos**. One can indeed start from a non chaotic initial data, in which case the Boltzmann hierarchy does not decouple. However, even in this case, we have seen in Section 3.1 that the limiting evolution is irreversible. We indeed have a Lyapunov functional, obtained by linear superposition of the entropy functionals with the Hewitt-Savage measure, which is strictly decreasing.
- Irreversibility is not due to **neglecting the interaction length in the collision process**. In the limit, we forget indeed about the relative (microscopic) positions of the particles at the time of collisions, but this information could be kept by introducing an intermediate description, i.e. a simple modification of the Boltzmann equation referred to as the Enskog hierarchy [28]. In this equation the collision operator is still of type (2.3). However, Arkeryd and Cercignani [2] (see also [5]) prove that the Enskog equation (and thus the Enskog hierarchy using the previous superposition principle) is irreversible.

5.2. **A very singular averaging process.** Neglecting spatial micro-translations in the limit induces a first loss of information. The second loss of information, which is actually responsible for the loss of reversibility, consists in neglecting pathological configurations, i.e. configurations leading to pseudo-trajectories involving recollisions. These sets $\mathcal{B}^{\pm}_{\varepsilon_0}$ defined in (2.10) have a simple geometric definition, and their measure converges to 0 in the limit. So apparently it seems rather natural not to care about them.

The point is that the marginals at time t can be computed as weighted averages of the initial marginals on very singular sets, which have exactly the same structure and the same measure. Recollisions of the backward dynamics are indeed exactly collisions of the forward pseudo-dynamics. We have therefore identified very precisely why time-concatenation is possible while reversing the arrow of time is not. This can be summarized as in Figure 7.

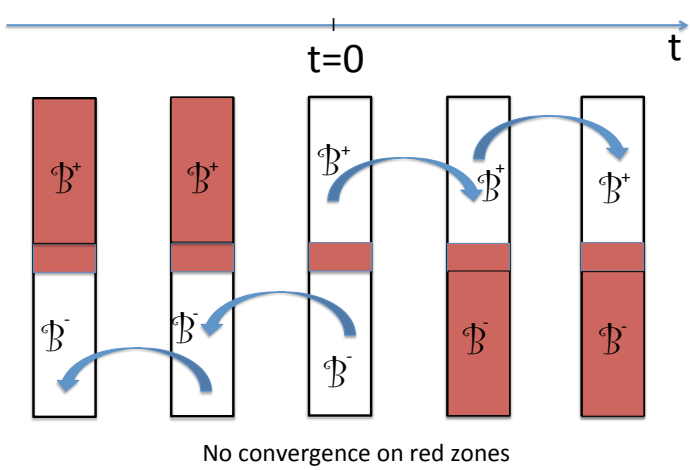

FIGURE 7. Convergence and lack of convergence over singular sets

Note that, for a better understanding of the Boltzmann dynamics, it is not enough to look at the specific initial data (1.2), as its particular form is not stable under the dynamics.

We would need a more systematic classification of the limiting dynamics depending on the microscopic structure of the n-particle distribution.

References

[1] R. Alexander, The Infinite Hard Sphere System, *Ph.D. dissertation,* Dept. Mathematics, Univ. California, Berkeley, 1975.

[2] L. Arkeryd, C. Cercignani. *Global existence in L^1 for the Enskog equation and convergence of the solutions to solutions of the Boltzmann equation*, J. Stat. Phys. **59**, 3–4, 845 (1990).

[3] C. Bardos. *Problèmes aux limites pour les équations aux dérivées partielles du premier ordre à coefficients réels; théorèmes d'approximation; application à l'équation de transport*, Ann. Sci. Ecole Norm. Sup. (4) 3, no.2, 185–233 (1970).

[4] H. van Beijeren, O.E. Lanford, J. Lebowitz, H. Spohn, *Equilibrium time correlation functions in the low-density limit*, J. Statist. Phys. **22**, 2, 237–257, (1980).

[5] N. Bellomo, M. Lachowicz, J. Polewczak, G. Toscani, *Mathematical Topics in Nonlinear Kinetic Theory II: The Enskog Equation*, Worldscientific, Singapore, 1991.

[6] T. Bodineau, I. Gallagher, L. Saint-Raymond, *The Brownian motion as the limit of a deterministic system of hard-spheres*, Inventiones, 1–61, (2015).

[7] T. Bodineau, I. Gallagher, L. Saint-Raymond, *From hard sphere dynamics to the Stokes-Fourier equations: an L^2 analysis of the Boltzmann-Grad limit*, to appear in Annals PDE (2016).

[8] L. Boltzmann, *Further Studies on the Thermal Equilibrium of Gas Molecules* (1872), The Kinetic Theory of Gases, History of Modern Physical Sciences **1**, 262–349, ed. S.G.Brush and N.S.Hall (2003).

[9] L. Boltzmann, *Vorlesungen über Gastheorie*, J.A. Barth, Leipzig (1896), english translation by S.G.Brush, Lectures on Gas Theory, University of California Press, Berkeley (1964).

[10] C. Cercignani, *On the Boltzmann equation for rigid spheres*, Transp. Theory. Stat. Phys. 2, 211–225 (1972).

[11] C. Cercignani, *Ludwig Boltzmann, The man who trusted atoms*, Oxford University Press, Oxford, 1998.

[12] F. Bouchet, L. Saint-Raymond. *Is Boltzmann's equation reversible? A new large deviation perspective on the irreversibility paradox*, in preparation.

[13] C. Cercignani, R. Illner, M. Pulvirenti, *The Mathematical Theory of Dilute Gases*, Springer Verlag, New York NY (1994).

[14] R. Denlinger, *The propagation of chaos for a rarefied gas of hard spheres in the whole space*, https://arxiv.org/abs/1605.00589.

[15] I. Gallagher, L. Saint-Raymond, B. Texier, From Newton to Boltzmann : the case of hard-spheres and short-range potentials, *ZLAM* (2014).

[16] H. Grad, Principles of the kinetic theory of gases, *Handbuch der Physik* **12**, Thermodynamik der Gase p. 205-294 Springer-Verlag, Berlin-Gottingen-Heidelberg, 1958.

[17] E. Hewitt, L. Savage, *Symmetric measures on Cartesian products.* Trans. Amer. Math. Soc. **80**, 470–501 (1955).

[18] F. King. BBGKY hierarchy for positive potentials. *Ph.D. Thesis*, Department of Math., Univ. California, Berkeley (1975).

[19] R. Kotecký. Cluster expansions. *Encycl. of Math. Phys.* **1**, 531–536, Oxford: Elsevier, 2006.

[20] O.E. Lanford, *Time evolution of large classical systems*, Lect. Notes in Physics **38**, J. Moser ed., 1–111, Springer Verlag (1975).

[21] O.E. Lanford, *On a derivation of the Boltzmann equation*, Asterisque 40, 117-137 (1976).

[22] J. Lebowitz, S. Goldstein, *On the (Boltzmann) Entropy of Non-equilibrium Systems*, Physica D, 193, 53-66, (2004).

[23] J. Lebowitz, *From Time-Symmetric Microscopic Dynamics to Time-Asymmetric Macroscopic Behavior: An Overview*, Boltzmann's Legacy-Proceedings of Vienna Conference on Boltzmann, 63–89, (2007).

[24] J. Loschmidt, Sitzungsber. Kais. Akad. Wiss. Wien, Math. Naturwiss. Classe 73, 128–142, (1876).

[25] S. Poghosyan and D. Ueltschi, *Abstract cluster expansion with applications to statistical mechanical systems,* J. Math. Phys. 50, 053509 (2009).

[26] M. Pulvirenti, C. Saffirio, S. Simonella, *On the validity of the Boltzmann equation for short range potentials.* Rev. Math. Phys. 26, no. 2 (2014).

[27] M. Pulvirenti, S. Simonella, *On the evolution of the empirical measure for the hard-sphere dynamics*, Bull. Inst. Math. Acad. Sin. 10(2), 171–204 (2015).

[28] M. Pulvirenti and S. Simonella, *The Boltzmann-Grad Limit of a Hard Sphere System: Analysis of the Correlation Error*, Inventiones (to appear).

[29] E. Pulvirenti and D. Tsagkarogiannis. *Cluster Expansion in the Canonical Ensemble.* Comm. in Math. Phys. **316**, 2, 289–306, (2012).
[30] S. Simonella. *Evolution of correlation functions in the hard sphere dynamics.* J. Stat. Phys., 155, 6, 1191–1221, (2014).
[31] H. Spohn, Boltzmann hierarchy and Boltzmann equation, in *Kinetic theories and the Boltzmann equation* (Montecatini, 1981), p. 207-220.
[32] H. Spohn, Large scale dynamics of interacting particles, *Springer-Verlag* **174** (1991).
[33] H. Spohn, *Loschmidt's Reversibility Argument and the H-Theorem*, Pioneering Ideas for the Physical and Chemical Sciences, ed. W. Fleischhacker and T. Schönfeld, 153-157, Springer US (1997).
[34] D. Ueltschi, *Cluster expansions and correlation functions*, Mosc. Math. J. **4**, 2, 511–522, (2004).
[35] S. Ukai, The Boltzmann-Grad Limit and Cauchy-Kovalevskaya Theorem, *Japan J. Indust. Appl. Math.,* **18**, 383–392, (2001).