Large scale dynamics of dilute gases

– Very preliminary version –

Thierry Bodineau

Laboratoire Alexander Grothendieck, IHES

September 27, 2022

Contents

Chapter 1

Introduction

1.1 Introduction

1.2 Equilibrium statistical mechanics

A gas or a liquid at equilibrium can be described in terms of a few macroscopic variables as the pressure P, the volume V and the temperature T. These parameters obey well established physical laws, for example $PV = nRT$ for an ideal gas, and the different phases of a system are determined by these parameters (see Figure 1.1).

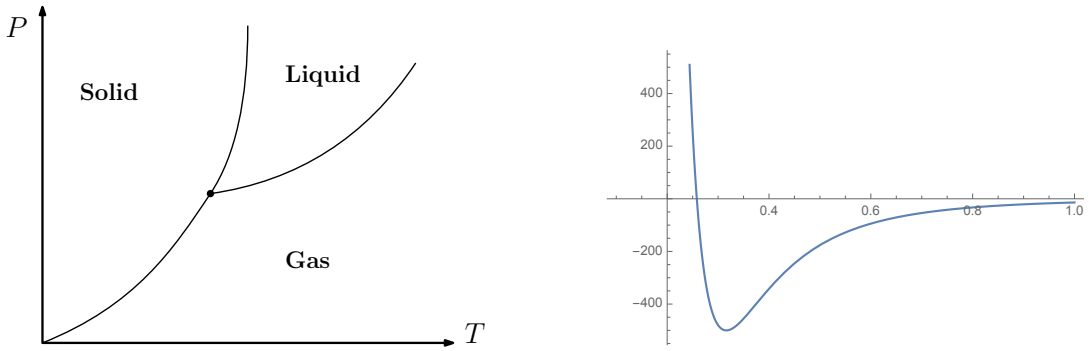


Figure 1.1: The left figure represents a phase diagram. The Lennard-Jones potential (1.2) is depicted on the right. The potential is tuned such that the particles repel at short distances and attract each other otherwise.

A gas is made of a huge number of interacting atoms $N \approx 10^{23}$ and the goal of statistical mechanics is to derive the macroscopic laws of physics starting from the microscopic interactions between the atoms. To model a gas, the positions and the velocities of the atoms can be encoded by the variables $\{(x_i, v_i)\}_{i \leqslant N}$ which are randomly distributed according to the Gibbs theory described next. The kinetic energy of a gas is a function of the velocities $V_N = \{v_i\}_{i \leqslant N}$

$$H_{kinetic}(V_N) = \frac{m}{2} \sum_{i=1}^{N} v_i^2, \tag{1.1}$$

where m stands for the mass of an atom. The particle interactions depend on an interaction potential Φ often represented by a Lennard-Jones potential of the form (see Figure 1.1)

$$r \geqslant 0, \qquad \Phi(r) = E_0 \left(\frac{1}{r^{12}} - \frac{1}{r^6} \right), \tag{1.2}$$

for some constant $E_0 > 0$. An additional parameter $\varepsilon > 0$ is introduced to fix the typical distance between the atoms. As a consequence, the potential Φ is rescaled in order to define the interaction energy which depends only on the positions $X_N = \{x_i\}_{i \leqslant N}$

$$H_{interaction}(X_N) = \sum_{\substack{i,j \\ i \neq j}} \Phi \left(\frac{x_i - x_j}{\varepsilon} \right). \tag{1.3}$$

The energy between two atoms is minimal when their distance is of order ε and their interaction becomes negligible if their distance is much larger than ε. To fix ideas, we assume that the gas is contained in a macroscopic vessel modelled by the unit domain $\mathbb{T}^3 = [0,1]^3$ so that each x_i takes values in $\mathbb{T}^3$. In order to disregard the boundary effects, the vessel is assumed to be periodic.

Gibbs theory is the cornerstone of statistical mechanics as it relates the macroscopic parameters to statistical averages of some microscopic observables. At equilibrium the atom coordinates $Z_N = \{X_N, V_N\}$ are distributed according to the *Gibbs measure* with density

$$\mathcal{G}_{N,\beta}(Z_N) = \frac{1}{\mathcal{Z}_{N,\beta}} \exp\Big(-\beta \Big(H_{kinetic}(V_N) + H_{interaction}(X_N) \Big) \Big), \tag{1.4}$$

with respect to the Lebesgue measure $dx_1 \dots dx_N \; dv_1 \dots dv_N$ in $(\mathbb{T}^3 \times \mathbb{R}^3)^N$. At equilibrium, the thermodynamic parameters can be recovered by averaging microscopic variables with respect to the Gibbs measure. Since the Gibbs measure is translation invariant, the density of the gas is simply given by

$$\rho = N\varepsilon^d. \tag{1.5}$$

The temperature is defined by

$$T = \frac{m}{3} \mathbb{E}_{\mathcal{G}_{N,\beta}}(v_i^2) = \frac{1}{\beta}. \tag{1.6}$$

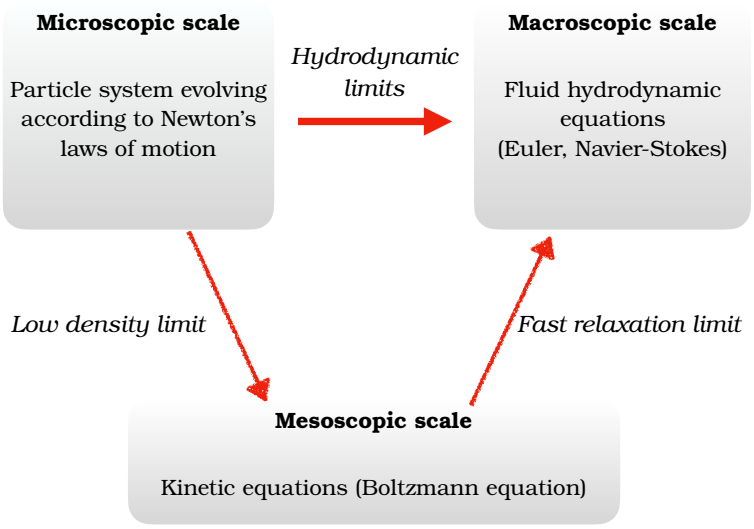
Further macroscopic parameters can be recovered by computing atom correlations under the Gibbs measure. The Gibbs theory relates the microscopic and macroscopic scales in the thermodynamic limit, i.e. when the particle number N tends to infinity and ε tends to 0 in such a way that the density is kept fixed

$$\rho = \lim_{\substack{N \to \infty, \\ \varepsilon \to 0}} N\varepsilon^d. \tag{1.7}$$

From a probabilist view point, this limit should be interpreted as a law of large numbers which ensures that averages of a random sequence are close to a mean behavior. As the particle number in a gas is typically of order 10^{23}, the limiting regime provides an accurate description of the gas. Physical laws at equilibrium can be deduced from the Gibbs theory and, in particular, the phase diagram can be predicted from the microscopic description (even so, this has not been achieved yet by a fully rigorous mathematical theory).

1.3 Hydrodynamic limits

In every day life, most of the physical systems are out-of equilibrium and one would like to derive the laws of non-equilibrium physics from the microscopic particle systems. In particular, one of the great challenge would be to recover the hydrodynamic equations from Newton's laws of motion governing the (classical) atom dynamics. In this section, we will review the different levels of description of a gas and their relations by scaling limits.



At a macroscopic scale, a gas can be seen as a continuous substance described by a few parameters varying in time and space: the local density $\rho(\tau, r)$, the local velocity $u(\tau, r)$ and the local temperature $T(\tau, r)$ (which is related to the local energy). Depending on the physical regimes, the evolution of the field (ρ, v, T) follows different hydrodynamic equations. For example, the Euler equations

$$\begin{cases} \partial_\tau \rho + \operatorname{div}(\rho v) = 0, \\ \partial_\tau v + (v \cdot \nabla) v + \frac{1}{\rho} \nabla(\rho T) = 0, \\ \partial_\tau T + (v \cdot \nabla) T + \frac{2}{3} T \nabla_x \cdot v = 0, \end{cases}$$

or the incompressible Navier-Stokes equations

$$\partial_\tau v + (v \cdot \nabla) v = -\nabla p + \nu \Delta v \quad \text{with} \quad \nabla_x \cdot v = 0,$$

in the latter case, p stands for the pressure which is determined by the incompressibility constraint.

At the microscopic scale, a gas is made of atoms whose coordinates $\{(x_i(t), v_i(t))\}_{i \leqslant N}$ evolve in time according to Newtonian dynamics

$$\frac{dx_i}{dt} = v_i, \qquad m \frac{dv_i}{dt} = -\frac{1}{\varepsilon} \sum_{j \neq i} \nabla \Phi\Big(\frac{x_i - x_j}{\varepsilon}\Big). \tag{1.8}$$

These equations are the Hamiltonian equations associated with the kinetic energy (1.1) and the potential energy (1.3). The Gibbs measure (2.18) is no longer relevant to describe the atom statistics which are now given by the evolution of a random initial data transported by the Hamiltonian flow (1.8). The macroscopic parameters (ρ, v, T) are related to the microscopic variables by local averaging, however the scaling limit is much more complicated than the thermodynamic limit (1.7) as it depends now on the space and time scales. In fact, the limiting hydrodynamic equations depend on the choice of the scaling and this explain why a fluid can be either modelled by the Euler or the Navier-Stokes equations. There is currently no mathematical derivation of the hydrodynamic equations from a purely deterministic dynamics. However by adding noise to the microscopic dynamics, the Euler equation has been obtained in [27] and the incompressible Navier-Stokes equation was derived in [30] for a class of stochastic particle systems on the cubic lattice.

Kinetic theory is a different perspective on the derivation of the hydrodynamic limits which describes the particle evolution at an intermediate scale, namely the mesoscopic scale. During a macroscopic time, each atom interacts, via the Hamiltonian flow (1.8), with a huge number of other atoms leading to intricate microscopic correlations between the atoms. Mathematically, it is then extremely difficult to recover the macroscopic parameters by averaging out microscopic coordinates with such a complex correlation structure. Kinetic theory describes the system at the mesoscopic scale defined such that typically, an atom encounters only a finite number of collisions. This corresponds to a dilute gas regime with rare collisions. At the mesoscopic scale, the system can be reduced to a probability density $f(t, x, v)$ which records the probability of finding, at time t, an atom at position x with velocity v. This density evolves according to the Boltzmann equation

$$\begin{aligned} &\partial_t f + v \cdot \nabla_x f = \ Q(f, f), \\ &Q(f, f)(x, v) := \int_{\mathbb{S}^2 \times \mathbb{R}^3} \Big(f(x, v') f(x, v_2') - f(x, v) f(x, v_2) \Big) \, b(v_2 - v, \nu) \, dv_2 d\nu, \end{aligned} \tag{1.9}$$

where ν is a unit vector in $\mathbb{R}^3$ and

$$v' = v - \big((v - v_2) \cdot \nu \big) \, \nu, \qquad v_2' = v_2 + \big((v - v_2) \cdot \nu \big) \, \nu.$$

The precise definition of the mesoscopic scaling as well as the heuristic explanation of the Boltzmann equation are postponed to Section 2.2. Compared to the macroscopic field (ρ, v, T), the kinetic theory provides a more detailed description of the system and the macroscopic parameters are deduced by averaging with respect to the velocity

$$\begin{aligned} &\rho(t, x) = \int_{\mathbb{R}^d} f(t, x, v) dv, \quad u(t, x) = \frac{1}{\rho(t, x)} \int_{\mathbb{R}^d} v f(t, x, v) dv, \\ &T(t, x) = \frac{1}{2\rho(t, x)} \left(\int_{\mathbb{R}^d} v^2 f(t, x, v) dv - \left(\int_{\mathbb{R}^d} v f(t, x, v) dv \right)^2 \right). \end{aligned}$$

As a consequence, the fluid hydrodynamic equations can then be recovered by rescaling the Boltzmann equation (see Figure 1.3). These limits, known as the *fast relaxation limits*, are well understood mathematically and we refer the reader to the book [35] for a survey.

To perform the fast relaxation limit, the Boltzmann equation has to be rescaled on large time scales. However the convergence to the Boltzmann equation has been only proven for short time scales so that an important step is currently missing in order to complete the program of kinetic theory and to justify fully the hydrodynamic equations.

1.4 Overview of these notes

In the first part of these notes, the set-up of the kinetic theory is introduced (see Chapter 2) and the Boltzmann equation is derived as the limit of deterministic microscopic dynamics (see Chapter 3). This derivation raises fundamental physical questions because the microscopic evolution is reversible but the limiting kinetic equation is irreversible. This paradox will be solved, in Chapter 3, by a detailed analysis of the propagation of chaos.

The second part of these notes is devoted to the study of the particle dynamics at large times in perturbative regimes. In Chapter 4, we consider a gas at equilibrium and show that the motion of a tagged particle, in this a gas, converges in the kinetic limit to a Brownian motion. This shows that a random process can be obtained as a limit of deterministic dynamics. The corresponding hydrodynamic equation, namely the heat equation, can be derived directly from the particle system.

Chapter 2

Hard-sphere dynamics

2.1 Microscopic description

2.1.1 Definition of hard-sphere dynamics

Hard-sphere dynamics are simply billiard dynamics where the atoms are modelled as balls moving in straight line and undergoing elastic collisions. The number N of balls in the dynamics is larger than in a standard billiard game, say that N is of the order of the Avogadro number 10^{23}. The billiard table is identified as the periodic domain $\mathbb{T}^d = [0,1]^d$ with dimension $d \geqslant 2$. Each particle has a label in $\{1, \dots, N\}$. Particle i is a ball of diameter $\varepsilon > 0$ centered at $x_i \in \mathbb{T}^d$ with velocity $v_i \in \mathbb{R}^d$. The coordinates of particle i are denoted by $z_i = (x_i, v_i)$. These balls are not allowed to overlap, so that the coordinates $Z_N = \{z_1, z_2, \dots, z_N\}$ of a configuration with N particles are restricted to the *phase space*, i.e. to the domain

$$\mathbb{D}_\varepsilon^N := \Big\{ Z_N = (X_N, V_N) \in \mathbb{T}^{dN} \times \mathbb{R}^{dN}, \qquad \forall i \neq j\,, \quad |x_i - x_j| > \varepsilon \Big\}, \tag{2.1}$$

with the notation $X_N = \{x_1, x_2, \dots, x_N\}$ and $V_N = \{v_1, v_2, \dots, v_N\}$. In the following, we will always consider $\varepsilon^d \ll \frac{1}{N}$ so that the phase space is not empty.

In dimension $d = 1$, the particles cannot cross and the dynamics is well understood [37, 12]. Nevertheless, very challenging questions remain if the particles have different masses. We refer to [39, 40] for an account of the open problems for one-dimensional systems. Throughout these notes, we will focus on the case $d \geqslant 2$.

For an initial configuration $Z_N(0)$ in the domain $\mathbb{D}_\varepsilon^N$, the exclusion constraint between the balls holds at any time and the microscopic evolution (1.8) can be reformulated as follows: for all $i \leqslant N$

$$\partial_t x_i(t) = v_i(t)\,, \quad \partial_t v_i(t) = 0, \qquad \text{if} \quad |x_i(t) - x_j(t)| > \varepsilon \quad \text{for all} \quad j \neq i, \tag{2.2}$$

with elastic reflection at collisions between two particles. In this case, if particles i, j with velocities v_i, v_j collide at time t, i.e. $|x_i(t) - x_j(t)| = \varepsilon$, then the outgoing velocities v_i', v_j' are given by

$$\begin{cases} v_i' = v_i - \big((v_i - v_j) \cdot \nu\big)\, \nu, \\ v_j' = v_j + \big((v_i - v_j) \cdot \nu\big)\, \nu, \end{cases} \tag{2.3}$$

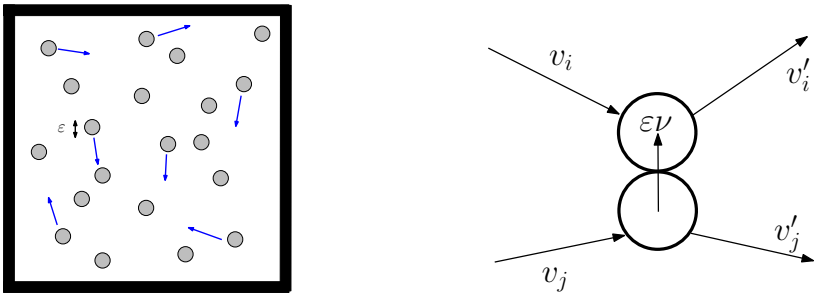


Figure 2.1: On the left, a gas with hard spheres of diameter ε. On the right, a collision between particles i, j is depicted with the notation (2.3).

where ν is the unit vector such that $x_i(t) = x_j(t) + \varepsilon\nu$ (see Figure 2.1). The conservation laws of the dynamics are :

$$\begin{cases} \text{Momentum conservation:} & v_i' + v_j' = v_i + v_j, \\ \text{Energy conservation:} & v_i'^2 + v_j'^2 = v_i^2 + v_j^2. \end{cases} \tag{2.4}$$

Collisions occur on configurations at the boundary $\partial\mathbb{D}_\varepsilon^N$ of the phase space $\mathbb{D}_\varepsilon^N$ (2.1). At a collision between particles i, j, we distinguish the pre-collisional configurations with incoming velocities

$$\partial\mathbb{D}_\varepsilon^{N-}(i,j) := \Big\{ Z_N \in \partial\mathbb{D}_\varepsilon^N, \quad |x_i - x_j| = \varepsilon\,, \quad (v_i - v_j)\cdot(x_i - x_j) < 0 \\ \text{and} \quad \forall (k,\ell) \neq (i,j), \quad |x_k - x_\ell| > \varepsilon \Big\},$$

from the post-collisional configurations with outgoing velocities

$$\partial\mathbb{D}_\varepsilon^{N+}(i,j) := \Big\{ Z_N \in \partial\mathbb{D}_\varepsilon^N, \quad |x_i - x_j| = \varepsilon\,, \quad (v_i - v_j)\cdot(x_i - x_j) > 0 \\ \text{and} \quad \forall (k,\ell) \neq (i,j), \quad |x_k - x_\ell| > \varepsilon \Big\}.$$

Using notation (2.3), we introduce the map

$$\mathcal{J} : (v_i, v_j, \nu) \mapsto (v_i', v_j', -\nu) \tag{2.5}$$

which is an involution $\mathcal{J} \circ \mathcal{J} = Id$. From this property, we deduce that $\mathcal{J}$ preserves the Lebesgue measure in $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{S}^{d-1}$, where $\mathbb{S}^{d-1}$ denotes the unit ball.

Let $\mathcal{R}$ be the map on $\mathbb{D}_\varepsilon^N$ which reverses the velocities

$$\mathcal{R}(Z_N) = (x_1, \dots, x_N, -v_1, \dots, -v_N) \quad \text{with} \quad Z_N = (x_1, \dots, x_N, v_1, \dots, v_N).$$

For any time $t \geqslant 0$, we denote by $\mathbf{T}_t$ the map of the flow at time t so that for an initial data Z_N in $\mathbb{D}_\varepsilon^N$

$$\mathbf{T}_t(Z_N) = Z_N(t). \tag{2.6}$$

A more accurate definition of $\mathbf{T}$ is given in Proposition 2.1. Given a configuration $Z_N(t)$ at time t, the inverse map can be obtained easily by running the dynamics backward, i.e.

by applying $\mathbf{T}_t$ to the configuration with reversed velocities $\mathcal{R}\big(Z_N(t)\big)$ (and reversing the velocities at the final step to recover the initial data). As a consequence, the following identity holds

$$\mathbf{T}_t \mathcal{R} \mathbf{T}_t = \mathcal{R}.$$

Thus the evolution can always be traced back and there is no loss of information by the dynamics. This amounts to say that the hard-sphere flow is *time reversible*. The inverse of $\mathbf{T}_t$ is denoted by $\mathbf{T}_{-t}$.

The particles interact only when they collide so that the corresponding interaction potential is degenerate

$$\Phi(r) = \begin{cases} \infty, & \text{if} \quad r < \varepsilon, \\ 0, & \text{if} \quad r \geqslant \varepsilon, \end{cases} \tag{2.7}$$

and the Hamiltonian of a configuration Z_N in $\mathbb{D}_\varepsilon^N$ depends only on the kinetic energy

$$H_N(Z_N) := \frac{1}{2} \sum_{i=1}^{N} |v_i|^2. \tag{2.8}$$

By (2.4), it is a conserved quantity for the dynamics. In particular, the sets of bounded energy

$$\forall R > 0, \qquad \mathbb{D}_\varepsilon^{N,R} = \Big\{ Z_N \in \mathbb{D}_\varepsilon^N, \qquad H_N(Z_N) \leqslant R^2 \Big\}, \tag{2.9}$$

are invariant by the dynamics.

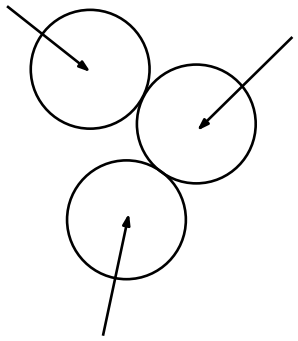


Figure 2.2: The hard-sphere flow is not well defined if 3 or more particles collide at the same time. However triple collisions never occur for almost all initial configurations.

As the potential (2.7) is degenerate the relation with (1.8) is not straightforward. In fact, the hard-sphere flow is not well defined for all initial data in $\mathbb{D}_\varepsilon^N$. For example, there is no dynamical rule to take into account the simultaneous collision of 3 particles (see Figure 2.2). However, the following proposition ensures that for almost all initial data (with respect to the Lebesgue measure in $\mathbb{D}_\varepsilon^N$), the hard-sphere flow is well defined at any time.

Proposition 2.1. *The hard-sphere flow is well defined for almost all configurations in $\mathbb{D}_\varepsilon^N$. Furthermore outside a set of measure 0, the map $\mathbf{T}$, introduced in* (2.6)*, is a one parameter group defined for all $t \in \mathbb{R}$ which preserves Lebesgue measure.*

This result has been first proved by Alexander [3], see also [10, 14]. The singularity of the hard-sphere dynamics is a serious technical drawback to prove the kinetic limit. However the microscopic collisions in the hard-sphere dynamics have a structure very similar to the collisions in the Boltzmann equation (see (2.41)) and this will greatly simplify the derivation of the convergence to the Boltzmann equation in Chapter 3.

Proof. To define the hard-sphere flow, we have to exclude the pathological configurations leading to one of the following events :

- simultaneous collisions of 3 or more particles,
- grazing collisions, i.e. collisions such that $(v_i - v_j) \cdot \nu = 0$ (with notation (2.3)),
- clustering, i.e. infinite number of collisions in a finite time.

Our goal is to build a dense open subset in $\mathbb{D}_\varepsilon^N$ consisting of configurations with only finitely many (pairwise) collisions in any time interval $[0, T]$, so that any of the previous events will be excluded.

On a given time interval $[0, T]$, we are going to show that the flow is well defined on the set $\mathbb{D}_\varepsilon^{N,R}$ of configurations with bounded energy (2.9) outside a set $\mathcal{S}_{T,R}$ of zero measure. Then taking diverging sequences of times T_n and energy cut-off R_n, the dynamics can be defined in $\mathbb{D}_\varepsilon^N$ for any time and any initial data outside the set $\bigcup_n \mathcal{S}_{T_n,R_n}$ of measure zero.

We split $[0, T]$ into small time intervals of size $\delta > 0$ and consider the set

$$\mathbf{P}_\delta = \Big\{ Z_N \in \mathbb{D}_\varepsilon^{N,R}, \quad \text{a particle collides twice during the time interval } [0, \delta] \quad (2.10)$$
$$\text{or has a grazing collision} \Big\}.$$

During $[0, \delta]$, the flow is well defined outside the set $\mathbf{P}_\delta$ as each particle undergoes at most one collision. For any time $t \in [0, \delta]$, the mapping $\mathbf{T}_t$ is well defined in the complement of $\mathbf{P}_\delta$ which turns out to be an open set. Furthermore, any initial data in $\mathbf{P}_\delta^c$ leads to a trajectory with a finite number of distinct collisions which depends smoothly on the initial data. We stress the fact that the configurations with grazing collisions have been neglected as the trajectories are not smooth in their vicinity. The mapping $\mathbf{T}_t$ is a diffeomorphism which leaves Lebesgue measure invariant. This latter point can be checked as the free flow (2.2) and the reflections (2.3) have both Jacobian equal to 1 (see [2] for details).

We are going to show that the measure of $\mathbf{P}_\delta$ is bounded from above by

$$|\mathbf{P}_\delta| = \int_{\mathbf{P}_\delta} dZ_N \leqslant C\delta^2, \quad (2.11)$$

where the constant C depends on R, N, ε. Suppose that particle i collides with the two particles j, k in a time smaller than δ. As the velocities are all bounded by R, the particles j, k have to be close to particle i initially

$$\varepsilon \leqslant |x_i - x_j| \leqslant \varepsilon + 2R\delta \quad \text{and} \quad \varepsilon \leqslant |x_i - x_k| \leqslant \varepsilon + 2R\delta,$$

where the lower bound is due to the hard-sphere constraint. The factor δ^2 in (2.11) comes from the integration with respect to $dx_j dx_k$ as both particles j,k are located in a shell of width $2R\delta$ around the ball i :

$$\forall x_i \in \mathbb{T}^d, \qquad \int dx_j dx_k \, \mathbf{1}_{\{\varepsilon \leqslant |x_i - x_j| \leqslant \varepsilon + 2R\delta\}} \, \mathbf{1}_{\{\varepsilon \leqslant |x_i - x_k| \leqslant \varepsilon + 2R\delta\}} \leqslant \left(cR\varepsilon^{d-1}\delta\right)^2,$$

The remaining variables in (2.11) are then integrated and the integral is estimated by the crude upper bound C.

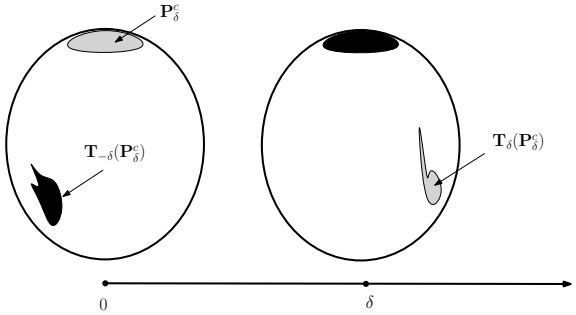


Figure 2.3: The phase space $\mathbb{D}_\varepsilon^{N,R}$ is schematically represented by an ellipse. At time 0, the subset $\mathbf{P}_\delta$ of configurations (represented in gray) is removed from $\mathbb{D}_\varepsilon^{N,R}$. At the later time δ, the subset $\mathbf{T}_\delta\big(\mathbf{P}_\delta^c\big) \cap \mathbf{P}_\delta \subset \mathbf{P}_\delta$ of configurations (represented in black) is removed again. This boils down to remove its pre-image $\mathbf{P}_\delta^{(1)} = \mathbf{T}_{-\delta}\big(\mathbf{T}_\delta\big(\mathbf{P}_\delta^c\big) \cap \mathbf{P}_\delta\big)$ at time 0. The flow can be constructed up to time T by removing subsets of the phase space at each time step δ.

The flow is well defined up to time δ for initial data in $\mathbf{P}_\delta^c$ and the configurations are mapped to the set $\mathbf{T}_\delta\big(\mathbf{P}_\delta^c\big)$ at time δ. To extend the dynamics on $[\delta, 2\delta]$, we consider only the configurations at time δ for which each particle undergoes at most one collision in $[\delta, 2\delta]$, i.e. that the configurations in the set $\hat{\mathbf{P}}_\delta^{(1)} = \mathbf{T}_\delta\big(\mathbf{P}_\delta^c\big) \cap \mathbf{P}_\delta$ at time δ are discarded. This amounts to neglecting also the initial configurations in $\mathbf{P}_\delta^{(1)} = \mathbf{T}_{-\delta}\big(\hat{\mathbf{P}}_\delta^{(1)}\big)$. As the mapping $\mathbf{T}$ leaves Lebesgue measure invariant, the measure of $\mathbf{P}_\delta^{(1)}$ will be less than $C\delta^2$ thanks to (2.11). For any initial data outside the set $\mathbf{P}_\delta \cup \mathbf{P}_\delta^{(1)}$, the dynamics is well defined up to time 2δ. One can then iterate this procedure and extend the flow on the interval $[k\delta, (k+1)\delta]$ by removing sets of initial configurations of the form $\mathbf{P}_\delta^{(k)}$ with measure less than $C\delta^2$. Splitting the time T into $K = \frac{T}{\delta}$ time intervals, the dynamics can be defined up to time T for initial configurations outside the set

$$\mathbf{P}_\delta\big(T, R\big) = \mathbf{P}_\delta \cup \mathbf{P}_\delta^{(1)} \cup \cdots \cup \mathbf{P}_\delta^{(K)}$$

which has a measure bounded by

$$\big|\mathbf{P}_\delta\big(T, R\big)\big| \leqslant (K+1)C\delta^2 \leqslant 2CT\delta.$$

The sets associated with the sequence $\delta_k = 2^{-k}$ are decreasing

$$\mathbf{P}_{\delta_{k+1}}\big(T, R\big) \subset \mathbf{P}_{\delta_k}\big(T, R\big).$$

Thus the set $\bigcap_k \mathbf{P}_{\delta_k}\big(T, R\big)$ has measure 0 and the flow is well defined outside this set. □

2.1.2 Liouville equation

We have seen in Proposition 2.1 that the hard-sphere flow is only defined for almost all initial configurations. From a physical point of view, this is not a serious issue, as the detailed description of a single configuration trajectory is irrelevant. The main goal is to model the statistical properties of the gas and therefore we will focus on the evolution of the probability measures on the configurations.

Initially the configurations Z_N are distributed in $\mathbb{D}_\varepsilon^N$ according to the density $f_N(0, Z_N)$ such that

$$\int_{\mathbb{T}^{dN}\times\mathbb{R}^{dN}} dX_N\, dV_N f_N(0, Z_N) = 1,$$

where dX_N stands for the Lebesgue measure $dx_1 \dots dx_N$ in $\mathbb{T}^{dN}$ and dV_N stands for the Lebesgue measure $dv_1 \dots dv_N$ in $\mathbb{R}^{dN}$. We will also use the notation $dZ_N = dX_N dV_N$. All the particles should behave in the same way, so that we will consider initial measures which are symmetric with respect to the particle labels. This symmetry will be preserved at any time.

According to Proposition 2.1, for almost all initial data Z_N, the configuration $Z_N(t) = \mathbf{T}_t(Z_N)$, at time t, is obtained by the mapping $\mathbf{T}_t$ defined in (2.6). Thus the particle density at time t is the image of the density at time 0 by the mapping $\mathbf{T}_t$

$$f_N(t, Z_N) = f_N(0, \mathbf{T}_{-t} Z_N). \tag{2.12}$$

To rephrase this relation in simpler terms, we suppose for a moment that $N = 1$ and we consider the evolution of a single particle $z_1(t)$. The particle moves in straight line according to

$$\forall t \geqslant 0, \qquad x_1(t) = x_1 + v_1 t, \quad v_1(t) = v_1, \tag{2.13}$$

where the coordinates of $x_1(t)$ should be understood modulo 1 so that the configuration remains in the periodic domain $\mathbb{T}^d$. Then the relation (2.12) reads

$$f_1(t, x_1, v_1) = f_1(0, x_1 - v_1 t, v_1).$$

Taking the time derivative, we see that the distribution obeys the transport equation

$$\partial_t f_1(t, x_1, v_1) = -v_1 \cdot \nabla_{x_1} f_1(0, x_1 - v_1 t, v_1) = -v_1 \cdot \nabla_{x_1} f_1(t, x_1, v_1),$$

where we used the notation

$$w \cdot \nabla_x \varphi = \sum_{k=1}^{d} w_k \partial_k \varphi \quad \text{with} \quad w = (w_1, \dots, w_d) \in \mathbb{R}^d$$

and ∂_k is the derivative with respect to the k^{th}-coordinate of $x \in \mathbb{T}^d$.

A similar computation holds also for a hard-sphere gas with N particles since the particles do not interact in between two collisions. Thus the density satisfies the *Liouville equation* in $\mathbb{D}_\varepsilon^N$

$$\partial_t f_N(t, Z_N) + \sum_{i=1}^{N} v_i \cdot \nabla_{x_i} f_N(t, Z_N) = 0. \tag{2.14}$$

However, one has also to prescribe boundary conditions on the density in $\partial\mathbb{D}_\varepsilon^N$ to take into account the collisions. We impose that on each set $\partial\mathbb{D}_\varepsilon^{N+}(i,j)$ of post-collisional configurations the density is given by

$$f_N(t, Z_N') = f_N(t, Z_N), \tag{2.15}$$

where Z_N' is obtained from the configuration $Z_N \in \partial\mathbb{D}_\varepsilon^{N-}(i,j)$ by changing the velocities v_i, v_j into v_i', v_j' according to the collision rule (2.3). Finally, we define $\mathbf{S}_N$ the group associated with the transport in $\mathbb{D}_\varepsilon^N$ by

$$\forall \varphi \in \mathbb{L}^\infty(\mathbb{D}_\varepsilon^N), \qquad \mathbf{S}_N(t)\varphi : Z_N \mapsto \varphi\big(\mathbf{T}_{-t}(Z_N)\big). \tag{2.16}$$

In particular, the density at time t defined in (2.12) can be rewritten

$$f_N(t, Z_N) = \Big(\mathbf{S}_N(t) f_N(0)\Big)(Z_N).$$

So far the Liouville equation was interpreted as the evolution equation for probability measures, but it can be solved for more general initial data. The following proposition will be used, later on, to derive uniform estimates on the density of the hard-sphere dynamics.

Proposition 2.2 (Maximum principle)**.** *Let g_N, h_N be two solutions of the Liouville equation with initial data such that $g_N(0, Z_N) \leqslant h_N(0, Z_N)$ for almost all configuration Z_N in $\mathbb{D}_\varepsilon^N$, then the order is preserved at any time $t \geqslant 0$*

$$g_N(t, Z_N) \leqslant h_N(t, Z_N),$$

for almost all configuration in $Z_N \in \mathbb{D}_\varepsilon^N$.

Proof. Thanks to (2.12), the solutions at time t can be expressed in terms of the initial data

$$h_N(t, Z_N) - g_N(t, Z_N) = h_N(0, \mathbf{T}_{-t} Z_N) - g_N(0, \mathbf{T}_{-t} Z_N) \geqslant 0.$$

This completes Proposition 2.2. □

2.1.3 Invariant measures

Before studying the evolution of the particle density, we are going to focus on the stationary solutions of the Liouville equation (2.14) which will play a key role in these notes. For $s \geqslant 1$ and $\beta > 0$, the Maxwellian distribution in $\mathbb{R}^{ds}$ is denoted by

$$M_\beta^{\otimes s}(V_s) := \prod_{i=1}^{s} M_\beta(v_i) \quad \text{with} \quad M_\beta(v) := \left(\frac{\beta}{2\pi}\right)^{\frac{d}{2}} \exp\left(-\frac{\beta}{2}|v|^2\right). \tag{2.17}$$

In physics, the parameter β is interpreted as the inverse of a temperature. The *Gibbs measure* on the particle configurations in $\mathbb{T}^{dN} \times \mathbb{R}^{dN}$ is defined by

$$M_{N,\beta}(Z_N) := \frac{1}{\mathcal{Z}_N}\left(\frac{\beta}{2\pi}\right)^{\frac{dN}{2}} \exp(-\beta H_N(V_N))\, \mathbf{1}_{\mathbb{D}_\varepsilon^N}(Z_N) = \frac{1}{\mathcal{Z}_N}\, \mathbf{1}_{\mathbb{D}_\varepsilon^N}(Z_N) M_\beta^{\otimes N}(V_N), \tag{2.18}$$

where the Hamiltonian H_N was introduced in (2.8) and the partition function $\mathcal{Z}_N$ is the normalization factor

$$\mathcal{Z}_N := \int_{\mathbb{T}^{dN}\times\mathbb{R}^{dN}} \mathbf{1}_{\mathbb{D}_\varepsilon^N}(Z_N) M_\beta^{\otimes N}(V_N)\, dZ_N = \int_{\mathbb{T}^{dN}} \prod_{1\leqslant i\neq j\leqslant N} \mathbf{1}_{|x_i-x_j|>\varepsilon}\, dX_N. \tag{2.19}$$

Under the Gibbs measure, only the positions are correlated by the exclusion constraint between the particles. We stress the fact that the particles play a symmetric role so that the measure is *exchangeable*.

Proposition 2.3. *The Gibbs measure $M_{N,\beta}$ is an* invariant measure *for the gas dynamics.*

Proof. The Gibbs measure is a solution of the Liouville equation (2.14) as it is constant in time and with respect to the positions of the configurations in $\mathbb{D}_\varepsilon^N$. Furthermore the boundary conditions (2.15) are satisfied thanks to the conservation of the kinetic energy (2.4) by the elastic collisions. □

Under the Gibbs measure the positions of the particles are uniformly distributed in the set $\{X_N, \quad i\neq j,\ |x_i-x_j|\geqslant\varepsilon\}$. The effect of the exclusion strongly depends on the particle density

$$\rho := N\kappa_d\varepsilon^d, \tag{2.20}$$

where $\kappa_d\varepsilon^d$ stands for the volume of a particle. At large density, the correlations between the particles are so strong that physicists expect the occurence of a phase transition [1, 42]. This has been observed numerically, but there is no mathematical proof of it yet. The behavior of the Gibbs measure is much simpler at low density. Indeed when the diameter ε of the particles tends to 0, the exclusion constraint between the particles becomes less relevant. At $\varepsilon = 0$, the positions are independent and uniformly distributed in $\mathbb{T}^{dN}$ under the measure $M_{N,\beta}$ as there is no interaction. Thus we expect that for ρ small enough, the Gibbs measure behaves as a perturbation of the product measure.

To quantify the properties of the Gibbs measure $M_{N,\beta}$, we define for any fixed $s\geqslant 1$, the marginal of order s by

$$M_{N,\beta}^{(s)}(Z_s) := \int M_{N,\beta}(Z_N)\, dz_{s+1}\dots dz_N. \tag{2.21}$$

If the particles were not interacting then the measure $M_{N,\beta}^{(s)}$ would be equal to the Gibbs measure $M_{s,\beta}$. The following proposition shows that both measures remain close at very low density, in particular the particles become asymptotically independent when $\rho\to 0$. Note that $M_{s,\beta}$ is equal to the measure $M_\beta^{\otimes s}\mathbf{1}_{\mathbb{D}_\varepsilon^s}$ up to the normalization factor $\mathcal{Z}_s$.

Proposition 2.4. *Given $\beta>0$, there is a constant $C>0$ such that for any $\rho\leqslant 1/2$ the following bound holds uniformly in N and $s\leqslant N$.*

$$\left|\left(M_{N,\beta}^{(s)} - M_\beta^{\otimes s}\right)\mathbf{1}_{\mathbb{D}_\varepsilon^s}\right| \leqslant C^s\,\rho\, M_\beta^{\otimes s}. \tag{2.22}$$

Better estimates can be obtained and we refer the reader to [34] for a detail account.

Proof.
Step 1. We are going to derive the following estimate on the partition function (2.19)

$$\forall s \leqslant N, \qquad 1 \leqslant \frac{\mathcal{Z}_{N-s}}{\mathcal{Z}_N} \leqslant \big(1-\rho\big)^{-s}. \tag{2.23}$$

The first inequality follows by removing the exclusion constraint on the last s particles

$$\mathcal{Z}_N \leqslant \mathcal{Z}_{N-s}\,.$$

Let us prove the second inequality in (2.23). Recall that

$$\mathcal{Z}_{s+1} = \int_{\mathbb{T}^{d(s+1)}} \Big(\prod_{1 \leqslant i \neq j \leqslant s+1} \mathbf{1}_{|x_i - x_j| > \varepsilon} \Big)\, dX_{s+1}\,.$$

By Fubini's theorem, we deduce that

$$\mathcal{Z}_{s+1} = \int_{\mathbb{T}^{ds}} \left(\int_{\mathbb{T}^d} \Big(\prod_{1 \leqslant i \leqslant s} \mathbf{1}_{|x_i - x_{s+1}| > \varepsilon} \Big)\, dx_{s+1} \right) \Big(\prod_{1 \leqslant i \neq j \leqslant s} \mathbf{1}_{|x_i - x_j| > \varepsilon} \Big)\, dX_s\,.$$

Since the volume excluded by the s particles is at most $s\kappa_d\varepsilon^d$, we get

$$\int_{\mathbb{T}^d} \Big(\prod_{1 \leqslant i \leqslant s} \mathbf{1}_{|x_i - x_{s+1}| > \varepsilon} \Big)\, dx_{s+1} \geqslant 1 - s\kappa_d\varepsilon^d,$$

where $\kappa_d\varepsilon^d$ stands for the volume of one particle. This implies the lower bound

$$\mathcal{Z}_{s+1} \geqslant \mathcal{Z}_s(1 - s\kappa_d\varepsilon^d) \geqslant \mathcal{Z}_s(1-\rho)\,,$$

where we used $s \leqslant N$ and $\rho = N\kappa_d\varepsilon^d$. Inequality (2.23) is then completed by induction

$$\mathcal{Z}_N \geqslant \mathcal{Z}_{N-s}\big(1-\rho\big)^s.$$

Step 2. We derive now Inequality (2.44). For $s \leqslant N$, the marginal is given by

$$\begin{aligned} M^{(s)}_{N,\beta}(Z_s) = \frac{1}{\mathcal{Z}_N} \mathbf{1}_{\{Z_s \in \mathbb{D}^s_\varepsilon\}}\, M_\beta^{\otimes s}\big(V_s\big) \int_{\mathbb{R}^{d(N-s)}} M_\beta^{\otimes(N-s)}\big(V_{s+1,N}\big)\, dV_{s+1,N} \\ \times \int_{\mathbb{T}^{d(N-s)}} \left(\prod_{s+1 \leqslant k \neq \ell \leqslant N} \mathbf{1}_{|x_k - x_\ell| > \varepsilon} \right) \left(\prod_{i \leqslant s < j} \mathbf{1}_{|x_i - x_j| > \varepsilon} \right) dX_{s+1,N}, \end{aligned}$$

with the notation

$$dX_{s+1,N} := dx_{s+1} \dots dx_N \quad \text{and} \quad dV_{s+1,N} := dv_{s+1} \dots dv_N.$$

As the velocities are decoupled from the positions, the measure on the velocities factorizes. It remains to estimate the effect of the interaction between the positions of the first s particles and the rest of the system. Using the symmetry, we get that

$$M^{(s)}_{N,\beta} = \frac{1}{\mathcal{Z}_N} \mathbf{1}_{\{Z_s \in \mathbb{D}^s_\varepsilon\}} M_\beta^{\otimes s} \Big(\mathcal{Z}_{N-s} - \mathcal{Z}^{\flat}_{(s+1,N)} \Big) \tag{2.24}$$

with the notation

$$\mathcal{Z}^{\flat}_{(s+1,N)} := \int_{\mathbb{T}^{d(N-s)}} \Big(1 - \prod_{i \leqslant s<j} \mathbf{1}_{|x_i - x_j|>\varepsilon}\Big) \prod_{s+1 \leqslant k \neq \ell \leqslant N} \mathbf{1}_{|x_k - x_\ell|>\varepsilon}\, dX_{s+1,N}.$$

This leads to the following decomposition

$$\mathbf{1}_{\{Z_s \in \mathbb{D}^s_\varepsilon\}} \left(M_\beta^{\otimes s} - M^{(s)}_{N,\beta}\right) = \Big(1 - \frac{\mathcal{Z}_{N-s}}{\mathcal{Z}_N}\Big)\mathbf{1}_{\{Z_s \in \mathbb{D}^s_\varepsilon\}} M_\beta^{\otimes s} + \frac{\mathcal{Z}^{\flat}_{(s+1,N)}}{\mathcal{Z}_N}\mathbf{1}_{\{Z_s \in \mathbb{D}^s_\varepsilon\}} M_\beta^{\otimes s}. \quad (2.25)$$

Inequality (2.23) implies that there is $C > 0$ such that for all $\rho \leqslant 1/2$

$$\forall s \leqslant N, \qquad \Big|1 - \frac{\mathcal{Z}_{N-s}}{\mathcal{Z}_N}\Big| \leqslant \big(1-\rho\big)^{-s} - 1 \leqslant C^s \rho.$$

Thus the first term in (2.25) is under control. To estimate the second term, we note that

$$0 \leqslant 1 - \prod_{i \leqslant s<j} \mathbf{1}_{|x_i - x_j|>\varepsilon} \leqslant \sum_{i \leqslant s<j} \mathbf{1}_{|x_i - x_j|<\varepsilon}$$

leads to the bound

$$\mathcal{Z}^{\flat}_{(s+1,N)} \leqslant \sum_{1 \leqslant i \leqslant s} \int_{\mathbb{T}^{d(N-s)}} \Big(\sum_{s+1 \leqslant j \leqslant N} \mathbf{1}_{|x_i - x_j|<\varepsilon}\Big) \prod_{s+1 \leqslant k \neq \ell \leqslant N} \mathbf{1}_{|x_k - x_\ell|>\varepsilon}\, dX_{s+1,N}.$$

Using the symmetry between the particles and Fubini's equality, we get

$$\begin{aligned}
&\int_{\mathbb{T}^{d(N-s)}} \Big(\sum_{s+1 \leqslant j \leqslant N} \mathbf{1}_{|x_i - x_j|<\varepsilon}\Big) \prod_{s+1 \leqslant k \neq l \leqslant N} \mathbf{1}_{|x_k - x_l|>\varepsilon}\, dX_{s+1,N} \\
&\quad \leqslant (N-s) \int_{\mathbb{T}^d} \mathbf{1}_{|x_i - x_{s+1}|<\varepsilon}\, dx_{s+1} \int_{\mathbb{T}^{d(N-s-1)}} \prod_{s+2 \leqslant k \neq l \leqslant N} \mathbf{1}_{|x_k - x_l|>\varepsilon}\, dX_{s+2,N} \\
&\quad = (N-s) \left(\int_{\mathbb{T}^d} \mathbf{1}_{|x_i - x_{s+1}|<\varepsilon}\, dx_{s+1}\right) \times \mathcal{Z}_{N-s-1},
\end{aligned}$$

where, in the second inequality, the interaction between $s+1$ and the rest of the particles has been removed and bounded from above by 1. We deduce that

$$\mathcal{Z}^{\flat}_{(s+1,N)} \leqslant s(N-s)\varepsilon^d \kappa_d\, \mathcal{Z}_{N-s-1} \leqslant s\rho \frac{\mathcal{Z}_{N-s-1}}{\mathcal{Z}_N} \mathcal{Z}_N \leqslant \frac{s}{(1-\rho)^s} \rho \mathcal{Z}_N\,, \quad (2.26)$$

where the last equality follows from (2.23). This shows that the second term in (2.25) is also bounded by $C^s\rho$, for some constant C.

The proof of Proposition 2.4 is complete. □

Before concluding this section, let us stress that the Gibbs measures $M_{N,\beta}$ (2.18) are not the only invariant measures : as the kinetic energy is conserved, any probability distribution of the form $F(H_N)M_{N,\beta}$ is also invariant. Nevertheless, in the large N limit, the Maxwellian turns out to be the only relevant measures for the marginal density. Indeed, for any $\beta > 0$, if one considers $\mathcal{U}_N$ the uniform measure on the set of codimension 1

$$\big\{(V_1, \dots, V_N) \in \mathbb{R}^{dN}; \qquad H_N(V_N) = \beta N\big\}$$

then the marginals converge to a product of Maxwellian when N tends to infinity (this result is known as Poincaré Lemma, see [24])

$$\forall s \geqslant 1, \qquad \mathcal{U}_N^{(s)} \xrightarrow[N\to\infty]{(law)} M_\beta^{\otimes s}.$$

For this reason, it is enough to focus on the measures of the form $M_{N,\beta}$.

2.1.4 BBGKY hierarchy

To describe the coarse grained behavior of a gas, it is enough to investigate the averaged evolution of a few particles as the distribution $f_N(t, Z_N)$, introduced in (2.12), is symmetric with respect to all the labels. Thus for $s < N$, we introduce the marginal

$$f_N^{(s)}(t, Z_s) := \int f_N(t, Z_N) dz_{s+1} \dots dz_N.$$

We are going to define the analog of the Liouville equation (2.14) for the marginals. This is known as the BBGKY hierarchy. The initials stand for Bogoliubov, Born, Green, Kirkwood and Yvon who discovered this hierarchy of equations independently.

In the case $s = 1$, a formal computation shows that

$$\begin{aligned} &\partial_t f_N^{(1)}(t, z_1) + v_1 \cdot \nabla_{x_1} f_N^{(1)}(t, z_1) \\ &\quad = (N-1)\varepsilon^{d-1} \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_N^{(2)}(t, x_1, v_1', x_1 + \varepsilon\nu, v_2') \big((v_2 - v_1)\cdot\nu\big)_+ d\nu dv_2 \\ &\qquad - (N-1)\varepsilon^{d-1} \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_N^{(2)}(t, x_1, v_1, x_1 + \varepsilon\nu, v_2) \big((v_2 - v_1)\cdot\nu\big)_- d\nu dv_2, \end{aligned} \tag{2.27}$$

with the collision rule

$$v_1' := v_1 - \big((v_1 - v_2)\cdot\nu\big)\,\nu, \qquad v_2' := v_2 + \big((v_1 - v_2)\cdot\nu\big)\,\nu. \tag{2.28}$$

Note that this collision rule is the inverse of the mapping (2.3) : v_1', v_2' should be interpreted as incoming velocities.

Equation (2.27) says that particle 1 travels in a straight line before colliding with another of the $(N-1)$ remaining particles. As all the particles play a symmetric role, we can assume that the collision occurs with particle 2 at position $x_2 = x_1 + \varepsilon\nu$. The contribution ε^{d-1} comes from the fact that the collision occurs at the surface of particle 1 (which is a sphere of diameter ε). Finally in the right hand side of the equation, the gain part corresponds to a collision leading to an outgoing velocity v_1, instead the loss part is associated with a collision with an incoming velocity v_1.

Before commenting further on equation (2.27), let us stress that solving the evolution of a single particle density $f_N^{(1)}$ requires the knowledge of the marginal $f_N^{(2)}$ to compute the collision term. Thus we are led to write the family of equations for all the marginals which is known as the *BBGKY hierarchy*. The marginal of order $s < N$ evolves in $\mathbb{D}_\varepsilon^s$ according to

$$\partial_t f_N^{(s)}(t, Z_s) + \sum_{i=1}^{s} v_i \cdot \nabla_{x_i} f_N^{(s)}(t, Z_s) = \big(C_{s,s+1} f_N^{(s+1)}\big)(t, Z_s) \tag{2.29}$$

with the boundary condition in $\partial\mathbb{D}^s_\varepsilon$ as in (2.15)

$$f_N^{(s)}(t, Z_s') = f_N^{(s)}(t, Z_s).$$

The structure of the collision term is similar to the source term in (2.27)

$$C_{s,s+1} = C^+_{s,s+1} - C^-_{s,s+1}, \tag{2.30}$$

where the gain and the loss part of the operator are defined for any smooth function f_{s+1} by

$$\begin{aligned}
&C^+_{s,s+1} f_{s+1}(Z_s) \\
&\qquad = (N-s)\varepsilon^{d-1} \sum_{i=1}^{s} \int f_{s+1}(\dots, x_i, v_i', \dots, x_i + \varepsilon\nu, v_{s+1}') \Big((v_{s+1} - v_i) \cdot \nu \Big)_+ d\nu dv_{s+1}\,, \\
&C^-_{s,s+1} f_{s+1}(Z_s) \\
&\qquad = (N-s)\varepsilon^{d-1} \sum_{i=1}^{s} \int f_{s+1}(\dots, x_i, v_i, \dots, x_i + \varepsilon\nu, v_{s+1}) \Big((v_{s+1} - v_i) \cdot \nu \Big)_- d\nu dv_{s+1}\,.
\end{aligned}$$

The hierarchy stops at $s = N$ which coincides with Liouville equation (2.14). Thus all the informations on the dynamics are encoded at the level $s = N$ and the other equations are simply consequences of Liouville equation. However the hierarchy will be extremely useful to describe the structure of the correlations between particles.

The collision operators (2.30) require to integrate functions on sets of codimension 1, i.e. on the surface of a particle. The flow is defined only for almost all initial data (see Proposition 2.1), thus the measure f_N and the marginals $f_N^{(s)}$ are also defined almost surely. This lack of regularity means that the equations (2.29) of the BBGKY hierarchy are formal. To make sense of these equations, they need to be rewritten in the mild sense. This has been achieved in [36, 14, 29, 10]. Below, we will follow an approach inspired by [36, 29].

Recall that for any integer s, the group $\mathbf{S}_s$ associated with the transport in $\mathbb{D}^s_\varepsilon$ was defined in (2.16) as

$$\forall \varphi \in \mathbb{L}^\infty(\mathbb{D}^s_\varepsilon), \qquad \mathbf{S}_s(t)\varphi : Z_s \mapsto \varphi\big(\mathbf{T}_{-t}(Z_s)\big). \tag{2.31}$$

We are going to derive a mild version of the formal equation (2.29).

Proposition 2.5. *Assume that the initial distribution $f_N(0)$ is bounded from above and supported in $\mathbb{D}^{N,R}_\varepsilon$ for some R. Then for any times $t, \delta \geqslant 0$, we get*

$$f_N^{(s)}(t+\delta) = \mathbf{S}_s(\delta) f_N^{(s)}(t) + \int_0^\delta d\tau\, \mathbf{S}_s(\delta - \tau) C_{s,s+1} \mathbf{S}_{s+1}(\tau) f_N^{(s+1)}(t) + O(\delta^2), \tag{2.32}$$

where the error term depends on $N, \varepsilon, R, \|f_N(0)\|_\infty$.

Before proving Proposition 2.5, let us first relate it to the BBGKY hierarchy. In formula (2.32), the collision operator $C_{s,s+1}$ does not apply directly to $f_N^{(s+1)}$ but to $\mathbf{S}_{s+1}(\tau) f_N^{(s+1)}$

so that the additional time parameter τ compensates the missing dimension. The role of the parameter τ will be made more transparent in the proof below.

Equation (2.29) can be recovered by taking the time derivative (i.e. by letting δ tend to 0 in (2.32)): the first term $\mathbf{S}_s(\delta)$ is associated with the transport part and the second term formally converges to the collision operator applied to $f_N^{(s+1)}(t)$.

Proof. For simplicity, we choose $t = 0$ as the proof does not depend on the structure of the initial data. In fact, the error term depends on $\|f_N(t)\|_\infty$, but one can deduce from Proposition 2.2 that $\|f_N(t)\|_\infty \leqslant \|f_N(0)\|_\infty$ thus there is no loss of generality to consider only the case $t = 0$. We are going to prove (2.32) in the case $s = 1$

$$f_N^{(1)}(\delta, z_1) = \Big(\mathbf{S}_1(\delta) f_N^{(1)}(0)\Big)(z_1) + \int_0^\delta d\tau \, \Big(\mathbf{S}_1(\delta-\tau) C_{1,2} \mathbf{S}_2(\tau) f_N^{(2)}(0)\Big)(z_1) + O(\delta^2). \tag{2.33}$$

The proof is split into several steps.

Step 1. *Decomposition of the trajectory.*

Let $\mathbb{E}$ be the expectation of the particle trajectories starting from the initial measure $f_N(0)$. In particular, for any function φ in $\mathbb{T}^d \times \mathbb{R}^d$, one can write

$$\mathbb{E}\big(\varphi(z_1(\delta))\big) = \int f_N^{(1)}(\delta, z_1)\varphi(z_1) dz_1.$$

In the previous expectation, the coordinates of the other particles are averaged and their exact positions are unknown. Thus, on a short time scale, the evolution of the first particle will be very similar to a Markov chain with random kicks at random times. By analogy with a Poisson process, the expectation can be decomposed as

$$\begin{aligned}\mathbb{E}\big(\varphi(z_1(\delta))\big) =& \mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has no collisions in } [0,\delta]\}}\Big) \\ &+ \mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has one collision in } [0,\delta]\}}\Big) + O(\delta^2),\end{aligned}$$

where the error term is bounded by $|O(\delta^2)| \leqslant \delta^2 C(N, \varepsilon, R)\|f_N(0)\|_\infty \|\varphi\|_1$. This error term comes from the probability that particle 1 has more than one collision in a time δ and it can be estimated as in the proof of Proposition 2.1. We are now going to evaluate the two other contributions independently.

Step 2. *Evaluating the probability of one collision during* $[0, \delta]$.

For any $i \in \{2, \dots, N\}$, let $\mathcal{A}_i \subset \mathbb{T}^{dN} \times \mathbb{R}^{dN}$ be the set of initial configurations such that particles 1 and i collide during the time interval $[0, \delta]$ and that they have no collision with the other particles. In the set $\mathcal{A}_2$, the trajectories of particles 1 and 2 are strongly constrained. Let $z_1 = (x_1, v_1)$ be the coordinates of particle 1 at time δ and suppose that it collides at time $\tau \in [0, \delta]$ with particle 2 (see Figure 2.4), then the coordinates of both particles at time 0 are given by

$$\Gamma'(z_1, v_2, \tau, \nu) = \Big((x_1 - (\delta-\tau)v_1 - \tau v_1', v_1'), (x_1 - (\delta-\tau)v_1 + \varepsilon\nu - \tau v_2', v_2')\Big).$$

In this expression, v_1, v_2 are outgoing velocities so that $(v_2 - v_1) \cdot \nu > 0$ and v_1', v_2' are the corresponding incoming velocities at time 0 given by (2.28) (see Figure 2.4). The mapping $\Gamma'(z_1, v_2, \tau, \nu)$ is the composition of

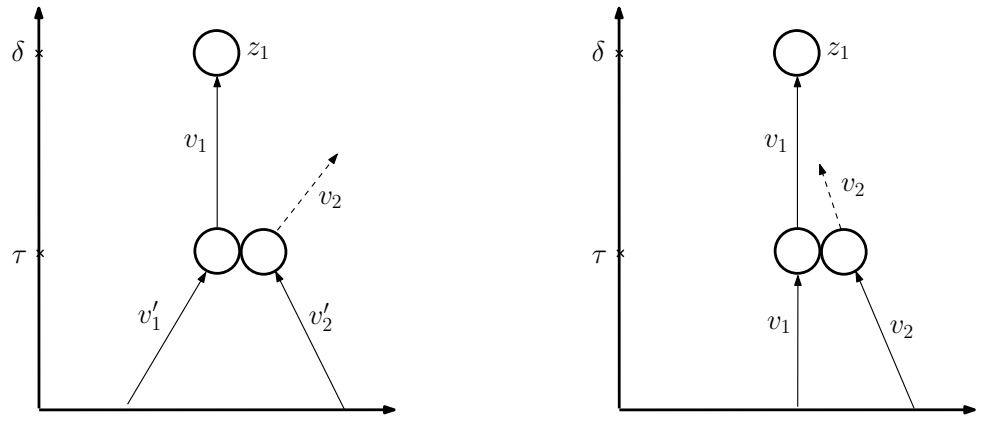


Figure 2.4: In both pictures, the vertical axis represents the time and the horizontal axis represents the space $\mathbb{T}^d$. Following the backward flow from the coordinate $z_1 = (x_1, v_1)$ of the first particle at time δ, a new particle is created at time τ. The left picture depicts the map Γ' associated with the scattering, instead the right picture corresponds to the map Γ.

- the backward transport for the first particle during a time $\delta - \tau$,
- the creation at time τ of a new particle at position $x_1 - (\delta - \tau)v_1 + \varepsilon\nu$ with velocity v_2' and the deflection of the velocity v_1 into v_1',
- the backward transport of both particles from time τ up to time 0.

Thus Γ' is defined on the set

$$\mathbb{G}^+ = \big\{(z_1, v_2, \tau, \nu) \in \mathbb{T}^d \times \mathbb{B}_R \times \mathbb{B}_R \times [0, \delta] \times \mathbb{S}^{d-1}; \qquad (v_2 - v_1) \cdot \nu > 0\big\},$$

where the velocities are restricted to the set $\mathbb{B}_R = \{v \in \mathbb{R}^d, \quad |v| \leqslant R\}$ as $f_N(0)$ is supported in $\mathbb{D}_\varepsilon^{N,R}$. Since the velocities are bounded, the particle trajectories will not wrap around the periodic domain $\mathbb{T}^d$ for δ small enough so that the map Γ' is a bijection from $\mathbb{G}^+$ to $\Gamma'\big(\mathbb{G}^+\big) \subset \mathbb{T}^{2d} \times \mathbb{R}^{2d}$. The following lemma allows us to rephrase the randomness of the initial data in terms of a random scattering of particle 1.

Lemma 2.6. *The change of variables Γ' from $\mathbb{G}^+$ to $\Gamma'\big(\mathbb{G}^+\big) \subset \mathbb{T}^{2d} \times \mathbb{R}^{2d}$ maps the measure $\varepsilon^{d-1}((v_2 - v_1) \cdot \nu)_+ \, dz_1 dv_2 d\tau d\nu$ onto $dz_1 dz_2$.*

This lemma is illustrated in Figure 2.5 and its proof is postponed to the end of this section. As a consequence, the initial configuration in $\mathcal{A}_2$ can be encoded by $\{z_1, v_2, \tau, \nu, Z_{3,N}\}$ where $Z_{3,N} = \{z_3, z_4, \dots, z_N\}$ stands for the remaining particles. Using this change of variables, we can write

$$\begin{aligned}\mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has one collision in } [0,\delta] \text{ with } 2\}}\Big) &= \int dZ_N f_N(0, Z_N)\ \varphi(z_1(\delta))\mathbf{1}_{\{Z_N \in \mathcal{A}_2\}} \\ = \varepsilon^{d-1} \int ((v_2 - v_1) \cdot \nu)_+ \, dz_1 dv_2 d\tau d\nu \int dZ_{3,N} \mathbf{1}_{\{(\Gamma'(z_1, v_2, \tau, \nu), Z_{3,N}) \in \mathcal{A}_2\}} \\ \times f_N\big(0, \Gamma'(z_1, v_2, \tau, \nu), Z_{3,N}\big)\ \varphi(z_1).\end{aligned}$$

The set $\mathcal{A}_2$ imposes two constraints on the initial data : first that particles 1 and 2 collide, second that the other particles do not collide with 1,2. Since the first constraint has been

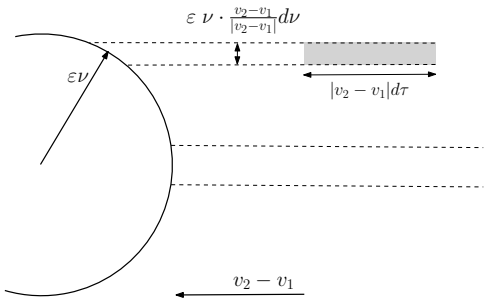


Figure 2.5: This picture illustrates the change of measure of Lemma 2.6 in dimension 2. The position x_2 is parametrized by a point on the surface of the particle 1 (determined by the vector $\varepsilon\nu$), the time τ and the relative velocity $v_2 - v_1$. Thus a small variation of time $d\tau$ and angle $d\nu$ means that x_2 will belong to a box of length $|v_2 - v_1|d\tau$ and width $\frac{(\nu\cdot(v_2-v_1))_+}{|v_2-v_1|}d\nu$.

taken into account by the representation $\Gamma'(z_1, v_2, \tau, \nu)$, the event $\mathcal{A}_2$ can be neglected as any further collision would lead to an event of probability of order δ^2. Thus we get

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has one collision in } [0,\delta] \text{ with } 2\}}\Big) \qquad (2.34)\\
&= \varepsilon^{d-1}\int ((v_2 - v_1)\cdot\nu)_+\, dz_1 dv_2 d\tau d\nu \int dZ_{3,N} f_N\big(0, \Gamma'(z_1, v_2, \tau, \nu), Z_{3,N}\big)\ \varphi(z_1) + O(\delta^2)\\
&= \varepsilon^{d-1}\int f_N^{(2)}\big(0, \Gamma'(z_1, v_2, \tau, \nu)\big)\ \varphi(z_1)((v_2 - v_1)\cdot\nu)_+\, dz_1 dv_2 d\tau d\nu + O(\delta^2),
\end{aligned}$$

where the second marginal is obtained by integrating over the remaining particles. All the particles play a symmetric role, thus we deduce that

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has one collision in } [0,\delta]\}}\Big)\\
&= (N-1)\varepsilon^{d-1}\int f_N^{(2)}\big(0, \Gamma'(z_1, v_2, \tau, \nu)\big)\ \varphi(z_1)((v_2 - v_1)\cdot\nu)_+\, dz_1 dv_2 d\tau d\nu + O(\delta^2).
\end{aligned}$$

The mapping Γ' can be rewritten in terms of the transport operator $\mathbf{S}_2$ and the creation operator $C_{1,2}^+$ introduced in (2.30)

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has one collision in } [0,\delta]\}}\Big) \qquad (2.35)\\
&\qquad = \int dz_1 \varphi(z_1) \int_0^\delta d\tau \Big(\mathbf{S}_1(\delta-\tau) C_{1,2}^+ \mathbf{S}_2(\tau) f_N^{(2)}(0)\Big)(z_1) + O(\delta^2).
\end{aligned}$$

This allows us to identify one of the contribution in (2.33).

Step 3. *Evaluating the probability of not colliding during* $[0, \delta]$.

We will proceed in a similar way to estimate the probability that particle 1 has no collision in $[0, \delta]$. Any initial data in $\bigcap_{i=2}^N \mathcal{A}_i^c$ would lead to one of the following event during the time interval $[0, \delta]$:

- particle 1 has no collision,
- particle 1 collides more than once,

- a particle collides before colliding with 1.

The last two events involve multiple collisions in the time interval $[0,\delta]$ and their probability can therefore be estimated from above by $O(\delta^2)$. Thus we can write

$$\begin{aligned}
&\mathbb{E}\Big(\varphi(z_1(\delta))\mathbf{1}_{\{1 \text{ has no collisions in } [0,\delta]\}}\Big)\\
&= \int dZ_N f_N(0,Z_N)\ \varphi\big((x_1+\delta v_1,v_1)\big)\ \prod_{i=2}^{N}\Big(1-\mathbf{1}_{\{Z_N\in\mathcal{A}_i\}}\Big)+O(\delta^2)\\
&= \int dZ_N f_N(0,Z_N)\ \varphi\big((x_1+\delta v_1,v_1)\big)-\sum_{i=2}^{N}\int_{\mathcal{A}_i} dZ_N f_N(0,Z_N)\varphi\big((x_1+\delta v_1,v_1)\big)+O(\delta^2)\\
&= \int dz_1 f_N^{(1)}(0,z_1)\ \varphi\big((x_1+\delta v_1,v_1)\big)-(N-1)\int_{\mathcal{A}_2} dZ_N f_N(0,Z_N)\varphi\big((x_1+\delta v_1,v_1)\big)+O(\delta^2),
\end{aligned}$$

where we used, in the second equality, that constraining more than one particle to hit particle 1 has a cost at least δ^2 and, in the third equality, we used the symmetry between the particles.

We are now going to change variable and parametrize the previous formula by the coordinates of particle 1 at time δ. The first term corresponds to the free transport part in (2.33)

$$\begin{aligned}
\int dz_1 f_N^{(1)}(0,z_1)\ \varphi\big((x_1+\delta v_1,v_1)\big) &= \int dz_1 f_N^{(1)}(0,(x_1-\delta v_1,v_1))\ \varphi\big(z_1\big)\\
&= \int dz_1 \varphi\big(z_1\big)\Big(\mathbf{S_1}(\delta) f_N^{(1)}(0)\Big)(z_1).
\end{aligned}$$

We turn now to the second term which involves $\mathcal{A}_2$. Let $z_1=(x_1,v_1)$ be the coordinates of particle 1 at time δ, then by analogy with the mapping Γ', we define

$$\Gamma(z_1,v_2,\tau,\nu)=\Big((x_1-\delta v_1,v_1),(x_1-(\delta-\tau)v_1+\varepsilon\nu-\tau v_2,v_2)\Big), \tag{2.36}$$

for configurations with incoming velocities, i.e. configurations in the set

$$\mathbb{G}^-=\big\{(z_1,v_2,\tau,\nu)\in\mathbb{T}^d\times\mathbb{B}_R\times\mathbb{B}_R\times[0,\delta]\times\mathbb{S}^{d-1};\qquad (v_2-v_1)\cdot\nu<0\big\}.$$

This map should be interpreted as the free transport for particle 1 and a direct collision between particles 1 and 2 at time τ (see Figure 2.4). Following the proof of Lemma 2.6, one can check that Γ is a bijection from $\mathbb{G}^-$ to $\Gamma'\big(\mathbb{G}^-\big)\ \subset\ \mathbb{T}^{2d}\times\mathbb{R}^{2d}$ which maps the measure $\varepsilon^{d-1}((v_2-v_1)\cdot\nu)_-\ dz_1dv_2dsd\nu$ onto dz_1dz_2. Thus we can write

$$\begin{aligned}
&\int_{\mathcal{A}_2} dZ_N f_N(0,Z_N)\varphi\big((x_1+\delta v_1,v_1)\big)\\
&\quad=\varepsilon^{d-1}\int((v_2-v_1)\cdot\nu)_-\ dz_1dv_2d\tau d\nu\int dZ_{3,N}\mathbf{1}_{\{(\Gamma(z_1,v_2,\tau,\nu),Z_{3,N})\in\mathcal{A}_2\}}\\
&\qquad\qquad\times f_N\big(0,\Gamma(z_1,v_2,\tau,\nu),Z_{3,N}\big)\ \varphi(z_1)\\
&\quad=\varepsilon^{d-1}\int f_N^{(2)}\big(0,\Gamma(z_1,v_2,\tau,\nu)\big)\ \varphi(z_1)((v_2-v_1)\cdot\nu)_-\ dz_1dv_2d\tau d\nu+O(\delta^2),
\end{aligned}$$

where the restriction in $\mathcal{A}_2$ has been removed by using the same argument as in (2.34). As in (2.35), this last term can be rewritten in terms of the collision operator

$$(N-1)\int_{\mathcal{A}_2} dZ_N f_N(0,Z_N)\varphi\big((x_1+\delta v_1,v_1)\big)$$
$$=\int dz_1\varphi(z_1)\int_0^\delta d\tau\Big(\mathbf{S}_1(\delta-\tau)C_{1,2}^-\mathbf{S}_2(\tau)f_N^{(2)}(0)\Big)(z_1)+O(\delta^2).$$

This completes the derivation of (2.33) for the case $s=1$.

Step 3. *The marginal of order s.*

The structure of formula (2.32) is similar to the one for a single particle (2.33). In a short time interval, the s particles evolve according to free transport and may collide at most once with one of the $(N-s)$ particles in the background. Thus one would like to consider the s particles as a cloud and decompose its evolution during $[0,\delta]$ into 3 events:

- the cloud is not impacted by background particles,
- one particle in the background collides with the cloud,
- at least two particles in the background collide with the cloud.

When implementing the previous procedure, new difficulties occur if a background particle collides successively with several of the s particles. This cannot be considered as an event of small probability because identity (2.32) holds for almost all the configurations Z_s and not only in average. These events are delicate to control and require to analyse subtle cancellations in (2.32). We refer to [36] for a complete proof. □

Proof of Lemma 2.6. We first derive the change of measure for the simpler map

$$\gamma(z_1,v_2,\tau,\nu)=\Big((x_1-\tau v_1,v_1),(x_1+\varepsilon\nu-\tau v_2,v_2)\Big). \tag{2.37}$$

For simplicity, we restrict to dimension $d=2$ and write the coordinates as

$$x_1=\begin{pmatrix}x_{1,1}\\x_{1,2}\end{pmatrix},\quad v_1=\begin{pmatrix}v_{1,1}\\v_{1,2}\end{pmatrix},\quad v_2=\begin{pmatrix}v_{2,1}\\v_{2,2}\end{pmatrix},\quad \nu=\begin{pmatrix}\cos(\vartheta)\\\sin(\vartheta)\end{pmatrix}.$$

As the velocities are unchanged by γ, we focus on the positions

$$\begin{pmatrix}x_{1,1}+\tau v_{1,1}\\x_{1,2}+\tau v_{1,2}\end{pmatrix}\quad\text{and}\quad\begin{pmatrix}x_{1,1}+\varepsilon\cos(\vartheta)+\tau v_{2,1}\\x_{1,2}+\varepsilon\sin(\vartheta)+\tau v_{2,2}\end{pmatrix}. \tag{2.38}$$

The Jacobian associated with this mapping is

$$\begin{vmatrix}1&0&v_{1,1}&0\\0&1&v_{1,2}&0\\1&0&v_{2,1}&-\varepsilon\sin(\vartheta)\\0&1&v_{2,2}&\varepsilon\cos(\vartheta)\end{vmatrix}=\varepsilon\big(\sin(\vartheta)(v_{1,2}-v_{2,2})+\cos(\vartheta)(v_{1,1}-v_{2,1})\big)=\varepsilon\nu\cdot(v_1-v_2).$$

The map Γ' is obtained from γ by composition with the free transport and then with the scattering map $(v_1,v_2,\nu)\mapsto(v_1',v_2',\nu)$. Both mapping are bijective and have unit Jacobian. This completes the derivation of Lemma 2.6 in dimension 2. A similar proof can be achieved in higher dimension leading to a surface factor ε^{d-1} instead of ε. □

2.2 Boltzmann-Grad limit

As explained in Section 1.3, the ultimate goal is to derive the fluid hydrodynamic equations by taking a scaling limit $N \to \infty$ while keeping the particle density $\rho = N\kappa_d\varepsilon^d$ of order 1. For large N, this leads to very complicated interactions between the particles and there is, so far, no mathematical derivation of the corresponding macroscopic limit (in fact the limit depends on the time and length scales at which the system is observed).

To simplify the problem, one would like to describe the statistical properties of the gas at the intermediate mesoscopic level so that a typical particle has only a finite number of collisions in a time scale of order 1. This can be achieved by considering a very dilute gas regime, i.e. by tuning N and ε according to the *Boltzmann-Grad scaling*

$$N \to \infty,\ \varepsilon \to 0 \qquad \text{such that} \qquad N\varepsilon^{d-1} = \alpha, \tag{2.39}$$

where $\alpha > 0$ is a fixed parameter. In the Boltzmann-Grad scaling, the density ρ scales like $\alpha\varepsilon$ and tends to 0 as ε vanishes (i.e. N when tends to infinity). A typical particle, with velocity of order 1, should collide with another particle in a time of order $1/\alpha$. Indeed a particle moving in a straight line during a time $t = 1/\alpha$ will cover a cylinder of volume of order ε^{d-1}/α (see Figure 2.6) and assuming that the N other particles are uniformly distributed in the domain $\mathbb{T}^d$, then the probability that one particle belongs to this domain is of order $N \times \varepsilon^{d-1}/\alpha = 1$ in the Boltzmann-Grad scaling.

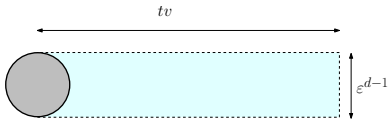


Figure 2.6: During a time t, a particle moving in a straight line covers a cylinder of volume $tv\varepsilon^{d-1}$.

It is expected that asymptotically, in the dilute regime, the density $f(t,x,v)$ of a typical particle in the hard-sphere gas follows an autonomous equation known as *Boltzmann equation*

$$\begin{aligned} &\partial_t f + v\cdot\nabla_x f = \alpha\, Q(f,f),\\ &Q(f,f)(x,v) := \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} \Big(f(x,v')f(x,v_2') - f(x,v)f(x,v_2)\Big)\, b(v_2-v,\nu)\, dv_2 d\nu, \end{aligned} \tag{2.40}$$

where the collision rules are the same as in (2.28)

$$v' = v - \big((v-v_2)\cdot\nu\big)\,\nu, \qquad v_2' = v_2 + \big((v-v_2)\cdot\nu\big)\,\nu.$$

For hard-sphere dynamics, the *cross section* is given by

$$b(v_2-v,\nu) = \big((v_2-v)\cdot\nu\big)_+.$$

More general Newtonian dynamics (1.8) with an interaction potential Φ would lead to a different cross-section b, but the structure of the collision term Q is expected to be unchanged (provided the potential Φ decays fast enough to 0 at infinity).

Before reviewing the mathematical results on the convergence of the particle system, let us give a heuristic justification of the Boltzmann equation. The density $f(t,x,v)$ of a typical particle should be understood as the first marginal $f_N^{(1)}$ in the hard-sphere gas. Equation (2.27) for the evolution of this marginal is strongly reminiscent of the Boltzmann equation

$$\begin{aligned}\partial_t f_N^{(1)}(t,z_1) + v_1 \cdot \nabla_{x_1} f_N^{(1)}(t,z_1) \\ = \underbrace{(N-1)\varepsilon^{d-1}}_{\simeq \alpha} \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} \Big(f_N^{(2)}(t,x_1,v_1',x_1+\varepsilon\nu,v_2') - f_N^{(2)}(t,x_1,v_1,x_1-\varepsilon\nu,v_2)\Big) \\ \times \big((v_2-v_1)\cdot\nu\big)_+ d\nu dv_2,\end{aligned} \tag{2.41}$$

where we used the change of variable $\nu \to -\nu$ in the loss term of the collision operator in order to factorize the cross-section. Note that the incoming particle is located at $x_1 - \varepsilon\nu$ after this change of variables. The Boltzmann-Grad scaling has been precisely tuned so that the prefactor in (2.41) converges to α. However, equation (2.41) keeps still a trace of the microscopic structure through the parameters N, ε. The main difficulty to recover the Boltzmann equation from (2.41) is to rewrite the marginal of order 2 in terms of $f_N^{(1)}$. Boltzmann's intuition was that in a very dilute gas, the particles should be almost independent and therefore the second marginal should factorize when ε is close to 0. In the article [7, 9] published in 1872, Boltzmann stated the *molecular chaos assumption*

$$\begin{aligned} f_N^{(2)}(t,x_1,v_1',x_1+\varepsilon\nu,v_2') &\approx f_N^{(1)}(t,x_1,v_1') f_N^{(1)}(t,x_1,v_2'), \\ f_N^{(2)}(t,x_1,v_1,x_1-\varepsilon\nu,v_2) &\approx f_N^{(1)}(t,x_1,v_1) f_N^{(1)}(t,x_1,v_2), \end{aligned} \tag{2.42}$$

which formally implies the Boltzmann equation from (2.41). Note also that in the limit $\varepsilon \to 0$, the position of both particles are identified. In fact Boltzmann's predictions were much less quantitative than the claim (2.42) and the precise Boltzmann-Grad scaling (2.39) was formalized much later by H. Grad in [17].

The molecular chaos assumption has strong implications as it relates the microscopic and the mesoscopic levels and formally allows to encode the behavior of the gas by a single equation (2.40). In Section 2.3.1, we will see that the Boltzmann equation does not keep track of all the microscopic properties of the system and for this reason, its validity has been challenged by several physicists at the time. Starting from the work by Grad [17], many mathematicians have been working to justify rigorously the molecular chaos assumption and the convergence of the particle system towards the Boltzmann equation. We will review below some of these works.

Let us start by checking the validity of the molecular chaos assumption at time 0 for a large class of initial data. We would like to prepare the microscopic system so that each particle is sampled in $\mathbb{T}^d \times \mathbb{R}^d$ according to a probability measure with a *smooth* density $f_0(x,v)M_\beta(v)$

$$\int_{\mathbb{T}^d\times\mathbb{R}^d} f_0(x,v)M_\beta(v)dxdv = 1.$$

To take into account the exclusion constraint between the hard spheres, we consider ini-

tial distributions of the form

$$f_N(0, Z_N) = \frac{1}{\mathcal{Z}_N^{f_0}} \Big(\prod_{i=1}^{N} f_0(z_i) \Big) \, M_{N,\beta}(Z_N), \tag{2.43}$$

where $\mathcal{Z}_N^{f_0}$ is the normalization of the probability measure. At low density, Proposition 2.4 ensures that the marginals of the Gibbs measures are very close to a product measure. In the Boltzmann-Grad limit, the density ρ vanishes as $\frac{\varepsilon}{\alpha}$ and the proof of Proposition 2.4 can be applied to show that the marginals of the initial data (2.43) factorize asymptotically.

Proposition 2.7. *Given $\beta > 0$, there is a constant $C > 0$ (depending linearly on $\|f_0\|_\infty$) such that in the Boltzmann-Grad limit, the following bound holds uniformly in $s \leqslant N$ for the marginals $f_N^{(s)}(0)$ of the initial data*

$$\left| \left(f_N^{(s)}(0) - f_0^{\otimes s} \, M_\beta^{\otimes s} \right) \mathbf{1}_{\mathcal{D}_\varepsilon^s} \right| \leqslant C^s \, \alpha\varepsilon \, M_\beta^{\otimes s}, \tag{2.44}$$

with the notation

$$f_0^{\otimes s}(Z_s) = \prod_{i=1}^{s} f_0(z_i).$$

This result shows that the molecular assumption (2.42) holds for the initial data (2.43) in the limit $\varepsilon \to 0$.

Remark 2.8. *We stress the fact that the upper bound in* (2.44) *is relevant only for $s \ll \log N$ and globally the particles remain correlated even in the Boltzmann-Grad limit. Indeed, the exclusion constraint amounts to conditioning the Lebesgue measure on $\mathbb{T}^{dN}$ by a set whose probability is given by*

$$\int_{\mathbb{T}^{dN}} \prod_{1 \leqslant i \neq j \leqslant N} \mathbf{1}_{|x_i - x_j| > \varepsilon} \, dX_N = \mathcal{Z}_N,$$

and this probability vanishes in the thermodynamic limit in dimension $d \geqslant 3$. It can be shown that the right hand side of estimate (2.23) *captures the correct scaling of $\mathcal{Z}_N$ in the Boltzmann-Grad limit, i.e.*

$$\mathcal{Z}_N \simeq \exp\left(-N\alpha\varepsilon\right) \simeq \exp\left(-\alpha^{\frac{d}{d-1}} N^{\frac{d-2}{d-1}}\right),$$

where we used that $\varepsilon^{1-d} = \frac{N}{\alpha}$. Thus if $d \geqslant 3$, then $\mathcal{Z}_N$ tends to 0 which means that in the thermodynamic limit, the Gibbs measure and the Lebesgue measure become singular.

In a breakthrough work [22], Lanford devised an amazing strategy of proof to establish the convergence of the hard-sphere dynamics in the low density limit. The proof has then been improved in a series of works by Cercignani, Illner and Pulvirenti [10], Uchiyama [43], Spohn [38], Gallagher, Saint-Raymond and Texier [14] leading to the following theorem.

Theorem 2.9. *Consider a gas of N hard spheres, initially distributed according to the distribution $f_N(0)$ introduced in* (2.43) *where $f_0 M_\beta$ is a* continuous *probability density. Then, there exists $T^* > 0$ (depending only on β and $\|f_0\|_\infty$) such that, in the Boltzmann-Grad limit* (2.39)*, the first marginal $f_N^{(1)}$ converges on the time interval $[0, T^*/\alpha]$ to the solution of the Boltzmann equation* (2.40).

The case of particles interacting with a short range potential Φ (1.8) was first investigated by King [20] and a complete proof can be found in [14, 28]. Theorem 2.9 will be derived in Chapter 3 as well as a partial form of the molecular chaos (2.42). Note that the convergence is valid only for short times, so that Theorem 2.9 cannot be used to recover the hydrodynamic equations as explained in Section 1.3.

By considering a modified particle dynamics with stochastic collisions, F. Rezakhanlou was able to derive the convergence to the Boltzmann equation for large times [32] as well as the stochastic fluctuations around the limit [31]. We refer to [33] for a review of these results. Boltzmann collision operator has a strong probabilist flavor and several stochastic particle systems have been devised to derive the homogeneous Boltzmann equation. We will not comment further on these approaches as they follow a different route from the one presented in these notes, but we refer the reader to [8, 26] for an overview.

2.3 The paradox of the irreversibility

In 1872, Boltzmann's ideas were revolutionary as they opened the way to the description of non-equilibrium phenomena by macroscopic equations. This change of paradigm in the atomistic description led ultimately to the kinetic theory which we know nowadays. However, the Boltzmann equation has first been heavily criticized as it seems to violate some basic physical principles. The first objection was that the time reversibility of the microscopic dynamics is broken by the Boltzmann equation which is irreversible. This is related to the famous $\mathbb{H}$-theorem which will be commented in Section 2.3.1. A second paradox, raised by Zermelo in 1896, comes from an apparent contradiction with the Poincaré recurrence Theorem. Indeed, the hard-sphere flow, which is Hamiltonian, is such that any microscopic trajectory of the configurations will come back arbitrarily close to its initial data after a recurrence time. This again appears to be incompatible with the irreversibility of the Boltzmann equation which relaxes towards an equilibrium. In Sections 2.3.2 and 2.3.3, we will introduce two simple dynamics which show that these apparent paradoxes can be solved provided that the microscopic system is considered at an appropriate time scale. We refer to the book [9] for a detailed historical account of the controversies triggered by Boltzmann's work.

2.3.1 $\mathbb{H}$-theorem

Boltzmann understood that the entropy, defined as

$$\mathbb{H}(t) = \int_{\mathbb{T}^d \times \mathbb{R}^d} dxdv\, f(t,x,v) \log f(t,x,v) \tag{2.45}$$

is monotonous along the flow of the Boltzmann equation (see Theorem 2.10 below) and he related this fact to the second law of thermodynamics. The physical consequence of the entropy dissipation is huge as it implies that Boltzmann equation is *irreversible* which is in contrast to the reversibility of the microscopic hard-sphere dynamics. This paradox was raised by Thomson [41] and Loschmidt [25].

The following theorem shows that the entropy decays in time (note that in physics, entropy is defined with the opposite sign and it is therefore increasing).

Theorem 2.10 (𝕀H-Theorem). *The entropy $t \mapsto \mathbb{H}(t)$ associated with a solution of Boltzmann equation is nonincreasing. The microscopic entropy*

$$\widehat{\mathbb{H}}(t) = \int_{\mathbb{D}_\varepsilon^N} dZ_N \, f_N(t, Z_N) \log f(t, Z_N),$$

associated with the hard-sphere dynamics is constant in time.

Proof.
Step 1. *The mesoscopic entropy.*

For any function φ and any x in $\mathbb{T}^d$, we are first going to derive the following preliminary result

$$\begin{aligned}\int Q(f,f)(x,v)\varphi(v)dv =& \frac{1}{4}\int \Big(f(x,v')f(x,v_1') - f(x,v)f(x,v_1)\Big) \\ &\times \Big(\varphi(v)+\varphi(v_1)-\varphi(v')-\varphi(v_1')\Big)\big((v_1-v)\cdot\nu\big)_+ d\nu dv_1 dv.\end{aligned} \tag{2.46}$$

This a consequence from the identities below

$$\begin{aligned}&\int Q(f,f)(x,v)\varphi(v)dv \\ &\quad= \int \Big(f(x,v')f(x,v_1') - f(x,v)f(x,v_1)\Big)\varphi(v)\big((v_1-v)\cdot\nu\big)_+ d\nu dv_1 dv \\ &\quad= \frac{1}{2}\int \Big(f(x,v')f(x,v_1') - f(x,v)f(x,v_1)\Big)\Big(\varphi(v)+\varphi(v_1)\Big)\big((v_1-v)\cdot\nu\big)_+ d\nu dv_1 dv \\ &\quad= -\frac{1}{2}\int \Big(f(x,v')f(x,v_1') - f(x,v)f(x,v_1)\Big)\Big(\varphi(v')+\varphi(v_1')\Big)\big((v_1-v)\cdot\nu\big)_+ d\nu dv_1 dv,\end{aligned}$$

where we used successively the symmetry between the variables v, v_1 and the fact that the mapping $(v, v_1, \nu) \mapsto (v', v_1', \nu)$ is an involution which preserves Lebesgue measure.

As x is fixed, (2.46) can be applied to the function $\varphi(v) = \log f(x,v)$, so that

$$\begin{aligned}\int Q(f,f)\log f(x,v)dv = -\frac{1}{4}\int \Big(f(x,v')f(x,v_1') - f(x,v)f(x,v_1)\Big)\log\frac{f(x,v')f(x,v_1')}{f(x,v)f(x,v_1)} \\ \big((v_1-v)\cdot\nu\big)_+ d\nu dv_1 dv \leqslant 0,\end{aligned} \tag{2.47}$$

where the inequality follows from the fact that

$$\forall a, b \in \mathbb{R}^+, \qquad (a-b)\log\frac{a}{b} \geqslant 0.$$

Taking the derivative of the entropy (2.45) along the flow of the Boltzmann equation, we get

$$\begin{aligned}\partial\mathbb{H}(t) &= \int_{\mathbb{T}^d\times\mathbb{R}^d} dx\, dv\; \partial_t f(t,x,v)\big(1+\log f(t,x,v)\big) \\ &= \int_{\mathbb{T}^d\times\mathbb{R}^d} dx\, dv\; \operatorname{div}\big(vf\log f\big) + Q(f,f)\log f(t,x,v) \leqslant 0,\end{aligned}$$

where the first term in the integral vanishes by integration by part and the second is non positive thanks to (2.47). This completes the first assertion.

Step 2. *The microscopic entropy.*

As the flow $\mathbf{T}_\cdot$ of the microscopic dynamics leaves the Lebesgue measure invariant, we get

$$\widehat{\mathbb{H}}(t) = \int_{\mathbb{D}_\varepsilon^N} dZ_N \, f_N\big(0, \mathbf{T}_{-t}(Z_N)\big) \log f\big(0, \mathbf{T}_{-t}(Z_N)\big) = \widehat{\mathbb{H}}(0).$$

Thus the microscopic entropy is time independent. □

The $\mathbb{H}$-Theorem shows that there is a discrepancy between the reversible microscopic dynamics with constant entropy and the entropy dissipation in the kinetic equation. As a consequence, the molecular chaos assumption (2.42) cannot be valid for the full distribution on N particles. Indeed, if the following approximation by a product measure

$$f_N(t, Z_N) \approx \prod_{i=1}^{N} f(t, x_i, v_i)$$

was correct then both entropies would be comparable

$$\widehat{\mathbb{H}}(t) = \int_{\mathbb{D}_\varepsilon^N} dZ_N \, f_N(t, Z_N) \log f(t, Z_N) \approx N \int_{\mathbb{T}^d \times \mathbb{R}^d} dx dv \, f(t, x, v) \log f(t, x, v) = N \mathbb{H}(t).$$

and this would contradict the $\mathbb{H}$-Theorem. As a consequence, the molecular chaos assumption should be questioned and a mathematical justification will be given in Section 3.5. A more in-depth physical discussion on the irreversibility and the arrow of time can be found in [23].

Even though the microscopic evolution is deterministic, the mesoscopic description of the system has a strong probabilist flavor. In particular, we will see in Chapter 4 that a tagged particle behaves as a brownian motion after rescaling. To emphasize the analogy with stochastic processes, we recall below the well known counterpart of the $\mathbb{H}$-theorem for Markov chains. For Markov chains, the decay of the relative entropy characterizes the relaxation towards the invariant measure.

For simplicity, we consider a Markov chain $\{X_t\}_{t \geqslant 0}$ taking values in $\mathbb{R}$ with semigroup at time t given by $\mathbf{P}_t(x, dy) = P_t(x, y)dy$. Suppose also that there exists an invariant measure $\pi(x)dx$ so that

$$\forall y \in \mathbb{R}, \qquad \int_{\mathbb{R}} dx \; \pi(x) P_t(x, y) = \pi(y).$$

For any measure μ on $\mathbb{R}$, the relative entropy of μ with respect to π is defined as

$$\mathbb{H}(\mu | \pi) = \int_{\mathbb{R}} \left(\frac{\mu(x)}{\pi(x)} \log \frac{\mu(x)}{\pi(x)} \right) \pi(x) \, dx \, .$$

If μ is not absolutely continuous with respect to π, the relative entropy is infinite. Note that the entropy defined in (2.45) is the relative entropy of $f(t, x, v)$ with respect to the measure $M_\beta(v)dxdv$ (which is invariant for the Boltzmann equation) with an additional energy term $\frac{\beta}{2} \int v^2 f(t, x, v) dxdv$. Since the energy is conserved by the Boltzmann equation, the entropy $\mathbb{H}(t)$ or the relative entropy are both nonincreasing.

The following theorem shows that the relative entropy associated with the distribution of the Markov chain is monotonous in time.

Theorem 2.11. *For any initial probability μ and any time $t \geqslant 0$, one has*

$$\mathbb{H}(\mu P_t | \pi) \leq \mathbb{H}(\mu | \pi).$$

Proof. We set $\phi(u) = u \log(u)$, then

$$\begin{aligned}\mathbb{H}(\mu P_t | \pi) &= \int_{\mathbb{R}} \phi\left(\frac{1}{\pi(x)} \int_{\mathbb{R}} dy\, \mu(y) P_t(y,x)\right) \pi(x)\, dx \\ &= \int_{\mathbb{R}} \phi\left(\int_{\mathbb{R}} dy\, \frac{\mu(y)}{\pi(y)} \frac{\pi(y) P_t(y,x)}{\pi(x)}\right) \pi(x)\, dx.\end{aligned}$$

Since π is an invariant measure, one can check that $y \mapsto \frac{\pi(y)P_t(y,x)}{\pi(x)}$ is a probability density on $\mathbb{R}$

$$\int_{\mathbb{R}} dy\, \frac{\pi(y)P_t(y,x)}{\pi(x)} = \frac{1}{\pi(x)} \int_{\mathbb{R}} dy\, \pi(y) P_t(y,x) = \frac{\pi(x)}{\pi(x)} = 1.$$

Thus by Jensen inequality, one has

$$\mathbb{H}(\mu P | \pi) \leqslant \int_{\mathbb{R}} dx \pi(x) \int_{\mathbb{R}} dy\, \frac{\pi(y)P_t(y,x)}{\pi(x)}\, \phi\left(\frac{\mu(y)}{\pi(y)}\right) = \int_{\mathbb{R}} dy \int_{\mathbb{R}} dx \pi(y) P_t(y,x) \phi\left(\frac{\mu(y)}{\pi(y)}\right).$$

Using that $\int_{\mathbb{R}} dx P_t(y,x) = 1$, we conclude Theorem 2.11

$$\mathbb{H}(\mu P_t | \pi) \leqslant \mathbb{H}(\mu | \pi).$$

□

2.3.2 Kac ring model

This section is devoted to a simple example of deterministic dynamics exhibiting both a relaxation towards an equilibrium and a recurrence property. This model was devised by Mark Kac in [19] and a very complete review of the model as well as of its physical implications can be found in the article [16].

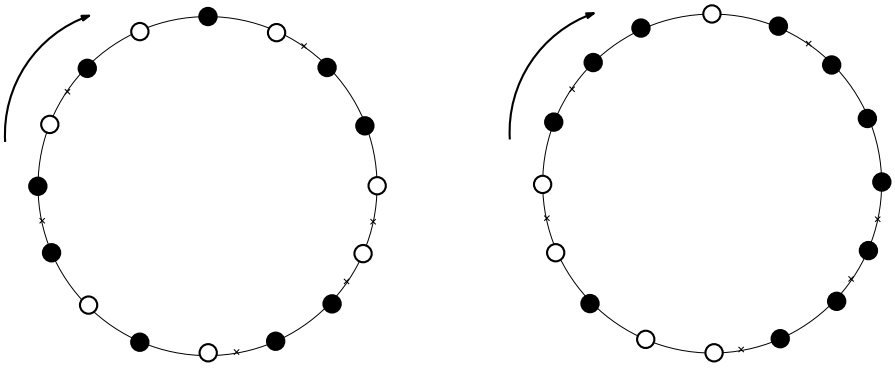


Figure 2.7: Each picture represents a ring with $N = 16$ colored sites and $n = 6$ markers depicted by crosses on some edges. The right configuration is deduced from the left configuration after 1 step of the dynamics.

We consider a system of N sites on a ring and on each site, there is a black or a white ball. Independently on each edge, a marker is placed with probability $p < 1$ or not with probability $1 - p$. The markers are fixed once for all and the initial configuration has N black balls. At each time step, the balls move clockwise to their neighboring sites. If during a move, a ball crosses an edge with a marker, then this ball changes color (see Figure 2.7).

These dynamical rules are deterministic and reversible: by applying the same rules, but rotating counterclockwise the evolution is reversed. Moreover the system is recurrent because after $2N$ steps, each ball will be in its initial position and it will have encountered an even number of markers so that its color will be the same as at time 0. In fact, if the number of markers is even initially, then after only N steps, the initial data is recovered. This second property plays an analogous role to the Poincaré recurrence theorem.

To model the system, it is easier to encode the color of the ball at site i and time t by $\eta_i(t) \in \{-1, 1\}$. The presence of a marker on the edge $(i, i+1)$ is represented by $m_i = -1$ and its absence by $m_i = 1$. The markers are initially placed randomly so that

$$\mathbb{P}(m_i = -1) = p < \frac{1}{2}, \qquad \mathbb{P}(m_i = 1) = 1 - p.$$

Thus the colors change according to the rule

$$\forall i \in \{1, \dots, N\}, \qquad \eta_{i+1}(t+1) = m_i \, \eta_i(t),$$

on the ring the index $i = N + 1$ is identified with $i = 1$. Initially all the balls have the same color $\eta_i(0) = 1$ and we are interested in the mean evolution of these colors

$$\forall t \geqslant 1, \qquad \mathcal{C}(t) = \sum_{i=1}^{N} \eta_i(t) = \sum_{i=1}^{N} m_{i-1} \eta_{i-1}(t-1) = \sum_{i=1}^{N} m_{i-1} m_{i-2} \dots m_{i-t} \eta_{i-t}(0),$$

where the indices should be understood modulo N. Using the periodicity of the domain, we recover that for $\mathcal{C}(2N) = \mathcal{C}(0)$.

Recall that initially $\eta_i(0) = 1$ for all site i. Averaging over the markers and using the fact that all the edges play a symmetric role, we get

$$\mathbb{E}\big(\mathcal{C}(t)\big) = N \mathbb{E}\big(m_1 m_2 \dots m_t\big).$$

For $t < N$, all the markers are independent and the expectation can be computed

$$\mathbb{E}\big(\mathcal{C}(t)\big) = N \mathbb{E}(m_1)^t = N(1 - 2p)^t.$$

Since $t < N$, the expectation decays exponentially fast and the memory of the initial condition is lost. However correlations start building up for $t \geqslant N$ and $\mathbb{E}\big(\mathcal{C}(0)\big) = N$. Thus this deterministic dynamics exhibits both a fast decay to some disordered state and a recurrence property on larger time scales. This example shows that there is no contradiction to describe an Hamiltonian evolution (as the hard-sphere dynamics) by an irreversible equation (as the Boltzmann equation) if the comparison is achieved on an appropriate time scale.

2.3.3 Ehrenfest model

This model has been proposed by Paul and Tatiana Ehrenfest in 1907 [13] in order to settle the controversies on the Boltzmann equation. A beautifully account on this model by Mark Kac can be found in the paper [18].

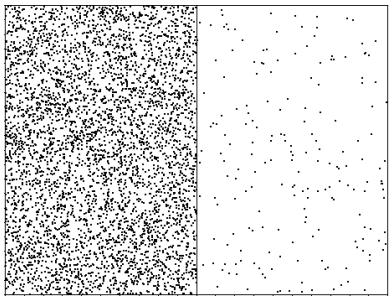


Figure 2.8: At an early stage in the Ehrenfest model, the atoms are mainly on the left side.

Consider a container split into two parts by a wall. The left part of the container is filled by a gas and the right part empty (see Figure 2.8). At time 0, a tiny hole is punctured in the wall so that the gas can reach both sides of the container. The aim of the Ehrenfest model is to describe the relaxation of the gas to its equilibrium state.

To simplify the model, we consider a discrete time and assume that at each time step one atom is chosen randomly among all the atoms and transferred from one side to the other. Let X_n be the number of atoms in the left compartment at time n and suppose that initially $X_0 = K$. Then $\{X_n\}_{n \geqslant 1}$ is a Markov chain taking values in $\{0, \dots, K\}$ with transition matrix

$$\mathbb{P}\big(X_{n+1} = \ell - 1 \big| X_n = \ell\big) = \frac{\ell}{K}, \qquad \mathbb{P}\big(X_{n+1} = \ell + 1 \big| X_n = \ell\big) = \frac{K - \ell}{K}.$$

When the system is at equilibrium, the atoms are uniformly distributed in both sides so that the invariant measure of this Markov chain is

$$\forall \ell \in \{0, \dots, K\}, \qquad \pi(\ell) = \frac{1}{2^K} \binom{K}{\ell}.$$

One can easily check that the Markov chain is reversible with respect to π

$$\pi(\ell) P(\ell, \ell + 1) = \pi(\ell + 1) P(\ell + 1, \ell).$$

This relation can be interpreted as the stochastic counterpart of the reversibility of the hard-sphere dynamics. Furthermore the Markov chain is irreducible which means that it will visit any configuration in $\{0, \dots, K\}$ an infinite number of times. If initially the left compartment contains all the particles, then the atoms will first diffuse in the whole container, but then after a very long time, they will eventually all return to the left side. This behavior is analogous to the one predicted by the Poincaré recurrence theorem for Hamiltonian systems.

Thus this simple stochastic model captures two important features of the gas dynamics which are not part of Boltzmann's theory. Fortunately, Ehrenfest model is much simpler to handle than a hard-sphere gas and exact computations can be achieved in order

to quantify the phenomena at play. If the system starts initially with $X_0 = \ell$, then the expectation of T_ℓ, the first return time of the Markov chain, at ℓ is given by

$$\mathbb{E}_\ell(T_\ell) = \frac{1}{\pi(\ell)} = 2^K \, \frac{\ell!(K-\ell)!}{K!}$$

where the subscript in the expectation $\mathbb{E}_\ell$ stands for the initial state $X_0 = \ell$.

When K is large, say of the order 10^{23}, then

$$\mathbb{E}_K(T_K) = 2^K \quad \text{et} \quad \mathbb{E}_{K/2}(T_{K/2}) \simeq \sqrt{2\pi K}\,.$$

As a consequence, the first return time to $K/2$ (which is the mean at equilibrium) will be infinitely shorter than the return time to K. This latter time is so large that it will never be observed in practice.

The Boltzmann equation describes the typical behavior of a gas and therefore it is conjectured to be valid only on much shorter time scales than the ones involved in the Poincaré recurrence Theorem. Thus the Ehrenfest model solves the apparent contradiction of the paradox raised by Zermelo.

Chapter 3

Convergence to the Boltzmann equation

3.1 Introduction

The goal of this chapter is to prove the convergence, stated in Theorem 2.9, of the hard-sphere dynamics towards the Boltzmann equation. The following result is a more complete version of Theorem 2.9.

Theorem 3.1. *Consider a gas of N hard spheres, initially distributed according to the distribution*

$$f_N(0, Z_N) = \frac{1}{\mathcal{Z}_N^{f_0}} \Big(\prod_{i=1}^{N} f_0(z_i) \Big) \, M_{N,\beta}(Z_N),$$

introduced in (2.43) *where the function f_0 satisfies*

$$\int_{\mathbb{T}^d \times \mathbb{R}^d} dx dv \, f_0(x,v) M_\beta(v) = 1, \qquad \|f_0\|_{\mathbb{L}^\infty} \leqslant \exp(-\mu) \quad \textit{and} \quad \|\nabla_x f_0\|_{\mathbb{L}^\infty} \leqslant C, \tag{3.1}$$

for some μ in $\mathbb{R}$ and $C > 0$.

There exists $T^* > 0$ (depending only on β, μ and on the dimension d) such that the Boltzmann equation (first introduced in (2.40)*)*

$$\partial_t g + v \cdot \nabla_x g = \alpha \, Q(g,g) \quad \textit{with initial data} \quad g(0,z) = f_0(z) M_\beta(v) \tag{3.2}$$

has a unique solution on the time interval $[0, T^*/\alpha]$ and the first marginal of the particle system converges to this solution in the Boltzmann-Grad limit $N\varepsilon^{d-1} = \alpha$ (2.39)

$$\forall t \in \left[0, \frac{T^*}{\alpha}\right], z \in \mathbb{T}^d \times \mathbb{R}^d, \qquad \lim_{N \to \infty} f_N^{(1)}(t,z) = g(t,z). \tag{3.3}$$

Furthermore, the propagation of chaos *holds for any time t in $[0, T^*/\alpha]$ and for almost all configurations Z_s in $\mathbb{T}^{ds} \times \mathbb{R}^{ds}$*

$$\lim_{N \to \infty} f_N^{(s)}(t, Z_s) = \prod_{i=1}^{s} g(t, z_i). \tag{3.4}$$

We stress the fact that the convergence holds only up to a short time of order $1/\alpha$ (depending on the initial data).

3.2 The series expansion

3.2.1 Duhamel representation

The marginals are the relevant quantities to describe the gas and they evolve according to the BBGKY hierarchy introduced in Section 2.1.4. Recall that, the evolution of the marginal $f_N^{(s)}$ of order s is related, formally, to the marginal $f_N^{(s+1)}$ of higher order by the following equation

$$\partial_t f_N^{(s)} + \sum_{i=1}^{s} v_i \cdot \nabla_{x_i} f_N^{(s)} = \alpha C_{s,s+1} f_N^{(s+1)}, \tag{3.5}$$

with specular reflection on the boundary $\partial \mathbb{D}_\varepsilon^s$ and the collision operator $C_{s,s+1}$ is given by

$$\begin{aligned}\big(C_{s,s+1} f_N^{(s+1)}\big)(Z_s) &:= \frac{(N-s)\varepsilon^{d-1}}{\alpha} \\ &\times \Big(\sum_{i=1}^{s} \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_N^{(s+1)}(\dots, x_i, v_i', \dots, x_i + \varepsilon\nu, v_{s+1}') \big((v_{s+1} - v_i)\cdot\nu\big)_+ d\nu dv_{s+1} \\ &- \sum_{i=1}^{s} \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_N^{(s+1)}(\dots, x_i, v_i, \dots, x_i + \varepsilon\nu, v_{s+1}) \big((v_{s+1} - v_i)\cdot\nu\big)_- d\nu dv_{s+1} \Big).\end{aligned} \tag{3.6}$$

Compared to the definition (2.30) of $C_{s,s+1}$, an additional factor α has been added to match the Boltzmann-Grad scaling.

The solution of a linear PDE with a source term can be represented by using Duhamel's principle which we recall next, in the case of a transport equation in $\mathbb{D}_\varepsilon^s$ with a source term U

$$\big(\partial_t + \sum_{i=1}^{s} v_i \cdot \nabla_{x_i}\big) g(t, Z_s) = U(t, Z_s) \quad \text{with} \quad g(t=0) = g_0. \tag{3.7}$$

The solution of the homogenous part of this equation

$$\big(\partial_t + \sum_{i=1}^{s} v_i \cdot \nabla_{x_i}\big) g^1(t, Z_s) = 0 \quad \text{with} \quad g^1(t=0) = g_0,$$

is given in terms of the group $\mathbf{S}_s$ associated with the transport in $\mathbb{D}_\varepsilon^s$ which was introduced in (2.31)

$$g^1(t, Z_s) = \big(\mathbf{S}_s(t) g_0\big)(Z_s) = g_0\big(\mathbf{T}_{-t}(Z_s)\big). \tag{3.8}$$

The part of the equation containing the source term is reduced to

$$\big(\partial_t + \sum_{i=1}^{s} v_i \cdot \nabla_{x_i}\big) g^2(t, Z_s) = U(t, Z_s) \quad \text{with} \quad g^2(t=0) = 0,$$

and the solution is given by

$$g^2(t, Z_s) = \int_0^t dt_1 \big(\mathbf{S}_s(t - t_1) U(t_1)\big)(Z_s).$$

Thus the solution of (3.7) is obtained by adding the expressions for g^1 and g^2

$$g(t, Z_s) = \big(\mathbf{S}_s(t) g_0\big)(Z_s) + \int_0^t dt_1 \big(\mathbf{S}_s(t - t_1) U(t_1)\big)(Z_s).$$

The solution of (3.5) can be rewritten by using Duhamel representation

$$f_N^{(s)}(t) = \mathbf{S}_s(t) f_N^{(s)}(0) + \int_0^t dt_1 \mathbf{S}_s(t-t_1) C_{s,s+1} f_N^{(s+1)}(t_1). \tag{3.9}$$

Note that this representation remains as formal as equation (3.5). Indeed the operator $C_{s,s+1}$ is defined by integrating on sets of codimension 1 and $f_N^{(s+1)}$ is not regular enough to be well defined on such sets. Besides this technical aspect, the main difficulty is that (3.9) relates the marginal $f_N^{(s)}$ at time t to the marginal $f_N^{(s+1)}$ at an intermediate time in $[0,t]$. Since $f_N^{(s+1)}$ is also unknown, the formula does not seem very helpful. The key idea is to iterate Duhamel formula up to time 0, in order to relate the solution of the BBGKY hierarchy to the initial data

$$\begin{aligned} f_N^{(s)}(t) = \sum_{n=0}^{N-s} \alpha^n \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} & \mathbf{S}_s(t-t_1) C_{s,s+1} \mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2} \\ & \dots \mathbf{S}_{s+n}(t_n) f_N^{(s+n)}(0)\, dt_n \dots dt_1, \end{aligned} \tag{3.10}$$

where the term $n=0$ stands for the transport part $\mathbf{S}_s(t) f_N^{(s)}(0)$ without any collision. The order of the marginals in the sum is always less or equal to N so that the sum is finite. The only source of randomness comes from the initial data and formula (3.10) will allow us to transfert this initial information at positive times. In Section 3.3.2 , we will derive uniform estimates in N in order to control the convergence of this sum.

As in Proposition 2.5, the integration on time in (3.10) provides the missing dimension needed to apply the collision operator $C_{s,s+1}$. The iterated Duhamel expansion is derived in the following proposition.

Proposition 3.2. *The marginals of the hard-sphere dynamics are given by*

$$\forall s \leqslant N, \qquad f_N^{(s)}(t) = \sum_{n=0}^{N-s} \alpha^n Q_{s,s+n}(t) f_N^{(s+n)}(0), \tag{3.11}$$

where the operators are defined as $Q_{s,s}(t) = \mathbf{S}_s(t)$ *and for* $n \geqslant 1$

$$Q_{s,s+n}(t) := \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}_s(t-t_1) C_{s,s+1} \mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2} \dots \mathbf{S}_{s+n}(t_n) dt_1 \dots dt_n. \tag{3.12}$$

Proof. By construction, $f_N(0)$ is bounded from above so that the density f_N is also bounded by $\|f_N(0)\|_\infty$ at any time (see Proposition 2.2 on the maximum principle). For the moment, we assume that $f_N(0)$ is supported in $\mathbb{D}_\varepsilon^{N,R}$ for some R. Thus we can use the short time result (2.32) obtained in Proposition 2.5 which is recalled below

$$f_N^{(s)}(t) = \mathbf{S}_s(\delta) f_N^{(s)}(t-\delta) + \alpha \int_{t-\delta}^t dt_1 \mathbf{S}_s(t-t_1) C_{s,s+1} \mathbf{S}_{s+1}\big(t_1-(t-\delta)\big) f_N^{(s+1)}(t-\delta) + O(\delta^2). \tag{3.13}$$

Compared to (2.32), the factor α comes from the new normalization (3.6) of the collision operator. As in the derivation of the existence result for the dynamics in Proposition 2.1,

we are going to iterate (3.13) and finally, let δ tend to 0. We stress the fact that the error term $O(\delta^2)$ depends on $N, \varepsilon, R, \|f_N(0)\|_\infty$.

Using twice (3.13), we get

$$
\begin{aligned}
f_N^{(s)}(t) &= \mathbf{S}_s(2\delta) f_N^{(s)}(t-2\delta) \\
&+ \int_{t-2\delta}^{t-\delta} dt_1 \mathbf{S}_s(\delta)\mathbf{S}_s(t-\delta-t_1) C_{s,s+1}\mathbf{S}_{s+1}(t_1-(t-2\delta)) f_N^{(s+1)}(t-2\delta) \\
&+ \int_{t-\delta}^{t} dt_1 \mathbf{S}_s(t-t_1) C_{s,s+1}\mathbf{S}_{s+1}(t_1-(t-\delta))\mathbf{S}_s(\delta) f_N^{(s+1)}(t-2\delta) \\
&+ \int_{t-\delta}^{t} dt_1 \int_{t-2\delta}^{t-\delta} dt_2 \mathbf{S}_s(t-t_1) C_{s,s+1}\mathbf{S}_{s+1}(t_1-(t-\delta)) \\
&\qquad \mathbf{S}_s((t-\delta)-t_2) C_{s+1,s+2}\mathbf{S}_{s+1}(t_2-(t-2\delta)) f_N^{(s+1)}(t-2\delta) + O(3\delta^2) \\
&= \mathbf{S}_s(2\delta) f_N^{(s)}(t-2\delta) \\
&+ \int_{t-2\delta}^{t} dt_1 \mathbf{S}_s(t-t_1) C_{s,s+1}\mathbf{S}_{s+1}(t_1-(t-2\delta)) f_N^{(s+1)}(t-2\delta) \\
&+ \int_{t-\delta}^{t} dt_1 \int_{t-2\delta}^{t-\delta} dt_2 \mathbf{S}_s(t-t_1) C_{s,s+1}\mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2}\mathbf{S}_{s+1}(t_2-t-2\delta) f_N^{(s+1)}(t-2\delta) \\
&+ O(3\delta^2).
\end{aligned}
$$

The factor 3 has been added in front of the error term to indicate that each iteration induces new error terms of the same type. Indeed (3.13) contains two terms involving $f_N^{(s)}(t-\delta)$ which lead to two additional contributions of order $O(\delta^2)$.

The time interval $[0,t]$ is split into $K := t/\delta$ intervals of length δ and (3.13) is iterated up to time 0. The order of the marginals has to remain less or equal to N so that there cannot be more than $N-s$ collisions. Once the marginal of order N is reached, it evolves up to time 0 by following the backward hard-sphere flow of the full microscopic dynamics. Note that after k iterations, $f_N^{(s)}(t)$ is decomposed into a sum of at most $N-s$ terms representing the number of collisions up to time $t-k\delta$. Thus the next iteration will produce an error of order at most $N\,O(\delta^2)$. Finally, we get

$$
\forall s \leqslant N, \qquad f_N^{(s)}(t) = \sum_{n=0}^{N-s} \alpha^n Q_{s,s+n}^{\delta}(t) f_N^{(s+n)}(0) + O(K N \delta^2), \tag{3.14}
$$

where the operators are defined as $Q_{s,s}^{\delta}(t) = \mathbf{S}_s(t)$ and for $n \geqslant 1$

$$
Q_{s,s+n}^{\delta}(t) := \int_{\mathcal{T}_n^{\delta}} dT_n \mathbf{S}_s(t-t_1) C_{s,s+1}\mathbf{S}_{s+1}(t_1-t_2) C_{s+1,s+2} \ldots \mathbf{S}_{s+n}(t_n), \tag{3.15}
$$

and the time integral is over the ordered collision times $T_n = (t_1, \ldots, t_n)$ taking values in the set

$$
\begin{aligned}
\mathcal{T}_n^{\delta} := \Big\{ T_n = (t_1, \ldots, t_n); \quad & t_i \in [t-(k_i-1)\delta, t-k_i\delta] \\
& \text{for a sequence of integers} \quad 0 < k_1 < k_2 < \cdots < k_n \leqslant K \Big\}.
\end{aligned}
$$

As in the proof of Proposition 2.1, the error term $O(KN\delta^2)$ vanishes when δ tends to 0 as $K = t/\delta$. Furthermore, the modified truncated operator $Q^{\delta}_{s,s+n}$ converge to the limit $Q_{s,s+n}$ (3.12).

Thus identity (3.11) holds for any R, so that the constraint on the initial data $f_N(0)$ can be relaxed. This completes the proof of Proposition 3.2. □

3.2.2 The limiting hierarchy and the Boltzmann equation

To study the convergence of the Duhamel representation (3.11) in the Boltzmann-Grad limit, we are going to define the limit of the BBGKY hierarchy which is known as the Boltzmann hierarchy.

In the dilute limit $\varepsilon \to 0$, the hard spheres are converging to points, thus the formal limit of the collision operator is given by

$$\begin{aligned}\big(C^0_{s,s+1}g^{(s+1)}\big)(Z_s) := &\sum_{i=1}^{s}\int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} g^{(s+1)}(\dots,x_i,v_i',\dots,x_i,v_{s+1}')\Big((v_{s+1}-v_i)\cdot\nu\Big)_+ d\nu dv_{s+1}\\ &-\sum_{i=1}^{s}\int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} g^{(s+1)}(\dots,x_i,v_i,\dots,x_i,v_{s+1})\Big((v_{s+1}-v_i)\cdot\nu\Big)_- d\nu dv_{s+1},\end{aligned} \tag{3.16}$$

where the collision between particles i and $s+1$ takes place at $x_{s+1} = x_i$. Note also that the prefactor $\frac{(N-s)\varepsilon^{d-1}}{\alpha}$ in the definition (3.6) of $C_{s,s+1}$ has been replaced by 1 in the Boltzmann-Grad limit.

As the number of particles N tends also to infinity, the limiting hierarchy involves an infinite sequence of functions $\{g^{(s)}(t,Z_s)\}_{s\geqslant 1}$. In the limit, the exclusion rule no longer applies so that the function $g^{(s)}(t)$ takes values in $\mathbb{T}^{ds}\times\mathbb{R}^{ds}$ and satisfy the *Boltzmann hierarchy*

$$\partial_t g^{(s)} + \sum_{i=1}^{s} v_i\cdot\nabla_{x_i} g^{(s)} = \alpha C^0_{s,s+1} g^{(s+1)}, \tag{3.17}$$

with initial data

$$g^{(s)}(0,Z_s) = \prod_{i=1}^{s} f_0(z_i)\; M_\beta^{\otimes s}(V_s). \tag{3.18}$$

Note that the functions $\{g^{(s)}(t,Z_s)\}_{s\geqslant 1}$ depend on the parameter α which controls the frequency of the collision operator in (3.17). As α is fixed throughout this section, we will omit this dependence in the notation.

In view of Proposition 2.7, the Boltzmann hierarchy is the natural limit for the marginals $f_N^{(s)}(0)$ at time 0. The solutions of (3.17) will be restricted to continuous function $g^{(s)}$ in $\mathbb{T}^{sd}\times\mathbb{R}^{sd}$ which is possible since the free transport preserves continuity on $\mathbb{T}^{sd}\times\mathbb{R}^{sd}$. The analog of the operator $Q_{s,s+n}$ defined in (3.12) is given by

$$Q^0_{s,s+n}(t) := \int_0^t\int_0^{t_1}\dots\int_0^{t_{n-1}} \mathbf{S}^0_s(t-t_1)C^0_{s,s+1}\mathbf{S}^0_{s+1}(t_1-t_2)C^0_{s+1,s+2}\dots\mathbf{S}^0_{s+n}(t_n)dt_1\dots dt_n, \tag{3.19}$$

where $\mathbf{S}^0_k$ is the group associated with the free flow of k particles on $\mathbb{T}^{dk}\times\mathbb{R}^{dk}$. Then the iterated Duhamel formula for the Boltzmann hierarchy takes the form

$$\forall s\geqslant 1,\qquad g^{(s)}(t) = \sum_{n\geqslant 0}\alpha^n Q^0_{s,s+n}(t) g^{(s+n)}(0). \tag{3.20}$$

The solution of the Boltzmann hierarchy can be explicitly related to the Boltzmann equation.

Lemma 3.3. *Let $g(t,z)$ be a continuous solution of the Boltzmann equation* (3.2) *with initial data $g(0,z) = f_0(z)M_\beta(v)$. Then a solution of the Boltzmann hierarchy* (3.17) *is given by*

$$\forall s \geqslant 1, \qquad g^{(s)}(t, Z_s) = \prod_{i=1}^{s} g(t, z_i). \tag{3.21}$$

We stress the fact that the product structure of the initial data (3.18) is kept at positive times in (3.21). Uniqueness of the solution will be derived in Proposition 3.7.

Proof. By construction the initial data (3.18) coincide with (3.21) at time 0 and (3.21) satisfies the equations (3.17) for all $s \geqslant 1$. □

3.3 Uniform bounds on the Duhamel series

3.3.1 The Cauchy-Kovalevsky argument

There is still no general theory to ensure existence and uniqueness of solutions of the Boltzmann equation (3.2)

$$\partial_t g + v \cdot \nabla_x g = \alpha\, Q(g,g)$$

starting from a general initial data $g(0,z) = f_0(z)M_\beta(v)$. The main difficulty is that the collision operator $Q(\cdot,\cdot)$ involves products of the density which would be naturally controlled by $\mathbb{L}^2$ estimates. Thus the a priori bound on the entropy obtained in $\mathbb{H}$-Theorem 2.10 is not sufficient to provide a control in time of the $\mathbb{L}^2$ norm. Nevertheless, the Cauchy problem can be solved globally for small perturbations around an equilibrium density and a notion of weak solutions has been develop by Di Perna and Lions to show the global existence starting from general initial conditions. Further references on the analytical aspects related to the Boltzmann equation can be found in [10].

Solving the Cauchy problem globally in time requires to understand the cancellations between the gain and loss parts of the collision operator. The derivation of the convergence stated in Theorem 3.1 relies on cruder controls which will ensure the existence and uniqueness only for finite time. Indeed, the collision operators in the hierarchies will be estimated by using operators of the form

$$|Q|(f,f)(x,v) := \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} \Big(f(x,v')f(x,v_1') + f(x,v)f(x,v_1) \Big) \left((v_1 - v)\cdot \nu \right)_+ dv_1 d\nu, \tag{3.22}$$

which do not take into account the cancellations. This procedure will be applied to both hierarchies in Sections 3.3.2 and 3.3.3, but we first discuss some consequences of this approach.

Neglecting the cancellations in the collision operator implies that the solutions cannot be controlled beyond a finite time. Indeed this boils down to considering the following Riccati equation

$$\dot{x}(t) = x(t)^2 \quad \text{with} \quad x(0) = x_0 > 0, \tag{3.23}$$

which blows up in finite time

$$t < \frac{1}{x_0}, \qquad x(t) = \frac{x_0}{1 - t x_0}.$$

Following [15] and [10] (page 104), we are going to explain how equation (3.23) is related to the Boltzmann hierarchy and use this correspondence to discuss the uniqueness of the solutions of the hierarchy. Let us start by considering the following hierarchy of ODE

$$\forall k \geqslant 1, \qquad \dot{y}_k(t) = k\, y_{k+1}(t) \quad \text{with initial data} \quad y_k(0) = y_k^0. \tag{3.24}$$

This hierarchy is related to (3.23) by noticing that

$$\partial_t x(t)^k = k\, x(t)^{k-1}\, \dot{x}(t) = k\, x(t)^{k+1}.$$

One should view the Riccati equation as the analogous of the Boltzmann equation (associated with the operator (3.22)) and the hierarchy (3.24) as a counterpart of the Boltzmann hierarchy introduced in (3.17). This toy model will serve as a benchmark to understand the conditions required for the uniqueness of the solutions of the Boltzmann hierarchy.

First of all, let us check that the uniqueness of the solutions of the hierarchy (3.24) cannot be determined only from the initial data. Indeed, the sequence of constant functions $y_k(t) = 0$ for all k is a solution of the hierarchy (3.24) and we are going to construct another solution with the same initial data. Consider the function

$$y_1(t) = \begin{cases} \exp\left(-\frac{1}{t}\right), & \text{for} \quad t \geqslant 0, \\ 0, & \text{for} \quad t \leqslant 0, \end{cases}$$

which belongs to $C^\infty(\mathbb{R})$ as all its derivatives at 0 satisfy $y_1^{(k)}(0) = 0$. Define recursively a solution of the hierarchy by

$$\forall k \geqslant 1, \qquad y_{k+1}(t) = \frac{1}{k}\dot{y}_k(t) \quad \text{with initial data} \quad y_{k+1}(0) = 0. \tag{3.25}$$

This sequence is also a solution of (3.24) with initial data $y_k^0 = 0$ for all k, but it is not real-analytic. Thus additional conditions on the growth of the solutions are required to recover the uniqueness.

We are going to look for solutions $\{y_k(t)\}_{k \geqslant 1}$ of (3.24) in the time interval $[0, T]$ satisfying the growth condition

$$\sup_{t \in [0,T]} \sup_{k \geqslant 1} \left| y_k(t) \right|^{1/k} < \infty. \tag{3.26}$$

In the spirit of Cauchy-Kovalevsky theory, we consider the series associated with a solution of (3.24)

$$U(t, z) = \sum_{k=1}^{\infty} y_k(t) z^{k-1},$$

which are well defined for z in a neighborhood of 0 under the growth condition (3.26). Differentiating in t, we get that

$$\partial_t U(t,z) = \sum_{k=1}^{\infty} \dot{y}_k(t) z^{k-1} = \sum_{k=1}^{\infty} k y_{k+1}(t) z^{k-1} = \partial_z U(t,z).$$

Thus the sequence of functions $\{y_k(t)\}_{k \geqslant 1}$ is a solution of (3.24) if and only if U solves the partial differential equation

$$\partial_t U(t,z) = \partial_z U(t,z). \tag{3.27}$$

Given an initial data such that

$$\sup_{k \geqslant 1} \left| y_k^0 \right|^{1/k} = \frac{1}{R} < \infty,$$

an explicit solution of (3.27) can be built as follows. Define

$$\forall |z| < R, \qquad U^0(z) = \sum_{k=1}^{\infty} y_k^0 \, z^{k-1},$$

then the unique solution real analytic on the domain $\{(t,z) \in \mathbb{R}^2, \quad |z| + |r| < R\}$ is given by $U(t,z) = U^0(t+z)$. The solution of (3.24) can then be obtained as the Taylor coefficients of the series. The solution of the Riccati equation (3.23) can also be recovered in this way. The growth condition (3.26) is essential to ensure the uniqueness of the solution.

3.3.2 Estimates on the BBGKY hierarchy

This section is devoted to the derivation of uniform estimates on the collision operators of the BBGKY hierarchy. We first start by defining an appropriate functional setting.

Proposition 2.7 implies that the marginals of the initial data are uniformly controlled by product measures

$$\forall k \leqslant N, \qquad \left| \left(f_N^{(k)}(0) - f_0^{\otimes k} \, M_\beta^{\otimes k} \right) \mathbf{1}_{\mathbb{D}_\varepsilon^k} \right| \leqslant C^k \, \alpha\varepsilon \, M_\beta^{\otimes k}.$$

We are going to require that a (weaker) uniform bound holds also at positive times (at least for short times)

$$\forall k \leqslant N, \qquad f_N^{(k)}(t, Z_k) \leqslant \exp(\gamma k) \, M_\lambda^{\otimes k}(V_k) \quad \text{for almost all } Z_k \text{ in } \mathbb{D}_\varepsilon^k, \tag{3.28}$$

where λ and γ are parameters to be determined. This growth condition is reminiscent of the condition (3.26) and it provides a control on the velocities (thanks to the parameter λ) as well as on the local density of the particles (thanks to the parameter γ). We are going to show that uniform bounds of the type (3.28) are enough to estimate the collision operators.

We start by introducing a functional setting with weighted norms compatible with (3.28). For $\lambda > 0$ and $k \geqslant 1$, let $\mathbb{X}_{\varepsilon,k,\lambda}$ be the space of measurable functions f_k defined almost everywhere on $\mathbb{D}_\varepsilon^k$ such that

$$\|f_k\|_{\varepsilon,k,\lambda} := \left\| f_k \, \exp\left(\lambda H_k\right) \right\|_{\mathbb{L}^\infty(\mathbb{D}_\varepsilon^k)} < \infty, \tag{3.29}$$

where the Hamiltonian $H_k(Z_k) = H_k(V_k) = \frac{1}{2}\sum_{i=1}^k |v_i|^2$ was introduced in (2.8).

Neglecting the cancellations between the gain and the loss part of the collision operators as in (3.22), we denote by $|Q|_{s,s+n}$ the operator obtained by summing the absolute values of all elementary contributions

$$|Q|_{s,s+n}(t) = \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}_s(t-t_1)\,|C_{s,s+1}|\,\mathbf{S}_{s+1}(t_1-t_2)\,|C_{s+1,s+2}|\dots\mathbf{S}_{s+n}(t_n)dt_n\dots dt_1 \tag{3.30}$$

where the operator $|C_{s,s+1}|$ is defined for any regular function f_{s+1} in $\mathbb{D}_\varepsilon^{s+1}$

$$\begin{aligned} &\left(|C_{s,s+1}|f_{s+1}\right)(Z_s) \qquad (3.31)\\ &:= \frac{(N-s)\varepsilon^{d-1}}{\alpha}\left(\sum_{i=1}^s \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_{s+1}(\dots,x_i,v_i',\dots,x_i+\varepsilon\nu,v_{s+1}')\Big((v_{s+1}-v_i)\cdot\nu\Big)_+ d\nu dv_{s+1}\right.\\ &\qquad\left. + \sum_{i=1}^s \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} f_{s+1}(\dots,x_i,v_i,\dots,x_i+\varepsilon\nu,v_{s+1})\Big((v_{s+1}-v_i)\cdot\nu\Big)_- d\nu dv_{s+1}\right). \end{aligned}$$

The proposition below provides controls on the norms of the collision operators.

Proposition 3.4. *There is a constant C depending only on the dimension d such that for all $s, n \geqslant 1$ and all $t \geqslant 0$, the operators $|Q|_{s,s+n}(t)$ satisfy the following continuity estimates. For all f_{s+n} in $\mathbb{X}_{\varepsilon,s+n,\lambda}$ then $|Q|_{s,s+n}(t)f_{s+n}$ belongs to $\mathbb{X}_{\varepsilon,s,\frac{\lambda}{2}}$ and*

$$\left\| |Q|_{s,s+n}(t) f_{s+n} \right\|_{\varepsilon,s,\frac{\lambda}{2}} \leqslant \exp(s)\,\left(\frac{Ct}{\lambda^{\frac{d+1}{2}}}\right)^n \|f_{s+n}\|_{\varepsilon,s+n,\lambda}\,. \tag{3.32}$$

Note that these estimates involve a loss in the parameter $\lambda > 0$ and that they deteriorate with time.

Proof. We start by estimating the norm of a single collision operator. Let f_{k+1} be a function in $\mathbb{X}_{\varepsilon,k+1,\lambda}$, we are going to prove that for almost all time τ and $k \leqslant s+n$

$$\left\| \left|C_{k,k+1}\right| \mathbf{S}_{k+1}(\tau) f_{k+1} \right\|_{\varepsilon,k+1,(1-\frac{1}{2n})\lambda} \leqslant \frac{C}{\lambda^{\frac{d+1}{2}}}(s+n)\|f_{k+1}\|_{\varepsilon,k+1,\lambda}. \tag{3.33}$$

The loss in the norm weight has been chosen of order $\frac{\lambda}{2n}$ in order to iterate n times this estimate.

The collision operator $|C_{k,k+1}|$ defined in (3.31) is made of two parts which will be evaluated separately. Since $N\varepsilon^{d-1} = \alpha$, the part $|C^+_{k,k+1}|$ with the scattering is bounded

from above by

$$\begin{aligned}
&\big(\big|C^+_{k,k+1}\big|\,\mathbf{S}_{k+1}(\tau)f_{k+1}\big)(Z_k)\\
&\quad:=\frac{(N-k)\varepsilon^{d-1}}{\alpha}\sum_{i=1}^{k}\int_{\mathbb{S}^{d-1}\times\mathbb{R}^d}d\nu dv_{k+1}\Big((v_{k+1}-v_i)\cdot\nu\Big)_+\mathbf{S}_{k+1}(\tau)f_{k+1}(Z'_{k+1})\\
&\quad\leqslant\|f_{k+1}\|_{\varepsilon,k+1,\lambda}\sum_{i=1}^{k}\int_{\mathbb{S}^{d-1}\times\mathbb{R}^d}d\nu dv_{k+1}\Big((v_{k+1}-v_i)\cdot\nu\Big)_+\mathbf{S}_{k+1}(\tau)\exp\big(-\lambda H_{k+1}(V'_{k+1})\big),
\end{aligned}$$

where Z'_{k+1} stands for the pre-collisional configuration $(x_1,v_1,\dots,x_i,v'_i,\dots,x_i+\varepsilon\nu,v'_{k+1})$ and V'_{k+1} for the corresponding velocities.

Using successively, the conservation of the kinetic energy by scattering $H_{k+1}(V'_{k+1})=H_{k+1}(V_{k+1})$ and by the hard-sphere flow, we get the following upper bounds

$$\begin{aligned}
&\big(\big|C^+_{k,k+1}\big|\,\mathbf{S}_{k+1}(\tau)f_{k+1}\big)(Z_k)\\
&\qquad\leqslant\|f_{k+1}\|_{\varepsilon,k+1,\lambda}\sum_{i=1}^{k}\int_{\mathbb{S}^{d-1}\times\mathbb{R}^d}d\nu dv_{k+1}\Big((v_{k+1}-v_i)\cdot\nu\Big)_+\exp\big(-\lambda H_{k+1}(V_{k+1})\big)\\
&\qquad\leqslant C\|f_{k+1}\|_{\varepsilon,k+1,\lambda}\int_{\mathbb{R}^d}dv_{k+1}\Big(k|v_{k+1}|+\sum_{i=1}^{k}|v_i|\Big)\exp\big(-\lambda H_{k+1}(V_{k+1})\big)\\
&\qquad\leqslant C\,\lambda^{-\frac{d}{2}}\Big(k\lambda^{-\frac12}+\sum_{i=1}^{k}|v_i|\Big)\exp\left(-\lambda H_k(Z_k)\right)\|f_{k+1}\|_{\varepsilon,k+1,\lambda}\,,
\end{aligned}\tag{3.34}$$

where the last term is obtained by integrating the velocity v_{k+1}. This upper bound holds for almost all time τ and the configuration Z_k.

As the velocities $\sum_{i=1}^k|v_i|$ are not bounded, the norm in $\mathbb{X}_{\varepsilon,k,\lambda}$ of the operator above is infinite. However, by loosing a factor $\frac{\lambda}{2n}$ on the exponential weight, the sum over the velocities in (3.34) can be controlled thanks to Cauchy-Schwarz estimate

$$\begin{aligned}
&\sum_{i=1}^{k}|v_i|\exp\Big(-\frac{\lambda}{4n}\sum_{1\leqslant j\leqslant k}|v_j|^2\Big)\\
&\qquad\leqslant\Big(k\frac{2n}{\lambda}\Big)^{\frac12}\left(\sum_{i=1}^{k}\frac{\lambda}{2n}|v_i|^2\exp\Big(-\frac{\lambda}{2n}\sum_{j=1}^{k}|v_j|^2\Big)\right)^{1/2}\leqslant\sqrt{\frac{2nk}{e\lambda}}\leqslant\sqrt{\frac{2}{e\lambda}}(s+n)\,,
\end{aligned}$$

where we used that $\|x\exp(-x)\|_\infty\leqslant e^{-1}$ and that $k\leqslant s+n$ in the last inequality. Plugging this inequality in (3.34), we deduce the upper bound (3.33).

The operator $|Q|_{s,s+n}(t)$ can be estimated by applying iteratively the upper bound (3.33) on the n collision operators. Each collision operator induces a loss $\frac{\lambda}{2n}$ on the exponential weight, thus the final estimate is with respect to the norm $\|\cdot\|_{\varepsilon,k+1,\frac{\lambda}{2}}$. This leads to an upper of the form $\frac{C^n}{\lambda^{\frac{(d+1)n}{2}}}(s+n)^n$ uniformly over the collision times $t_1>\cdots>t_n$. Integrating then over the times provides an additional factor

$$\int_0^t\int_0^{t_1}\dots\int_0^{t_{n-1}}dt_n\dots dt_1=\frac{t^n}{n!}.$$

Finally, we deduce that

$$\left\| |Q|_{s,s+n}(t) f_{s+n} \right\|_{\varepsilon,s,\frac{\lambda}{2}} \leqslant \frac{(s+n)^n}{n!} \left(\frac{Ct}{\lambda^{\frac{d+1}{2}}} \right)^n \|f_{s+n}\|_{\varepsilon,s+n,\lambda} .$$

Stirling's formula implies that

$$\frac{(s+n)^n}{n!} \leqslant \exp\left(n \log \frac{n+s}{n} + n \right) \leqslant \exp(s+n).$$

This completes the upper bound (3.32) on the operator $|Q|_{s,s+n}(t)$. □

The marginals of the BBGKY hierarchy are represented by the iterated Duhamel formula (3.11) in terms on the initial data

$$\forall s \leqslant N, \qquad f_N^{(s)}(t) = \sum_{n=0}^{N-s} \alpha^n Q_{s,s+n}(t) f_N^{(s+n)}(0) .$$

For sufficiently small times, Proposition 3.4 leads to uniform controls in N of each term in this sum.

Proposition 3.5. *There exists $T^* > 0$ and a constant $C > 0$ such that the following upper bound holds uniformly in N and $t \in [0, \frac{T^*}{\alpha}]$*

$$\left\| \alpha^n Q_{s,s+n}(t) f_N^{(s+n)}(0) \right\|_{\varepsilon,s,\frac{\beta}{2}} \leqslant \left\| \alpha^n |Q|_{s,s+n}(t) f_N^{(s+n)}(0) \right\|_{\varepsilon,s,\frac{\beta}{2}} \leqslant \frac{C^s}{2^n}, \tag{3.35}$$

for any $s \geqslant 1$ and $n \geqslant 0$ such that $s + n \leqslant N$. Both constants T^*, C depend on $\|f_0\|_\infty$ and β.

Proof. Recall that the marginals of the initial data are uniformly controlled in Proposition 2.7. In particular, the following bound holds for any $s \geqslant 1$

$$\forall Z_s \in \mathbb{D}_\varepsilon^s, \qquad f_N^{(s)}(0, Z_s) \leqslant c^s \prod_{i=1}^{s} f_0(z_i)\ M_\beta(v_i),$$

for some $c \geqslant 1$. By Assumption (3.1), the function f_0 is bounded from above so that

$$\left\| f_N^{(s)}(0) \right\|_{\varepsilon,s,\beta} \leqslant C_1^s \exp(-s\mu),$$

for some constant C_1. Combined with Estimate (3.32) of Proposition 3.4 , this implies that

$$\left\| \alpha^n |Q|_{s,s+n}(t)\, f_N^{(s+n)}(0) \right\|_{\varepsilon,s,\frac{\beta}{2}} \leqslant \alpha^n \exp(s) \left(\frac{Ct}{\beta^{\frac{d+1}{2}}} \right)^n C_1^{s+n} \exp(-(s+n)\mu).$$

Thus (3.35) holds for $t \in [0, \frac{T^*}{\alpha}]$ with

$$T^* = \frac{\beta^{\frac{d+1}{2}}}{C\, C_1} \exp(\mu). \tag{3.36}$$

□

3.3.3 Estimates on the Boltzmann hierarchy

We introduce now the counterpart of the norm (3.29) for the Boltzmann hierarchy. Let $\mathbb{X}_{0,k,\lambda}$ be the space of continuous functions g_k defined on $\mathbb{T}^{dk} \times \mathbb{R}^{dk}$ such that

$$\|g_k\|_{0,k,\lambda} := \sup_{Z_k \in \mathbb{T}^{dk} \times \mathbb{R}^{dk}} \Big| g_k(Z_k) \exp\left(\lambda H_k(Z_k)\right) \Big| < \infty.$$

The absolute value of the collision operator is denoted by

$$|Q^0|_{s,s+n}(t) = \int_0^t \int_0^{t_1} \dots \int_0^{t_{n-1}} \mathbf{S}_s^0(t-t_1)\, |C_{s,s+1}^0|\, \mathbf{S}_{s+1}^0(t_1-t_2)\, |C_{s+1,s+2}^0| \\ \dots \mathbf{S}_{s+n}^0(t_n)\, dt_n \dots dt_1$$

with

$$\left(|C_{s,s+1}^0| g_{s+1}\right)(Z_s) := \sum_{i=1}^s \int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} g_{s+1}(\dots, x_i, v_i', \dots, x_i, v_{s+1}') \Big((v_{s+1} - v_i) \cdot \nu \Big)_+ d\nu dv_{s+1} \\ + \sum_{i=1}^s \int_{\mathbb{S}^{d-1} \times \mathbb{R}^d} g_{s+1}(\dots, x_i, v_i, \dots, x_i, v_{s+1}) \Big((v_{s+1} - v_i) \cdot \nu \Big)_- d\nu dv_{s+1}\,.$$

The structure of the operators $|Q|_{s,s+n}(t)$ and $|Q^0|_{s,s+n}(t)$ is similar so that the continuity estimates of Proposition 3.4 and 3.5 hold also for the Boltzmann hierarchy.

Proposition 3.6. *There is a constant C depending only on the dimension d such that for all $s, n \geqslant 1$ and all $t \geqslant 0$, the operators $|Q^0|_{s,s+n}(t)$ satisfy the following continuity estimates. For all g_{s+n} in $\mathbb{X}_{0,s+n,\lambda}$, then $|Q^0|_{s,s+n}(t) g_{s+n}$ belongs to $\mathbb{X}_{0,s,\frac{\lambda}{2}}$ and*

$$\Big\| |Q^0|_{s,s+n}(t) g_{s+n} \Big\|_{0,s,\frac{\lambda}{2}} \leqslant \exp(s) \left(\frac{Ct}{\lambda^{\frac{d+1}{2}}} \right)^n \|g_{s+n}\|_{0,s+n,\lambda}. \tag{3.37}$$

Furthermore, there exists $T^* > 0$ and a constant $C > 0$ such that uniformly in $s \geqslant 1, n \geqslant 0$ and $t \in [0, \frac{T^*}{\alpha}]$, one has

$$\Big\| \alpha^n Q_{s,s+n}^0(t) g_\alpha^{(s+n)}(0) \Big\|_{\varepsilon,s,\frac{\beta}{2}} \leqslant \Big\| \alpha^n |Q^0|_{s,s+n}(t) g_\alpha^{(s+n)}(0) \Big\|_{\varepsilon,s,\frac{\beta}{2}} \leqslant \frac{C^s}{2^n}. \tag{3.38}$$

To investigate the uniqueness of the solutions of the Boltzmann hierarchy, we are going to introduce an appropriate functional setting which takes into account a growth condition as in (3.26). For $\gamma \in \mathbb{R}$, we define a norm on the whole sequence $G = \{g_k\}_{k \geqslant 1}$ of continuous functions g_k in $\mathbb{X}_{0,k,\lambda}$

$$\|G\|_{0,\lambda,\gamma} := \sup_{k \geqslant 1} \Big(\|g_k\|_{0,k,\lambda} \exp(k\gamma) \Big)$$

and denote by $\hat{\mathbb{X}}_{0,\lambda,\gamma}$ the Banach space of functions such that $\|G\|_{0,\lambda,\gamma} < \infty$. Given $T > 0$ and two functions $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$ in $[0, T]$, we also define $\hat{\mathbb{X}}_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},T}$ the space of time continuous functions

$$G : t \in [0,T] \mapsto G(t) = \{g_k(t)\}_{k \geqslant 1} \in \hat{\mathbb{X}}_{0,\boldsymbol{\lambda}(t),\boldsymbol{\gamma}(t)}$$

endowed with the norm

$$|||G|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},T} = \sup_{t\in[0,T]} \|G(t)\|_{0,\boldsymbol{\lambda}(t),\boldsymbol{\gamma}(t)}\,. \tag{3.39}$$

Existence and uniqueness of a solution of the Boltzmann hierarchy follows from the proposition below.

Proposition 3.7. *Consider the Boltzmann hierarchy* (3.17) *starting from the initial data* $G(0) = \{g^{(s)}(0)\}_{s\geqslant 1}$ *in* $\hat{\mathbb{X}}_{0,\lambda_0,\gamma_0}$ *with* $\lambda_0 > 0$. *Then there exists a time* $T^* > 0$ *and two decreasing functions* $\boldsymbol{\lambda} > 0$ *and* $\boldsymbol{\gamma}$ *in the time interval* $[0, \frac{T^*}{\alpha}]$ *such that the Boltzmann hierarchy gives rise to a unique solution in* $\hat{\mathbb{X}}_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}$.

Note that an explicit solution of the Boltzmann hierarchy was given in (3.21) when the initial data is a product (3.18).

Proof. The derivation follows the argument devised in [44] and [14] which is based on Cauchy-Kovalevsky type estimates.

Let $T^* > 0$ be a parameter to be tuned later and consider the time dependent weights defined in $[0, \frac{T^*}{\alpha}]$ by

$$\boldsymbol{\lambda}(t) = \lambda_0 - ct \geqslant \frac{\lambda_0}{2} \quad \text{and} \quad \boldsymbol{\gamma}(t) = \gamma_0 - ct \geqslant \gamma_0 - \frac{\lambda_0}{2} \quad \text{with} \quad c = \frac{\alpha}{T^*}\frac{\lambda_0}{2}. \tag{3.40}$$

Mild solutions $G = \{g_k\}_{k\geqslant 1}$ of the Boltzmann hierarchy are defined for t in $[0, \frac{T^*}{\alpha}]$ by

$$G(t) = \mathbf{S}^0(t)g(0) + \alpha\int_0^t dt_1\, \mathbf{S}^0(t-t_1)\mathbf{C}^0 G(t_1)\, dt_1, \tag{3.41}$$

where the transport and the collision operators are acting on each coordinates

$$\forall s \geqslant 1, \quad \big(\mathbf{S}^0(t)G\big)_s = \mathbf{S}^0_s(t)g_s \quad \text{and} \quad \big(\mathbf{C}^0 G\big)_s = \mathbf{C}^0_{s,s+1}\, g_{s+1}.$$

Proposition 3.7 will follow from a fixed-point theorem. Define the map on $\hat{\mathbb{X}}_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}$ as

$$\mathcal{F}(G) : t \in \left[0, \frac{T^*}{\alpha}\right] \mapsto \alpha\int_0^t dt_1\, \mathbf{S}^0(t-t_1)\mathbf{C}^0 G(t_1)\, dt_1.$$

We are going to check that for some value of $T^* > 0$ (which will be determined in (3.43)) then $\mathcal{F}$ is a contraction in $\hat{\mathbb{X}}_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}$

$$|||\mathcal{F}(G)|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}} \leqslant \frac{1}{2}|||G|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}. \tag{3.42}$$

As a consequence, there exists a unique fixed point in $\hat{\mathbb{X}}_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}$ which solves (3.41). This completes Proposition 3.7.

We turn now to the proof of (3.42). Fix t in $[0, \frac{T^*}{\alpha}]$. As the free transport preserves the $\mathbb{L}^\infty$ norm, we get

$$\begin{aligned}
&\Big\| \int_0^t dt_1 \, \mathbf{S}_k^0(t-t_1) \big| C_{k,k+1}^0 \big| \, g_{k+1}(t_1) \Big\|_{0,k,\boldsymbol{\lambda}(t)} \\
&\leqslant \sup_{Z_k} \int_0^t dt_1 \, \exp\Big(\boldsymbol{\lambda}(t) H_k(Z_k) \Big) \, \big| C_{k,k+1}^0 \big| \, \big| g_{k+1}(t_1) \big| (Z_k) \\
&\leqslant \sup_{Z_k} \int_0^t dt_1 \, \exp\Big(\big(\boldsymbol{\lambda}(t) - \boldsymbol{\lambda}(t_1) \big) H_k(Z_k) \Big) \frac{C}{\boldsymbol{\lambda}(t_1)^{\frac{d}{2}}} \Big(\frac{k}{\boldsymbol{\lambda}(t_1)^{\frac{1}{2}}} + \sum_{i=1}^k |v_i| \Big) \| g_{k+1} \|_{0,k+1,\boldsymbol{\lambda}(t_1)},
\end{aligned}$$

where the last inequality is derived as in (3.34).

Using the norm (3.39) defined on the whole sequence $G(t) = \{g_k(t)\}_{k \geqslant 1}$, the following holds for any t_1 in $[0, \frac{T^*}{\alpha}]$

$$\| g_{k+1} \|_{0,k+1,\boldsymbol{\lambda}(t_1)} \leqslant \exp\big(-\boldsymbol{\gamma}(t_1)(k+1) \big) \, |||G|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}.$$

This implies

$$\begin{aligned}
&\Big\| \int_0^t dt_1 \, \mathbf{S}_k(t-t_1) \big| C_{k,k+1}^0 \big| \, g_{k+1}(t_1) \Big\|_{0,k,\boldsymbol{\lambda}(t)} \\
&\qquad \leqslant |||G|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}} \, \exp\big(-\boldsymbol{\gamma}(t)(k+1) \big) \frac{C}{\boldsymbol{\lambda}(t)^{\frac{d}{2}}} \\
&\qquad\qquad \sup_{Z_k} \int_0^t dt_1 \, \exp\big(-c(t-t_1)k \big) \Big(\frac{k}{\boldsymbol{\lambda}(t)^{\frac{1}{2}}} + \sum_{i=1}^k |v_i| \Big) \exp\Big(-c(t-t_1) H_k(Z_k) \Big),
\end{aligned}$$

where we used that the functions $\boldsymbol{\lambda}$ and $\boldsymbol{\gamma}$, defined in (3.40), are linear and decreasing. Noticing that $\boldsymbol{\lambda}(t) \geqslant \frac{\lambda_0}{2}$ and integrating over t_1 leads to the uniform estimate

$$\Big\| \int_0^t dt_1 \, \mathbf{S}_k(t-t_1) \big| C_{k,k+1}^0 \big| \, g_{k+1}(t_1) \Big\|_{0,\boldsymbol{\lambda}(t),\boldsymbol{\gamma}(t)} \leqslant |||G|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}} \frac{C \exp\big(-\boldsymbol{\gamma}(t)\big)}{(\lambda_0/2)^{\frac{d}{2}}} \Big(\frac{1}{(\lambda_0/2)^{\frac{1}{2}}} + 1 \Big) \frac{1}{c}.$$

Replacing c by its value, we deduce that

$$|||\mathcal{F}(G)|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}} \leqslant \frac{C \exp\big(-\gamma_0 - \lambda_0/2\big)}{(\lambda_0/2)^{\frac{d}{2}} \, \lambda_0} \Big(\frac{1}{(\lambda_0/2)^{\frac{1}{2}}} + 1 \Big) T^* \, |||G|||_{0,\boldsymbol{\lambda},\boldsymbol{\gamma},\frac{T^*}{\alpha}}.$$

This implies (3.42) when

$$T^* = \frac{(\lambda_0/2)^{\frac{d+3}{2}}}{C\big(1 + (\lambda_0/2)^{\frac{1}{2}}\big)} \exp\big(\gamma_0 + \lambda_0/2\big). \tag{3.43}$$

□

3.4 Convergence to the Boltzmann equation

In this section, we are going to derive the first part of Theorem 3.1, i.e. that the typical distribution of a gas particle follows the Boltzmann equation in the Boltzmann-Grad limit. The strategy is to show that the Duhamel series obtained from the BBGKY hierarchy converges to the Duhamel series associated with the Boltzmann hierarchy. The tails of both series can be controlled by the bounds on the collision operators obtained in Section 3.3, but the convergence of the first terms of the series (namely the dominant contribution) requires a more careful analysis. This will be achieved by first interpreting the collision operators in terms of branching processes and then coupling the branching processes associated with both hierarchies.

3.4.1 Collision trees viewed as branching processes

The Duhamel representation (3.11) of the marginal $f_N^{(1)}$ has an analytic flavor but it can also be interpreted as a branching process of particles evolving according to the hard-sphere dynamics. This is reminiscent of the derivation of Proposition 2.5 where the first order of the Duhamel representation was identified as a collision between two particles.

We will analyze separately each terms in the sum (3.11) of the representation of $f_N^{(1)}(t)$ We start with the term of order n

$$Q_{1,n+1}(t) f_N^{(n+1)}(0) = \int_{\mathcal{T}_n} dT_n \, \mathbf{S}_1(t-t_1) C_{1,2} \mathbf{S}_2(t_1-t_2) C_{2,3} \dots \mathbf{S}_n(t_n) f_N^{(n+1)}(0), \qquad (3.44)$$

where the time integral is over the ordered collision times $T_n = (t_1, \dots, t_n)$ taking values in the set

$$\mathcal{T}_n := \Big\{ T_n = (t_1, \dots, t_n); \quad 0 < t_n < \dots < t_2 < t_1 < t \Big\}. \qquad (3.45)$$

The root of the collision tree is associated with the coordinates $z_1 = (x_1, v_1) \in \mathbb{T}^d \times \mathbb{R}^d$ of the first particle at time t. The operator $\mathbf{S}_1(t-t_1)$ is interpreted as the motion of the first particle following the backward flow up to time t_1 where a collision takes place

$$\forall u \in [t_1, t], \qquad z_1(u) = \mathbf{T}_{u-t}(z_1).$$

As in the proof of Proposition 2.5, the collision operator $C_{1,2}$ is interpreted as the adjunction at time t_1 of a new particle at position $x_1(t_1) + \varepsilon \nu_2$ with a deflection angle $\nu_2 \in \mathbb{S}^{d-1}$ and with a velocity $v_2 \in \mathbb{R}^d$. The new configuration $Z_2(t_1)$ depends on the type of the collision (see Figure 2.4):

- In the pre-collisional case, i.e. if $(v_2 - v_1(t_1^+)) \cdot \nu_2 < 0$, then the pair of particles is
$$Z_2(t_1) = \big((x_1(t_1), v_1), (x_1(t_1) + \varepsilon \nu_2, v_2) \big). \qquad (3.46)$$
This corresponds to the loss part of the operator $C_{1,2}$.
- In the post-collisional case, i.e. if $(v_2 - v_1(t_1^+)) \cdot \nu_2 > 0$, then the pair of particles is defined as the pre-collisional configuration
$$Z_2(t_1) = \big((x_1(t_1), v_1'), (x_1(t_1) + \varepsilon \nu_2, v_2') \big), \qquad (3.47)$$
by using the scattering (2.28). Note that the velocity $t \mapsto v_1(t)$ is discontinuous at time t_1. This corresponds to the gain part of the operator $C_{1,2}$.

Once created, the new pair of particles evolves according to the backward 2-particle flow during the time interval $[t_2, t_1]$

$$\forall u \in [t_2, t_1], \qquad Z_2(u) = \mathbf{T}_{t_1-u}(Z_2(t_1)).$$

Iterating this procedure, a branching process is built inductively by adding, at time t_i, a particle labelled $i+1$ to the particle with label $a(i) \leqslant i$ chosen randomly among the already existing i particles. Given a deflection angle ν_{i+1} and a velocity v_{i+1}, the velocity of the particle $a(i)$ and of the new particle $i+1$ at time t_i are updated according to the pre-collisional (3.46) or post-collisional (3.47) rule

$$\begin{cases} Z_{i+1}(t_i) = \big(\{z_j(t_i)\}_{j\neq a(i)}, (x_{a(i)}(t_i), v_{a(i)}(t_i)), (x_{a(i)}(t_i) + \varepsilon\nu_{i+1}, v_{i+1})\big), \\ \qquad\qquad \text{in the pre-collisional case, i.e. if } \ (v_{i+1} - v_{a(i)}(t_i)) \cdot \nu_{i+1} < 0, \\ Z_{i+1}(t_i) = \big(\{z_j(t_i)\}_{j\neq a(i)}, (x_{a(i)}(t_i), v'_{a(i)}(t_i^+)), (x_{a(i)}(t_i) + \varepsilon\nu_{i+1}, v'_{i+1})\big), \\ \qquad\qquad \text{in the post-collisional case, i.e. if } \ (v_{i+1} - v_{a(i)}(t_i^+)) \cdot \nu_{i+1} > 0\,. \end{cases}$$

The velocities of the other particles are unchanged. It may happen that the new particle created at position $x_{a(i)}(t_i) + \varepsilon\nu_{i+1}$ overlaps with another particle with label in $\{1, \dots, i\}$. In this case, the creation is cancelled and the pseudo-trajectory keeps evolving backward with only i particles. Otherwise, the new configuration follows the backward hard-sphere flow starting from $Z_{i+1}(t_i)$

$$\forall u \in [t_{i+1}, t_i], \qquad Z_{i+1}(u) = \mathbf{T}_{u-t_i}\big(Z_{i+1}(t_i)\big).$$

The marginal $f_N^{(n+1)}(0)$ at time 0 will then be evaluated on the configuration $Z_{n+1}(0)$ which represents the coordinates of the leaves of the tree.

In between two branching times, the particles in the tree follow the backward hard-sphere flow so that they rebound elastically when they meet. We will call *recollision* these events and the name *collision* will only refer to the creation of a particle at a branching time.

In this way, the operator $Q_{1,n+1}(t)$ is interpreted in terms of collision trees with n collisions and the corresponding random process is described by the *pseudo-trajectories* (see Figure 3.1). The pseudo-trajectories are not physical trajectories of the gas particles. They are a geometric interpretation of the iterated Duhamel formula in terms of a branching process flowing backward in time and determined by

- the coordinates z_1 of the initial particle at time t,
- the collision times $T_n = (t_1, \dots, t_n) \in \mathcal{T}_n$ which are interpreted as branching times,
- the labels of the colliding particles $\mathbf{a} = \big(a(1), \dots, a(n)\big)$ from which the branchings occur and which take values in the set

$$\mathcal{A}_n := \big\{\mathbf{a} = \big(a(1), \dots, a(n)\big), \qquad 1 \leqslant a(i) \leqslant i\big\},$$

- the velocities $V_{2,n+1} = \{v_2, \dots, v_{n+1}\}$ in $\mathbb{R}^{dn}$ and deflection angles $\Theta_{2,n+1} = \{\nu_2, \dots, \nu_{n+1}\}$ in $\mathbb{S}^{(d-1)n}$ for each additional particle.

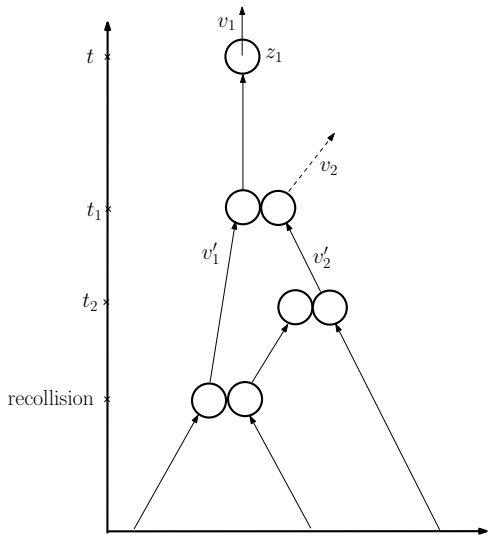


Figure 3.1: A pseudo-trajectory in a collision tree of the BBGKY hierarchy is depicted. The branching process is initialized at time t with the coordinates $z_1 = (x_1, v_1)$ of the first particle and then it evolves backwards. At time t_1 the velocity of particle 1 has been changed after the creation of particle 2. Particle 2 has then generated a third particle at time t_2 in a collision without scattering. This third particle recollides with particle 1 and both velocities are modified according to the scattering rule of the hard-sphere dynamics.

For a given collision tree $\mathbf{a}$ in $\mathcal{A}_n$, the variables associated with a pseudo-trajectory will be denoted by

$$\Omega_n = \{T_n, V_{2,n+1}, \Theta_{2,n+1}\} \in \mathcal{T}_n \times \mathbb{R}^{dn} \times \mathbb{S}^{(d-1)n}, \tag{3.48}$$

in order to simplify the notation. We stress the fact that Ω_n cannot take any value in $\mathcal{T}_n \times \mathbb{R}^{dn} \times \mathbb{S}^{(d-1)n}$: the parameters Ω_n leading to an overlap between two particles at a creation time have to be discarded. Indeed, the corresponding pseudo-trajectories are ill-defined.

The contribution (3.44) for n collisions during $[0, t]$ can be computed by integrating $f_N^{(n+1)}(0)$ on the value of the pseudo-trajectories $Z_{n+1}(0)$ at time 0

$$Q_{1,n+1}(t) f_N^{(n+1)}(0) = \left[\frac{\varepsilon^{(d-1)n}}{\alpha^n} \prod_{i=1}^{n} (N-i) \right] \sum_{\mathbf{a} \in \mathcal{A}_n} \int d\Omega_n \, \mathcal{C}(z_1, \Omega_n) \, f_N^{(n+1)}(0, Z_{n+1}(0)), \tag{3.49}$$

where the contribution of the cross sections involved in each collision is denoted by

$$\mathcal{C}(z_1, \Omega_n) = \prod_{k=1}^{n} (v_{k+1} - v_{a(k)}(t_k^+)) \cdot \nu_{k+1}. \tag{3.50}$$

Note that the contributions of the gain and loss terms in the collision operator $C_{k,k+1}$ are taken into account by the sign of $(v_{k+1} - v_{m_k}(t_k^+)) \cdot \nu_{k+1}$.

We turn now to the branching process associated with the Boltzmann hierarchy (3.20) and start with the geometric interpretation of the term representing the occurence of n

collisions

$$Q^0_{1,n+1}(t)g^{(n+1)}(0) = \int_{\mathcal{T}_n} dT_n \mathbf{S}^0_s(t-t_1)C^0_{1,2}\mathbf{S}^0_2(t_1-t_2)C^0_{2,3}\dots\mathbf{S}^0_{n+1}(t_n)g^{(n+1)}(0), \quad (3.51)$$

where we used the previous notation $T_n = (t_1, \dots, t_n) \in \mathcal{T}_n$.

The operator $Q^0_{1,n+1}(t)$ can be represented by collision trees indexed by the coordinates $z^0_1 = (x^0_1, v^0_1)$ of the first particle at time t, a collection of branching times $T_n = (t_1, \dots, t_n)$, the labels of the colliding particles $\mathbf{a} = \big(a(1), \dots, a(n)\big) \in \mathcal{A}_n$ as well as a collection of velocities $V_{2,n+1} = \{v_2, \dots, v_{n+1}\}$ in $\mathbb{R}^{dn}$ and deflection angles $\Theta_{2,n+1} = \{\nu_2, \dots, \nu_{n+1}\}$ in $\mathbb{S}^{(d-1)n}$. If a tree has been built up to time t_k, then the $(k+1)^{\text{th}}$ particle with coordinates z^0_{k+1} is added at time t_k at the same position $x^0_{a(k)}(t_k)$ of the particle $a(k)$ and their velocities are updated according to the type of the collision:

- if $(v_{k+1} - v_{a(k)}(t_k)) \cdot \nu_{k+1} < 0$, then the new configuration is

$$Z^0_{k+1}(t_i) = \big(\{z^0_j(t_k)\}_{j\neq a(k)}, (x^0_{a(k)}(t_k), v_{a(k)}(t_k)), (x^0_{a(k)}(t_k), v_{k+1})\big),$$

- if $(v_{k+1} - v_{a(k)}(t_k^+)) \cdot \nu_{k+1} > 0$, then the new configuration is

$$Z^0_{k+1}(t_i) = \big(\{z_j(t_k)\}_{j\neq a(k)}, (x^0_{a(k)}(t_k), v'_{a(k)}(t_k^+)), (x^0_{a(k)}(t_k), v'_{k+1})\big),$$

  the velocity of $a(k)$ is also updated.

During the time interval $[t_{k+1}, t_k]$, the corresponding pseudo-trajectory Z^0_{k+1} evolves according to the backward *free* flow denoted by $\mathbf{T}^0_{-u}$ until the next particle creation. The collision tree is built backward starting from time t up to the configuration $Z^0_{n+1}(0)$ at time 0. We stress the fact that the particles are represented by points which never recollide (almost surely) during the backward evolution. This is a key difference with the pseudo-trajectories in the BBGKY hierarchy. Notice also that when a new particle is created, there is no constraint due to a possible overlap with another particle. Thus the pseudo-trajectories can be parametrized by the variable Ω_n, introduced in (3.48), which can take any value in $\mathcal{T}_n \times \mathbb{R}^{dn} \times \mathbb{S}^{(d-1)n}$.

The contribution of $Q^0_{1,n+1}(t)$ can be computed by integrating $g^{(n+1)}(0)$ on the values of the pseudo-trajectories $Z^0_{n+1}(0)$ at time 0

$$Q^0_{1,n+1}(t)g^{(n+1)}(0) = \sum_{\mathbf{a}\in\mathcal{A}_n} \int d\Omega_n \, \mathcal{C}^0(z_1, \Omega_n) \, g^{(n+1)}(0, Z^0_{n+1}(0)), \quad (3.52)$$

where the contribution of the cross sections involved in each collision is denoted by

$$\mathcal{C}^0(z_1, \Omega_n) = \prod_{k=1}^{n} (v_{k+1} - v_{a(k)}(t_k^+)) \cdot \nu_{k+1}. \quad (3.53)$$

Once again the signs of the gain and loss terms in the collision operator $C^0_{k,k+1}$ are taken into account by the sign of $(v_{k+1} - v_{a(k)}(t_k^+)) \cdot \nu_{k+1}$. Note that $\mathcal{C}$ and $\mathcal{C}^0$ have an identical structure, except that $\mathcal{C}$ is evaluated on the BBGKY hierarchy pseudo-trajectories and $\mathcal{C}^0$ on the Boltzmann hierarchy pseudo-trajectories.

3.4.2 Coupling the pseudo-trajectories

The collision trees associated with each hierarchy are constructed in a very similar way, and they are encoded by the same set of parameters $\{z_1, \mathbf{a}, \Omega_n\}$. However there are two important differences between these trees :

- In the Boltzmann hierarchy, particles are created at the same position of their parent, instead in the BBGKY hierarchy, they are created with a shift of order εv_i which depends on the deflection angle.
- In the Boltzmann hierarchy, particles are points transported by the free flow, thus no recollision occurs. The particles in the pseudo-trajectories of the BBGKY hierarchy bounce elastically when they touch so that the recollisions may modify their trajectories.
- The creation of a particle in the BBGKY hierarchy is discarded if it leads to an overlap with another particle. The overlaps are not a constraint for point particles, so that the occurence of an overlap induces a discrepancy between the pseudo-trajectories of the two hierarchies.

The discrepancy due to the shifts occuring at each branching times is very small, as the shifts are of order ε. A recollision may have much stronger consequences as it can lead to an important modification of the pseudo-trajectory (see Figure 3.2). The overlaps can be treated as local recollisions.

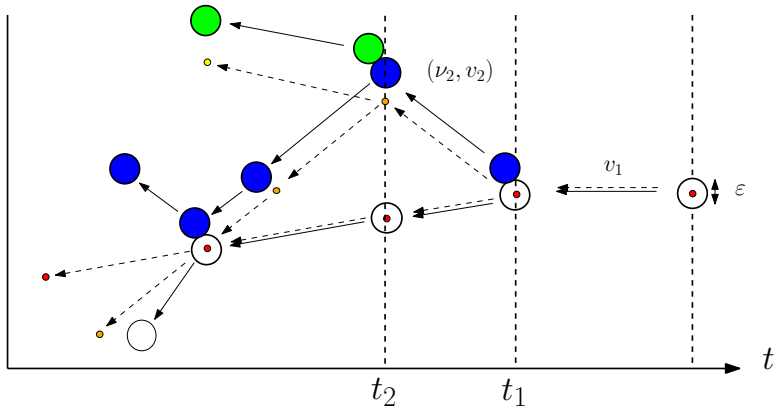


Figure 3.2: The first stages of both pseudo-trajectories are depicted up to the occurence of a recollision. The BBGKY pseudo-trajectories are represented with plain arrows, whereas the Boltzmann pseudo-trajectories correspond to the dashed arrows. At time t, the particle with label 1 in the BBGKY hierarchy is a ball of radius ε centered at position x_1 and the particle in the Boltzmann hierarchy is depicted as a point located at $x_1^0 = x_1$. At time t_1 the second particle is added and at time t_2 the third. Both hierarchies are coupled, but a small error in the particle positions of order ε can occur at each collision. In this figure, a recollision between the first and the second particle of the BBGKY pseudo-trajectories occurs and after this recollision the Boltzmann and the BBGKY pseudo-trajectories are no longer close to each other. Indeed the BBGKY trajectories are deflected after the recollision, instead the point particles do not collide and follow a straight line (see the dashed arrows). Note that before the recollision, the trajectories of z_1 and z_1^0 are identical and therefore the plain and the dashed arrows overlap.

In the following proposition, we will show that the iterated Duhamel series of both

hierarchies converge term by term. To do this, it will be enough to check that the recollisions are extremely rare when ε tends to 0, so that the two types of pseudo-trajectories can be coupled in order to remain close to each other.

Proposition 3.8. *For any time $t \geqslant 0$ and any integer n, the following convergence holds in the Boltzmann-Grad limit*

$$\forall z_1 \in \mathbb{T}^d \times \mathbb{R}^d, \qquad \lim_{N\to\infty} \left(Q_{1,n+1}(t) f_N^{(n+1)}(0)\right)(z_1) = \left(Q^0_{1,n+1}(t) g^{(n+1)}(0)\right)(z_1). \quad (3.54)$$

This proposition will be derived by following the approach from [10, 28] and its proof is postponed to Section 3.4.4. Note that a quantitative version of Proposition 3.8 was derived in [14]. We will first use Proposition 3.8 to show the convergence (3.3) of the first marginal towards the solution of the Boltzmann equation.

3.4.3 Proof of Theorem 3.1 : convergence to the Boltzmann equation

Recall that from (3.11) and (3.20), the first marginals in the Boltzmann and the BBGKY hierarchies can be rewritten

$$f_N^{(1)}(t) = \sum_{n=0}^{N-1} \alpha^n Q_{s,s+n}(t) f_N^{(s+n)}(0),$$
$$g^{(1)}(t) = \sum_{n \geqslant 0} \alpha^n Q^0_{s,s+n}(t) g^{(s+n)}(0).$$

Furthermore by Proposition 3.7, the solution of the Boltzmann hierarchy is unique in the time interval $[0, \frac{T^*}{\alpha}]$ and it coincides with the solution of the Boltzmann equation $g^{(1)}(t, z_1) = g(t, z_1)$ thanks to Lemma 3.3. Thus it is enough to show that $f_N^{(1)}(t)$ converges to $g^{(1)}(t)$ in order to prove that it converges to the solution of the Boltzmann equation.

For any $\delta > 0$, one can find an integer K, thanks to Propositions 3.5 and 3.6, such that for any $t \in [0, \frac{T^*}{\alpha}]$ and uniformly in $N > K$

$$\sum_{n=K}^{N-1} \left\| \alpha^n Q_{1,n+1}(t) f_N^{(n+1)}(0) \right\|_{\varepsilon,1,\frac{\beta}{2}} + \sum_{n=K}^{\infty} \left\| \alpha^n Q^0_{1,n+1}(t) g^{(n+1)}(0) \right\|_{0,1,\frac{\beta}{2}} \leqslant \delta. \quad (3.55)$$

Using (3.11) and (3.20), we deduce

$$f_N^{(1)}(t) - g^{(1)}(t) = \sum_{n=0}^{K} \alpha^n \left(Q_{1,n+1}(t) f_N^{(n+1)}(0) - Q^0_{1,n+1}(t) g^{(n+1)}(0) \right) + O(\delta).$$

Thus the convergence of the first marginal for any $t \in [0, \frac{T^*}{\alpha}]$ and z_1 in $\mathbb{T}^d \times \mathbb{R}^d$

$$\lim_{N\to\infty} f_N^{(1)}(t, z_1) = g^{(1)}(t, z_1)$$

follows from the convergence of the first K terms proved in Proposition 3.8 (and then by letting δ tend to 0). This completes the first part of Theorem 3.1.

3.4.4 Proof of Proposition 3.8

We first check that instead of $Q_{1,n+1}(t)f_N^{(n+1)}(0)$ introduced in (3.49), it is enough to consider

$$\widehat{Q}_{1,n+1}(t)f_N^{(n+1)}(0) = \sum_{\mathbf{a}\in\mathcal{A}_n} \int d\Omega_n \, \mathcal{C}(z_1,\Omega_n) \, f_N^{(n+1)}(0, Z_{n+1}(0)), \tag{3.56}$$

as the prefactor

$$\frac{\varepsilon^{(d-1)n}}{\alpha^n} \prod_{i=1}^{n}(N-i) = \prod_{i=1}^{n}\Big(1-\frac{i}{N}\Big) = 1 + O\Big(\frac{n^2}{N}\Big), \tag{3.57}$$

is asymptotically close to 1 thanks to the Boltzmann-Grad scaling $N\varepsilon^{(d-1)} = \alpha$.

Using the representations (3.52) and (3.56) in terms of the pseudo-trajectories, the difference can be decomposed as

$$\begin{aligned}
&\widehat{Q}_{1,n+1}(t)f_N^{(n+1)}(0) - Q^0_{1,n+1}(t)g^{(n+1)}(0) \qquad (3.58)\\
&= \sum_{\mathbf{a}\in\mathcal{A}_n} \Big[\int d\Omega_n \, \Big(\mathcal{C}(z_1,\Omega_n) \, g^{(n+1)}(0, Z_{n+1}(0)) - \, \mathcal{C}^0(z_1,\Omega_n) \, g^{(n+1)}(0, Z^0_{n+1}(0)) \Big) \\
&+ \int d\Omega_n \, \mathcal{C}(z_1,\Omega_n) \, \Big(f_N^{(n+1)}(0, Z_{n+1}(0)) - g^{(n+1)}(0, Z_{n+1}(0)) \Big) \Big] ,
\end{aligned}$$

where the pseudo-trajectories in each integral are evaluated with the same branching structure given by the collision tree **a**. Note that overlaps may occur in the BBGKY hierarchy, in this case the pseudo-trajectories are discarded and the corresponding terms in the integrals are set equal to 0.

Using the definition (3.18) of $g^{(n+1)}(0)$, the last term in (3.58) can be easily estimated by Proposition 2.7 which implies that the initial densities are close

$$\begin{aligned}
&\Big| \int d\Omega_n \, \mathcal{C}(z_1,\Omega_n) \, \Big(f_N^{(n+1)}(0, Z_{n+1}(0)) - g^{(n+1)}(0, Z_{n+1}(0)) \Big) \Big| \\
&\qquad \leqslant C^n \, \alpha\varepsilon \int d\Omega_n \, \Big| \mathcal{C}(z_1,\Omega_n) \Big| \; M_\beta^{\otimes(n+1)}(Z_{n+1}(0)) \\
&\qquad \leqslant C^n \, \alpha\varepsilon \, \Big(|Q|_{1,1+n}(t) M_\beta^{\otimes(n+1)} \Big)(z_1).
\end{aligned}$$

Given z_1 and n, the last term is linear in ε. Thus when ε tends to 0, the last term in (3.58) vanishes.

We turn now to the estimation of the difference of the first two terms in (3.58). This is a key point of the proof as it boils down to controlling the recollisions and the overlaps which may occur in the pseudo-trajectories. Indeed, two pseudo-trajectories Z_{n+1}, Z^0_{n+1} encoded by the same parameters $\{z_1, \Omega_n\}$ are close if the BBGKY pseudo-trajectory has no recollision and no overlap :

$$\forall i \leqslant n+1, \forall t \in [0, t_{i-1}], \qquad |x_i(t) - x_i^0(t)| \leqslant (i-1)\varepsilon \quad \text{and} \quad v_i(t) = v_i^0(t), \tag{3.59}$$

with the notation $t_0 = t$. When there are no recollisions, the velocities are updated only at the collision times and besides the shift of the positions by ε at the creation times, both

pseudo-trajectories evolve exactly in the same way. At the creation time of particle i the cumulated error of these shifts is at most $(i-1)\varepsilon$.

If (3.59) holds, we deduce from the expression (3.18) of the initial data that

$$\left|g^{(n+1)}(0,Z_{n+1}(0)) - g^{(n+1)}(0,Z^0_{n+1}(0))\right| = M_\beta^{\otimes n+1}(V_{n+1})\left|\prod_{i=1}^{n+1} f_0(z_i) - \prod_{i=1}^{n+1} f_0(z_i^0)\right| \quad (3.60)$$
$$\leqslant \|\nabla_x f_0\|_\infty \, \|f_0\|_\infty^n \, M_\beta^{\otimes n+1}(V_{n+1})(n+1)^2\varepsilon,$$

where we used that $\nabla_x f_0$ is bounded to estimate from above the small shifts of the positions of the pseudo-trajectories. Therefore, for the collision parameters Ω_n such that the pseudo-trajectories have no overlap and no recollisions, the right hand side of (3.58) tends to 0. It remains to show that the occurence of the recollisions and of the overlaps are so rare that they do not contribute to the difference in (3.58).

For a given collision tree **a** and an initial data z_1, we are going to compare the pseudo-trajectories from both hierarchies associated with the same collection of parameters Ω_n introduced in (3.48). In probabilistic terms, both pseudo-trajectories are coupled with the same source of randomness. We define below $\mathcal{G}_\varepsilon$ a set of *good* parameters for which there will be no recollisions or overlaps

$$\mathcal{G}_\varepsilon = \Big\{\Omega_n = \{T_n, V_{2,n+1}, \Theta_{2,n+1}\}, \quad (3.61)$$
$$\forall i,j \in \{1,\dots,n+1\},\ i \not\sim j,\ \forall u \in [0,t], \qquad \left|x_i^0(u) - x_j^0(u)\right| \geqslant 4n\varepsilon,$$
$$\forall i \in \{2,\dots,n+1\},\ \forall u \in \left[0, t_{i-1} - \frac{4n\varepsilon}{|v_{i+1}(t_{i-1}) - v_{a(i)}(t_{i-1})|}\right], \quad \left|x_i^0(u) - x_{a(i)}^0(u)\right| \geqslant 4n\varepsilon,$$
$$\forall i \in \{1,\dots,n-1\}, \quad t_{i+1} < t_i - \frac{4n\varepsilon}{|v_{i+1}(t_i) - v_{a(i)}(t_i)|}\Big\},$$

where the symbol $i \not\sim j$ means that i is different from j and that i and j are not related by a branching (i.e $i \neq a(j)$ or $j \neq a(i)$). Note that $\mathcal{G}_\varepsilon$ depends on z_1 which is fixed throughout the proof.

The set $\mathcal{G}_\varepsilon$ is indexed by the pseudo-trajectories of the Boltzmann hierarchy and the Lebesgue measure of its complement $\mathcal{G}_\varepsilon^c$ will be shown to converge to 0. We are first going to check that a parameter Ω_n in $\mathcal{G}_\varepsilon$ generates a BBGKY pseudo-trajectory without recollision or overlap. The idea is that if the particles in a Boltzmann pseudo-trajectory remain far apart from each other then the corresponding BBGKY pseudo-trajectories cannot have any recollision because both trajectories have to be close thanks to (3.59).

Given Ω_n a set of parameters in $\mathcal{G}_\varepsilon$, we are going to build the BBGKY and Boltzmann pseudo-trajectories inductively. Suppose that both pseudo-trajectories have evolved during the time interval $[s,t]$ and that no recollisions or overlaps have occurred so far. We are going to check that the conditions on the parameters in $\mathcal{G}_\varepsilon$ implies that the next creations will not lead to a recollision or an overlap:

- The first condition in (3.61) implies that particles in a Boltzmann pseudo-trajectory with labels $i \not\sim j$ are at distance at least $4n\varepsilon$ thus the corresponding particles in the Boltzmann pseudo-trajectories will remain at distance $2n\varepsilon$ thanks to (3.59). Notice also that when particles are far enough, then no overlap can occur when a new particle is created.

- When a particle is created, it is close to its parent for some time (depending on their relative velocity), but then both particles depart from each other. However, due to a future collision (which would deflect the trajectory) or due to the periodicity of $\mathbb{T}^d$, a particle and its parents might get closer at a later time and the second condition in (3.61) has been added to forbid this eventuality.
- Finally the third condition in (3.61) implies that two particle creations do not occur in a very short amount of time. Indeed when a new particle i is created, it is very close to its parent $a(i)$, thus if a third particle is created, by i or $a(i)$, immediately after then it may lead to an overlap or a recollision between these 3 particles. The third condition implies that particles i and $a(i)$ drift at distance $4n\varepsilon$ before any creation of a new particle.

As the pseudo-trajectories of the BBGKY hierarchy indexed by parameters in $\mathcal{G}_\varepsilon$ have no recollisions, the difference in (3.58) can be decomposed as

$$\begin{aligned}
&\int d\Omega_n \left(\mathcal{C}(z_1,\Omega_n)\, g^{(n+1)}(0,Z_{n+1}(0)) - \mathcal{C}^0(z_1,\Omega_n)\, g^{(n+1)}(0,Z^0_{n+1}(0))\right)\\
&\quad = \int d\Omega_n\, \mathbf{1}_{\{\Omega_n\notin\mathcal{G}_\varepsilon\}} \left(\mathcal{C}(z_1,\Omega_n)\, g^{(n+1)}(0,Z_{n+1}(0)) - \mathcal{C}^0(z_1,\Omega_n)\, g^{(n+1)}(0,Z^0_{n+1}(0))\right)\\
&\quad + \int d\Omega_n\, \mathbf{1}_{\{\Omega_n\in\mathcal{G}_\varepsilon\}}\, \mathcal{C}(z_1,\Omega_n) \left(g^{(n+1)}(0,Z_{n+1}(0)) - g^{(n+1)}(0,Z^0_{n+1}(0))\right),
\end{aligned}$$

where we used the fact that for $\Omega_n \in \mathcal{G}_\varepsilon$, there is no recollision or overlap and the velocities of both pseudo-trajectories coincide so that $\mathcal{C}(z_1,\Omega_n) = \mathcal{C}^0(z_1,\Omega_n)$.

Applying inequality (3.60) to control the trajectory without recollisions, we get

$$\begin{aligned}
&\left|\int d\Omega_n \left(\mathcal{C}(z_1,\Omega_n)\, g^{(n+1)}(0,Z_{n+1}(0)) - \mathcal{C}^0(z_1,\Omega_n)\, g^{(n+1)}(0,Z^0_{n+1}(0))\right)\right|\\
&\quad \leqslant \|f_0\|_\infty^{n+1} \int d\Omega_n\, \mathbf{1}_{\{\Omega_n\notin\mathcal{G}_\varepsilon\}} \left(\left|\mathcal{C}(z_1,\Omega_n)\right| M_\beta^{\otimes n+1}(V_{n+1}) + \left|\mathcal{C}^0(z_1,\Omega_n)\right| M_\beta^{\otimes n+1}(V^0_{n+1})\right)\\
&\quad + \|\nabla_x f_0\|_\infty\, \|f_0\|_\infty^n\, (n+1)^2\varepsilon \int d\Omega_n\, \mathbf{1}_{\{\Omega_n\in\mathcal{G}_\varepsilon\}} \left|\mathcal{C}(z_1,\Omega_n)\right| M_\beta^{\otimes n+1}(V_{n+1}). \qquad (3.62)
\end{aligned}$$

When ε tends to 0, the sets $\mathcal{G}_\varepsilon^c$ are going to converge to

$$\begin{aligned}
\mathcal{N}_0 = \Big\{\Omega_n = &\{T_n, V_{2,n+1}, \Theta_{2,n+1}\}, \qquad (3.63)\\
&\exists i,j\in\{1,\dots,n+1\},\ i\not\sim j,\ \exists u\in[0,t], \quad \left|x_i^0(u) - x_j^0(u)\right| = 0,\\
&\exists i\in\{2,\dots,n+1\},\ \exists u < t_{i-1}, \qquad \left|x_i^0(u) - x_{a(i)}^0(u)\right| = 0,\\
&t_n < t_{n-1} < \cdots < t_2 < t_1\Big\}.
\end{aligned}$$

For any fixed z_1, $\mathcal{N}_0$ has measure 0. Indeed, suppose that $x_i^0(u) = x_j^0(u)$ at some time u and denote by $u' \in [u,t]$ the time of the last deflection of the particle i or j. Then the constraint $x_i^0(u) = x_j^0(u)$ forces the collision parameters at time u' to be in a set of measure 0. Thus we deduce that for almost all Ω_n

$$\lim_{\varepsilon\to 0} \mathbf{1}_{\mathcal{G}_\varepsilon^c}(\Omega_n) = \mathbf{1}_{\mathcal{N}_0}(\Omega_n) = 0.$$

The dominated convergence theorem implies the following convergence for any z_1

$$\lim_{\varepsilon\to 0}\int d\Omega_n\,\mathbf{1}_{\{\Omega_n\notin\mathcal{G}_\varepsilon\}}\,\Big(\big|\mathcal{C}(z_1,\Omega_n)\big|\,M_\beta^{\otimes n+1}(V_{n+1})+\big|\mathcal{C}^0(z_1,\Omega_n)\big|\,M_\beta^{\otimes n+1}(V_{n+1}^0)\Big)=0, \tag{3.64}$$

where the domination relies on the comparison of the term in parenthesis with the collision operators which are both finite (Propositions 3.4 and 3.6)

$$\big|Q\big|_{1,1+n}(t)M_\beta^{\otimes(n+1)}(z_1)<\infty\quad\text{and}\quad\big|Q^0\big|_{1,1+n}(t)M_\beta^{\otimes(n+1)}(z_1)<\infty.$$

The last term in (3.62) can be bounded in the same way.

$$\varepsilon\int d\Omega_n\,\mathbf{1}_{\{\Omega_n\in\mathcal{G}_\varepsilon\}}\,\big|\mathcal{C}(z_1,\Omega_n)\big|\,M_\beta^{\otimes n+1}(V_{n+1})\leqslant\varepsilon\big|Q\big|_{1,1+n}(t)M_\beta^{\otimes(n+1)}(z_1).$$

Thus this term also tends to 0 and this concludes the proof of Proposition 3.8.

3.5 The molecular chaos assumption

3.5.1 Convergence of the marginals

In this section, we complete the derivation of Theorem 3.1 and show that the marginals asymptotically factorize (3.4). Using the explicit solution of the Boltzmann hierarchy derived in Lemma 3.3, this boils down to checking that for any time t in $[0,T^*/\alpha]$ and for almost all configurations Z_s in $\mathbb{T}^{ds}\times\mathbb{R}^{ds}$

$$\lim_{N\to\infty}f_N^{(s)}(t,Z_s)=g^{(s)}(t,Z_s). \tag{3.65}$$

As for the convergence of the first marginal, the proof is based on the comparison of the Duhamel series (3.11) and (3.20)

$$f_N^{(s)}(t)-g^{(s)}(t)=\sum_{n=0}^{N-s}\alpha^n Q_{s,s+n}(t)f_N^{(s+n)}(0)-\sum_{n=0}^{\infty}\alpha^n Q^0_{s,s+n}(t)g^{(s+n)}(0).$$

Using Propositions 3.5 and 3.6, one can show as in (3.55) that for any $\delta>0$, there is K large enough such that for any $t\in[0,\frac{T^*}{\alpha}]$ and uniformly in $N\geqslant s+K$

$$\sum_{n=K}^{N-s}\Big\|\alpha^n Q_{s,n+s}(t)f_N^{(n+s)}(0)\Big\|_{\varepsilon,s,\frac{\beta}{2}}+\sum_{n=K}^{\infty}\Big\|\alpha^n Q^0_{s,n+s}(t)g^{(n+s)}(0)\Big\|_{0,s,\frac{\beta}{2}}\leqslant\delta. \tag{3.66}$$

Thus it is enough to generalize Proposition 3.8 and to show that for any time $t\geqslant 0$ and any integer n, then the following convergence holds for almost all initial data Z_s in $\mathbb{T}^{ds}\times\mathbb{R}^{ds}$

$$\forall t\in\left[0,\frac{T^*}{\alpha}\right],\qquad\lim_{N\to\infty}\big(Q_{s,n+s}(t)f_N^{(n+s)}(0)\big)(Z_s)=\big(Q^0_{s,n+s}(t)g^{(n+s)}(0)\big)(Z_s). \tag{3.67}$$

The convergence (3.65) can then be deduced by combining (3.66) and (3.67).

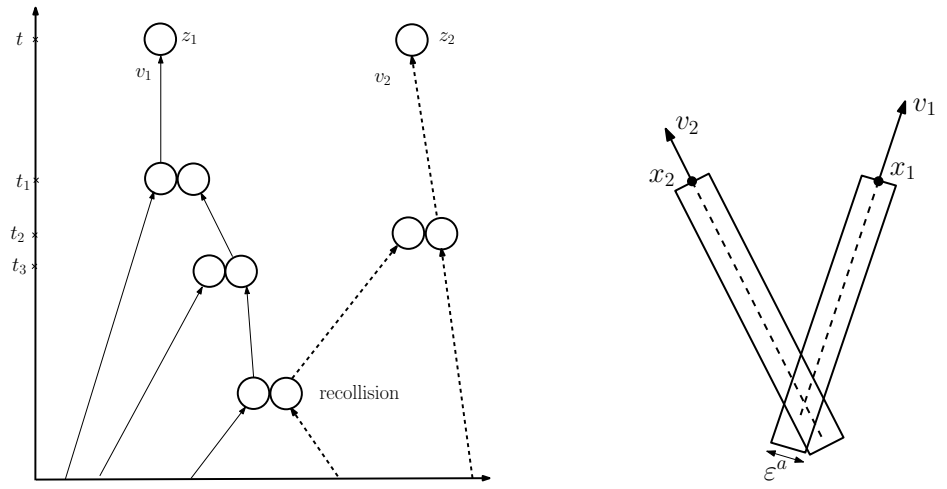


Figure 3.3: Computing the marginal $f_N^{(2)}(t, z_1, z_2)$ boils down to controlling the pseudo-trajectories with 2 roots at time t (see the figure on the left). The collision tree associated with the second particle is in dashed lines to distinguish it from the tree associated with the first particle. The correlations between the two particles are related to the recollisions between the two collision trees. The configuration (z_1, z_2), depicted on the right, belongs to the set $\mathcal{B}_0^2$ as the trajectories intersect in the past. The enlarged set of pathological configurations $\mathcal{B}_\varepsilon^2$ should be understood as the set of configurations such that the tubes of width ε^a around the backward trajectories intersect.

The proof of (3.67) follows the same pattern as the argument of Proposition 3.8, nevertheless there is an important twist as the convergence does not hold for all configurations Z_s. The main steps of the derivation are described below. The operators $Q_{s,s+n}(t)$ and $Q^0_{s,s+n}(t)$ can be also represented by collision trees. The main difference is that the pseudo-hierarchies are now encoded by s trees with roots at time t indexed by the coordinates of each of the s particles Z_s or Z_s^0. These trees are built inductively by evolving backward in time and new particles are created at each branching time. Once again a major difference between both hierarchies comes from the recollisions. In the Boltzmann hierarchy, the particles evolve according to the free flow and therefore do not interact with each other except possibly at the creation of a new particle. Thus the collision trees associated with the configuration Z_s^0 are the union of the s non-interacting collision trees starting from the roots $z_1^0, z_2^0, \dots, z_s^0$ at time t. For the BBGKY hierarchy the s trees are no longer independent as the pseudo-trajectories can be modified by the recollisions either between particles belonging to the same collision tree or between particles belonging to collision trees generated at time t by different roots (see Figure 3.3). To control the recollisions, one has to show that the pseudo-trajectories encoded by the parameters Ω_n in the set $\mathcal{G}_\varepsilon^c$ (3.61) have a vanishing contribution when ε tends to 0. We stress the fact that $\mathcal{G}_\varepsilon^c$ depends on the configuration Z_s at time t. Contrary to the case $s = 1$, this will have important consequences. Indeed for $s = 1$, the limiting set $\mathcal{N}_0$ introduced in (3.63) has measure 0 for any given configuration z_1. This is no longer the case for $s > 1$ if the configuration Z_s belongs to the set

$$\mathcal{B}_0^s = \Big\{ Z_s \in \mathbb{T}^{ds} \times \mathbb{R}^{ds}, \qquad \exists i \neq j, \exists u \in [0,t]\,, \quad \big|(x_i - uv_i) - (x_j - uv_j)\big| = 0 \Big\} \tag{3.68}$$

such that two particles will encounter along the backward free flow starting from Z_s

(see Figure 3.3). Indeed, for any configuration Z_s in $\mathcal{B}_0^s$ there is a set of parameters Ω_n of positive measure such that the pseudo-trajectories starting from Z_s will have a recollision for any $\varepsilon > 0$. Nevertheless, one can check that, for Z_s outside $\mathcal{B}_0^s$, the limiting set

$$\begin{aligned}\mathcal{N}_0^s = \Big\{ \Omega_n = &\{T_n, V_{2,n+1}, \Theta_{2,n+1}\},\\ &\exists i,j \in \{1,\dots,n+s\},\ i \not\sim j,\ \exists u \in [0,t], \quad \left|x_i^0(u) - x_j^0(u)\right| = 0,\\ &\exists i \in \{s+1,\dots,n+s\},\ \exists u < t_{i-s}, \qquad \left|x_i^0(u) - x_{a(i)}^0(u)\right| = 0,\\ &t_n < t_{n-1} < \cdots < t_2 < t_1 \Big\}.\end{aligned}$$

has measure 0 and the dominated convergence theorem can be applied as in (3.64). This completes the convergence (3.67) for any configuration Z_s outside $\mathcal{B}_0^s$. Since $\mathcal{B}_0^s$ has measure 0 in $\mathbb{T}^{ds} \times \mathbb{R}^{ds}$, the convergence (3.65) of the s marginal holds for almost all configurations Z_s in $\mathbb{T}^{ds} \times \mathbb{R}^{ds}$.

3.5.2 Quantitative estimates

The convergence in Theorem 3.1 can be improved in order to quantify the distance between the marginal $f_N^{(s)}$ and its limit $g^{(s)}$. This was first initiated in [14] and developed in [28, 29, 11, 6].

Before stating the convergence result, we introduce the sets which can lead to recollisions in the pseudo-trajectories starting from configurations with s particles. This boils down to enlarging the set $\mathcal{B}_0^s$ introduced in (3.68). Given $a \in (0,1)$ and T^*, we set

$$\mathcal{B}_\varepsilon^s = \left\{ Z_s \in \mathbb{T}^{ds} \times \mathbb{R}^{ds}, \qquad \exists i \neq j, \exists u \in \left[0, \frac{T^*}{\alpha}\right], \quad \left|(x_i - uv_i) - (x_j - uv_j)\right| \leqslant \varepsilon^a \right\}. \tag{3.69}$$

The following theorem summarizes the results derived in [14, 11, 6]. An extension of this result to soft potentials has been proven in [14, 28].

Theorem 3.9. *Under the assumptions of Theorem 3.1 on the initial data, the marginal $f_N^{(s)}$ converges, in the Boltzmann-Grad limit, to $g^{(s)}$ uniformly on the configurations Z_s in $\mathbb{D}_\varepsilon^s \setminus \mathcal{B}_\varepsilon^s$, i.e. there exists $T^* > 0$ such that*

$$\forall t \in \left[0, \frac{T^*}{\alpha}\right], \qquad \left| (f_N^{(s)}(t, Z_s) - g^{(s)}(t, Z_s)) \mathbf{1}_{\{Z_s \notin \mathcal{B}_\varepsilon^s\}} \right| \leqslant C^s M_{\frac{\beta}{2}}^{\otimes s}(V_s)\ \gamma(\varepsilon)\,,$$

where the function $\gamma(\varepsilon)$ converges to 0 as ε tends to 0. The rate of convergence depends on the parameter a which controls the set of initial data (3.69)*.*

For any given ε, the configurations in the set $\mathcal{B}_\varepsilon^s$ are excluded as they could potentially lead to recollisions. When ε tends to 0, these sets shrink to the set $\mathcal{B}_0^s$ defined in (3.68) and therefore Theorem 3.9 implies the factorization (3.4).

Let us briefly review the new ideas involved in the proof of Theorem 3.9. In the derivation of Theorem 3.1, all the estimates were explicit in ε with the exception of the control on the recollisions in Proposition 3.8 which relied on the dominated convergence

theorem. Thus the key step is to estimate the size of $\mathcal{N}_\varepsilon$ (3.61) in terms of ε. Restricting to configurations Z_s in the complement of $\mathcal{B}_\varepsilon^s$ means that the particles remains at distance $\varepsilon^a \gg \varepsilon$ from each other when they evolve according to backward transport. When a new particle is created, one has to prove that the probability that a recollision occurs with another particle is small. Indeed, a recollision requires to aim, at the creation time, very precisely at a small ball of size ε at distance at least ε^a and this imposes strong constraints on the collision parameters at the creation time of a particle. By evaluating at each new collision the range of the parameters which would lead to a recollision, upper bounds on the measure of $\mathcal{N}_\varepsilon$ can be obtained [14, 28].

We will not derive Theorem 3.9, nevertheless the ideas mentioned previously will be key for the study of the large time asymptotics and they will be detailed in Chapters 4 and **??**.

The function $\gamma(\varepsilon)$, introduced in Theorem 3.9, can be estimated and one can show that $\gamma(\varepsilon)$ decays to 0 as $\varepsilon^{a'}$ for some $a' < 1$. Thus Theorem 3.9 implies the convergence of all the marginals of order $s \ll \log N$. By studying the cumulants instead of the marginals, Pulvirenti and Simonella [29] were able to prove that the correlations of much higher order tend to 0. Let $\{\varphi_k\}_{k \geqslant 1}$ be a sequence of smooth functions in $\mathbb{T}^d \times \mathbb{R}^d$ uniformly bounded and define the empirical averages at time t of these functions as well as their limits by

$$\forall k \geqslant 1, \qquad F_k\big(Z_N(t)\big) = \frac{1}{N}\sum_{i=1}^{N} \varphi_k(z_i(t)), \qquad \widehat{\varphi}_k(t) = \int_{\mathbb{T}^d \times \mathbb{R}^d} dz\, \varphi_k(z) g_\alpha(t,z),$$

where g_α is the solution of the Boltzmann equation. The idea is to consider the correlations for some $p > 0$

$$\sup_{j < \varepsilon^{-p}} \mathbb{E}_N \left(\prod_{k=1}^{j} \big(F_k\big(Z_N(t)\big) - \widehat{\varphi}_k(t)\big) \right), \tag{3.70}$$

which take into account many cancelations. In a slightly different setting than the one presented here (i.e. under a grand-canonical initial measure and with the hard-sphere dynamics in $\mathbb{R}^d$) it is proven in [29], that there exists $p > 0$ and a time $T^* > 0$ such that (3.70) convergences to 0 when ε tends to 0.

3.5.3 Time reversal and propagation of chaos

The Boltzmann equation was derived heuristically under the molecular chaos assumption stated in (2.42) which requires a very quantitative version of the factorization when two particles are in contact with each other

$$f_N^{(2)}(t, x_1, v_1, x_1 - \varepsilon\nu, v_2) \approx f_N^{(1)}(t, x_1, v_1') f_N^{(1)}(t, x_1, v_2'). \tag{3.71}$$

Lanford's strategy does not rely on the molecular chaos assumption and the proof of the convergence does not provide any clue for closing directly the equations of the BBGKY hierarchy : the convergence holds for the whole hierarchy. The factorization of the marginals is only recovered in the limit (3.65) by using the fact that the solution of the

Boltzmann hierarchy is itself factorized

$$\lim_{N\to\infty} f_N^{(s)}(t, Z_s) = g^{(s)}(t, Z_s) = \prod_{i=1}^{s} g(t, z_i). \tag{3.72}$$

In fact, this convergence does not hold for all the configurations Z_s. In particular, the configurations in the set $\mathcal{B}_0^s$ (3.68) have been excluded as they lead to recollisions. Furthermore, the propagation of chaos (3.72) is valid only for fixed Z_s in the limit $N \to \infty$ and more quantitative bounds on the factorization would be needed to be used in the proof of the convergence towards the Boltzmann equation. For a given ε, the convergence could not be proved in Theorem 3.9 for the configurations in the set $\mathcal{B}_\varepsilon^s$. This shows that the convergence derived in Theorem 3.9 is very far from establishing the kind of estimate required by the molecular chaos assumption (3.71). In fact, excluding the configurations leading to recollisions is not a technical necessity, but a necessary condition to prove the convergence towards an irreversible equation.

Theorem 3.9 implies the convergence of the marginals on the time interval $[0, \frac{T^*}{\alpha}]$. The microscopic hard-sphere dynamics is reversible, thus if all the velocities are flipped at the intermediate time $\tau = \frac{T^*}{2\alpha}$ then the particle system returns to its initial state. As the Boltzmann equation is not reversible, the reversed equation is the backward Boltzmann equation and Lanford's strategy cannot apply to the microscopic dynamics starting from the density at time τ with the reversed velocities, i.e. from the distribution

$$\widehat{f}_N(0, Z_N) = f_N(\tau, \mathcal{R}(Z_N)).$$

The iterated Duhamel formula (3.11) applied to this initial data implies that for $t \in [0, \tau]$

$$\widehat{f}_N^{(1)}(t) = \sum_{n=0}^{N-1} \alpha^n Q_{1,n+1}(t) \widehat{f}_N^{(n+1)}(0). \tag{3.73}$$

By the reversibility of the microscopic dynamics

$$\widehat{f}_N^{(1)}(t, x, v) = f_N^{(1)}(\tau - t, x, -v).$$

Evaluating the term $Q_{1,n+1}(t)\widehat{f}_N^{(n+1)}(0)$ in (3.73) amounts to integrate $\widehat{f}_N^{(n+1)}(0, Z_{n+1})$ over the configurations Z_{n+1} of the pseudo-trajectories. Reversing the velocities, one has

$$\widehat{f}_N^{(n+1)}(0, Z_{n+1}) = f_N^{(n+1)}\big(\tau, \mathcal{R}(Z_{n+1})\big).$$

By construction of the pseudo-trajectories (see Figure 3.4), the configuration $\mathcal{R}(Z_{n+1})$ arises from a branching process thus it belongs to the set $\mathcal{B}_\varepsilon^s$ (with $a = 1$) as the particles have to recollide in the backward hard-sphere flow. Therefore, the configurations needed to compute the reverse dynamics are exactly those which were discarded in Theorem 3.9. This means that the molecular chaos assumption cannot hold for the configurations Z_s in $\mathcal{B}_\varepsilon^s$ (with $a = 1$)

$$f_N^{(s)}(\tau, Z_s) \not\approx g^{(s)}(\tau, Z_s),$$

otherwise the system would converge towards the Boltzmann equation and not the backward Boltzmann equation. The marginal $f_N^{(s)}(t)$ is almost factorized, but it lacks regularity on very specific sets which have a vanishing measure with ε. Once the rare events

leading to recollisions have been neglected, some information on the reversibility has been lost so that the initial state cannot be recovered by running the dynamics backward. Thus the irreversibility of the limiting process comes from this loss of memory. The key role of the recollision and the importance of the sets $\mathcal{B}_0^s$ defined in (3.68) has been first emphasized in [45] and a more quantitative discussion on the size of the pathological sets $\mathcal{B}_\varepsilon^s$ can be found in [6].

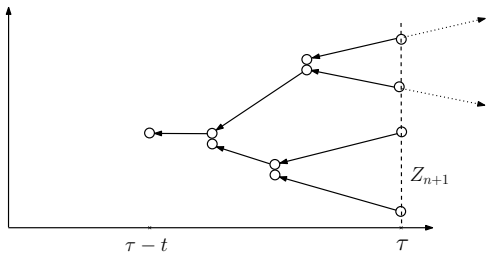


Figure 3.4: At time τ, all the velocities are flipped and the system restarts from the density $\widehat{f}_N(0, Z_N) = f_N(\tau, \mathcal{R}(Z_N))$. The pseudo-trajectory associated with the reversed evolution $\widehat{f}_N^{(1)}(t, x, v)$ form a branching process with final configurations Z_{n+1} in the set $\mathcal{B}_\varepsilon^s$.

Chapter 4

Tagged particle

4.1 Introduction

The main goal of this chapter is to show that a stochastic process can arise from purely deterministic dynamics.

4.2 The ideal gas

4.2.1 A Markov chain representation

The simplest model of a tagged particle in an ideal gas is a stochastic process such that the particle moves in a straight line and at random times changes its direction according to random kicks. In this setting, the gas molecules do not interact which each other and they simply act on the tagged particle as a thermal bath at a given temperature $1/\beta$.

The tagged particle velocity $\{v(t)\}_{t\geqslant 0}$ is then a Markov chain evolving according to the following stochastic rule. If the tagged particle has a velocity v, it will keep this velocity up to an exponential time $\mathcal{T}$ with rate $\alpha a(v)$ defined as

$$\mathbb{P}(\mathcal{T}\geqslant u)=\exp\big(-\alpha a(v)\big)\quad\text{and}\quad a(v)=\int_{\mathbb{S}^{d-1}\times\mathbb{R}^d}\big((w-v)\cdot\nu\big)_+M_\beta(w)dwd\nu,$$

then at time $\mathcal{T}$, the velocity is updated to

$$v'=v-\big((v-w)\cdot\nu\big)\,\nu,$$

where the parameters w and ν are chosen according to the probability density

$$\frac{\big((w-v)\cdot\nu\big)_+M_\beta(w)}{a(v)}dwd\nu.$$

This amounts to say that a particle with velocity w, chosen randomly according to the Maxwellian distribution M_β, collides with the tagged particle according to a law depending on the relative velocity $w-v$ and on a random impact parameter ν.

This procedure is then iterated to build the random sequence of collisions and the path of the Markov chain $\{v(t)\}_{t\geqslant 0}$. The ideal gas is uniformly distributed therefore the

collision mechanism is independent of the position of the tagged particle which evolves simply according to

$$x(t) = \int_0^t ds\ v(s). \tag{4.1}$$

This stochastic evolution mimics the collisions encountered by a tagged particle in a dilute gas. Nevertheless contrary to the deterministic gas dynamics, the spatial structure is neglected and the Markov chain is memoryless. This means that the trajectory of an incoming particle at velocity w will not influence the tagged particle beyond the collision. As a consequence, the outgoing velocity w' plays no role in the evolution of the tagged particle : the momentum and the energy are no longer conserved by the stochastic process which is dissipative.

Rephrasing the previous discussion, we say that the Markov chain $\{v(t)\}_{t\geqslant 0}$ has a generator given by $\alpha\mathcal{L}$ with

$$\mathcal{L}\varphi(v) := \int_{\mathbb{S}^{d-1}\times\mathbb{R}^d} \Big(\varphi(v') - \varphi(v)\Big)\ \big((w-v)\cdot\nu\big)_+\, M_\beta(w)dwd\nu, \tag{4.2}$$

and $\mathcal{L}$ is known as the *linear Boltzmann operator*. The proposition below summarizes some properties of $\mathcal{L}$.

Proposition 4.1. *The operator $\mathcal{L}$ is self-adjoint in $\mathbb{L}^2(M_\beta)$, which amounts to say that the stochastic process $\{v(t)\}_{t\geqslant 0}$ is reversible with respect to its invariant measure M_β. Furthermore, $-\mathcal{L}$ has a spectral gap, i.e. that there exists $c>0$ such that for any φ in $\mathbb{L}^2(M_\beta)$*

$$\mathbb{E}_{M_\beta}(\varphi^2) - \Big(\mathbb{E}_{M_\beta}(\varphi)\Big)^2 \leqslant -c\mathbb{E}_{M_\beta}(\varphi\,\mathcal{L}\varphi), \tag{4.3}$$

for some $c>0$.

Proof. Recall that the map $\mathcal{J} : (v,w,\nu)\mapsto(v',w',-\nu)$ introduced in (2.5) is an involution and preserves the Lebesgue measure. Noticing that

$$M_\beta^{\otimes 2}(v',w') = M_\beta^{\otimes 2}(v,w) \quad\text{and}\quad \big((w-v)\cdot\nu\big)_+ = \big((w'-v')\cdot\nu\big)_-,$$

we get, for any functions ψ and φ in $\mathbb{L}^2(M_\beta)$, that

$$\begin{aligned}
&\int_{\mathbb{S}^{d-1}\times\mathbb{R}^{2d}} d\nu dv dw\ M_\beta^{\otimes 2}(v,w)\ \psi(v)\varphi(v')\ \big((w-v)\cdot\nu\big)_+ &(4.4)\\
&\qquad = \int_{\mathbb{S}^{d-1}\times\mathbb{R}^{2d}} d\nu dv dw\ M_\beta^{\otimes 2}(v',w')\ \psi(v)\varphi(v')\ \big((w'-v')\cdot\nu\big)_- &(4.5)\\
&\qquad = \int_{\mathbb{S}^{d-1}\times\mathbb{R}^{2d}} d\nu dv dw\ M_\beta^{\otimes 2}(v,w)\ \psi(v')\varphi(v)\ \big((w-v)\cdot\nu\big)_+, &(4.6)
\end{aligned}$$

where second equality is obtained by the change of variables $\mathcal{J}$. Identity (4.4) implies that $\mathcal{L}$ is self-adjoint in $\mathbb{L}^2(M_\beta)$

$$\mathbb{E}_{M_\beta}\big(\psi\,\mathcal{L}\varphi\big) = \int M_\beta^{\otimes 2}(v,w)\,\psi(v)\big(\varphi(v')-\varphi(v)\big)\ \big((w-v)\cdot\nu\big)_+ = \mathbb{E}_{M_\beta}\big(\varphi\,\mathcal{L}\psi\big).$$

Thus the stochastic process is reversible with respect to the measure M_β and M_β is the invariant measure.

Using the identity (4.4), we get also that $\mathcal{L}$ has a negative spectrum

$$\mathbb{E}_{M_\beta}(\varphi\,\mathcal{L}\varphi) = -\frac{1}{2}\int M_\beta^{\otimes 2}(v,w)\,\big(\varphi(v')-\varphi(v)\big)^2\,\big((w-v)\cdot\nu\big)_+ \leqslant 0.$$

We refer to [35] for a proof of the spectral gap. □

The semigroup of the Markov process is denoted by

$$P_t = \exp(t\alpha\mathcal{L}). \tag{4.7}$$

The number of random kicks per unit of time is of order $1/\alpha$. We stress that the prefactor α plays a role similar to the factor α in the Boltzmann-Grad scaling $\alpha = N\varepsilon^{d-1}$. When α tends to infinity, the collision frequency increases which boils down to saying that the density of the particles in the bath is increasing.

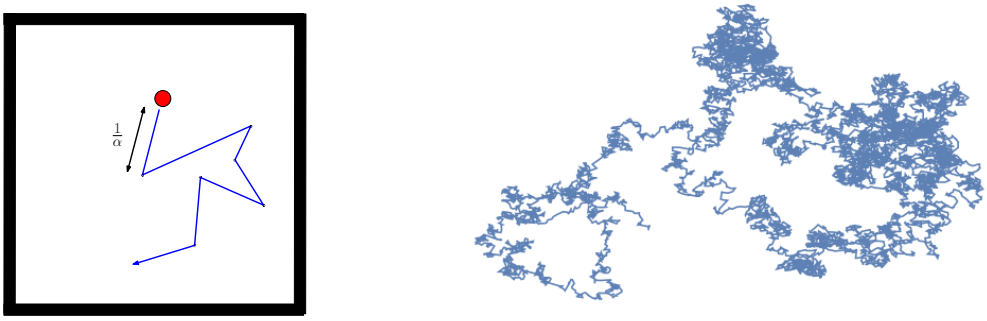


Figure 4.1: The typical time between two collisions in the stochastic process is $1/\alpha$. Thus in a time $t = \alpha\tau$, the position is the sum of roughly $\alpha^2\tau$ random displacements of size $1/\alpha$. A Brownian path is depicted on the right.

The density $g_\alpha(t,x,v) = \varphi_\alpha(t,x,v)M_\beta(v)$ of the tagged particle is given in terms of φ_α, the solution of the *linear Boltzmann equation*

$$\partial_t\varphi_\alpha + v\cdot\nabla_x\varphi_\alpha = \alpha\mathcal{L}\varphi_\alpha, \tag{4.8}$$

with initial data $\varphi_0(x,v)$. We stress the fact that the structure of the linear Boltzmann equation is reminiscent to the one of the Boltzmann equation. We refer to [8, 26] for a survey on the probabilist representations of the Boltzmann equation.

We are interested in the large time behavior of the position of the tagged particle $x(t)$ (4.1). After an appropriate space/time rescaling, the process is expected to converge to a Brownian motion. Instead of rescaling space, it will be more convenient to consider the limit $\alpha\to\infty$ of the following processes

$$\forall\alpha>0, \qquad \chi_\alpha(\tau) = x(\alpha\tau) \tag{4.9}$$

starting at $\chi_\alpha(0) = 0$ with initial velocity distributed according to M_β. In this limit, τ stands for the macroscopic time and the number of random kicks scales as $\alpha t = \alpha^2\tau$ and each increment is of order $\frac{1}{\alpha}$. Thus this is the correct scaling in α to observe a central limit theorem and it is not necessary to rescale the position (see Figure 4.1).

Theorem 4.2. *The sequence of processes $\{\chi_\alpha\}_{\alpha>0}$ converges in law to a Brownian motion in $\mathbb{R}^d$ with diagonal diffusion matrix $\mathbf{D}\,Id$ where the diffusion coefficient is given by*

$$\mathbf{D} = \frac{1}{d}\mathbb{E}_{M_\beta}\left(v\cdot(-\mathcal{L})^{-1}v\right). \tag{4.10}$$

The proof is standard and the main arguments are sketched in Section 4.2.2.

In the rest of this section, we justify the form of the diffusion coefficient $\mathbf{D}$ and check that it is given in terms of the limiting variance of the process. Given an orthonormal basis $\{e_1,\dots,e_d\}$, the coordinates of the process in $\mathbb{R}^d$ are denoted by

$$\forall i\leqslant d,\qquad x_{e_i}(t)=x(t)\cdot e_i,\quad v_{e_i}(t)=v(t)\cdot e_i.$$

We are going to show that for any $\tau>0$

$$\lim_{\alpha\to\infty}\mathbb{E}_{M_\beta}\left(x_{e_1}(\alpha\tau)^2\right)=2\mathbf{D}\tau. \tag{4.11}$$

The other covariances can be computed in the same way.

To prove (4.11), we first use the fact that M_β is the invariant measure in order to rewrite

$$\begin{aligned}\mathbb{E}_{M_\beta}\left(x_{e_1}(\alpha\tau)^2\right)&=\mathbb{E}_{M_\beta}\left(\int_0^{\alpha\tau}ds\int_0^{\alpha\tau}ds'\,v_{e_1}(s)v_{e_1}(s')\right)\\&=2\int_0^{\alpha\tau}ds\int_s^{\alpha\tau}ds'\,\mathbb{E}_{M_\beta}\left(v_{e_1}(0)v_{e_1}(s'-s)\right).\end{aligned}$$

Changing variables and integrating by parts, we deduce that

$$\begin{aligned}\mathbb{E}_{M_\beta}\left(x_{e_1}(\alpha\tau)^2\right)&=2\int_0^{\alpha\tau}ds\int_0^{\alpha\tau-s}du\,\mathbb{E}_{M_\beta}\left(v_{e_1}(0)v_{e_1}(u)\right)\\&=2\alpha\tau\int_0^{\alpha\tau}ds\frac{s}{\alpha\tau}\mathbb{E}_{M_\beta}\left(v_{e_1}(0)v_{e_1}(\alpha\tau-s)\right)\\&=2\alpha\tau\int_0^{\infty}ds\left(1-\frac{s}{\alpha\tau}\right)_+\mathbb{E}_{M_\beta}\left(v_{e_1}(0)v_{e_1}(s)\right).\end{aligned}$$

This can be rephrased in terms of the semigroup $P_s=\exp(\alpha s\mathcal{L})$

$$\begin{aligned}\mathbb{E}_{M_\beta}\left(x_{e_1}(\alpha\tau)^2\right)&=2\alpha\tau\int_0^{\infty}ds\left(1-\frac{s}{\alpha\tau}\right)_+\mathbb{E}_{M_\beta}\left(v_{e_1}P_sv_{e_1}\right)\\&=2\tau\int_0^{\infty}du\left(1-\frac{u}{\alpha^2\tau}\right)_+\mathbb{E}_{M_\beta}\left(v_{e_1}\exp(u\mathcal{L})v_{e_1}\right),\end{aligned} \tag{4.12}$$

where the scaling of the generator has been absorbed in the change of variable $u=\alpha s$. As the process is reversible, the right hand side of the identity is nonnegative

$$\mathbb{E}_{M_\beta}\left(v_{e_1}\exp(u\mathcal{L})v_{e_1}\right)=\mathbb{E}_{M_\beta}\left(\left(\exp(\frac{u}{2}\mathcal{L})v_{e_1}\right)^2\right)\geqslant 0.$$

By the monotone convergence theorem, the limiting variance converges to

$$\mathbb{E}_{M_\beta}\left(x_{e_1}(\alpha\tau)^2\right)=2\tau\int_0^{\infty}du\mathbb{E}_{M_\beta}\left(v_{e_1}\exp(u\mathcal{L})v_{e_1}\right)=2\tau\mathbb{E}_{M_\beta}\left(v_{e_1}(-\mathcal{L})^{-1}v_{e_1}\right). \tag{4.13}$$

Recall, from Proposition 4.1, that $-\mathcal{L}$ has a spectral gap so that the inverse $(-\mathcal{L})^{-1}v_{e_1}$ is well defined. This completes the proof of (4.11).

4.2.2 Proof of Theorem 4.2

Fix $\mathcal{T} > 0$. To prove the convergence, in Theorem 4.2, of the processes χ_α in the time interval $[0, \mathcal{T}]$ towards a Brownian motion B with variance $\mathbf{D}\,\mathrm{Id}$, it is enough (see [4], Chapter 2) to check that

- the convergence of the marginals at different times $\tau_1 < \tau_2 < \cdots < \tau_\ell \leqslant \mathcal{T}$

$$\lim_{\alpha\to\infty} \mathbb{E}_{M_\beta}\Big(h_1\big(\chi_\alpha(\tau_1)\big)\dots h_\ell\big(\chi_\alpha(\tau_\ell)\big)\Big) = \mathbb{E}\Big(h_1\big(B(\tau_1)\big)\dots h_\ell\big(B(\tau_\ell)\big)\Big)\,, \tag{4.14}$$

  where $\{h_1, \dots, h_\ell\}$ is a collection of continuous functions in $\mathbb{R}^d$.

- the tightness of the sequence, that is for any $\tau \in [0, \mathcal{T}]$

$$\forall \delta > 0, \qquad \lim_{\eta\to 0}\lim_{\alpha\to\infty} \mathbb{P}_{M_\beta}\left(\sup_{\tau<\sigma<\tau+\eta} \big|\chi_\alpha(\sigma) - \chi_\alpha(\tau)\big| \geqslant \delta\right) = 0\,. \tag{4.15}$$

The main ingredients for deriving Theorem 4.2 are recalled below and a complete proof can be found in [21].

For simplicity, we are going to focus on one of the position coordinates, say the one along the axis e_1, and derive the invariance principle in this case. From Proposition 4.1, the operator $\mathcal{L}$ is invertible on the space $\mathbb{L}^2(M_\beta)$ of functions of mean 0, thus there exists a unique solution to the Poisson equation associated with the velocity

$$-\mathcal{L}\psi(v) = v_{e_1} \quad \text{with} \quad \mathbb{E}_{M_\beta}(\psi) = 0. \tag{4.16}$$

The stochastic process

$$\mathcal{M}_t = \psi\big(v(t)\big) - \psi\big(v(0)\big) - \alpha\int_0^t ds\, \mathcal{L}\psi\big(v(s)\big)$$

is a martingale. Since ψ is a solution of the Poisson equation, the invariance principle for the position $x_{e_1}(\alpha\tau)$ will be deduced from the invariance principle for the martingale $\tau \to \mathcal{M}_{\alpha\tau}$

$$x_{e_1}(\alpha\tau) = \int_0^{\alpha\tau} ds\, v_{e_1}(s) = \frac{1}{\alpha}\mathcal{M}_{\alpha\tau} - \frac{1}{\alpha}\left(\psi\big(v(\alpha\tau)\big) - \psi\big(v(0)\big)\right). \tag{4.17}$$

Indeed, the second term can be neglected as it converges to 0 in $\mathbb{L}^2$ by (4.16)

$$\frac{1}{\alpha^2}\mathbb{E}_{M_\beta}(\psi^2) = \frac{1}{\alpha^2}\mathbb{E}_{M_\beta}\left(\left(\mathcal{L}^{-1}v_{e_1}\right)^2\right) \xrightarrow[\alpha\to\infty]{} 0.$$

The inequality

$$\forall r \in \mathbb{R}, \qquad |\exp(ir) - 1 - ir| \leqslant r^2$$

implies that the characteristic function of $x_{e_1}(\alpha\tau)$ is well approximated by the one of $\frac{1}{\alpha}\mathcal{M}_{\alpha\tau}$.

The square of the martingale can be decomposed as

$$\mathcal{M}_t^2 = \alpha\int_0^t ds\Big(\mathcal{L}\psi^2\big(v(s)\big) - 2\psi\big(v(s)\big)\mathcal{L}\psi\big(v(s)\big)\Big) + N_t,$$

where N_t is another martingale. Thus the variance (4.11) can be recovered by computing the variance of the martingale term in (4.17)

$$\mathbb{E}_{M_\beta}\left(\frac{\mathcal{M}^2_{\alpha\tau}}{\alpha^2}\right) = -2\frac{\alpha^2\tau}{\alpha^2}\mathbb{E}_{M_\beta}(\psi\mathcal{L}\psi) = 2\tau\mathbb{E}_{M_\beta}\left(v_{e_1}(-\mathcal{L})^{-1}v_{e_1}\right) = 2\mathbf{D}\tau, \tag{4.18}$$

where we used (4.16) in the last equality. More precisely, the ergodic theorem implies the convergence of the quadratic variation almost surely and in $\mathbb{L}^1(M_\beta)$

$$\frac{1}{\alpha}\int_0^{\alpha\tau} ds\Big(\mathcal{L}\psi^2 - 2\psi\mathcal{L}\psi\Big)\big(v(s)\big) = \frac{1}{\alpha}\int_0^{\alpha\tau} ds\Big(\mathcal{L}\big(\mathcal{L}^{-1}v_{e_1}\big)^2 - 2v_{e_1}\,\mathcal{L}^{-1}v_{e_1}\Big)\big(v(s)\big) \xrightarrow[\alpha\to\infty]{\mathbb{L}^1} 2\mathbf{D}\tau.$$

Thus the assumptions of Theorem 2.1 in [21] are satisfied and this leads to the convergence in law of $x_{e_1}(\alpha\tau)$ to a Gaussian variable. A similar approach (see Theorem 2.32 in [21]) can be implemented to derive the joint convergence in law (4.14) of the marginals.

It remains to justify the tightness criterion (4.15) which will be a consequence of the following lemma.

Lemma 4.3. *Fix $\eta > 0$ then*

$$\mathbb{E}_{M_\beta}\left(\sup_{0<\sigma<\eta}\left(\int_0^{\alpha\sigma} v_{e_1}(s)ds\right)^2\right) \leqslant 18\eta\,\mathbb{E}_{M_\beta}\big(v_{e_1}(-\mathcal{L})^{-1}v_{e_1}\big).$$

This is Lemma 2.4 from [21] and its proof is derived at the end of this section for the sake of completness. As the measure M_β is invariant, it is enough to prove (4.15) for $\tau = 0$. Lemma 4.3 implies a uniform upper bound in α

$$\mathbb{P}_{M_\beta}\left(\sup_{0<\sigma<\eta}\Big|\int_0^{\alpha\sigma} v_{e_1}(s)\,ds\Big| \geqslant \delta\right) \leqslant \frac{18\eta}{\delta^2}\,\mathbb{E}_{M_\beta}\big(v_{e_1}(-\mathcal{L})^{-1}v_{e_1}\big). \tag{4.19}$$

Letting η tend to 0, this completes (4.15) for the coordinate v_{e_1}. By symmetry of the coordinates of the velocity, the tightness criterion (4.15) follows.

Proof of Lemma 4.3. Recall that in (4.17), the additive process has been rewritten in terms of a martingale

$$\alpha\int_0^{\alpha\sigma} v_{e_1}(s)\,ds = \mathcal{M}_{\alpha\tau} - \big(\psi\big(v(\alpha\sigma)\big) - \psi\big(v(0)\big)\big)\,,$$

and the martingale fluctuations can be estimated by Doob's Theorem. To control the fluctuations of the second term with ψ, we are going also to replace it by a martingale.

Let $T = \alpha\eta$ and define the backward filtration $\{\mathcal{F}_t^-;\quad t\in[0,T]\}$ generated by the backward process $\{v(T-t);\quad t\in[0,T]\}$. The Markov chain is reversible, so that the generator of the backward process is also $\alpha\mathcal{L}$. Thus the following process is also a martingale

$$\mathcal{M}_t^- = \psi\big(v(T-t)\big) - \psi\big(v(T)\big) - \alpha\int_0^t \mathcal{L}\psi\big(v(T-s)\big)\,ds.$$

In particular, using the Poisson equation (4.16), we get

$$\mathcal{M}^-_{\alpha\eta} - \mathcal{M}^-_{\alpha(\eta-\sigma)} = \psi\big(v(0)\big) - \psi\big(v(\alpha\sigma)\big) + \alpha\int_0^{\alpha\sigma} v_{e_1}(s)\,ds.$$

Combined with (4.17), we deduce that

$$2\alpha \int_0^{\alpha\sigma} v_{e_1}(s)\, ds = \mathcal{M}_{\alpha\sigma} + \mathcal{M}^-_{\alpha\eta} - \mathcal{M}^-_{\alpha(\eta-\sigma)}.$$

Thus the derivation of Lemma 4.3 boils down to controlling the martingale fluctuations by Doob's inequality

$$\begin{aligned}
&\mathbb{E}_{M_\beta}\left(\sup_{0<\sigma<\eta}\left(\int_0^{\alpha\sigma} v_{e_1}(s)ds\right)^2\right) \\
&\quad = \frac{3}{4\alpha^2}\left(\mathbb{E}_{M_\beta}\left(\sup_{0<\sigma<\eta}\mathcal{M}^2_{\alpha\sigma}\right) + \mathbb{E}_{M_\beta}\left(\sup_{0<\sigma<\eta}\left(\mathcal{M}^-_{\alpha(\eta-\sigma)}\right)^2\right) + \mathbb{E}_{M_\beta}\left(\left(\mathcal{M}^-_{\alpha\eta}\right)^2\right)\right) \\
&\quad \leqslant \frac{3}{\alpha^2}\left(\mathbb{E}_{M_\beta}\left(\mathcal{M}^2_{\alpha\eta}\right) + 2\mathbb{E}_{M_\beta}\left(\left(\mathcal{M}^-_{\alpha\eta}\right)^2\right)\right) = 18\eta\mathbb{E}_{M_\beta}\left(v_{e_1}(-\mathcal{L})^{-1}v_{e_1}\right),
\end{aligned}$$

where we used the computation (4.18) of the variance. □

4.3 Hard-sphere dynamics

We are going to study the asymptotic behavior of a tagged particle in the hard-sphere dynamics at equilibrium and show that after rescaling its trajectory converges to a Brownian motion.

4.3.1 The convergence result

In the hard-sphere gas, all the particles are exchangeable and we are going to break this symmetry in order to follow the evolution of a given particle which will play a specific role. This tagged particle will be labelled by 1 with coordinates $z_1 = (x_1, v_1)$. The initial data is a perturbation of the equilibrium density (2.18) with respect to the position x_1 of the tagged particle

$$\forall Z_N \in \mathbb{D}_\varepsilon^N, \qquad f_N(0, Z_N) := \rho^0(x_1)\, M_{N,\beta}(Z_N), \tag{4.20}$$

where $\rho^0 \in C^0(\mathbb{T}^d)$ is a continuous density of probability on $\mathbb{T}^d$. Note that the distribution $f_N(0)$ is normalized by 1 in $\mathbb{L}^1(\mathbb{T}^{dN} \times \mathbb{R}^{dN})$ thanks to the translation invariance of $\mathbb{T}^d$ and that $\displaystyle\int_{\mathbb{T}^d} \rho^0(x)dx = 1$.

In the macroscopic limit $\alpha \to \infty$, the trajectory of the tagged particle is rescaled as

$$\widehat{\chi}_\alpha(\tau) := x_1\big(\alpha\tau\big) \in \mathbb{T}^d, \tag{4.21}$$

where τ stands for the macroscopic time scale. The distribution of $\widehat{\chi}_\alpha(\tau)$ is given by the first marginal $f_N^{(1)}(\alpha\tau, x, v)$ of the particle system. The following theorem, derived in [5], describes the macroscopic behavior of the tagged particle and of its density.

Theorem 4.4 ([5]). *Consider a tagged particle in a hard-sphere gas initially distributed according to the density defined in* (4.20)*. Then, in the limit*

$$N \to \infty, \quad \varepsilon \to 0, \quad \alpha \to \infty \quad \text{with} \quad \alpha = N\varepsilon^{d-1} \text{ and } \alpha \ll \sqrt{\log\log N}, \tag{4.22}$$

the process $\{\widehat{\chi}_\alpha(\tau)\}_{\tau\geq 0}$ *associated with the tagged particle* (4.21) *converges in law towards a Brownian motion initially distributed under the measure* ρ^0 *and with variance* $2\mathbf{D}$ *given by* (4.10)*.*

In particular, the distribution $f_N^{(1)}(\alpha\tau, x, v)$ *of* $\widehat{\chi}_\alpha(\tau)$ *is well approximated, in the limit* (4.22)*, by*

$$\left\| f_N^{(1)}(\alpha\tau, x, v) - \rho(\tau, x) M_\beta(v) \right\|_{\mathbb{L}^\infty([0,T]\times\mathbb{T}^d\times\mathbb{R}^d)} \to 0, \tag{4.23}$$

where $\rho(\tau, x)$ *is the solution of the linear heat equation*

$$\partial_\tau \rho - \mathbf{D}\Delta_x \rho = 0 \quad \text{in} \quad \mathbb{T}^d \quad \text{with} \quad \rho_{|\tau=0} = \rho^0. \tag{4.24}$$

The heat equation should be understood as a (simple) hydrodynamic limit because it arises in a regime such that the tagged particle undergoes a diverging number of collisions per unit time. The derivation of this hydrodynamic equation relies on the kinetic approximation of the microscopic dynamics, however contrary to the strategy described in Section 1.3, the hydrodynamic equation is not obtained in a two-step limit but in the joint limit $N, \alpha \to \infty$ (4.22).

The invariance principle will be proved by coupling the process $\{\widehat{\chi}_\alpha(\tau)\}_{\tau\geq 0}$ with the rescaled Markov chain $\{\widehat{\chi}_\alpha(\tau)\}_{\tau\geq 0}$ introduced in (4.9). The probability that both processes coincide in the coupling is shown to converge to 1 with an explicit error depending on N, ε, α. Then the invariance principle for the Markov chain (Theorem 4.2) implies, in the limit (4.22), the convergence of the tagged particle in the mechanical process to a Brownian motion. The following theorem is a simplified version of this coupling showing that the distance between the densities at time t of both processes can be estimated quantitatively in terms of N, α, t.

Theorem 4.5. *The distribution* $f_N^{(1)}(t, x, v)$ *of the tagged particle is close to* $M_\beta(v)\varphi_\alpha(t, x, v)$*, where* $\varphi_\alpha(t, x, v)$ *is the solution of the linear Boltzmann equation* (4.8) *with initial data* ρ^0*. More precisely, for all* $t > 0$ *and all* $\alpha > 1$*, in the limit* $N \to \infty$*, with* $N\varepsilon^{d-1} = \alpha$*, one has*

$$\left\| f_N^{(1)}(t, x, v) - M_\beta(v)\varphi_\alpha(t, x, v) \right\|_{L^\infty(\mathbb{T}^d\times\mathbb{R}^d)} \leqslant C \left[\frac{t\alpha}{(\log\log N)^{\frac{A-1}{A}}} \right]^{\frac{A^2}{A-1}}, \tag{4.25}$$

where $A \geqslant 2$ *can be taken arbitrarily large, and* C *depends on* A, β, d *and* $\|\rho^0\|_{\mathbb{L}^\infty}$*.*

A proof can be found in [5] ...

Bibliography

[1] B. J. Alder and T. E. Wainwright. Phase transition for a hard sphere system. *The Journal of Chemical Physics*, 27(5):1208–1209, 1957.

[2] Roger Alexander. *The infinite hard-sphere system*. 1975. Thesis (Ph.D.)–University of California, Berkeley.

[3] Roger Alexander. Time evolution for infinitely many hard spheres. *Comm. Math. Phys.*, 49(3):217–232, 1976.

[4] Patrick Billingsley. *Convergence of probability measures*. Wiley Series in Probability and Statistics: Probability and Statistics. John Wiley & Sons, Inc., New York, second edition, 1999. A Wiley-Interscience Publication.

[5] Thierry Bodineau, Isabelle Gallagher, and Laure Saint-Raymond. The Brownian motion as the limit of a deterministic system of hard-spheres. *Invent. Math.*, 203(2):493–553, 2016.

[6] Thierry Bodineau, Isabelle Gallagher, Laure Saint-Raymond, and Sergio Simonella. One-sided convergence in the boltzmann-grad limit. *Annales de la Faculté des Sciences de Toulouse (in press)*, 2018.

[7] Ludwig Boltzmann. Weitere studien über das wärmegleichgewicht unter gasmolekülen. *Sitzungsberichte der Akademie der Wissenschaften, Wien*, 66(II):275–370, 1872.

[8] D. L. Burkholder, É. Pardoux, and A. Sznitman. *École d'Été de Probabilités de Saint-Flour XIX—1989*, volume 1464 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1991. Papers from the school held in Saint-Flour, August 16–September 2, 1989, Edited by P. L. Hennequin.

[9] Carlo Cercignani. *Ludwig Boltzmann: the man who trusted atoms*. Oxford University Press, 1998.

[10] Carlo Cercignani, Reinhard Illner, and Mario Pulvirenti. *The mathematical theory of dilute gases*, volume 106 of *Applied Mathematical Sciences*. Springer-Verlag, New York, 1994.

[11] Ryan Denlinger. The propagation of chaos for a rarefied gas of hard spheres in the whole space. *Preprint, arXiv:1605.00589*, 2016.

[12] Detlef Dürr, Sheldon Goldstein, and Joel L Lebowitz. Asymptotics of particle trajectories in infinite one-dimensional systems with collisions. *Communications on pure and applied mathematics*, 38(5):573–597, 1985.

[13] Paul Ehrenfest and Tatiana Ehrenfest. Uber zwei bekannte einwände gegen das boltzmannsche h-theorem. *Zeitschrift füür Physik*, 8:311–314, 1907.

[14] Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier. *From Newton to Boltzmann: hard spheres and short-range potentials*. Zurich Lectures in Advanced Mathematics. European Mathematical Society (EMS), Zürich, 2013.

[15] François Golse. Boltzmann-grad limit. *Scholarpedia*, 8(10):9141, 2013.

[16] Georg A. Gottwald and Marcel Oliver. Boltzmann's dilemma: An introduction to statistical mechanics via the kac ring. *SIAM Review*, 51(3):613–635, 2009.

[17] Harold Grad. Principles of the kinetic theory of gases. In *Handbuch der Physik (herausgegeben von S. Flügge), Bd. 12, Thermodynamik der Gase*, pages 205–294. Springer-Verlag, Berlin-Göttingen-Heidelberg, 1958.

[18] Mark Kac. Random walk and the theory of brownian motion. *The American Mathematical Monthly*, 54(7):369–391, 1947.

[19] Mark Kac. Some remarks on the use of probability in classical statistical mechanics. *Bull. Acad. Roy. Belg*, 42:356–361, 1956.

[20] Francis Gordon King. *BBGKY Hierarchy for positive potentials*. ProQuest LLC, Ann Arbor, MI, 1975. Thesis (Ph.D.)–University of California, Berkeley.

[21] Tomasz Komorowski, Claudio Landim, and Stefano Olla. *Fluctuations in Markov processes*, volume 345 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer, Heidelberg, 2012. Time symmetry and martingale approximation.

[22] Oscar E. Lanford, III. Time evolution of large classical systems. pages 1–111. Lecture Notes in Phys., Vol. 38, 1975.

[23] Joel L. Lebowitz. Macroscopic laws, microscopic dynamics, time's arrow and Boltzmann's entropy. *Phys. A*, 194(1-4):1–27, 1993. Statistical physics (Berlin, 1992).

[24] M. Ledoux. *Isoperimetry and Gaussian analysis*, volume 1648 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1996. Lectures from the 24th Saint-Flour Summer School held July 7–23, 1994, Edited by P. Bernard.

[25] J. Loschmidt. Uber den zustand des wärmegleichgewichtes eines systems von körpern mit rücksicht auf die schwerkraft. *Wien. Ber.*, 73:128, 1876.

[26] Sylvie Méléard. Asymptotic behaviour of some interacting particle systems; McKean-Vlasov and Boltzmann models. In *Probabilistic models for nonlinear partial differential equations (Montecatini Terme, 1995)*, volume 1627 of *Lecture Notes in Math.*, pages 42–95. Springer, Berlin, 1996.

[27] S. Olla, S. R. S. Varadhan, and H. T. Yau. Hydrodynamical limit for a hamiltonian system with weak noise. *Communications in Mathematical Physics*, 155(3):523–560, 1993.

[28] M. Pulvirenti, C. Saffirio, and S. Simonella. On the validity of the Boltzmann equation for short range potentials. *Rev. Math. Phys.*, 26(2):1450001, 64, 2014.

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

[30] J. Quastel and H.-T. Yau. Lattice gases, large deviations, and the incompressible navier-stokes equations. *Annals of Mathematics*, 148(1):51–108, 1998.

[31] Fraydoun Rezakhanlou. Equilibrium fluctuations for the discrete Boltzmann equation. *Duke Math. J.*, 93(2):257–288, 1998.

[32] Fraydoun Rezakhanlou. Boltzmann-Grad limits for stochastic hard sphere models. *Comm. Math. Phys.*, 248(3):553–637, 2004.

[33] Fraydoun Rezakhanlou and Cédric Villani. *Entropy methods for the Boltzmann equation*, volume 1916 of *Lecture Notes in Mathematics*. Springer, Berlin, 2008. Lectures from a Special Semester on Hydrodynamic Limits held at the Université de Paris VI, Paris, 2001, Edited by François Golse and Stefano Olla.

[34] David Ruelle. *Statistical mechanics: Rigorous results*. W. A. Benjamin, Inc., New York-Amsterdam, 1969.

[35] Laure Saint-Raymond. *Hydrodynamic limits of the Boltzmann equation*, volume 1971 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 2009.

[36] Sergio Simonella. Evolution of correlation functions in the hard sphere dynamics. *J. Stat. Phys.*, 155(6):1191–1221, 2014.

[37] Frank Spitzer. Uniform motion with elastic collision of an infinite particle system. *Journal of Mathematics and Mechanics*, 18(10):973–989, 1969.

[38] Herbert Spohn. *Large scale dynamics of interacting particles*. Springer Science, 2012.

[39] Domokos Szász and Bálint Tóth. Bounds for the limiting variance of the "heavy particle" in $\mathbf{R}^1$. *Comm. Math. Phys.*, 104(3):445–455, 1986.

[40] Domokos Szász and Bálint Tóth. Towards a unified dynamical theory of the brownian particle in an ideal gas. *Comm. Math. Phys.*, 111(1):41–62, 1987.

[41] W. Thomson. The kinetic theory of the dissipation of energy. *Proceedings of the Royal Society of Edinburgh*.

[42] Salvatore Torquato. *Random heterogeneous materials*, volume 16 of *Interdisciplinary Applied Mathematics*. Springer-Verlag, New York, 2002. Microstructure and macroscopic properties.

[43] Kōhei Uchiyama. Derivation of the Boltzmann equation from particle dynamics. *Hiroshima Math. J.*, 18(2):245–297, 1988.

[44] Seiji Ukai. The boltzmann-grad limit and cauchy-kovalevskaya theorem. *Japan Journal of Industrial and Applied Mathematics*, 18(2):383, 2001.

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

Index