Lanford's Theorem and the Emergence of Irreversibility

Jos Uffink · Giovanni Valente

Abstract It has been a longstanding problem to show how the irreversible behaviour of macroscopic systems can be reconciled with the time-reversal invariance of these same systems when considered from a microscopic point of view. A result by Lanford (Dynamical systems, theory and applications, 1975, Asterisque 40:117–137, 1976, Physica 106A:70–76, 1981) shows that, under certain conditions, the famous Boltzmann equation, describing the irreversible behaviour of a dilute gas, can be obtained from the time-reversal invariant Hamiltonian equations of motion for the hard spheres model. Here, we examine how and in what sense Lanford's theorem succeeds in deriving this remarkable result. Many authors have expressed different views on the question which of the ingredients in Lanford's theorem is responsible for the emergence of irreversibility. We claim that these interpretations miss the target. In fact, we argue that there is no time-asymmetric ingredient at all.

1 Introduction

The Boltzmann equation is one of the most important tools of statistical physics. It describes the evolution of a dilute gas towards equilibrium, and serves as the key to the derivation of further hydrodynamical equations. A striking aspect of this equation

J. Uffink
Department of Philosophy, University of Minnesota, Minneapolis, MN, USA
e-mail: jbuffink@umn.edu

G. Valente (✉)
Pittsburgh, USA
e-mail: valente@pitt.edu

is that it is irreversible, i.e., it is not invariant under time reversal. Indeed, when Boltzmann [2] presented this equation, he immediately derived from it a celebrated theorem, now commonly known as the H-theorem, which shows that a certain quantity H of the gas can only change monotonically in time, so that the gas displays an evolution towards equilibrium.

Despite its long-standing legacy, the status of the H-theorem has remained controversial. The reversibility objection by Loschmidt [19] questioned the validity of the H-theorem by constructing a counterexample. Essentially, this objection raised the problem of how an irreversible macro-evolution equation can be obtained from the time-reversal invariant micro-evolution equations governing molecular motion. More than twenty years later, Culverwell [10] posed the same problem and inaugurated a famous debate in *Nature* with a provocative question: "Will anyone say exactly what the H-theorem proves?".

In his responses to the reversibility objection, [4,5] suggested an alternative approach and reading of the H-theorem, which the [12] called the "modified formulation of the H-theorem", and which we will refer to as the statistical H-theorem. Yet, the problem of providing a rigorous statistical counterpart of the Boltzmann equation and the H-theorem was left unsolved. It is widely believed that a theorem by Oscar [15–17] provides the best available candidate for a rigorous derivation of the Boltzmann equation and the H-theorem from statistical mechanics, in the limiting case of an infinitely diluted gas system described by the hard spheres model, at least for a very brief time.

The proof of Lanford's result is cast in the formalism developed by Bogolyubov, Born, Green, Kirkwood and Yvon (BBGKY). This formalism provides, departing from the Hamiltonian formulation of statistical mechanics, a hierarchy of equations for the time-evolution of macroscopic systems, called the BBGKY hierarchy. On the other hand, the Boltzmann equation itself can also be reformulated in the form of a hierarchy (the Boltzmann hierarchy). Lanford's theorem then shows how the Boltzmann hierarchy can be obtained from the BBGKY hierarchy for the hard spheres model in the Boltzmann–Grad limit under specific conditions. To be sure, the technical assumptions needed in this rigorous derivation present on several points severe limitations. In particular, the convergence obtained in this Boltzmann–Grad limit holds for a very brief time only, and the Boltzmann–Grad limit itself implies that the density of the gas-model goes to zero, which is quite incompatible with the hydrodynamic limit where the Boltzmann equation is actually supposed to work. These clauses of course imply that the theorem will hardly apply to realistic circumstances.

Still, Lanford's theorem has recently been called "maybe the most important mathematical result of kinetic theory" [28]. The importance of this theorem is that it claims to show how the conceptual gap between macroscopic irreversibility and microscopic reversibility can in principle be overcome, at least in simple cases. However, Lanford's papers suggested various answers to the question exactly how the irreversibility embodied in the Boltzmann equation or the ensuing H-theorem arises in this rigorous statistical mechanical setting. Later authors on Lanford's theorem (e.g.: [8,9,18,22–24,26]) have also expressed mutually incompatible views on this particular issue. So, one may well ask: "Will anyone say exactly what Lanford's theorem proves?".

The present paper addresses this question. We analyse the problem of how Lanford's theorem gives rise to the irreversible behaviour of the Boltzmann equation and show that most previous interpretations of the emergence of irreversibility in this theorem miss the target. In fact, we argue that there is no genuine irreversibility in Lanford's theorem. We begin by reviewing the Boltzmann equation and the H-theorem in the kinetic theory of gases for the hard spheres model, along with the quest for a statistical H-theorem (Sect. 2). Section 3 discusses the connection between the BBGKY hierarchy for the hard spheres model and the Boltzmann hierarchy. Lanford's theorem is then stated in Sect. 4. We take up the issue of irreversibility in Sect. 5, and present our conclusions in Sect. 6. To summarize, we conclude that Lanford's theorem does not contain any time asymmetric ingredient. In fact, we argue that all the assumptions of the theorem are time-reversal invariant. We also show that, contrary to claims by Lanford [15], Cercignani et al. [9] and Cercignani [8] the technical procedure of rewriting all collision integrals in terms of an incoming representation, which is used in the proof, does not introduce time asymmetry. In particular, while the initial conditions allowed by the theorem allow one to derive the Boltzmann equation for positive times, they also allow one to derive the so-called 'anti-Boltzmann equation' (the time-reversal transform of the Boltzmann equation) for negative times. Whereas the solutions of the Boltzmann equation lead to an increase of entropy, solutions of the anti-Boltzmann equation lead to a decrease of entropy. The upshot of our analysis is that Lanford's result is time-reversal invariant, and thus it is neutral with respect to the arrow of time. As a consequence, there cannot be any source of irreversibility in the theorem. These conclusions mirror observations that have been made many times concerning Boltzmann's statistical H-theorem. Thus, although Lanford's theorem does not give rise to irreversibility, it does nevertheless provide a mathematically rigorous underpinning of the statistical H-theorem.

2 Boltzmann's Derivation of the Boltzmann Equation and the H-Theorem

In the kinetic theory of gases, one considers a gas as a system consisting of a very large number N of molecules, moving in accordance with the laws of classical mechanics, enclosed in a container Λ with perfectly elastic reflecting and smooth walls. In the hard spheres model, these molecules are further idealized as rigid and impenetrable spheres of diameter a interacting only by collisions. The instantaneous state of the gas system at time t is represented by a *distribution function* $f_t(\vec{q}, \vec{p})$, such that $f_t(\vec{q}, \vec{p})d\vec{q}d\vec{p}$ is supposed to give the relative number of molecules in the gas with positions between $\vec{q}$ and $\vec{q} + d\vec{q}$ inside the container Λ and momenta between $\vec{p}$ and $\vec{p} + d\vec{p}$.

Of course, the question exactly how such a smooth function f_t is meant to represent the distribution of a finite number of particles is somewhat tricky, and cannot be literally true. We shall come back to it later (Sect. 4). For the moment, we notice that, for each time t, f_t formally defines a probability density on the so-called μ-space $\mu = \Lambda \times \mathbb{R}^3$, i.e.:

$$f_t \geq 0 \quad \text{and} \quad \int_\Lambda d\vec{q} \int_{\mathbb{R}^3} d\vec{p} \, f_t(\vec{q}, \vec{p}) = 1, \tag{1}$$

assigning probabilities to molecular positions and momenta —which thus play the role of stochastic variables. However, in kinetic theory, the distribution function itself is thought to represent, in some sense, the relative number of particles over their various possible positions and momenta in the actual microstate of the gas. The distribution function should therefore be sharply distinguished from probability densities as they arise from some probability measure on phase space in statistical mechanics.

In order to describe the evolution of the gas system, one needs to consider how the distribution function $f_t(\vec{q}, \vec{p})$ evolves in time. The crucial assumption in Boltzmann's heuristic derivation of this evolution equation is the *Stoßzahlansatz*, or "assumption about the number of collisions", also often referred to as the Hypothesis of Molecular Chaos, which provides a constraint on the way in which collisions between the particles take place. There are (at least) two distinct versions of this assumption in the literature, which we would like to distinguish:

- Factorization
  The relative number of pairs of particles, with positions within $d\vec{q}_1$ and momenta within $d\vec{p}_1$, and within $d\vec{q}_2$ and $d\vec{p}_2$, respectively, is given by

$$f_t^{(2)}(\vec{q}_1, \vec{p}_1; \vec{q}_2, \vec{p}_2)d\vec{p}_1 d\vec{q}_1 d\vec{p}_2 d\vec{q}_2 = f_t(\vec{q}_1, \vec{p}_1) f_t(\vec{q}_2, \vec{p}_2) d\vec{p}_1 d\vec{q}_1 d\vec{p}_2 d\vec{q}_2. \quad (2)$$

- Pre-collision
  The relative number $N(\vec{q}, \vec{p}_1; \vec{q}, \vec{p}_2)$ of pairs of particles which are *about to collide* in a region $d\vec{q}$ and within a time span dt is proportional to the product $f_t(\vec{q}, \vec{p}_1) f_t(\vec{q}, \vec{p}_2)$ and the volume dV of the "collision cylinder", i.e. the spatial region around the position q at which the particles are located when colliding, i.e.

$$N(\vec{q}, \vec{p}_1; \vec{q}, \vec{p}_2) = N f_t(\vec{q}, \vec{p}_1) f_t(\vec{q}, \vec{p}_2) dV d\vec{p}_1 d\vec{p}_2, \quad (3)$$

where

$$dV = a^2 \pi \, \vec{\omega}_{12} \cdot \left(\frac{\vec{p}_1 - \vec{p}_2}{m} \right) dt d\vec{\omega}_{12}. \quad (4)$$

Here, $\vec{\omega}_{12}$ is a unit vector pointing form the center of particle 1 to the center of particle 2 (See Fig. 1). The condition that the particles are "about to collide" can be expressed mathematically by the condition

$$\vec{\omega}_{12} \cdot (\vec{p}_1 - \vec{p}_2) \geq 0. \quad (5)$$

These two versions of the *Stoßzahlansatz* have an importantly different interpretation. The factorization condition (2) may be interpreted as saying that particle pairs are uncorrelated, i.e. finding a first particle at a particular position $\vec{q}_1$ and moving with momentum $\vec{p}_1$ gives no information about whether we find the second particle at position $\vec{q}_2$ and with momentum $\vec{p}_2$. Note that the condition (2) cannot be literally true

for all $(\vec{q}_1, \vec{p}_1; \vec{q}_2, \vec{p}_2)$: due to their finite diameter, no two particles can have positions such that $\|\vec{q}_1 - \vec{q}_2\| < a$. It is plausible, therefore that Boltzmann implicitly assumed some limit in which $a \longrightarrow 0$.

The pre-collision condition, on the other hand, says that, when we focus on those pairs of particles that are just about to undergo a collision, they are to be regarded as uncorrelated. It is this condition, and not the factorization condition that is actually used in deriving the Boltzmann equation and the H-theorem. Of course, one could note that this condition, again, cannot be literally true for the hard-spheres model for the same reason as mentioned before: as long as the particles have a finite diameter, no two particles can be found at the same position $\vec{q}$ (which would indeed raise questions about the definition of ω_{12}). But again, Boltzmann's derivation of the Boltzmann equation should be, at best, regarded as a heuristic argument, rather than a rigorous proof.

We also note that the literature is somewhat confusing in the terminology here. Many authors use the name "Molecular Chaos" for the factorization condition (2) alone, rather than the pre-collision condition.

Whenever a collision occurs, molecular velocities change. If the particles have momenta $\vec{p}_1$, $\vec{p}_2$ just before the collision, their outgoing momenta will be denoted as $\vec{p}_1{}'$ and $\vec{p}_2{}'$, respectively. In the hard-spheres model, these outgoing momenta are simple functions of $\vec{p}_1$ and $\vec{p}_2$ and $\vec{\omega}_{12}$. Indeed:

$$\begin{aligned}
\vec{p}_1{}' &= \vec{p}_1 - (\vec{\omega}_{12} \cdot (\vec{p}_1 - \vec{p}_2)) \, \vec{\omega}_{12} \\
\vec{p}_2{}' &= \vec{p}_2 + (\vec{\omega}_{12} \cdot (\vec{p}_1 - \vec{p}_2)) \, \vec{\omega}_{12},
\end{aligned} \tag{6}$$

which can be written more compactly in terms of a linear collision operator $T_{\vec{\omega}_{12}}$, defined by (6):

$$(\vec{p}_1, \vec{p}_2) \longrightarrow (\vec{p}_1{}', \vec{p}_2{}') = T_{\vec{\omega}_{12}}(\vec{p}_1, \vec{p}_2). \tag{7}$$

By standard arguments, one obtains from the pre-collision version of the *Stosszahlansatz* the Boltzmann equation, which describes the change of the distribution function in the course of time:

$$\begin{aligned}
&\frac{\partial}{\partial t} f_t(\vec{q}, \vec{p}_1) + \frac{\vec{p}_1}{m} \cdot \frac{\partial}{\partial \vec{q}} f_t(\vec{q}, \vec{p}_1) \\
&= N a^2 \int_{\mathbb{R}^3} d\vec{p}_2 \int_{\vec{\omega}_{12} \cdot (\vec{p}_1 - \vec{p}_2) \geq 0} d\vec{\omega}_{12} \, (\frac{\vec{p}_1 - \vec{p}_2}{m}) \\
&\quad \cdot \vec{\omega}_{12} \left[f_t(\vec{q}, \vec{p}_1') f_t(\vec{q}, \vec{p}_2') - f_t(\vec{q}, \vec{p}_1) f_t(\vec{q}, \vec{p}_2) \right]
\end{aligned} \tag{8}$$

The second term in the left-hand side of the equation accounts for the change of the distribution function through free motion of particles, whereas the right-hand side is the collision term. Here, the variables $\vec{p}_i'$ are to be thought of as implicit functions $\vec{p}_i'(\vec{p}_1, \vec{p}_2)$ given by (7). Note that the collision term is not linear in f_t. Hence, the overall Boltzmann equation is non-linear, and this is a major obstacle

Fig. 1 Geometry of a collision between two hard spheres
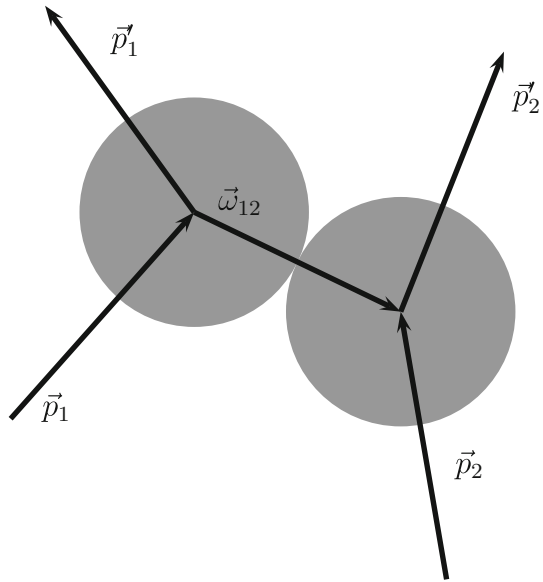
in attempts towards solving the equation. In fact, the question whether the equation does have physically meaningful solutions for all times for some given f_0 as initial condition remains hard even today and has only been answered in special cases.

Boltzmann circumvented this problem by showing that a general theorem could nevertheless be obtained. To derive this result, now commonly known as the H-theorem, he introduced a functional of the distribution function defined as

$$H[f_t] \equiv \int f_t(\vec{q}, \vec{p}) \ln f_t(\vec{q}, \vec{p}) \, d\vec{p} d\vec{q}, \tag{9}$$

and proved that, under the assumption that the Boltzmann equation holds at all times, and f_t is a solution to this equation, then this quantity cannot increase, i.e.

$$\frac{dH[f_t]}{dt} \leq 0 \tag{10}$$

for all t. In the case that the distribution function is and remains spatially uniform, i.e. $f_t(\vec{q}, \vec{p}) = f(\vec{p})$, equality obtains only for a Maxwell distribution $f(\vec{p}) = Ae^{-\vec{p}^2/B}$, which describes the equilibrium distribution. If the negative of the H-function is associated with the entropy of the system, Boltzmann's result means that this entropy increases monotonically through non-equilibrium distirbutionss until the systems reaches equilibrium and then remains constant. Thus, the H-theorem seems to capture the spontaneous approach to equilibrium for gas systems, at least for the hard-spheres gas model.

However, the validity of the H-theorem was called into question soon after its formulation. Loschmidt's reversibility objection, as rephrased by Boltzmann [4] goes

basically as follows. Take a non-equilibrium initial distribution of state f_0 for which the H-theorem holds and let it evolve for a certain amount of time t, so that $H[f_t] < H[f_0]$. Then, suddenly reverse the velocities of all particles. The particles will now simply retrace all their previous motions back to their original spatial configuration at time $2t$. If at that point we reverse their velocities again the distribution of state at time $2t$ will be identical to f_0. But since H, as defined by (9) is invariant under a velocity reversal, this means that under the evolution from t to $2t$, H must have been increasing. In other words, for every dynamically allowed evolution for the particles during which H decreases, one can construct another for which H increases, but also allowed by the dynamics.

This argument relies on the tension between the time-reversal invariance of the dynamics governing the motion of the particles and the explicit time-reversal non-invariance of the H-theorem. In fact, one can trace this time-reversal non-invariance back to the Boltzmann equation from which the H-theorem has been derived. To be explicit, the behaviour under time reversal of the Boltzmann equation may be checked by implementing the usual transformations: (i) replace $\partial/\partial t$ by $-\partial/\partial t$; (ii) reverse the direction of all momenta. It is easy to verify that under these transformations, the left-hand side of (8) changes sign, but the collision integral does not, so that the equation is indeed not time-reversal invariant. This is shown explicitly in *Proposition* 1 in the Appendix. The equation one obtains by applying a time-reversal transformation to the Boltzmann equation is commonly called the *anti-Boltzmann equation*, i.e. the version of the Boltzmann equation with an additional minus sign in front of the collision term.

The upshot of the reversibility objection is that the irreversible time-evolution of macroscopic systems cannot be a consequence of the laws of Hamiltonian mechanics alone. There must be some additional non-dynamical ingredient in the H-theorem, or indeed in the Boltzmann equation from which it follows, that picks out a preferred direction in time. As we now know, the Stoßzahlansatz is the culprit. The pre-collision condition introduces a time-asymmetric element, since it is assumed to hold only for particle pairs immediately *before* collisions, but not for pairs immediately *after* they collided. This is responsible for the failure of the Boltzmann equation to be time-reversal invariant. Indeed, if we had supposed, instead of the pre-collision condition, a similar condition for the momenta immediately after collision, we would, by the same argument, have obtained the anti-Boltzmann equation, and accordingly, we would have derived an anti-H theorem, i.e. $dH[f_t]/dt \geq 0$. Hence the irreversible behaviour in the macro-evolution of non-equilibrium distributions towards equilibrium is due to the preference of this pre-collision rather than a corresponding post-collision condition. But this preference cannot be grounded in the dynamics.

Boltzmann [4] response to Loschmidt already argued that one cannot prove that every initial distribution function should always evolve towards the equilibrium distribution function, but rather that there are infinitely many more initial states that do evolve, in a given time, towards equilibrium than do evolve away from equilibrium, and that even these latter states will evolve towards equilibrium after an even longer time. However, Boltzmann did not provide proofs for these claims. A more detailed argument can be found in [5,6]. To any microstate one can associate a curve (the H-curve), representing the behavior of $H[f_t]$ in the course of time. Boltzmann claimed

that, with the exception of certain 'regular' microstates, the curve exhibits the following properties: (i) for most of the time, $H[f_t]$ is very close to its minimal value H_{min}; (ii) occasionally the H-curve rises to a peak well above the minimum value; (iii) higher peaks are extremely less probable than lower ones. If at time $t = 0$ the curve takes on a value $H[f_0]$ much greater than H_{min}, the function may evolve only in two alternative ways. Either $H[f_0]$ lies in the neighborhood of a peak, and hence $H[f_t]$ decreases in both directions of time; or it lies on an ascending or descending slope of the curve, and hence $H[f_t]$ would correspondingly decrease or increase. However, statement (iii) entails that the first case is much more probable than the second. One would thus conclude that there is a very high probability that at time $t = 0$ the entropy of the system, associated with the negative of the H-function, would increase for positive time; likewise there is a very high probability that the entropy would increase for negative time. It is this conclusion that is sometimes called the *statistical H-theorem*.

Nevertheless, Boltzmann gave no proof of these claims, nor did he indicate whether or how they might still depend on the Boltzmann equation, or the *Stoßahlansatz*. In fact, the statistical H-theorem is hardly a theorem at all. The problem of finding an analogue of Boltzmann's H-theorem in statistical mechanics thus remained unsettled. In order to make progress upon this problem, many authors have called upon Lanford's theorem. Indeed, this theorem is often presented as providing a rigorous derivation of the Boltzmann equation and the associated H-theorem from statistical mechanics.

3 Lanford's Derivation of the Boltzmann Equation from Hamiltonian Statistical Mechanics

3.1 From the Hamiltonian Framework to the BBGKY Hierarchy

In this section we briefly describe the general form of the BBGKY hierarchy. Again, we consider a classical mechanical system consisting of N particles, each with the same mass m. In order to alleviate a bit the notation of the equations to follow, we will set $m = 1$. The particles are contained in a vessel $\Lambda \subset \mathbb{R}^3$ with a finite volume and smooth wall $\partial\Lambda$. But we now approach this system from statistical mechanics, rather than kinetic theory. Its $6N$-dimensional phase space is given by $\Gamma_N = (\Lambda \times \mathbb{R}^3)^N$ and its evolution is governed by a Hamiltonian

$$H_N(x) = \sum_{i=1}^{N} \vec{p}_i{}^2 + \sum_{i<j}^{N} \phi(\vec{q}_i - \vec{q}_j). \tag{11}$$

Here, x denotes the microstate $x = (x_1, \ldots, x_N) = (\vec{q}_1, \vec{p}_1, \ldots, \vec{q}_N, \vec{p}_N)$.

Strictly speaking, the Hamiltonian should also contain a term corresponding to the elastic wall potential, describing the interaction when individual particles collide with the boundary $\partial\Lambda$ of the vessel. However, there are ways to suppress this complication. The easiest way is to suppose that each particle i undergoes specular reflection when it hits the wall and identify the values $(\vec{q}_i, \vec{p}_i)$ just before such a collision and the values $(\vec{q}_i, \vec{p}_i')$ immediately after. In this move, the phase-space Γ_N is endowed with

the topology of a torus, and the dynamics under wall collisions becomes smooth. Indeed, a collision with the wall becomes indistinguishable from free motion, and consideration of the wall potential becomes redundant.

Now, although we will eventually focus on the hard-sphere model, i.e. the special case when

$$\phi(\vec{q}_i - \vec{q}_j) = \begin{cases} \infty & \text{when} \|(\vec{q}_i - \vec{q}_j)\| \leq a. \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

We assume, for now, that ϕ is a smooth function obeying the Lipshitz condition. The virtue of this assumption is that, in this case, the Hamiltonian (11) is known to be integrable, so that there exists a smooth one-parameter group of transformations, $\{T_t, t \in \mathbb{R}\}$ on Γ_N, called the Hamiltonian flow, $T_t : \Gamma_N \longrightarrow \Gamma_N, \Gamma_N \ni x \mapsto x_t = T_t(x)$ that characterizes the dynamics.

The statistical state of the system is given by a probability measure μ over Γ_N. We assume that μ is absolutely continuous with respect to the Lebesgue measure on Γ, so one can write

$$\mu(A) = \int_A \mu(x) dx, \tag{13}$$

in terms of a density function $\mu(x)$ with respect to the Lebesgue measure on Γ.

The evolution of such a statistical state $\mu(x)$ at any instant t is defined by

$$\mu_t(x) = \mu(T_{-t}x), \tag{14}$$

in terms of the Hamiltonian flow or, equivalently, by means of the Liouville equation

$$\frac{\partial \mu}{\partial t} = \{H, \mu\} = \sum_{i=1}^{N} \frac{\partial H}{\partial \vec{q}_i} \frac{\partial \mu}{\partial \vec{p}_i} - \frac{\partial H}{\partial \vec{p}_i} \frac{\partial \mu}{\partial \vec{q}_i} =: \mathcal{H}_N \mu. \tag{15}$$

The BBGKY approach exploits the fact that the above Hamiltonian (11) is invariant under permutation of the particles, and that, moreover, the inter-particle potential ϕ only contains pair-interactions. Furthermore, we assume that μ is permutation invariant as well:

$$\mu(x_1, \ldots, x_i \ldots, x_j, \ldots x_N) = \mu(x_1, \ldots, x_j \ldots, x_i, \ldots x_N)$$
$$\forall i, j \in \{1, \ldots, N\}, i \neq j. \tag{16}$$

Obviously, for a permutation invariant Hamiltonian such as (11), this property of μ_t will be conserved under the dynamical evolution (15).

With the above symmetry assumptions in place, it is clear that macroscopic quantities of physical interest will only depend on *how many* particles have certain molecular properties, or how many pairs have certain relations to each other, etc., but not on their particle labels. It thus becomes attractive to study the dynamics in terms of reduced

probability densities obtained by conveniently integrating out most of the variables. For this purpose, one defines a hierarchy of reduced or marginal probability densities:

$$\mu_1(x_1) := \int \mu(x_1, \ldots x_N) dx_2 \cdots x_N,$$

$$\vdots$$

$$\mu_k(x_1, \ldots, x_k) := \int \mu(x) dx_{k+1} \cdots dx_N, \qquad (17)$$

$$\vdots$$

$$\mu_N(x_1, \ldots, x_N) := \mu(x_1, \ldots, x_N).$$

Here, for instance, μ_k is the probability density that k particles occupy positions $\vec{q}_1, \ldots, \vec{q}_k$ and move with momenta $\vec{p}_1, \ldots, \vec{p}_k$, while the remaining $N - k$ particles possess arbitrary positions and momenta. Note that, although $\mu_1(x)$ is thus a marginal probability density on μ-space, just like the distribution function $f(\vec{q}, \vec{p})$ discussed in Sect. 2, the conceptual status of these two density functions is very different.

With a somewhat different normalization convention, one defines rescaled reduced probability densities:

$$\rho_k(x_1, \ldots, x_k) = \frac{N!}{(N-k)!} \frac{1}{N^k} \mu_k(x_1, \ldots, x_k). \qquad (18)$$

It remains, of course, to specify the time evolution of these rescaled reduced probability densities.

Now, the N-particle Liouville operator $\mathcal{H}_N$ in the Liouville equation (15) can be expanded as

$$\mathcal{H}_N = -\sum_{i=1}^{N} \vec{p}_i \cdot \frac{\partial}{\partial \vec{q}_i} + \sum_{i \neq j}^{N} \mathcal{L}_{ij}, \qquad (19)$$

where

$$\mathcal{L}_{ij} := \frac{\partial \phi(\vec{q}_i - \vec{q}_j)}{\partial \vec{q}_i} \cdot \frac{\partial}{\partial \vec{p}_i}. \qquad (20)$$

The evolution of ρ_1 is therefore given by

$$\frac{\partial \rho_{1,t}(x_1)}{\partial t} = \vec{p}_1 \cdot \frac{\partial}{\partial \vec{q}_1} \rho_{1,t}(x_1) + N \int dx_2 \mathcal{L}_{12} \rho_{2,t}(x_1, x_2), \qquad (21)$$

and for the higher-order rescaled reduced probability densities:

$$\frac{\partial \rho_{k,t}}{\partial t} = -\sum_{i=1}^{N} \vec{p}_i \cdot \frac{\partial}{\partial \vec{q}_i} \rho_{k,t} + \sum_{i \neq j}^{k} \mathcal{L}_{ij} \rho_{k,t} + N \sum_{i=1}^{k} \int dx_{k+1} \frac{\partial \phi(\vec{q}_i - \vec{q}_{k+1})}{\partial \vec{q}_i}$$

$$\cdot \frac{\partial}{\partial \vec{p}_i} \rho_{k+1,t}. \tag{22}$$

Or, in abbreviated form:

$$\frac{\partial \rho_{k,t}}{\partial t} = \mathcal{H}_k \rho_{k,t} + \mathcal{C}^{\phi}_{k,k+1} \rho_{k+1,t}, \quad k = 1, \dots, N, \tag{23}$$

where the superscript on the operator $\mathcal{C}$ is intended to remind one that it depends on the smooth inter-particle potential ϕ in (11).

These dynamical equations for the rescaled reduced densities of the statistical state μ constitute the BBGKY hierarchy. Note that, taken together, they are strictly equivalent to the Hamiltonian evolution, i.e. nothing else has been assumed yet, except for the rather harmless permutation invariance of ρ and the specific form of the Hamiltonian (11). As one might expect, therefore, solving these equations is just as hard as for the original Hamiltonian equations. Indeed, to find the time-evolution of ρ_1 from 21, we need to know $\rho_{2,t}$. But to solve the dynamical equation for ρ_2, we need to know $\rho_{3,t}$ etc. Moreover, the equations of the BBGKY hierarchy are still perfectly time-reversal invariant.

Nevertheless, the above might already make one hopeful that a counterpart of the Boltzmann equation can be obtained from the exact Hamiltonian dynamics. Indeed, if we tentatively identify Boltzmann's f function with ρ_1, (21) looks somewhat similar to the Boltzmann equation (8). Of course, much work still remains to be done: first of all, the Boltzmann equation pertains to the hard-sphere model, whereas the Eq. 21 assumes a smooth pair-potential $\phi(\vec{q}_i - \vec{q}_j)$. More importantly, we would have to justify the tentative relationship between ρ_1 and f. These tasks will be addressed in the following subsections.

3.2 From Smooth Potentials to the Hard-spheres Gas Model

While the BBGKY hierarchy provides a generally useful format for studying the evolution of a statistical state for a system of indistinguishable particles interacting by smooth pair potentials, it is our purpose here to apply it to the hard-spheres potential 12.

There are several caveats when applying Hamiltonian dynamics or the BBKGY hierarchy to the case of a hard-spheres model, in particular, because the potential 12 of this model does not obey the Lipshitz condition. First of all, we have to remove configurations in which particles overlap, i.e. restrict our original phase space Γ_N to:

$$\Gamma^{(a)}_{N,\neq} := \{ x \in (\Lambda \times \mathbb{R}^3)^N : \|\vec{q}_i - \vec{q}_j\| \geq a, \ i \neq j, \ i, j \in \{1, \dots, N\} \}. \tag{24}$$

More importantly, the dynamical evolution of the microstate of a collection of N hard spheres enclosed in a vessel might lead to (i) grazing collisions (ii) more than two particles colliding simultaneously or (iii) an infinite number of collisions (either between the particles mutually or between some particle and the wall) occurring within a finite lapse of time. In all of these cases, the Hamiltonian equations cannot be solved,

and the trajectory in phase space cannot be extended for all times. Fortunately, it has been shown by Alexander [1] that the subset consisting of microstates x showing such anomalous evolutions has a Lebesgue measure zero in $\Gamma_{N,\neq}^{(a)}$. Therefore, if, as we assumed, the statistical state μ is absolutely continuous with respect to the Lebesgue measure, these unwanted microstates make up a set of probability zero, and can be ignored for the purpose of our analysis.

That is to say, we can either delete this unwanted set Δ of measure zero from our phase space $\Gamma_{N,\neq}^{(a)}$, and in doing so guarantee that there is a Hamiltonian flow $\{T_t, t \in \mathbb{R}\}$ defined on the smaller phase space $\Gamma_{N,\neq} \setminus \Delta$, or continue with the original space, with the provision that its Hamiltionian flow is defined only almost everywhere, i.e. outside of the above set Δ.

Thus, the hard-sphere dynamics is such that if we consider any given phase point $x = (x_1, \ldots x_N)$ ($x \notin \Delta$) and consider how it will move under the flow in the next sufficiently small time increment δt, then either all particles persist in free motion (or perhaps some collide with the wall); or else some pair of particles, say i and j, collide. In the latter case, at the moment of collision, they touch, i.e., their positions obey

$$\vec{q}_j = \vec{q}_i + a\vec{\omega}_{ij} \quad \text{for} \quad \vec{\omega}_{ij} = \vec{q}_j - \vec{q}_i, \tag{25}$$

which implies that the microstate x lies on the boundary of $\Gamma_{N,\neq}$, and in the collision their momenta undergo an instantaneous transition, cf. (7):

$$(\vec{p}_i, \vec{p}_j) \longrightarrow (\vec{p}_i', \vec{p}_j') = T_{\vec{\omega}_{ij}}(\vec{p}_i, \vec{p}_j). \tag{26}$$

Note that $T_{\vec{\omega}}$ is measure preserving, and an involution, i.e. $T_{\vec{\omega}} \circ T_{\vec{\omega}} = \mathbf{1}$. In other words, whenever the incoming momenta before a collision between particles i, j happen to take values $(\vec{p}_i', \vec{p}_j')$, they are transformed into $(\vec{p}_i, \vec{p}_j)$:

$$(\vec{p}_i', \vec{p}_j') \longrightarrow (\vec{p}_i, \vec{p}_j) = T_{\vec{\omega}_{ij}}(\vec{p}_i', \vec{p}_j'). \tag{27}$$

Now, although this momentum transfer during collision is clearly discontinuous, one can nevertheless maintain the idea that the dynamics is smooth, by mimicking a procedure already applied to deal with collisions with the wall, i.e., by adopting a topology in which the pre-collision coordinates $(\vec{q}_i, \vec{p}_i; \vec{q}_j, \vec{p}_j)$ and the post-collision coordinates $(\vec{q}_i, \vec{p}_i'; \vec{q}_j, \vec{p}_j')$ are identified. We will discuss this procedure in greater detail in Sect. 5.

With these caveats taken care of, we thus recover a smooth dynamics for the hard spheres model, and indeed one can show that the Eq. (23) go over in

$$\frac{\partial \rho_{k,t}^{(a)}}{\partial t} = \mathcal{H}_k \rho_{k,t}^{(a)} + \mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,t}^{(a)} \quad k \in \{1, \ldots, N\}, \tag{28}$$

where now

$$
\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,t}^{(a)} = Na^2 \sum_{i=1}^{k} \int_{\mathbb{R}^3} d\vec{p}_{k+1} \int_{S^2} d\vec{\omega}_{i,k+1},
$$
$$
\left(\vec{\omega}_{i,k+1} \cdot (\vec{p}_{k+1} - \vec{p}_i) \right) \rho_{k+1,t}^{(a)}(x_1, \ldots, x_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}), \qquad (29)
$$

and the superscript (a) is intended to remind one that these operators and the rescaled probability densities refer to a hard spheres model with a sphere diameter $a > 0$. Of course, for each value of k, these rescaled probability densities $\rho_k^{(a)}$ are defined only on the domains

$$
\Gamma_{k,\neq}^a = \{ (x_1, \ldots x_k) \in (\Lambda \times \mathbb{R}^3)^k : \ \|\vec{q}_i - \vec{q}_j\| \geq a, \ i \neq j, \ i, j \in \{1, \ldots k\} \}. \quad (30)
$$

We emphasize that the resulting form (28) of the BBGKY hierarchy for the hard sphere model is still time-reversal invariant. This is shown explicitly in *Proposition* 2 in the Appendix.

However, one more step is needed in order to obtain Lanford's theorem. Let us split the integral over the unit sphere S^2 into two parts: the hemisphere $\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \geq 0$, and the hemisphere $\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \leq 0$. In the first hemisphere, the collision configuration $(\vec{q}_i, \vec{p}_i; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ represents a collision between particles i and $k+1$ with incoming momenta $\vec{p}_i, \vec{p}_{k+1}$, and we leave the integrand as it is.

In the other hemisphere, characterized by $\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \leq 0$, the momenta in the configuration $(\vec{q}_i, \vec{p}_i; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ appear as outgoing momenta. In this hemisphere, these momenta are replaced by the corresponding incoming momenta, which, according to (27), gives the configuration:

$$
(\vec{q}_i, \vec{p}_i{}'; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}') \quad \text{where} \quad (\vec{p}_i{}', \vec{p}_{k+1}') = T_{\omega_{i,k+1}}(\vec{p}_i, \vec{p}_{k+1}). \quad (31)
$$

Also, we replace the integration variable $\vec{\omega}_{i,k+1}$ by $-\vec{\omega}_{i,k+1}$. The result of these operations is that we obtain from (29) the collision term:

$$
\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,t}^{(a)}(x_1, \ldots, x_k) = Na^2 \sum_{i=1}^{k} \int_{\mathbb{R}^3} d\vec{p}_{k+1} \int_{\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \geq 0} d\vec{\omega}_{i,k+1}
$$
$$
\left(\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \right) \left[\rho_{k+1,t}^{(a)}(x_1, \ldots, \vec{q}_i, \vec{p}_i{}', \ldots x_k, \vec{q}_i - a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}') \right.
$$
$$
\left. - \rho_{k+1,t}^{(a)}(x_1, \ldots, \vec{q}_i, \vec{p}_i, \ldots x_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}) \right] \quad (32)
$$

Summing up Lanford's argument so far, the general BBGKY hierarchy has been applied to the particular case of the hard-spheres model. An Eq. (28) for the time-evolution of the relevant reduced probability densities including the details of both collisions and rectilinear motion of the particles is thus obtained. In the last passage from (29) to (32) a particular step was made, namely to rewrite the integrands in terms of pre-collision rather than the post-collision configurations. This step was accompanied by the argument that one may identify these configurations as representing

the same physical phase point, as Lanford himself suggested. Actually, as we shall see in Sect. 5, this step is regarded by some authors as crucial for the emergence of irreversibility in Lanford's theorem, although this is disputed by others.

3.3 From the Boltzmann Equation to the Boltzmann Hierarchy

In this section, we start from the other side of the bridge that we aim to cross. That is, we take the Boltzmann equation as given, and reformulate it in a mathematically equivalent hierarchy of distribution functions. This idea is captured by the lemma below, which is spelled out by Lanford [15], p. 88.

First, define a hierarchy of multi-particle distribution functions by

$$f_{k,t}(x_1,\ldots,x_k) = \prod_{i=1}^{k} f_t(x_i) \quad k \in \mathbb{N}, \tag{33}$$

where $x_i = (\vec{q}_i, \vec{p}_i)$. Then we have:

Lemma 3.1 *The following two statements are equivalent: (i): f_t is a solution of the Boltzmann equation and (ii) the functions $f_{k,t}$ obey the equations*

$$\frac{\partial f_{k,t}}{\partial t} = \mathcal{H}_k f_{k,t} + \mathcal{C}_{k,k+1} f_{k+1,t} \quad k \in \mathbb{N}, \tag{34}$$

where:

$$\mathcal{H}_k := \sum_{i=1}^{k} \mathcal{L}_i := -\sum_{i=1}^{k} \vec{p}_i \cdot \frac{\partial}{\partial \vec{q}_i}, \tag{35}$$

and

$$
\begin{aligned}
&C_{k,k+1} f_{k+1,t}(x_1,\ldots,x_k) \\
&= Na^2 \sum_{i=1}^{k} \int_{\vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})\geq 0} d\vec{\omega}_{i,k+1} d\vec{p}_{k+1} \ \left(\vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})\right) \\
&\quad \times \left[f_{k+1,t}(x_1,\ldots,\vec{q}_i,\vec{p}_i{\,}',\ldots,\vec{q}_i,\vec{p}_{k+1}') - f_{k+1,t}(x_1,\ldots,\vec{q}_i,\vec{p}_i,\ldots,\vec{q}_i,\vec{p}_{k+1}) \right],
\end{aligned}
\tag{36}
$$

and

$$(\vec{p}_i{\,}', \vec{p}_{k+1}') = T_{\omega_{i,k+1}}(\vec{p}_i, \vec{p}_{k+1}). \tag{37}$$

In other words, the problem of solving the Boltzmann equation for a distribution function f is equivalent to the problem of solving a hierarchy of evolution Eq. (34), called the *Boltzmann hierarchy*, for the functions $(f_1, f_2, \ldots)$ under the assumption of a factorization condition (33). One can write this hierarchy more compactly by regarding the f_k as components of a vector: $\boldsymbol{f} = (f_1, f_2, \ldots)$. Then we can write (34) schematically as:

$$\frac{\partial}{\partial t} \boldsymbol{f} = \mathcal{H}\boldsymbol{f} + \mathcal{C}\boldsymbol{f}, \tag{38}$$

where $\mathcal{H}$ is a diagonal matrix with diagonal elements $\mathcal{H}_k$ and $\mathcal{C}$ a matrix with elements $\mathcal{C}_{k,k+1}$ and zero elsewhere.

The lemma has two virtues: First, and most important is the point that while the original non-linear Boltzmann equation is notoriously hard to solve, the Boltzmann hierarchy (38) is linear. This contrast arises, of course, because the non-linearity is put, so to say, in the factorization constraint (33). As a consequence, it is easier to write down (at least formal) solutions to the Boltzmann hierarchy. A formal solution to this hierarchy of equations is obtained by writing down an expansion familiar from Dyson's time-dependent perturbation theory:

$$\begin{aligned} \boldsymbol{f}_t &= S(t)\boldsymbol{f}_0 \\ &+ \sum_{i=1}^{\infty} \int_0^t dt_1 \int_0^{t_1} dt_2 \ldots \int_0^{t_{i-1}} dt_m S(t-t_1)\mathcal{C}S(t_1-t_2)\mathcal{C} \cdots \mathcal{C}S(t_m)\boldsymbol{f}_0, \end{aligned} \tag{39}$$

where the operator $\mathcal{S}(t)$ represents the collisionless time evolution, i.e.:

$$\mathcal{S}(t)f_k(x_1,\ldots x_k) := f_k(\vec{q}_1 - t\vec{p}_1, \ \vec{p}_1, \vec{q}_2 - t\vec{p}_2, \ \vec{p}_2, \ \ldots, \ \vec{q}_N - t\vec{p}_N, \vec{p}_N). \tag{40}$$

Obviously, the above formal way of writing a general solution to the Boltzmann hierarchy does not alleviate the original problems in solving the Boltzmann equation entirely; these problems are merely transposed into a further problem of showing that the series expansion in (39) actually converge.

The second virtue of the lemma is that it brings the Boltzmann equation in a form which is more similar to the results from the BBGKY formalism discussed above, which likewise take the form of a hierarchy, and this alleviates the effort to build a rigorous bridge between them.[1]

As we have remarked above, the factorization condition (33), taken as a generalization of Boltzmann's condition (2), is sometimes called 'molecular chaos'. This is an unfortunate habit because (33) does not contain an accompanying condition to single out the pre-collision coordinates, as a generalization of (3). Nevertheless, it is worth noting that if the initial data of the Boltzmann hierarchy at time $t = 0$ take the form (33), then this factorization is preserved through time, i.e. it holds for the solution of (34) for all time t, with f_t being a solution of the Boltzmann equation. This important property of the Boltzmann hierarchy is commonly known as 'propagation of chaos'.[2] We emphasize, however, that this factorization, and its preservation in time, has nothing to do with the pre-collision condition (3) mentioned in Sect. 2 as a crucial ingredient of the molecular chaos hypothesis Boltzmann used to obtain the Boltzmann equation.

---

[1] The similarity is not complete, of course. For example, the BBGKY hierarchy is finite while the Boltzmann hierarchy is not.

[2] See ([23], p. 45; [9], p. 85) for a more extensive discussion of this property.

Finally, we stress that the Boltzmann hierarchy, being an equivalent way of expressing the Boltzmann equation, is just as time-reversal non-invariant as the original Boltzmann equation. In fact, by applying a time-reversal transformation to it, one obtains a hierarchy of evolution equations which has the same form as (34) except for a minus sign in front of the collision term $C_{k,k+1}$. We refer to the latter as the *anti-Boltzmann hierarchy*. The 'propagation of chaos' property which we just noted for the Boltzmann hierarchy is also valid for this anti-Boltzmann hierarchy. Also, notice that both the collision operators and the distribution functions in (36) resemble those involved in (32), except that they do not depend on the diameter a of the particles. The crucial point in Lanford's theorem is to demonstrate that all relevant terms in the BBGKY hierarchy tend to their counterparts in the Boltzmann hierarchy in the Boltzmann–Grad limit, whereby $a \longrightarrow 0$. That would establish that the Boltzmann equation can be obtained from Hamiltonian mechanics.

4 Lanford's Theorem

So far, we have seen how the Hamiltonian dynamics for the hard-spheres model leads, under a relatively harmless assumption of permutation invariance, to a hierarchy of BBGKY equations describing the evolution of reduced density functions of a statistical state. And we have also seen how the Boltzmann equation can be reformulated as a hierarchy of equations in close resemblance to the BBGKY hierarchy. The question still remains how to bridge the gap between these two descriptions. Lanford's theorem establishes the convergence of the BBGKY hierarchy to the Boltzmann hierarchy in the so-called Boltzmann–Grad limit.

This Boltzmann–Grad limit defines a particular limiting regime within the hard spheres model. In this limit, one not only lets the number of particles grow to infinity, i.e. $N \to \infty$, but also requires that their diameter goes to zero, i.e. $a \to 0$, while keeping the volume $|\Lambda|$ of the container fixed. The limit is taken in such a way that the quantity Na^2 remains constant, or at least approaches a finite non-zero value. This guarantees that the collision term in the Boltzmann equation or Boltzmann hierarchy, which is proportional to this quantity, does not vanish. Accordingly, the 'mean free path' $\lambda := \frac{|\Lambda|}{2\pi Na^2}$, which is the typical scale-distance traveled by any particle between two subsequent collisions in an equilibrium state, also remains of order one. The same holds for the 'mean free time', i.e. the typical duration between collisions in equilibrium, which is of the order $\sqrt{(\beta/3)}\pi a^2 N/|\Lambda|$, where β is the inverse temperature.

There is one final technical point we need to mention. Recall that the rescaled probability densities $\rho_k^{(a)}$ of the BBGKY hierarchy have as their domains the sets (28). As one takes the Boltzmann–Grad limit, these sets converge to

$$\Gamma_{k,\neq} := \bigcup_{a>0} \Gamma_{k,\neq}^{(a)} = \{(x_1, \ldots x_k) \in (\Lambda \times \mathbb{R}^3)^k : \vec{q}_i \neq \vec{q}_j, \ i \neq j, \ i, j \in \{1, \ldots k\} \}.$$

(41)

Obviously, we cannot expect the convergence of $\rho_k^{(a)} \longrightarrow f_k$ everywhere in $\Gamma_k := (\Lambda \times \mathbb{R}^3)^k$, but at most on $\Gamma_{k,\neq}$, i.e., away from the hypersurface $\Gamma_{k,=} := \Gamma_k \setminus \Gamma_{k,\neq}$

of phase points for which two particles (now considered as point particles), coincide in space. Actually, we need to be even a little bit more restrictive. Let

$$\Gamma_{k,\neq}(s) := \{(x_1, \ldots x_k) \in (\Lambda \times \mathbb{R}^3)^k : \vec{q}_i - s\vec{p}_i \neq \vec{q}_j - s\vec{p}_j,$$
$$i \neq j, \ i, j \in \{1, \ldots k\}, \forall t : 0 \leq t \leq s \}. \tag{42}$$

In words, this is the set of k point particle configurations for which no particle pairs collide at time 0, but also have not collided within a time span $[-s, 0]$.

We are now ready to state the precise version of Lanford's theorem, as given by Spohn [23], Theorem 4.5. Here, when we write $\lim_{a \to 0}$, the Boltzmann–Grad limit is understood, i.e. it is assumed that $N \longrightarrow \infty$ simultaneously, while keeping Na^2 a fixed non-zero constant.

LANFORD'S THEOREM
With the notation introduced in Sect. 3, take $0 < a < a_0$ and let $\rho_{k,t}^{(a)}$ be a family of functions defined on $\Gamma_{k,\neq}^{(a)}$, and assume that for all such a, the following conditions hold at time $t = 0$.

(i) There exists positive real constants z, β, M, independent of a, such that

$$\rho_{k,0}^{(a)}(x_1, \ldots, x_k) \leq Mz^k \prod_{i=1}^{k} h_\beta(\vec{p}_i), \tag{43}$$

for any $k = 1, 2, \ldots$, where $h_\beta(\vec{p}_i)$ denotes the normalized Maxwellian distribution over momenta: $h_\beta(\vec{p}_i) = (\frac{\beta}{2\pi})^{\frac{3}{2}} \cdot e^{-\frac{\beta \vec{p}_i^2}{2}}$ at inverse temperature β, and the spatial distribution is constant inside the vessel Λ with density z.

(ii) There exist continuous functions $f_{k,0}$ on Γ_k, for $k = 1, 2 \ldots$ such that

$$\lim_{a \to 0} \underset{(x_1, \ldots x_k) \in K}{\text{ess sup}} |\rho_{k,0}^{(a)}(x_1, \ldots, x_k) - f_{k,0}(x_1, \ldots x_k)| = 0, \tag{44}$$

for all compact subsets $K \subset \Gamma_{k,\neq}(s)$ for some $s \geq 0$.
Then, there exists a strictly positive time τ, such that for all times $0 \leq t \leq \tau$

$$\lim_{a \to 0} \underset{(x_1, \ldots, x_k) \in K}{\text{ess sup}} |\rho_{k,t}^{(a)}(x_1, \ldots, x_k) - f_{k,t}(x_1, \ldots, x_k)| = 0, \tag{45}$$

for any $k = 1, 2, \ldots$ and compact subset $K \subset \Gamma_{k,\neq}(s + t)$.
Here, $\rho_{k,t}^{(a)}$ are solutions of the BBGKY hierarchy with initial conditions $\rho_{k,0}^{(a)}$ and $f_{k,t}$ solutions of the Boltzmann hierarchy with initial conditions $f_{k,0}$.

Let us make some comments on the theorem. Assumption (i) admits only initial conditions for the rescaled reduced densities $\rho(a)_{k,0}$ of the BBGKY hierarchy bounded by a k-fold product of uniform spatial density z and a Maxwellian distribution over the momenta with inverse temperature β. This assumption is described by Lanford [16] as a regularity condition which prevents building in very strong correlations into

the sequence of initial densities $\rho_{k,0}^{(a)}$ of the BBGKY hierarchy, while at the same time preventing a significant probability for particles of very high energy. In fact, [23] has shown that condition (i) implies that if this sequence of densities converges in the limit $a \longrightarrow 0$ (as it does in the sense of assumption (ii)), it must be to a sequence of convex combination of factorizing densities:

$$\lim_{a \longrightarrow 0} \rho_{k,0}^{(a)} = \int_{L_+^\infty} P(df) \prod_{i=1}^{k} f_0(x_i), \tag{46}$$

where P denotes a probability measure on density functions. This means that, in the limit $a \longrightarrow 0$, correlations are indeed severely restricted by assumption (i), since the densities $\rho_{k,0}^{(a)}$ will then become conditionally independent.

By assumption (ii), the sequence of initial conditions for the BBGKY hierarchy converge to functions $f_{k,0}$ that serve as initial conditions of the Boltzmann hierarchy. The theorem then states that this convergence is maintained through time, at least for $t \in [0, \tau]$, so that solutions $\rho_{k,t}^{(a)}$ of the BBGKY hierarchy converge to solutions $f_{k,t}$ of the Boltzmann hierarchy as $a \to 0$ for all k, except for the phase-points comprised in the set $\Gamma_{k,=}(s+t)$. The size of such exceptional sets increases in time, i.e. $\Gamma_{k,=}(s+t) \subset \Gamma_{k,=}(s+t')$ if $0 < t < t'$. It follows that the type of convergence obtained for $\rho_{k,t}^{(a)}$ in (45) is weaker than the convergence assumed for the initial conditions $\rho_{k,0}^{(a)}$ in (55). This is actually to be expected due to the fact that the BBGKY hierarchy is time-reversal invariant, while the Boltzmann hierarchy is not. Note, however, that, being hypersurfaces of codimension one, the exceptional sets $\Gamma_{k,=}(s+t)$ all have Lebesgue measure zero for any time t, and hence they also have probability zero for any statistical state which is absolutely continuous with respect to the Lebesgue measure.

Now, if we add the further assumption that the functions $f_{k,0}$ in equation (55) factorize according to (33), that is

$$f_{k,0}(x_1, \ldots, x_k) = \prod_{i=1}^{k} f_0(x_i), \tag{47}$$

then we can infer by the Lemma of Sect. 3.3, and the propagation of chaos, that this factorization property is maintained in time:

$$f_{k,t}(x_1, \ldots, x_k) = \prod_{i=1}^{k} f_t(x_i), \tag{48}$$

where f_t is a solution of the Boltzmann equation. In that case, the Lanford theorem not only obtains convergence of $\rho_k^{(a)}$ to solutions of the Boltzmann hierarchy, but also obtains convergence to solutions of the Boltzmann equation in the sense of (48), for the duration $t \in [0, \tau]$. This completes the derivation of the Boltzmann equation from the BBGKY hierarchy for hard spheres. Thus, under the factorization condition

(47), Lanford's result establishes the limiting validity of the Boltzmann equation in the Boltzmann–Grad regime. But note that such an assumption is not actually needed in the above theorem.

Nevertheless, even if the factorization condition (47) is not needed in the Lanford theorem, the result by Spohn just mentioned, i.e., that the assumptions of the theorem imply (46), show that these assumptions *almost* imply factorization of the limiting densities. Moreover, since the Boltzmann hierarchy is linear and propagates chaos, it will preserve this convex combination for later times, i.e., if the initial conditions for the Boltzmann hierarchy take the form of the right-hand side of Eq. (46), they will evolve in time to

$$f_{k,t}(x_1, \ldots, x_k) = \int_{L_+^\infty} P(df) \prod_{i=1}^{k} f_t(x_i).$$

(49)

It is worth pointing out a few limitations on the applicability and physical relevance of Lanford's theorem. A rather serious drawback lies in the usage of the Boltzmann–Grad limit. Indeed, the fact that one lets the number N of particles in the gas go to infinity, while at the same time letting their diameter a go to zero in such a way that the quantity Na^2 remains constant, implies that $Na^3 \to 0$. This means that the gas becomes infinitely diluted in the Boltzmann–Grad limit, and hence the result would apply just to infinitely diluted gases. So, even though it seems reasonable to impose this limit in order to give the Boltzmann equation a fighting chance to be valid, the theorem can hardly be relevant to real-life gases in which the density is not close to zero. The main merit of Lanford's theorem is therefore conceptual, in that it makes a case that, under precise conditions on the initial data, the Boltzmann equation can be derived from the Hamiltonian equations of motion, although just in rather idealized circumstances.

Another issue, that is emphasized by Lanford himself and nearly all subsequent commentators, concerns the limited time of validity of the theorem. In fact, the time bound τ, for which the sought-after convergence of solutions of the BBGKY hierarchy to solutions of the Boltzmann hierarchy is assured, proves extremely short. An explicit estimate given by Spohn [23], p. 62 shows that $\tau \leq 0.2\sqrt{(\beta/3)}\pi a^2 z$, and hence the result holds only during one-fifth of the mean free time between collisions for the given Maxwellian. Since such a time-scale, for realistic gas systems under ordinary circumstances, will be of the order of milliseconds, the theorem will hardly be enough ammunition to provide a justification of the Boltzmann equation through macroscopic time scales, or even the time scale in which equilibration sets in.[3] Yet, this is not too short to make irreversibility unobservable: in a duration of $1/5^{th}$ of the mean collision

---

[3] It is true that Illner and Pulvirenti [14], have derived a longer validity but only under much more stringent conditions, i.e. for a gas cloud expanding into a vacuum. As a matter of fact, this repeated attention to time scale has deluded attention from more serious problems. Indeed, Lanford already pointed out that there is a simple, if merely technical, "fix" to the above problem: one would only need to require that assumption (i) of the theorem holds for arbitrary times, and not just at $t = 0$, and Lanford's result may be extended to all times (see [27] for a discussion of this issue).

time, one expects that about 20 % of the particles will have collided, and this can be sufficient for a significant increase of entropy of the gas. As Lanford put it,

> Although these results apply only to small positive times, the times involved are large enough for Boltzmann's H-function to decrease a strictly positive amount. Thus our results show unambiguously that there is no contradiction between the reversibility of molecular dynamics and irreversibility implied by the H-theorem. [15], p. 99.

But is there really irreversibility in Lanford's theorem? And, if so, where does it come from? This question seems particularly pressing since this theorem does not explicitly rely on the *Stosszahlansatz* or an analogous statement. We take up these questions in the next section.

5 Irreversibility in Lanford's Theorem

Lanford's theorem shows how one can derive the Boltzmann equation from the Hamiltonian equations of motion under precise assumptions. As a statistical version of Boltzmann's H-theorem, it seems to account for the approach to equilibrium for a general class of non-equilibrium initial conditions characterized by the regularity condition (i), at least during the time-interval $[0, \tau]$. The most important question is then how the implied irreversibility of this macro-evolution arises. On this point Lanford and other authors on his theorem made remarks that are not quite univocal. We first survey and criticize these different views and then present our own argument on the emergence of irreversibility.

5.1 Views on the Emergence of Irreversibility in the Literature

In the final pages of his first paper, Lanford offers a diagnosis for the emergence of irreversibility. There, he stresses that the factorization condition cannot be the time-asymmetric ingredient needed to derive the Boltzmann equation.

> The Boltzmann hierarchy, like the Boltzmann equation is not invariant under time reversal. That is, irreversibility appears in passing to the limit $a \longrightarrow 0$, not in the assumption that the rescaled correlation functions factorize. [15], p. 110

Indeed, Lanford's result does not require a factorization condition to get convergence towards a solution of the Boltzmann hierarchy, but only to guarantee that the latter becomes equivalent to the Boltzmann equation. In fact, the factorization condition, at least in the version adopted by Lanford, (i.e. (47), as distinguished from the Stoßzahlansatz) is itself time-reversal invariant. Therefore, it surely does not yield an explanation for irreversibility. This point is fortified by the fact that, as we saw in section 4.1, factorization is also invoked in the version of theorem holding for negative times to derive the anti-Boltzmann equation. Instead, since irreversible behaviour already appears at the level of the Boltzmann hierarchy (or the anti-Boltzmann hierarchy), Lanford puts the blame on the procedure to take the limit from the BBGKY

hierarchy to the Boltzmann hierarchy. (In the above quote, the notation $a \longrightarrow 0$ ought to be understood as equivalent to $N \longrightarrow \infty$, as the Boltzmann–Grad limit prescribes.) We obviously agree here with [15] that any irreversibility is not due to the factorization assumption, but should we thereby conclude that it is due to the Bolzmann-Grad limiting procedure?

In a subsequent paper, Lanford [17] himself argues that this limiting procedure is not sufficient for the appearance of irreversibility. He illustrated this by the Vlasov equation. This equation describes a system in which the interaction between particles is given by a sum of two-body potentials of the form

$$\phi^{(N)}(\vec{q}_1 - \vec{q}_2) = \frac{1}{N}\phi_0(\vec{q}_1 - \vec{q}_2), \tag{50}$$

where $\phi_0(\vec{q}_1 - \vec{q}_2)$ is a fixed continuous potential, rather than the hard-spheres potential (12). The macroscopic distribution function $f(\vec{q}, \vec{p})$ is in this case determined from the microstate of the system in a similar way as in the Boltzmann–Grad limit. However, the time-evolution of f as given by the Vlasov equation, is time-reversal invariant. Indeed, one can show that the H-function remains constant through time for this equation, and hence taking the limit $N \longrightarrow \infty$ here does not lead to irreversibility.[4] The lesson Lanford draws from this is the following:

> None of this, however, really implies that irreversible behavior *must* occur in the limiting regime; it merely makes this behavior plausible. For a really compelling argument in favor of irreversibility, it seems to be necessary to rely on some version of Boltzmann's original proof of the H-theorem [17], p. 75.

Unfortunately, Lanford did not specify how appealing to (some version of) Boltzmann's derivation of the H-theorem would provide a compelling argument in favor of irreversibility. But it seems reasonable to suppose that he meant to go back to the Stoßzahlansatz, and its distinction between pre- and post-collision configurations.

An explanation of what Lanford intended in the last quote may perhaps be traced back to the analysis he develops in his 1975 work. When presenting his own re-examination of the derivation of the Boltzmann equation from the BBGKY hierarchy, i.e. (32) specialized to the case $k = 1$, Lanford comments:

> We obtained [the BBGKY hierarchy for the hard spheres model with the collision term expressed by Eq. (32) in the present paper] by systematically writing collision phase points in their incoming representations. We could have equally well have written them in their outgoing representations; if we then assumed factorization we would have obtained the Boltzmann collision term with its sign reversed. It is thus essential, in order to get the Boltzmann equation, to assume
>
> $$\rho_2(x_1, x_2) = \rho_1(x_1)\rho_1(x_2), \tag{51}$$
>
> for *incoming* collision points (x_1, x_2) and not for outgoing ones. [15], p. 88

---

[4] This can actually be seen as a counterexample to [20], who claimed that the irreversible approach to equilibrium would follow from taking the limit for $N \longrightarrow \infty$.

Here, Lanford indicates the distinction between adopting the incoming representation for collision phase points as opposed to the outgoing representation as being responsible for the sign of the collision term in the Boltzmann hierarchy. In fact, one can derive a form of the BBGKY hierarchy in which the collision term is rewritten by systematically replacing the configurations with incoming momenta by the outgoing momenta, instead of *vice versa* as we did in (32). Let us make this point explicit. In analogy to what we did in Sect. 3.2, going back to Eqn. (29), we again split the integral over S^2 into two hemispheres, but on the hemisphere with $\vec{\omega}_{i,k+1} \cdot (\vec{p}_{k+1} - \vec{p}_i) \geq 0$ we would leave the configuration for the pair $(\vec{q}_i, \vec{p}_i, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ as it is (i.e. outgoing). On the hemisphere characterized by $\vec{\omega}_{i,k+1} \cdot (\vec{p}_{k+1} - \vec{p}_i) \leq 0$ we replace the coordinates $(\vec{q}_i, \vec{p}_i, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ by $(\vec{q}_i, \vec{p}_i{}', \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}{}')$ as well as change the sign of the integration variable $\vec{\omega}_{i,k+1}$ so that the two hemisphere integrals have a common domain. The result is that Eq. (29) now goes over in

$$
\mathcal{C}^{(a)}_{k,k+1}\rho^{(a)}_{k+1,t}(x_1, \ldots, x_k) = Na^2 \sum_{i=1}^{k} \int_{\mathbb{R}^3} d\vec{p}_{k+1} \int_{\vec{\omega}_{i,k+1}\cdot(\vec{p}_i - \vec{p}_{k+1})\geq 0} d\vec{\omega}_{i,k+1}
$$
$$
(\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}))[\rho^{(a)}_{k+1}(x_1, \ldots, \vec{q}_i, \vec{p}_i, \ldots x_k, \vec{q}_i - a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})
$$
$$
-\rho^{(a)}_{k+1}(x_1, \ldots, \vec{q}_i, \vec{p}_i{}', \ldots, \vec{q}_i + a\vec{\omega}_{i,k+1}, \ldots \vec{p}_{k+1}{}')] \tag{52}
$$

as an alternative to (32). Notice that the primed variables now appear in the second term in the integrand, rather than the first. It is thus natural to expect that, if one takes the Boltzmann–Grad limit $a \longrightarrow 0$, this collision term for the BBGKY hierarchy converges to the collision term $-\mathcal{C}_{k,k+1}$ in the anti-Boltzmann hierarchy, rather than the collision term $\mathcal{C}_{k,k+1}$ in the Boltzmann hierarchy. So, according to the above quote from [15], it is the preference for the incoming representation for collision points over the outgoing representation that would yield the time-asymmetric ingredient needed to obtain the sought-after result. This indeed seems analogous to Boltzmann's original derivation of the H-theorem, where the irreversibility of the Boltzmann equation was introduced by the Pre-collision version of the Stoßzahlansatz.

What is puzzling about this view, though, is that in an earlier passage [15], p. 86 had argued for the *identification* of phase points which differ only by having an incoming collision configuration replaced by the corresponding outgoing collision configuration. This suggests that the origin of irreversibility in the Boltzmann equation would now lie in a conventional choice of representation of the *same* phase point. An obvious objection to this view is that it is not clear at all how physical irreversibility can be due to a conventional choice of representation. Since the question of whether we derive $dH/dt \leq 0$ or $dH/dt \geq 0$ is a substantive issue, such a difference cannot be a matter of mere convention.

Cercignani et al. [9] and Cercignani [8] have argued that there actually is a non-conventional justification for adopting the incoming representation, and that the choice for this representation indeed has a dynamical underpinning:

> We are compelled to ask whether the representation in terms of ingoing configurations is the right one, i.e. physically meaningful. As we shall later see, in a more careful analysis of the validity problem, the representation in terms of

ingoing configurations follows automatically from hard-spheres dynamics and is, indeed, not a matter of an a priori choice [9], p. 74.

So, for Cercignani et al., the preference for the incoming configuration over the ongoing one is an automatic consequence of the hard-spheres dynamics. This view might actually seem to settle the issue in a non-conventional way. However, in Sect. 5.2 below, we will argue in detail against the claim that irreversibility is due to the adoption of one representation rather than another.

An entirely different argument for the emergence of irreversibility is presented by Spohn [22], [23] and Lebowitz [18]. A similar argument is also sketched by Lanford [16]. Spohn [22], p. 596 devotes a paragraph to the question "how Lanford's theorem escapes the conflict between the reversible character and the irreversible character of the Boltzmann hierarchy see also [23], p. 66. He points out how Lanford's theorem will not sustain the construction of a counterexample as in the original reversibility objection to the H-theorem. Recall that in that construction we assumed an initial distribution function $f_0(x)$ that evolves in accordance with the Boltzmann equation from the initial time 0 to some positive time t, when the distribution function is $f_t(x)$, and then suddenly reverse all the velocities of all particles. Due to the time-reversal invariance of the microdynamics, $H[f]$ would have to increase during the interval $[t, 2t]$. Spohn discusses what happens if we try to run this same argument on the basis of Lanford's theorem (a more elaborate version of his reasoning is given by Lebowitz [18]). The crucial point in his analysis is that the set $\Gamma_{k,=}(s + t)$ of phase points for which the convergence need not hold increases with time. Hence, if we consider the rescaled densities $\rho_{k,t}^{(a)}$ at time t, with $0 < t < \tau$, and then reverse the velocities, the ensuing evolution of these functions will no longer be guaranteed by Lanford's result: in fact, for the theorem to be applicable at the new initial time t after the velocity-reversal, the convergence of solutions of the BBGKY hierarchy to solutions of the Boltzmann hierarchy would have be assured again over the domain of convergence $\Gamma_{k,\neq}(s)$ as demanded by the convergence condition (ii). Yet, from the result of Lanford's theorem from the evolution during $[0, t]$, we would only have convergence on $\Gamma_{k,\neq}(s + t)$. Given that $\Gamma_{k,\neq}(s) \supset \Gamma_{k,\neq}(s + t)$, we cannot apply Lanford's theorem (after velocity-reversal) to obtain an evolution during the time $[t, 2t]$ convergent to a solution of the Boltzmann equation with increasing H on the same domain of rescaled density functions as we started out with. An argument along the lines of the reversibility objection against Lanford's theorem is thereby blocked.

This suggests a very different view on the emergence of irreversibility than the previous ones. The time-asymmetric ingredient in the theorem is now identified in its assumptions, specifically in the convergence condition (ii). Indeed, the prescribed domain of convergence is not invariant under time reversal in the sense that $\Gamma_{k,\neq}(s) \neq \Gamma_{k,\neq}(-s)$. Therefore, according to such an interpretation, the source of irreversibility would lie in time-asymmetric assumptions of Lanford's theorem.

To summarize the different views on the emergence of irreversibility in Lanford's theorem available in the literature: Lanford first identified the Boltzmann–Grad limit as the source of irreversibility, but later he mitigated this claim; a next argument of Lanford is that irreversibility arises by adopting the incoming representation for collision points, instead of the outgoing representation. We have argued against this

view that it would make the appearance of irreversibility seem to be a matter of mere conventional choice of representation; Cercignani et al. claimed that a privileged role of the incoming representation follows from the hard-spheres dynamics; Spohn and Lebowitz instead pointed out a time-asymmetric ingredient in the domain of assumption (ii) of the theorem. Below, we argue that all these views fail to provide a satisfactory account of the status of irreversibility in Lanford's result. Moreover, we will argue that, in our opinion, there is no genuine irreversibility in the theorem.

5.2 Is there Really Irreversibility Embodied in Lanford's Theorem?

Let us begin by discussing the notion of the incoming representation, as opposed to the outgoing representation, of a collision point. As [15], p. 87 puts it "[T]hese two are really just different representations of the same phase point." Even [18], p. 8 argues similarly when he writes about the incoming and outgoing momenta as being "just two different representations of the same phase point." However, this very identification of the two configurations as representations of the same phase point seems to make the distinction between incoming and outgoing collisions, which is allegedly at the heart of the issue of irreversibility—at least according to some of the quotes we have just discussed—quite hard to maintain.

Recall that the appeal to a topology identifying the pre-collision and the post-collision coordinates was introduced in order to assure the technical point that the hard-sphere dynamics becomes smooth. This point may perhaps be elucidated by considering what happens if one takes a series of smooth spherically symmetrical pair potentials ϕ in (11) that approaches the hard-spheres model (12). In such a case, whenever a collision between two particles, say i and j, occurs, the momenta of the particles do not change instantaneously from incoming values $(\vec{p}_i, \vec{p}_j)$ to outgoing values $(\vec{p}_i{}', \vec{p}_j{}')$ as given by 6, but by some smooth trajectory in a non-zero time span (cf. Fig. 2). When we take the hard-sphere limit for such a collision (i.e., if we let the pair potential ϕ approach the hard-sphere potential (12)), this interval goes to zero, and the trajectory would jump instantaneously from x_{in} to x_{out}. Now, we can then still regard this hard-sphere collision as a continuous process by adopting a topology on $\Gamma_{\neq}^{(a)}$ in which the holes in this phase space are, so to say, stitched up, so that incoming collision coordinates and the outgoing collision coordinates become, as it were, adjacent to each other in phase space, and a trajectory that jumps from $x_{\text{in}} = (x_1, \ldots, x_{i-1}, \vec{q}_i, \vec{p}_i, \ldots, \vec{q}_j, \vec{p}_j, x_{j+1}, \ldots x_n)$ to $x_{\text{out}} = (x_1, \ldots, x_{i-1}, \vec{q}_i, \vec{p}_i{}', \ldots, \vec{q}_j, \vec{p}_j{}', x_{j+1}, \ldots x_n)$ is regarded as continuous (Fig. 3). One may express this procedure colloquially as an "identification" of these two points.

Yet, taking this way of speaking too literally may expose one to the risk of a misleading interpretation of the theorem. Indeed, we are free in adopting any topology we like on the boundary of $\Gamma_{\neq}^{(a)}$ (as long as it extends the Euclidean topology on its interior), and in particular we can choose a topology to make an instantaneous transition from x_{in} to x_{out} appear as a smooth trajectory. Such a choice of topology entails that every metric, or distance function, d on $\Gamma_{N\neq}^{(a)}$, compatible with it would have the property that $d(x_{\text{in}}, x_{\text{out}}) = 0$, and hence (by the usual definition of a metric) it would follow that $x_{\text{in}} = x_{\text{out}}$, i.e. those points are identified. But, when choosing a topology,

we are not forced to introduce a metric. Moreover, the topological identification of phase-space coordinates ought not to be granted physical significance. Even if we identify the incoming and outgoing points x_{in} and x_{out} for the purpose of topological or metrical considerations, it does not follow that they thereby are *physically* identical. Indeed, that would overlook the distinctive and physically relevant fact that the momenta are quite different in these two points. Furthermore, if the diameter a of the particles is not null, as it happens in real physical situations described by the hard spheres model, their positions are also necessarily distinct, even if they are colliding. In other words, all that this choice of topology enforces is that a trajectory connecting points like x_{in} and x_{out} becomes smooth, but not that these points are physically one and the same. In fact, the very use of the terminology "representation" appears quite inappropriate in this context. We shall therefore use the term "configuration" instead.

Moreover, Spohn [25] has argued that one can also obtain the smoothness of the dynamics for solutions of the BBGKY hierarchy for hard spheres without assuming the identification of the phase points representing the pre-collision and post-collision configurations. In its stead, Spohn adopts a weaker condition which we refer to as *continuity across collisions*: Denote the phase point in which two particles, say i and $k+1$, are touching each other, i.e., $(x_1, \ldots, x_{i-1}, \vec{q}_i, \vec{p}_i, x_{i+1}, \ldots x_k, \vec{q}_i + a\vec{\omega}, \vec{p}_{k+1})$

Fig. 2 A region of the phase space $\Gamma_{N,\neq}^a$ showing a "hole" (gray area) due to the forbidden overlap of hard spheres i and j. The points x_{in} and x_{out} represent the microstate immediately before and after the collision between particles i and j. The dashed curve between them denotes the continuous trajectory connecting these points in a smooth potential approximation to the hard spheres potential
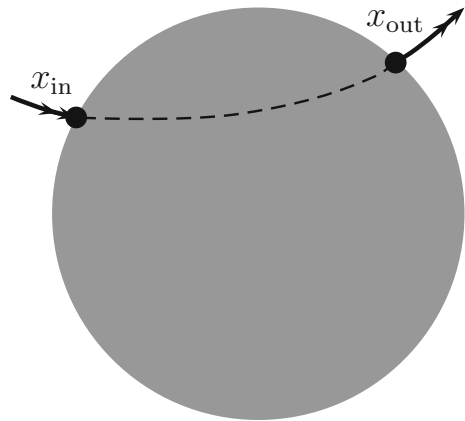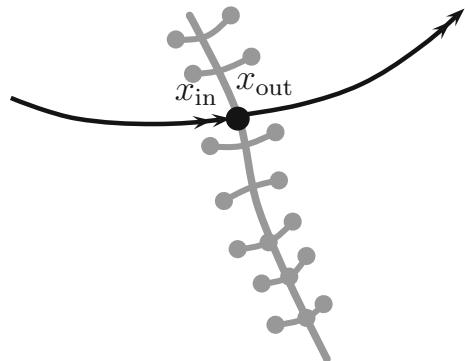

Fig. 3 The same region of the phase space $\Gamma_{N,\neq}^a$ with the hole sown up, and the points x_{pre} and x_{out} identified. The phase space trajectory is now smooth even during the hard spheres collision

by the abbreviation $(\vec{q}_i, \vec{p}_i; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$. Continuity across collisions then requires that, if $\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \geq 0$, then

$$\rho_{k+1}^{(a)}(\vec{q}_i, \vec{p}_i; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}) = \rho_{k+1}^{(a)}(\vec{q}_i, \vec{p}_i{}'; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}'), \quad (53)$$

for all points $(\vec{q}_i, \vec{p}_i; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ in $\Gamma_{N,\neq}^{(a)}$. In other words, when any pair of particles labeled by i and $k + 1$, respectively, undergo a mutual collision, the values of the rescaled reduced density functions at the pre-collision point $(\vec{q}_i, \vec{p}_i; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ should be equal to the values of those same functions at the post-collision point $(\vec{q}_i, \vec{p}_i{}'; \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}')$. One can easily verify that this continuity condition allows one to obtain the collision term in the form (32) of the BBGKY hierarchy in essentially the same way as it was shown in Sect. 3.2. (and likewise one obtains expression for the collision term (52) in section 4.1), and thus it is sufficient for the purpose of deriving Lanford's theorem. Hence, one does not really need to appeal to a literal identification of pre-collision and post-collision phase points; it is sufficient that the functions $\rho_{k+1}^{(a)}$ take the same value at each such pairs of points. Note that even if one would adopt the view that the pre- and post-collision points ought to be identified, one would still need to rely on condition (53) to be fulfilled to make $\rho_{k+1}^{(a)}$ continuous and obtain a smooth dynamics for the BBGKY hierarchy.

The claim we wish to argue against here is that the choice of the incoming configuration over the outgoing configuration could be the source of irreversibility in Lanford's theorem. The basis for our argument rests on two main points. On the one hand, as we have just seen, these two configurations are in fact different from each other in a physically relevant sense. On the other hand, to counter the claim of Cercigiani *et al.*, we argue that the hard-spheres dynamics will not provide a preference for writing the collision term of the BBGKY hierarchy in the incoming configuration rather than the outgoing configuration. Indeed, *Proposition* 3 in the Appendix shows that, if (53) holds, then the BBGKY hierarchy with the collision term expressed by (32) and the BBGKY hierarchy with the collision term expressed by (52) are perfectly equivalent. Thus, the choice of either one of the two collision configurations does not make any difference at the level of the BBGKY hierarchy. In particular, one can derive the Boltzmann hierarchy, as well as the anti-Boltzmann hierarchy, from the BBGKY hierarchy rewritten in terms of either the incoming or the outgoing configurations without having to choose the "right" one.

Furthermore, *Proposition* 3 proves that continuity across collisions is sufficient to guarantee the time-reversal invariance of the BBGKY hierarchy with the collision term expressed by (32), or equivalently by (52). In other words, this choice is neutral with respect to the direction of time. This indicates that the source of irreversibility in Lanford's theorem does not lie in the adoption of either one between the incoming and the outgoing configurations of collision points.

Let us emphasize that continuity across collisions is a peculiar condition on the BBGKY hierarchy for hard spheres, which does not carry over when we take the Boltzmann–Grad limit, at least if one wishes to derive a genuinely irreversible behaviour. In fact, such a condition would have no analogue for the Boltzmann hierarchy, or the anti-Boltzmann hierarchy. For, let us restrict to the simple case of two par-

ticles, i.e. $k = 1$, in order to not overburden the notation. Suppose that the solution of the Boltzmann hierarchy $\rho_2(\vec{q}_1, \vec{p}_1; \vec{q}_1, \vec{p}_2)$ to which $\rho_2^{(a)}$ converges at the pre-collision point $(\vec{q}_1, \vec{p}_1; \vec{q}_1 + a\vec{\omega}_{1,2}, \vec{p}_2)$ is equal to the solution of the Boltzmann hierarchy $\rho_2(\vec{q}_1, \vec{p}_1'; \vec{q}_1, \vec{p}_2')$ to which $\rho_2^{(a)}$ converges at the post-collision point $(\vec{q}_1, \vec{p}_1'; \vec{q}_1 + a\vec{\omega}_{1,2}, \vec{p}_2')$, as an analogue of (53) for the Boltzmann hierarchy would presumably require. Then, since the factorization condition holds for solutions of the Boltzmann hierarchy at all phase points, one would obtain

$$\rho_1(\vec{q}_1, \vec{p}_1)\rho_1(\vec{q}_1, \vec{p}_2) = \rho_1(\vec{q}_1, \vec{p}_1')\rho_1(\vec{q}_1, \vec{p}_2'). \tag{54}$$

However, that can only be true if the gas is already at equilibrium.[5] An obvious lesson one should draw from this argument is that, for an irreversible approach to equilibrium to obtain, the continuity at collisions condition, which maintains time-reversal invariance of the BBGKY hierarchy written in any collision configuration, has no analogue in the Boltzmann–Grad limit.

We now turn to the view offered by Spohn and Lebowitz. In our opinion, they convincingly showed that the reversibility objection cannot be run against Lanford's theorem in the same way as is it was used by Loschmidt and Culverwell against Boltzmann's original presentation of the H-theorem. However, we believe that it is one thing to show how the reversibility objection is evaded, but it is quite another thing to explain the emergence of irreversibility in Lanford's theorem. And although the Spohn-Lebowitz argument is successful in the first objective, we feel it does little to offer the sought-after explanation. In particular, the suggestion that the source of irreversibility is to be traced back to time-asymmetric initial conditions employed in the theorem seems unconvincing. First of all, Lanford's theorem holds also if one sets $s = 0$ in assumption (ii). That is actually how [15] first formulated his theorem. The domain of convergence $\Gamma_{k,\neq}(0)$ in condition (ii) then corresponds to the largest set of initial configurations one can admit. The key point is to recognize that this domain is clearly invariant under velocity-reversal, and thus condition (ii) is now time-reversal invariant too.[6] Yet, the theorem still implies the irreversible Boltzmann hierarchy for positive times, as well as the irreversible anti-Boltzmann hierarchy for negative times. The problem concerning the emergence of irreversibility thus still presents itself, even when the time-reversal non-invariance of assumption (ii) is avoided by choosing $s = 0$.

Secondly, even if we choose $s > 0$ in conditions (ii) and (ii'), recall that the set of exceptional states $\Gamma_{k,=}(s + t) \cap \Gamma_{k,\neq}(s)$, comprising those microstates in the initial domain of convergence $\Gamma_{k,\neq}(s)$ for which the solution $\rho_{k,t}^{(a)}$ of the BBGKY hierarchy for hard spheres may not converge to a solution $f_{k,t}$ of the Boltzmann hierarchy, has Lebesgue measure zero for all times t. Now, in the spirit in which Lanford's theorem has been formulated, Lebesgue mejasure zero sets in phase-space are not held to be physically significant, and we have already neglected several such measure-zero sets from the outset. Thus, it would seem that the sets of exceptional states, even if they are

---

[5] This argument is really an adaptation to the present context of a discussion of the notion of pre-collisional chaos contained in [28], p.35.

[6] In fact, in this case, that is when $s = 0$, the condition coincides with its time-reversal.

not invariant under time-reversal, ought to be neglected as physically irrelevant too. However, in order to obtain an emergence of irreversibility, one would like to see that the overwhelming majority of initial phase space points will evolve in the course of time in such a way to obtain the Boltzmann hierarchy, rather than the anti-Boltzmann hierarchy. Considerations of the time-reversal non-invariance of measure zero sets will not be helpful in this regard.

So far, we have criticized the available views on the emergence of irreversibility in Lanford's theorem. None of them, in our opinion, really succeeds in identifying an ingredient responsible for the irreversible behaviour of the Boltzmann equation. The claim we wish to make now is that there is no such a time-asymmetric ingredient at all.

In order to substantiate this claim, let us stress that Lanford's theorem can be proven also for negative times. In fact, while the statement of the theorem in Sect. (4) is formulated for positive times, one can derive an analogous result for negative times $-\tau < t < 0$, as shown by Lanford [15], p. 109–110, and more explicitly by Lebowitz [18], p. 9–10. For this purpose, one ought to take the time-reversal of the assumptions of the theorem and verify that one obtains the time-reversal of the conclusion. The regularity assumption (i) is time-independent, and as such it is time-reversal invariant. Therefore, one leaves it in the same form. Assumption (ii), instead, makes an explicit reference to time in the domain of convergence $\Gamma_{k,\neq}(s)$. Hence, if one keeps $s \geq 0$, one can rewrite it as

(ii') There exist continuous functions $f_{k,0}$ on Γ_k, for $k = 1, 2 \ldots$ such that

$$\lim_{a \longrightarrow 0} \operatorname*{ess\,sup}_{(x_1, \ldots x_k) \in K} |\rho_{k,0}^{(a)}(x_1, \ldots, x_k) - f_{k,0}(x_1, \ldots x_k)| = 0, \tag{55}$$

for all compact subsets $K \subset \Gamma_{k,\neq}(-s)$ for some $s \geq 0$.

One can then prove[7] that there exists a strictly positive time τ, such that equation (45) holds for any $k = 1, 2, \ldots$ and compact subset $K \subset \Gamma_{k,\neq}(-s-t)$ during the time interval $t \in [\tau, 0]$, where now the solutions $\rho_{k,t}^{(a)}$ of the BBGKY hierarchy with initial conditions $\rho_{k,0}^{(a)}$ are taken to converge to solutions $f_{k,t}$ of the anti-Boltzmann hierarchy with initial conditions $f_{k,0}$. Since the domain of convergence $\Gamma_{k,\neq}(-s-t)$ is the time-reversal of $\Gamma_{k,\neq}(s+t)$ and the anti-Boltzmann hierarchy is the time-reversal of the Boltzmann hierarchy, this conclusion is just the time-reversal of the conclusion of the theorem as stated for positive times. Further, if one additionally assumes the factorization condition (46), one would derive a solution f_t of the anti-Boltzmann equation.

As a consequence, if we consider a time t such that $-\tau \leq t \leq \tau$, Lanford's theorem is clearly neutral with respect to time reversal: that is, for positive times we obtain convergence to solutions of the Boltzmann equation, and hence a decrease of H, just as the H-theorem requires, but for negative times we obtain convergence towards a

---

[7] The proof of this result proceeds in the same way as that for the theorem for positive times, except that here one explicitly appeals to the outgoing configuration instead of the incoming one. However, in light of *Proposition* 3 establishing the equivalence of the BBGKY hierarchy written in terms of each configuration, the latter step is not an independent ingredient.

solution of the anti-Boltzmann equation, and hence an increase of H. This is analogous to Boltzmann's 1897 argument based on the H-curve. Indeed, in this understanding, the theorem proves that for most initial microstates the H-function lies at a local peak of the H-function. So, at the initial time instant $t = 0$, $H[f_t]$ is expected to decrease in both directions of time. This offers a mathematical formalization of Boltzmann's claim that, apart from equilibrium, the most probable case is that the H-function is at a maximum of the curve. As such, Lanford's result provides a rigorous version of the statistical H-theorem sketched by Boltzmann. More importantly, with respect to the issue of emergence of irreversibility, the thus-established time-reversal invariance of the theorem does supply evidence that there is no time-asymmetric ingredient in the theorem.

On the other hand, though, one should notice that the behavior of the rescaled probability densities $\rho_{k,t}^{(a)}$ implied by the theorem for negative times has a serious drawback. In fact, this behavior conflicts with the expectation from thermodynamics that entropy of an isolated gas system should increase rather than decrease even during the interval $[-\tau, 0]$. This issue has already been discussed many times in the literature cf. [21], Feynman *et al.* 1964, [24], [11] and [27].

6 Conclusion

We discussed the problem of the emergence of irreversibility in Lanford's theorem. We argued that all the different views on the issue presented in the literature miss the target, in that they fail to identify a time-asymmetric ingredient that, added to the Hamiltonian equations of motion, would obtain the Boltzmann equation. More to the point, we argued that there is no such an ingredient at all, as one can infer from the fact that the theorem is indeed time-reversal invariant.

7 Appendix

In this appendix we consider in more detail the issue of time-reversal invariance for the Boltzmann equation and two versions of the BBGKY hierarchy equations and prove the claims concerning this issue in the body of the paper. The commonly accepted criterium for judging time-reversal invariance of such evolution equations, describing a density function (either f or ρ_k) over particle configurations is as follows: the equation determines a class S of allowed solutions, where each solution can be seen as a 'history' of the density function, either $\mathcal{H} := \{f_t, \ t \in \mathbb{R}\} \in S$ or $\mathcal{H} := \{\rho_{k,t}, \ t \in \mathbb{R}\}$, carving out, so to say, a trajectory in their respective spaces of all density functions that solve the evolution equation.

Now consider the time-reversal of such a history, defined as $\mathcal{TH} := \{\tilde{f}_t \ t \in \mathbb{R}\}$ or $\mathcal{TH} := \{\tilde{\rho}_{k,t} \ t \in \mathbb{R}\}$, where $\tilde{f}_t(\vec{q}, \vec{p}) = f_{-t}(\vec{q}, -\vec{p})$ and $\tilde{\rho}_{k,t}(\vec{q}_1, \vec{p}_1; \ldots; \vec{q}_k, \vec{p}_k) = \rho_{k,-t}(\vec{q}_1, -\vec{p}_1; \ldots; \vec{q}_k, -\vec{p}_k)$ are obtained from f_t and $\rho_{k,t}$ by reversing t and the momenta in their arguments. The question is then whether such a time-reversed history

$\mathcal{T}\mathcal{H}$ is a solution of the equation too, whenever H is a solution of the equation in question. In other words: Is $\mathcal{T}\mathcal{H} \in S$ whenever $T \in S$? If the answer is yes, the equation is time-reversal invariant, otherwise not.

Our strategy will be the same in all three cases. We consider an arbitrary solution of the equation and construct from this a time-reversed history, and derive the equation it obeys. If the resulting equation is equivalent to the original equation we have proved that the original equation is time-reversal invariant. But if it is not, the original equation is not time-reversal invariant. To be sure, our first two propositions below are not surprising: every author in the field recognizes that the Boltzmann equation is not time-reversal invariant, whereas the BBGKY hierarchy is, even if detailed proofs of these assertions are hard to find. We have added these proofs mainly to show that that by the same standard of rigour, one can obtain Proposition 3, which we believe to be of more interest.

Proposition 1 *The Boltzmann equation (8) is not time-reversal invariant.*

Proof Since nothing interesting happens to the position variables in the Boltzmann equation (8) we will suppress them in the notation below, and also put the mass $m = 1$. Note that $\vec{p}_1$ is the only independent momentum variable in the equation: $\vec{p}_2$ appears only in the right-hand side as a mere integration variable and the outgoing momenta variables $\vec{p}_1{}', \vec{p}_2{}'$ in the collision integral are functions of $\vec{p}_i$: that is, $\vec{p}_i{}' = \vec{p}_i{}'(\vec{p}_1, \vec{p}_2) = T_{\omega_{12}}(\vec{p}_1, \vec{p}_2)$, where $T_{\omega_{12}}$ is defined by (7). So, let $f_t(\vec{q}, \vec{p})$, $t \in \mathbb{R}$, by an arbitrary solution of the Boltzmann equation (8), and consider the joint transformation of $t \longrightarrow -t$ and $\vec{p}_1 \longrightarrow -\vec{p}_1$. Applying this transformation to the left-hand side of (8) leads to:

$$- \frac{\partial}{\partial t} f_{-t}(-\vec{p}_1) - \vec{p}_1 \cdot \frac{\partial}{\partial \vec{q}} f_{-t}(-\vec{p}_1) = \frac{\partial}{\partial t} \tilde{f}_t(\vec{p}_1) + \vec{p}_1 \cdot \frac{\partial}{\partial \vec{q}} \tilde{f}_t(\vec{p}_1), \qquad (56)$$

while the right-hand side of (8) becomes

$$N a^2 \int d\vec{p}_2 \int_{\vec{\omega}_{12} \cdot (\vec{p}_1 + \vec{p}_2) \leqslant 0} d\vec{\omega}_{12} \, (-\vec{p}_1 - \vec{p}_2) \cdot \vec{\omega}_{12} \\ \left[f_{-t}(\vec{p}_1{}') f_{-t}(\vec{p}_2{}') - f_{-t}(-\vec{p}_1) f_{-t}(\vec{p}_2) \right] \qquad (57)$$

Here, the notation $\vec{p}_i{}''$ is used to indicate that these momenta have to be thought of as functions of $(-\vec{p}_1, \vec{p}_2)$: $(\vec{p}_1{}'', \vec{p}_2{}'') = T_{\omega_{12}}(-\vec{p}_1, \vec{p}_2)$.

If we now perform an additional (cosmetic) transformation of the integration variables $\vec{p}_2 \longrightarrow -\vec{p}_2$ and $\vec{\omega}_{12} \longrightarrow -\vec{\omega}_{12}$, (57) can also be written as

$$N a^2 \int d\vec{p}_2 \int_{\vec{\omega}_{12} \cdot (\vec{p}_1 - \vec{p}_2) \geqslant 0} d\vec{\omega}_{12} \, (\vec{p}_1 - \vec{p}_2) \cdot \vec{\omega}_{12} \\ \left[f_t(\vec{p}_1{}''') f_t(\vec{p}_2{}''') - f_t(-\vec{p}_1) f_t(-\vec{p}_2) \right], \qquad (58)$$

where $(\vec{p}_1{}''', \vec{p}_2{}''') := T_{\omega_{12}}(-\vec{p}_1, -\vec{p}_2)$. But $T_{\omega_{12}}$ is a linear operator, and therefore $\vec{p}_i{}''' = -\vec{p}_i{}'$. If we now substitute $f_{-t}(-\vec{p}) = \tilde{f}_t(\vec{p})$ in (58) and equate the transformed

left-hand side (55) of (8) to the transformed right-hand side (58) of (8), we find that $\tilde{f}_t$ satisfies the equation

$$
\frac{\partial}{\partial t} \tilde{f}_t(\vec{p}_1) + \vec{p}_1 \cdot \frac{\partial}{\partial \vec{q}} \tilde{f}_t(\vec{p}_1) =
$$
$$
-Na^2 \int d\vec{p}_2 \int_{\vec{\omega}_{12} \cdot (\vec{p}_1 - \vec{p}_2) \geqslant 0} d\vec{\omega}_{12} (\vec{p}_1 - \vec{p}_2) \cdot \vec{\omega}_{12} \left[\tilde{f}_t(\vec{p}_1') \tilde{f}_t(\vec{p}_2') - \tilde{f}_t(\vec{p}_1) \tilde{f}_t(\vec{p}_2) \right],
$$

$$(59)$$

also known as the anti-Boltzmann equation. We conclude: whenever the solution $\{f_t , t \in \mathbb{R}\}$ satisfies the Boltzmann equation, the time reversed solution $\{\bar{f}_{-t} , t \in \mathbb{R}\}$ solves the (inequivalent) anti-Boltzmann equation, and therefore, the Boltzmann equation is not time-reversal invariant.   □

Proposition 2 *The BBGKY hierarchy with the collision term given by (29) is time-reversal invariant.*

Proof Recall that the BBGKY hierarchy has the form, for $k = 1, \ldots N$:

$$
\frac{\partial \rho_{k,t}^{(a)}(x_1, \ldots, x_k)}{\partial t} - \mathcal{H}_k \rho_{k,t}^{(a)}(x_1, \ldots, x_k) = \left(\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,t}^{(a)} \right) (x_1, \ldots, x_k).   \quad (60)
$$

In this equation, we deal with k particles (the momentum of the $k + 1$th particle appears in (29) only as an integration variable). If we reverse sign of the momenta $\vec{p}_1, \ldots \vec{p}_k$ and the sign of t, it is easy to see the the left hand side of (66) changes sign. But here, the right-hand side (29) clearly changes sign too when we change sign of all momenta $\vec{p}_1, \ldots, \vec{p}_{k+1}$, due to the fact that the integration over the antisymmetric factor $\left(\vec{\omega}_{i,k+1} \cdot (\vec{p}_{k+1} - \vec{p}_i) \right)$ in the integrand is extended over the entire surface of theunit sphere. More explicitly, if we use the notation $\bar{x}_i = (\vec{q}_i, -\vec{p}_i)$ along $x_i = (\vec{q}_i, \vec{p}_i)$, the transformed version of the left-hand side of Eq. (60) is

$$
-\frac{\partial}{\partial t} \rho_{k,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_k) + \mathcal{H}_k \rho_{k,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_k) = -\frac{\partial}{\partial t} \tilde{\rho}_{k,t}^{(a)}(x_1, \ldots, x_k)
$$
$$
+ \mathcal{H}_k \tilde{\rho}_{k,t}^{(a)}(x_1, \ldots, x_k),   \quad (61)
$$

while the right-hand side transforms into:

$$
\left(\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,-t}^{(a)} \right) (\bar{x}_1, \ldots \bar{x}_k) = Na^2 \sum_{i=1}^{k} \int_{\mathbb{R}^3} d\vec{p}_{k+1} \int_{S^2} d\vec{\omega}_{i,k+1}
$$
$$
\left(\vec{\omega}_{i,k+1} \cdot (\vec{p}_{k+1} + \vec{p}_i) \right) \rho_{k+1,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}).   \quad (62)
$$

Hence, if we rewrite the integration variable $\vec{p}_{k+1}$ as $-\vec{p}_{k+1}$, we obtain from (62):

$$
\begin{aligned}
&\left(\mathcal{C}_{k,k+1}^{(a)}\rho_{k+1,-t}^{(a)}\right)(\bar{x}_1,\ldots,\bar{x}_k) = \\
&-Na^2\sum_{i=1}^{k}\int_{\mathbb{R}^3}d\vec{p}_{k+1}\int_{S^2}d\vec{\omega}_{i,k+1}\left(\vec{\omega}_{i,k+1}\cdot(\vec{p}_{k+1}-\vec{p}_i)\right)\rho_{k+1,-t}^{(a)}(\bar{x}_1,\ldots,\bar{x}_k,\vec{q}_i \\
&+a\vec{\omega}_{i,k+1},-\vec{p}_{k+1}) = \\
&-Na^2\sum_{i=1}^{k}\int_{\mathbb{R}^3}d\vec{p}_{k+1}\int_{S^2}d\vec{\omega}_{i,k+1}\left(\vec{\omega}_{i,k+1}\cdot(\vec{p}_{k+1}-\vec{p}_i)\right)\tilde{\rho}_{k+1,-t}^{(a)}(x_1,\ldots,x_k,\vec{q}_i \\
&+a\vec{\omega}_{i,k+1},\vec{p}_{k+1}).
\end{aligned}
\tag{63}
$$

Comparing this with (29), we conclude

$$
\left(\mathcal{C}_{k,k+1}^{(a)}\rho_{k+1,-t}^{(a)}\right)(\bar{x}_1,\ldots\bar{x}_k) = -(\mathcal{C}_{k,k+1}^{(a)}\tilde{\rho}_{k+1,t}^{(a)})(x_1,\ldots,x_k).
\tag{64}
$$

Putting (61) and (64) together, we see that the time-reversed version $\tilde{\rho}_{k,t}^{(a)}$ of an arbitrary solution $\rho_{k,t}^{(a)}$ of (60) obeys the equivalent equation

$$
-\frac{\partial}{\partial t}\tilde{\rho}_{k,t}^{(a)} + \mathcal{H}_k\tilde{\rho}_{k,t}^{(a)} = -\mathcal{C}_{k,k+1}^{(a)}\tilde{\rho}_{k+1,t}^{(a)}.
\tag{65}
$$

This shows that if $\{\rho_{k+1,t}^{(a)}, t \in \mathbb{R}\}$ solves Eq. (60), then $\{\tilde{\rho}_{k+1,t}^{(a)}, t \in \mathbb{R}\}$ is a solution of the same equation, so that we can conclude that (60) is time-reversal invariant. □

Proposition 3 *Let the continuity across collisions condition (53) hold. Then, the BBGKY hierarchy with the collision term expressed by (32) is equal to the BBGKY hierarchy with the collision term expressed by (52). Furthermore, it is time-reversal invariant.*

Proof Recall that, after adopting the incoming configuration for collision points, the BBGKY hierarchy takes the form

$$
\frac{\partial\rho_{k,t}^{(a)}(x_1,\ldots,x_k)}{\partial t} - \mathcal{H}_k\rho_{k,t}^{(a)}(x_1,\ldots,x_k) = \left(\mathcal{C}_{k,k+1}^{(a)}\rho_{k+1,t}^{(a)}\right)(x_1,\ldots,x_k),
\tag{66}
$$

where the left-hand side is the same as in (60), but the collision operator is now expressed by (32). On the other hand, when adopting the outgoing configuration for collision points, the collision operator is expressed by (52).

We first show that the BBGKY hierarchy with the collision term expressed by (32) is equal to the BBGKY hierarchy with the collision term expressed by (52). Since the only apparent difference between these hierarchies lies in the collision term, one just needs to show that the operator (32) is the same as (52). Let us focus on the former. Notice that the phase point $(x_1,\ldots,x_{i-1},\vec{q}_i,\vec{p}_i{}',x_{i+1},\ldots x_k,\vec{q}_i - a\vec{\omega}_{i,k+1},\vec{p}_{k+1}')$ as

well as the phase point $(x_1, \ldots, x_{i-1}, \vec{q}_i, \vec{p}_i, x_{i+1}, \ldots x_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})$ are pre-collision coordinates in the hemisphere $\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \geqslant 0$. Therefore, if the continuity at collisions condition holds, it follows that (32) can be written as

$$
\left(\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,t}^{(a)} \right) (x_1, \ldots, x_k)
$$

$$
= Na^2 \sum_{i=1}^{k} \int_{\mathbb{R}^3} d\vec{p}_{k+1} \int_{\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \geqslant 0} d\vec{\omega}_{i,k+1} \; \vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1})
$$

$$
\times \big[\rho_{k+1,t}^{(a)}(x_1, \ldots, x_{i-1}, \vec{q}_i, \vec{p}_i, x_{i+1}, \ldots x_k, \vec{q}_i - a\vec{\omega}_{i,k+1}, \vec{p}_{k+1})
$$

$$
- \rho_{k+1,t}^{(a)}(x_1, \ldots, x_{i-1}, \vec{q}_i, \vec{p}_i{}', x_{i+1}, \ldots x_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}{}') \big], \qquad (67)
$$

which is just equal to the collision term (52).

Next, we demonstrate that the BBGKY hierarchy with the collision term expressed by (32) is time-reversal invariant. Here, we are again dealing with k particles, but in this case both incoming and outgoing momenta appear in the same formula, just as in the Boltzmann equation. And just as in the Boltzmann equation, one ought to take the outgoing momenta variables here as (implicit) functions of the incoming momenta: $(\vec{p}_i{}', \vec{p}_{k+1}{}') = T_{\omega_{i,k+1}}(\vec{p}_i, \vec{p}_{k+1})$. We now apply a combination of the arguments we used above to judge the time-reversal invariance of the Boltzmann equation and the BBGKY hierarchy in the version (60): we replace t by $-t$ and $(x_1, \ldots, x_k)$ by $(\bar{x}_1, \ldots, \bar{x}_k)$. Since the left-hand side is the same as in (60), we draw the same conclusion: this side transforms into (61). But we have to scrutinize the behaviour of the right-hand side in more detail. This side transforms to:

$$
\left(\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,-t}^{(a)} \right) (\bar{x}_1, \ldots, \bar{x}_k)
$$

$$
= d\vec{p}_{k+1} d\vec{\omega}_{i,k+1} \; \vec{\omega}_{i,k+1} \cdot (\vec{p}_i + \vec{p}_{k+1})
$$

$$
\times \big[\rho_{k+1,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_{i-1}, \vec{q}_i, \vec{p}_i{}', \bar{x}_{i+1}, \ldots \bar{x}_k, \vec{q}_i - a\vec{\omega}_{i,k+1}, \vec{p}_{k+1}{}')
$$

$$
- \rho_{k+1,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_{i-1}, \vec{q}_i, -\vec{p}_i, \bar{x}_{i+1}, \ldots \bar{x}_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, -\vec{p}_{k+1}) \big], \quad (68)
$$

where, as before, the variables $(\vec{p}_i{}', \vec{p}_{k+1}{}')$ are defined as $(\vec{p}_i{}', \vec{p}_{k+1}{}') = T_{\omega_{i,k+1}}(-\vec{p}_i, \vec{p}_{k+1})$. Repeating a similar step of our first argument, we perform a conventional transformation on the integration variables $\vec{p}_{k+1} \longrightarrow -\vec{p}_{k+1}$ and $\vec{\omega}_{i,k+1} \longrightarrow -\vec{\omega}_{i,k+1}$ and use that the primed momenta transform as $(\vec{p}_i{}', \vec{p}_{k+1}{}') \longrightarrow T_{\omega_{i,k+1}}(-\vec{p}_i, -\vec{p}_{k+1}) = (-\vec{p}_i{}', -\vec{p}_{k+1}{}')$ to rewrite the integral (68) as

$$
\left(\mathcal{C}_{k,k+1}^{(a)} \rho_{k+1,-t}^{(a)} \right) (\bar{x}_1, \ldots, \bar{x}_k)
$$

$$
= Na^2 \sum_{i=1}^{k} \int_{\mathbb{R}^3} d\vec{p}_{k+1} \int_{\vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1}) \geqslant 0} d\vec{\omega}_{i,k+1} \; \vec{\omega}_{i,k+1} \cdot (\vec{p}_i - \vec{p}_{k+1})
$$

$$
\times \big[\rho_{k+1,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_{i-1}, \vec{q}_i, -\vec{p}_i{}', \bar{x}_{i+1}, \ldots \bar{x}_k, \vec{q}_i + a\vec{\omega}_{i,k+1}, -\vec{p}_{k+1}{}')
$$

$$
- \rho_{k+1,-t}^{(a)}(\bar{x}_1, \ldots, \bar{x}_{i-1}, \vec{q}_i, -\vec{p}_i, \bar{x}_{i+1}, \ldots \bar{x}_k, \vec{q}_i - a\vec{\omega}_{i,k+1}, -\vec{p}_{k+1}) \big]. \quad (69)
$$

Now, in analogy to the previous case (cf. Eqn. (64)), we must show that

$$\left(\mathcal{C}^{(a)}_{k,k+1}\rho^{(a)}_{k+1,-t}\right)(\bar{x}_1,\ldots,\bar{x}_k) = -\left(\mathcal{C}^{(a)}_{k,k+1}\tilde{\rho}^{(a)}_{k+1,t}\right)(x_1,\ldots,x_k). \qquad (70)$$

For this purpose, notice that the phase point $(\bar{x}_1,\ldots,\bar{x}_{i-1},\vec{q}_i,-\vec{p}_i{}',\bar{x}_{i+1},\ldots\bar{x}_k,\vec{q}_i+a\vec{\omega}_{i,k+1},-\vec{p}'_{k+1})$ as well as the phase point $(\bar{x}_1,\ldots,\bar{x}_{i-1},\vec{q}_i,-\vec{p}_i,\bar{x}_{i+1},\ldots\bar{x}_k,\vec{q}_i+a\vec{\omega}_{i,k+1},-\vec{p}_{k+1})$ are pre-collision coordinates in the hemisphere $\vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})\geqslant 0$, and thus one can apply the continuity at collisions condition (53) to both such points. The collision term in (67) can thus be written as

$$\left(\mathcal{C}^{(a)}_{k,k+1}\rho^{(a)}_{k+1,-t}\right)(\bar{x}_1,\ldots,\bar{x}_k)$$

$$= Na^2\sum_{i=1}^{k}\int_{\mathbb{R}^3}d\vec{p}_{k+1}\int_{\vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})\geqslant 0}d\vec{\omega}_{i,k+1}\ \vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})$$

$$\times\left[\rho^{(a)}_{k+1,-t}(\bar{x}_1,\ldots,\bar{x}_{i-1},\vec{q}_i,-\vec{p}_i,\bar{x}_{i+1},\ldots\bar{x}_k,\vec{q}_i+a\vec{\omega}_{i,k+1},-\vec{p}_{k+1})\right.$$
$$\left.-\rho^{(a)}_{k+1,-t}(\bar{x}_1,\ldots,\bar{x}_{i-1},\vec{q}_i,-\vec{p}_i{}',\bar{x}_{i+1},\ldots\bar{x}_k,\vec{q}_i-a\vec{\omega}_{i,k+1},-\vec{p}'_{k+1})\right]. \quad (71)$$

If one then performs the (cosmetic) transformation of the integration variable $\vec{\omega}_{i,k+1}$ into $-\vec{\omega}_{i,k+1}$, one obtains

$$\left(\mathcal{C}^{(a)}_{k,k+1}\rho^{(a)}_{k+1,-t}\right)(\bar{x}_1,\ldots,\bar{x}_k)$$

$$= Na^2\sum_{i=1}^{k}\int_{\mathbb{R}^3}d\vec{p}_{k+1}\int_{\vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})\geqslant 0}d\vec{\omega}_{i,k+1}\,\vec{\omega}_{i,k+1}\cdot(\vec{p}_i-\vec{p}_{k+1})$$

$$\times\left[\rho^{(a)}_{k+1,-t}(\bar{x}_1,\ldots,\bar{x}_{i-1},\vec{q}_i,-\vec{p}_i,\bar{x}_{i+1},\ldots\bar{x}_k,\vec{q}_i-a\vec{\omega}_{i,k+1},-\vec{p}_{k+1})\right.$$
$$\left.-\rho^{(a)}_{k+1,-t}(\bar{x}_1,\ldots,\bar{x}_{i-1},\vec{q}_i,-\vec{p}_i{}',\bar{x}_{i+1},\ldots\bar{x}_k,\vec{q}_i+a\vec{\omega}_{i,k+1},-\vec{p}'_{k+1})\right], \quad (72)$$

which is equal to $-\left(\mathcal{C}^{(a)}_{k,k+1}\rho^{(a)}_{k+1,-t}\right)(\bar{x}_1,\ldots,\bar{x}_k)$, as one can see by contrasting the above equation with (67). By recalling the definition of $\tilde{\rho}^{(a)}_{k,t}$, we conclude that

$$\left(\mathcal{C}^{(a)}_{k,k+1}\rho^{(a)}_{k+1,-t}\right)(\bar{x}_1,\ldots,\bar{x}_k) = -\left(\mathcal{C}^{(a)}_{k,k+1}\tilde{\rho}^{(a)}_{k+1,t}\right)(x_1,\ldots,x_k). \qquad (73)$$

It has thus been shown that the BBGKY hierarchy with the collision term expressed by (32) is time-reversal invariant. Equivalently, the BBGKY hierarchy with the collision term expressed by (52) is time-reversal invariant too. $\qquad\square$

References

1. Alexander, R.K.: The infinite hard-sphere system, Ph.D. Thesis, University of California at Berkeley (1975)

2. Boltzmann, L.: Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen, Wien. Ber. 66, 275–370 (1872); in Boltzmann (1909) vol.I, pp. 316–402

3. Boltzmann, L.: Über das Wärmegleichgewicht Von Gasen, auf welche äußere Kräfte wirken, Wien. Ber. 72, 427–457 (1872); in Boltzmann (1909) vol.II, pp. 1–30

4. Boltzmann, L.: Bemerkungen über einige probleme der mechanische Wärmetheorie, Wien. Ber. 66, 275–370 (1877); in Boltzmann (1909) vol.II, pp. 112–148

5. Boltzmann, L.: On certain questions of the theory of gases. Nature, 51, 413–415 (1895); in Boltzmann (1909) vol.III, pp. 535–544

6. Boltzmann L.: Zu Hrn Zermelos Abhandlung Über die mechanische Erklärung irreversibler Vorgänge, Wied. Ann. 60, 392–398 (1897); in Boltzmann (1909), Vol III, paper 119

7. Boltzmann, L.: Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. In: Hasenöhrl, F. (ed.) Wissenschaftliche Abhandlungen. Chelsea, Leipzig (1909)

8. Cercignani, C.: 134 years of Boltzmann equation. In: Gallavotti, G., Gallavotti, W.L., Yngvason, J. (eds.) Boltzmann's Legacy, pp. 107–128. Zürich: European Mathematical Society, Zürich (2008)

9. Cercignani, C., Illner, R., Pulvirenti, M.: The Mathematical Theory of Diluted Gases. Springer, New York (1994)

10. Culverwell, E.P.: Dr. Watson's proof of Boltzmann's theorem on permanence of distributions. Nature **50**, 617 (1894)

11. Drory, A.: Is there a reversibility paradox? Recentering the debate on the thermodynamic time arrow. Stud. Hist. Philos. Mod. Phys. **39**, 889–913 (2008)

12. Ehrenfest, P., Ehrenfest-Afanassjewa, T.: The Conceptual Foundations of the Statistical Approach in Mechanics. Cornell University Press, New York (1912)

13. Feynman, R., Leighton, R., Sands, M.: The Feynman Lectures on Physics. Addison Wesley, Reading (1964)

14. Illner, R., Pulvirenti, M.: Global validity of the Boltzmann equation for a two-dimensional rare gas in a vacuum. Commun. Math. Phys. 105, 189–203 (1986); erratum ibidem 121 143–146 (1989)

15. Lanford, O.E.: Time evolution of large classical systems. In: Moser, J. (ed.) Dynamical Systems, Theory and Applications. Lecture Notes in Theoretical Physics, vol. 38, pp. 1–111. Springer, Berlin (1975)

16. Lanford, O.E.: On the derivation of the Boltzmann equation. Asterisque **40**, 117–137 (1976)

17. Lanford, O.E.: The hard sphere gas in the Boltzmann-Grad limit. Physica **106A**, 70–76 (1981)

18. Lebowitz, J.L.: Microscopic dynamics and macroscopic laws. In: Horton Jr, C.W., Reichl, L.E., Szebehely, V.G. (eds.) Long-Time Prediction in Dynamics, pp. 220–233. Wiley, New York (1983)

19. Loschmidt, J.: Über die Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft. Wien. Ber. **73**, 139 (1877)

20. Goldstein, S.: Boltzmann's approach to statistical mechanics. In: Bricmon, J., et al. (eds.) Chance in Physics: Foundations and Perspectives. Lecture Notes in Physics, pp. 39–54. Springer, Berlin (2001)

21. Schrödinger, E.: Irreversibility. Proc. R. Ir. Acad. **53**, 189–195 (1950)

22. Spohn, H.: Kinetic equations from Hamiltonian dynamics: markovian limits. Rev. Mod. Phys. **53**, 569–615 (1980)

23. Spohn, H.: Large Scale Dynamics of Interacting Particles. Springer, Berlin (1991)

24. Spohn, H.: Loschmidt's reversibility argument and the H-theorem. In: Fleischhacker, W., Schlonfeld, T. (eds.) Pioneering Ideas for the Physical and Chemical Sciences, Loschmidt's Contributions and Modern Developments in Structural Organic Chemistry, Atomistics and Statistical Mechanics, pp. 153–158. Plenum Press, New York (1997)

25. Spohn, H.: On the integrated form of the BBGKY hierarchy for hard spheres. http://arxiv.org/pdf/math-ph/0605068 (2006). Accessed 25 May 2006

26. Uffink, J.: Compendium of the foundations of classical statistical physics. In: Butterfield, J., Earman, J. (eds.) Handbook for the Philosophy of Physics, pp. 923–1074. Elsevier, Amsterdam (2008)

27. Valente, G.: The approach towards equilibrium in Lanford's theorem. Eur. J. Philos. Sci. **4**(3), 309–335 (2014)

28. Villani, C.: (Ir)reversibilité et entropie. Séminaire Poincaré XV, Le Temps, pp. 17–75. Institut Henri Poincaré, Paris (2010)