on this issue and that it was a matter of taste.

4.2   The Boltzmann equation and H-theorem (1872)

In 1872 Boltzmann published one of his most important papers. It contained two celebrated results nowadays known as the Boltzmann equation and the H-theorem. The latter result was the basis of Boltzmann's renewed claim to have obtained a general theorem corresponding to the Second Law. This paper has been studied and commented upon by numerous authors, and an entire translation of the text has been provided by (Brush 1966). Thus, for the present purposes, a succinct summary of the main points might have been sufficient. However, there is still dispute among modern commentators about its actual content.

The issue at stake in this dispute is the question whether the results obtained in this paper are presented as necessary consequences of the mechanical equations of motion, or whether Boltzmann explicitly acknowledged that they would allow for exceptions. Klein has written:

> I can find no indication in his 1872 memoir that Boltzmann conceived of possible excep-
> tions to the H-theorem, as he later called it (Klein 1973, p. 73).

Klein argues that Boltzmann only came to acknowledge the existence of such exceptions thanks to Loschmidt's critique in 1877. An opposite opinion is expressed by von Plato (1994). Calling Klein's view a "popular image", he argues that, already in 1872, Boltzmann was well aware that his H-theorem had exceptions, and thus "already had a full hand against his future critics". Indeed, von Plato states that

> Contrary to a widely held opinion, Boltzmann is not in 1872 claiming that the Second
> Law and the Maxwellian distribution are *necessary* consequences of kinetic theory (von
> Plato 1994, p. 81).

So it might be of some interest to try and settle this dispute.

Boltzmann (1872) starts with an appraisal of the role of probability theory in the context of gas theory. The number of particles in a gas is so enormous, and their movements are so swift that we can observe nothing but average values. The determination of averages is the province of probability calculus. Therefore, "the problems of the mechanical theory of heat are really problems in probability calculus" (Abh. I, p. 317). But, Boltzmann says, it would be a mistake to believe that the theory of heat would therefore contain uncertainties.

He emphasizes that one should not confuse incompletely proven assertions with rigorously derived theorems of probability theory. The latter are necessary consequences of their premisses, just like in any other theory. They will be confirmed by experience as soon as one has observed a sufficiently large number of cases. This last condition, however, should be no significant problem in

the theory of heat because of the enormous number of molecules in macroscopic bodies. Yet, in this context, one has to make doubly sure that we proceed with the utmost rigour.

Thus, the message expressed in the opening pages of this paper seems clear enough: the results Boltzmann is about to derive are advertised as doubly checked and utterly rigorous. Still, they are theoretical. Their relationship with experience might be less secure, since any probability statement is only reproduced in observations by sufficiently large numbers of independent data. Thus, Boltzmann would have allowed for exceptions in the relationship between theory and observation, but not in the relation between premises and conclusion.

He continues by saying what he means by probability, and repeats its equivocation as a fraction of time and the relative number of particles that we have seen earlier in 1868:

> If one wants [...] to build up an exact theory [...] it is before all necessary to determine the probabilities of the various states that one and the same molecule assumes in the course of a very long time, and that occur simultaneously for different molecules. That is, one must calculate how the number of those molecules whose states lie between certain limits relates to the total number of molecules (Abh. I p. 317).

However, this equivocation is not vicious. For most of the paper the intended meaning of probability is always the relative number of molecules with a particular molecular state. Only at the final stages of his paper (Abh. I, p. 400) does the time-average interpretation of probability (suddenly) recur.

Boltzmann says that both Maxwell and he had attempted the determination of these probabilities for a gas system but without reaching a complete solution. Yet, on a closer inspection, "it seems not so unlikely that these probabilities can be derived on the basis of the equations of motion alone..." (Abh. I, p. 317). Indeed, he announces, he has solved this problem for gases whose molecules consist of an arbitrary number of atoms. His aim is to prove that whatever the initial distribution of state in such a system of gas molecules, it must inevitably approach the distribution characterized by the Maxwellian form (ibid. p. 320).

The next section specializes to the simplest case of monatomic gases and also provides a more complete specification of the problem he aims to solve. The gas molecules are contained in a fixed vessel with perfectly elastic walls. They interact with each other only when they approach each other at very small distances. These interactions can be mimicked as collisions between elastic bodies. Indeed, these bodies are modeled as hard spheres (Abh I, p. 320). Boltzmann represents the state of the gas by a time-dependent distribution function $f_t(\vec{v})$, called the "distribution of state", which gives us, at each time t, the relative number of molecules with velocity between $\vec{v}$ and $\vec{v} + d^3\vec{v}$.[29]

He also states two more special assumptions:

---

[29]Actually Boltzmann formulated the discussion in terms of a distribution function over kinetic energy rather than velocity. I have transposed this into the latter, nowadays more common formulation.

1. Already in the initial state of the gas, each direction of velocity is equally probable. That is:

$$f_0(\vec{v}) = f_0(v). \tag{47}$$

It is assumed as obvious that this will also hold for any later time.

2. The gas is spatially uniform within the container. That is, the relative number of molecules with their velocities in any given interval, and their positions in a particular spatial region R does not depend on the location of R in the available volume.

The next and crucial assumption used by Boltzmann to calculate the change in the number of particles with a velocity $\vec{v}_1$ per unit time, is the *Stoßzahlansatz*, (29) and (30).

For modern readers, there are also a few unstated assumptions that go into the construction of this equation. First, the number of molecules must be large enough so that the (discrete) distribution of their velocities can be well approximated by a continuous and differentiable function f. Secondly, f changes under the effect of binary collisions only. This means that the density of the gas should be low (so that three-particle collisions can be ignored) but not too low (which would make collisions too infrequent to change f at all). These two requirements are already hard enough to put in a mathematically precise form. The modern explicitation is that of taking the so-called Boltzmann-Grad limit (cf. paragraph 6.4). The final (unstated) assumption is that all the above assumptions remain valid in the course of time.

He addresses his aim by constructing a differentio-integral evolution equation for f_t, by taking the difference of (29) and (30) and integrating over all variables except $\vec{v}_1$ and t. The result (in a modern notation) is the *Boltzmann equation*:

$$\frac{\partial f_t(\vec{v}_1)}{\partial t} = N \int_0^d b\,db \int_0^{2\pi} d\phi \int_{\mathbb{R}^3} d^3\vec{v}_2 \,\|\vec{v}_2 - \vec{v}_1\| \left(f_t(\vec{v}_1{}') f_t(\vec{v}_2{}') - f_t(\vec{v}_1) f_t(\vec{v}_2) \right) \tag{48}$$

which describes the change of f in the course of time, when this function at some initial time is given. (Recall from paragraph 3.3 that the primed velocities are to be thought of as functions of the unprimed velocities and the geometrical parameters of the collision: $\vec{v}_i' = \vec{v}_i{}'(\vec{v}_1, \vec{v}_2, b, \phi)$, and d denotes the diameter of the hard spheres.)

4.2.1 The H-theorem

Assuming that the Boltzmann equation (48) is valid for all times, one can prove, after a few well-known manipulations, that the following quantity

$$H[f_t] := \int f_t(\vec{v}) \ln f_t(\vec{v}) d^3\vec{v} \tag{49}$$

decreases monotonically in time, i.e.

$$\frac{dH[f_t]}{dt} \leq 0;$$ (50)

as well as its stationarity for the Maxwell distribution, i.e.:

$$\frac{dH[f_t]}{dt} = 0 \ (\forall t) \ \text{iff} \ f_t(v) = Ae^{-Bv^2}.$$ (51)

Boltzmann concludes Section I of the paper as follows:

> It has thus been rigorously proved that whatever may have been the initial distribution of kinetic energy, in the course of time it must necessarily approach the form found by Maxwell. [...] This [proof] actually gains much in significance because of its applicability to the theory of multi-atomic gas molecules. There too, one can prove for a certain quantity [H] that, because of the molecular motion, this quantity can only decrease or in the limiting case remain constant. Thus, one may prove that because of the atomic movement in systems consisting of arbitrarily many material points, there always exists a quantity which, due to these atomic movements, cannot increase, and this quantity agrees, up to a constant factor, exactly with the value that I found in [(Boltzmann 1871c)] for the well-known integral $\int dQ/T$.

> This provides an analytical proof of the Second Law in a way completely different from those attempted so far. Up till now, one has attempted to proof that $\int dQ/T = 0$ for a reversible (*umkehrbaren*) cyclic[30] process, which however does not prove that for an irreversible cyclic process, which is the only one that occurs in nature, it is always negative; the reversible process being merely an idealization, which can be approached more or less but never perfectly. Here, however, we immediately reach the result that $\int dQ/T$ is in general negative and zero only in a limit case... (Abh. I, p. 345)

Thus, as in his 1866 paper, Boltzmann claims to have a rigorous, analytical and general proof of the Second Law. From our study of the paper until now, (i.e. section I) it appears that Klein's interpretation is more plausible than von Plato's. I postpone a further discussion of this dispute to paragraph 4.2.3, after a brief look at the other sections of the paper.

4.2.2 *Further sections of Boltzmann (1872)*

Section II is entitled "Replacement of integrals by sums" and devoted to a repetition of the earlier arguments, now assuming that the kinetic energies of the molecules can only take values in a discrete

---

[30]The term "cyclic" is missing in Brush's translation, although the original text does speak of "Kreisprozeß". The special notation $\oint$ for cyclic integrals was not introduced until much later.

set $\{0, \epsilon, 2\epsilon, \ldots, p\epsilon\}$. Boltzmann shows that in the limit $\epsilon \longrightarrow 0$, $p\epsilon \longrightarrow \infty$ the same results are recovered.

Many readers have been surprised by this exercise, which seems rather superfluous both from a didactic and a logical point of view. (However, some have felt that it foreshadowed the advent of quantum theory.) Boltzmann offers as motivation for the detour that the discrete approach is clearer than the previous one. He argues that integrals only have a symbolic meaning, as a sum of infinitely many infinitesimal elements, and that a discrete calculation yields more understanding. He does not argue, however, that it is closer to physical reality. Be that as it may, the section does eventually take the limit, and recovers the same results as before.

The third section treats the case where the gas is non-uniform, i.e., when condition 2 above is dropped. For this case, Boltzmann introduces a generalized distribution function $f_t(\vec{r}, \vec{v})$, such that $f_t d^3\vec{r} d^3\vec{v}$ represents the relative number of particles with a position in a volume element $d^3\vec{r}$ around $\vec{r}$ and a velocity in an element $d^3\vec{v}$ around $\vec{v}$.

He obtains a corresponding generalized Boltzmann equation:

$$\frac{\partial f_t(\vec{r}, \vec{v})}{\partial t} + \vec{v} \cdot \nabla_x f_t + \frac{\vec{F}}{m} \cdot \nabla_v f_t =$$
$$N \int b\,db\,d\phi\,d^3\vec{v}_2 \, \|\vec{v}_2 - \vec{v}_1\| \left(f_t(\vec{r}, \vec{v}_1') \right) f_t(\vec{r}, \vec{v}_2') - f_t(\vec{r}, \vec{v}_1)) f_t(\vec{r}, \vec{v}_2) \right) \tag{52}$$

where $\vec{F}$ denotes an external force field on the gas. The quantity H now takes the form $H[f_t] := \int f_t(\vec{r}, \vec{v}) d^3\vec{r} d^3\vec{v}$; and a generalization of the H-theorem $dH/dt \leq 0$ is obtained.

The last three sections are devoted to polyatomic molecules, and aim to obtain generalized results for this case too. The key ingredient for doing so is, of course, an appropriately generalized *Stoßzahlansatz*. The formulation of this assumption is essentially the same as the one given in his paper on poly-atomic molecules (1871a), which was later shown wrong and corrected by Lorentz. I will not go into this issue (cf. Lorentz 1887, Boltzmann 1887b, Tolman 1938).

An interesting passage occurs at the very end of the paper, where he expands on the relationship between H and entropy. He considers a monatomic gas in equilibrium. The stationary distribution of state is given as:

$$f^*(\vec{r}, \vec{v}) = V^{-1} \left(\frac{3m}{4\pi T} \right)^{3/2} \exp(\frac{-3mv^2}{4T}) \tag{53}$$

where V is the volume of the container. (Note that in comparison with (27), Boltzmann adopts units for temperature that make $k = 2/3$.) He shows that

$$H[f^*] := \int f^* \log f^* dx dv = -N \log V \left(\frac{4\pi T}{3m} \right)^{3/2} - \frac{3}{2} N; \tag{54}$$

which agrees (assuming $S = -kNH[f^*]$) with the thermodynamical expression for the ideal gas

47

(16) up to an additive constant. A similar result holds for the polyatomic gas.

4.2.3 Remarks and problems

1. The role of probability.   As we have seen, the H-theorem formed the basis of a renewed claim by Boltzmann to have obtained a theorem corresponding to the full Second Law (i.e. including both parts) at least for gases. A main difference from his 1866 claim, is that he now strongly emphasizes the role of probability calculus in his derivation. It is clear that the conception of probability expounded here is thoroughly frequentist and that he takes 'the laws of probability' as empirical statements. Furthermore, probabilities can be fully expressed in mechanical terms: the probability distribution f is nothing but the relative number of particles whose molecular states lie within certain limits. Thus, there is no conflict between his claims that on the one hand, "the problems of the mechanical theory of heat are really problems in probability calculus" and that the probabilities themselves are derived on the basis of the equations of motion alone, on the other hand. Indeed, it seems to me that Boltzmann's emphasis on the crucial role of probability in this paper is only intended to convey that probability theory provides a particularly useful and appropriate language for discussing mechanical problems in gas theory. There is no indication in this paper yet that probability theory could play a role by furnishing assumptions of a non-mechanical nature, i.e., independent of the equations of motion (cf. Boltzmann & Nabl 1904, p. 520).

2. The role of the Stoßzahlansatz.   Note that Boltzmann stresses the generality, rigour and "analyticity" of his proof. He puts no emphasis on the special assumptions that go into the argument. Indeed, the *Stoßzahlansatz*, later identified as the key assumption that is responsible for the time-asymmetry of the H-theorem, is announced as follows

> The determination [of the number of collisions] can only be obtained in a truly tedious manner, by consideration of the relative velocities of both particles. But since this consideration has, apart from its tediousness, not the slightest difficulty, nor any special interest, and because the result is so simple that one might almost say it is self-evident I will only state this result." (Abh. I, p. 323)

It thus seems natural that Boltzmann's contemporaries must have understood him as claiming that the H-theorem followed necessarily from the dynamics of the mechanical gas model.[31] I can find no evidence in the paper that he intended this claim to be read with a pinch of salt, as (von Plato 1991, p.. 81) has argued.

---

[31] Indeed this is *exactly* how Boltzmann's claims were understood. For example, the recommendation written in 1888 for his membership of the Prussian Academy of Sciences mentions as his main feat that Boltzmann had proven that, whatever its initial state, a gas must necessarily approach the Maxwellian distribution (Kirsten & Körber 1975, p.109).

Is there then no evidence at all for von Plato's reading of the paper? Von Plato refers to a passage from Section II, where Boltzmann repeats the previous analysis by assuming that energy can take on only discrete values, and replacing all integrals by sums. He recovers, of course, the same conclusion, but now adds a side remark, which touches upon the case of non-uniform gases:

> Whatever may have been the initial distribution of states, there is one and only one distribution which will be approached in the course of time. [...] This statement has been proved for the case where the distribution of states was already initially uniform. It must also be valid when this is not the case, i.e. when the molecules are initially distributed in such a way that in the course of time they mix among themselves more and more, so that after a very long time the distribution of states becomes uniform. This will always be the case, with the exception of very special cases, e.g. when all molecules were initially situated along a straight line, and were reflected by the walls onto this line (Abh. I, p. 358).

It is this last remark that, apparently, led to the view that after all Boltzmann did already conceive of exceptions to his claims. However, I should say that this passage does not convince me. True enough, Boltzmann in the above quote indicates that there are exceptions. But he mentions them only in connection with an *extension* of his results to the case when the gas is not initially uniform, i.e. when condition (2) above is dropped. There can be no doubt that under the assumption of the conditions (1) and (2), Boltzmann claimed the rigorous validity of the H-theorem. (Curiously, his more systematic treatment of the non-uniform gas (Section III of (1872)) does not mention any exception to the claim that "H can only decrease" (Abh. I p. 362).

As a matter of fact, when Loschmidt formulated the objection, it happened to be by means of an example of a non-uniform gas (although nothing essential depended on this). Thus, if Boltzmann had in 1872 a "full hand against his future critics", as von Plato claims, one would expect his reply to Loschmidt's objection to point out that Loschmidt was correct but that he had already anticipated the objection. Instead, he accused Loschmidt of a fallacy (see paragraph 4.3 below).

But apart from the historical issue of whether Boltzmann did or did not envisage exceptions to his H-theorem, it seems more important to ask what kind of justification Boltzmann might have adduced for the *Stoßzahlansatz*. An attempt to answer this question must be somewhat speculative, since, as we have seen, Boltzmann presented the assumption as "almost self-evident" and "having no special interest", and hence presumably as not in need of further explanation. Still the following remarks may be made with some confidence.

First, we have seen that Maxwell's earlier usage of the assumption was never far away from an argument from insufficient reason. Thus, in his approach, one could think of the *Stoßzahlansatz* as expressing that we have no reason to expect any influence or correlation between any pair of particles

that are about to collide. The assumption would then appear as a probabilistic assumption, reflecting a 'reasonable judgment', independent from mechanics.

In contrast, Boltzmann's critique of Maxwell's approach (cf. footnote 16) suggests that he did not buy this arguments from insufficient reason. But since the *Stoßzahlansatz* clearly cannot be conceived of as an assumption about dynamics —like the ergodic hypothesis—, this leaves only the option that it must be due to a special assumption about the mechanical state of the gas. Indeed, in the years 1895-6, when Boltzmann acknowledged the need for the *ansatz* in the proof of his H-theorem more explicitly —referring to it as "Assumption A" (Boltzmann 1895) or "the hypothesis of molecular disorder" (Boltzmann 1896)—, he formulated it as an assumption *about* the state of the gas.

Yet, even in those years, he would also formulate the hypothesis as expressing that "haphazard governs freely" (Boltzmann 1895, Abh. III, p. 546) or "that the laws of probability are applicable for finding the number of collisions" (Boltzmann 1895b). Similarly, he describes states for which the hypothesis fails as contrived "so as to intentionally violate the laws of probability"(Boltzmann 1896, §3). However, I think these quotations should not be read as claims that the *Stoßzahlansatz* was a consequence of probability theory itself. Rather, given Boltzmann's empirical understanding of "the laws of probability", they suggest that Boltzmann thought that, as a matter of empirical fact, the assumption would 'almost always' hold, even if the gas was initially very far from equilibrium.

3. The H-theorem and the Second Law. Note that Boltzmann misconstrues, or perhaps understates, the significance of his results. Both the Boltzmann equation and the H-theorem refer to a body of gas in a fixed container that evolves in isolation from its environment. There is no question of heat being exchanged by the gas during a process, let alone in an irreversible cyclic process. His comparison in the quotation on page 46 with Clausius' integral $\int dQ/T$ (i.e. $\oint dQ/T$ in equation (18) above) is therefore really completely out of place.

The true import of Boltzmann's results is rather that they provide (i) a generalization of the entropy concept to non-equilibrium states,[32] and (ii)a claim that this non-equilibrium entropy $-kH$ increases monotonically as the isolated gas evolves for non-equilibrium towards an equilibrium state. The relationship with the Second Law is, therefore, somewhat indirect: On the one hand, Boltzmann proves much more than was required, since the second law does not speak of non-equilibrium entropy, nor of monotonic increase; on the other hand it proves also less, since Boltzmann does not consider the increase of entropy in general adiabatic processes.

---

[32]Boltzmann emphasized that his expression for entropy should be seen as an *extension* of thermodynamic entropy to non-equilibrium states in (1877b, Abh. II, p. 218; 1896, §5) . Of course there is no guarantee that this generalization is the *unique* candidate for a non-equilibrium entropy.

6.4 Lanford's approach to the Boltzmann equation

We now turn to consider some modern approaches to non-equilibrium statistical mechanics. Of these, the approach developed by Lanford and others (cf. Lanford 1975, Lanford 1976, Lanford 1981, Spohn 1991, Cercignani, Illner &Pulvirenti 1994) deserves special attention because it stays conceptually closer to Boltzmann's 1872 work on the Boltzmann equation and the H-theorem than any other modern approach to statistical physics. Also, the problem Lanford raised and tried to answer is one of no less importance than the famous reversibility and recurrence objections. Furthermore, the results obtained are the best efforts so far to show that a statistical reading of the Boltzmann equation or the H-theorem might hold for the hard spheres gas.

The question Lanford raised is that of the consistency of the Boltzmann equation and the underlying Hamiltonian dynamics. Indeed, if we consider the microstate of a mechanical system such as a dilute gas, it seems we can provide two competing accounts of its time evolution.

(1) On the one hand, given the mechanical microstate x_0 of a gas, we can form the distribution of state $f(\vec{r}, \vec{v})$, such that $f(\vec{r}, \vec{v})d^3\vec{v}d^3\vec{r}$ gives the relative number of molecules with a position between $\vec{r}$ and $\vec{r} + d^3\vec{r}$ and velocity between $\vec{v}$ and $\vec{v} + d^3\vec{v}$. Presumably, this distribution should be uniquely determined by the microstate x_0. Let us make this dependence explicit by adopting the notation $f^{[x_0]}$. This function, then, should ideally serve as an initial condition for the Boltzmann equation (48), and solving this equation —assuming, that is, that it, that it has a unique solution— would give us the shape of the distribution function at a later time, $f_t^{[x_0]}(\vec{r}, \vec{v})$.

(2) On the other hand, we can evolve the microstate x_0 for a time t with the help of the Hamiltonian equations. That will give us $x_t = T_t x_0$. This later state x_t will then also determine a distribution of state $f^{[x_t]}(\vec{r}, \vec{v})$.

It is a sensible question whether these two ways of obtaining a later distribution of state from an initial microstate are the same, i.e. whether the two time evolutions are consistent. In other words, the problem is whether the diagram below commutes:

$$
\begin{array}{ccc}
x_0 & \overset{\text{Hamilton}}{\longrightarrow} & x_t \\
\downarrow & & \downarrow \\
f^{[x_0]} & \overset{\text{Boltzmann}}{\longrightarrow} & f_t^{[x_0]} \overset{?}{=} f^{[x_t]}
\end{array}
\tag{136}
$$

The first issue that has to be resolved here is the precise relation between a microstate and the distribution of state f. It is obvious that, in so far as this function represents the physical property of a gas system, it should be determined by the momentary microstate x. It is also clear, that in so far as it is assumed to be continuous and differentiable in time in order to obey the Boltzmann equation, this cannot be literally and exactly true.

So let us assume, as Boltzmann did, that the gas consists of N hard spheres, each of diameter

d and mass m, contained in some fixed bounded spatial region Λ with volume $|\Lambda| = V$. Given a microstate x of the system one can form the 'exact' distribution of state:

$$F^{[x]}(\vec{r}, \vec{v}) := \frac{1}{N} \sum_i^N \delta^3(\vec{r} - \vec{q}_i)\delta^3(\vec{v} - \frac{\vec{p}_i}{m}). \tag{137}$$

This distribution is, of course, not a proper function, and being non-continuous and non-differentiable, clearly not a suitable object to plug into the Boltzmann equation. However, one may reasonably suppose that one ought to be able to express Boltzmann's ideas in a limit in which the number of particles, N, goes to infinity. However, this limit clearly must be executed with care.

On the one hand, one ought to keep the gas dilute, so that collisions involving three or more particles will be rare enough so that they can safely be ignored in comparison to two-particle collisions. On the other hand, the gas must not be so dilute that collisions are altogether too rare to contribute to a change of f. The appropriate limit to consider, as Lanford argues, is the so-called Boltzmann-Grad limit in which $N \longrightarrow \infty$, and:[67]

$$\frac{Nd^2}{V} = \text{constant} > 0. \tag{138}$$

Denote this limit as "$N \xrightarrow{\text{BG}} \infty$", where it is implicitly understood that $d \propto N^{-1/2}$. The hope is then that in this Boltzmann-Grad limit, the exact distribution $F^{[x^N]}$ will tend to a continuous function that can be taken as an appropriate initial condition for the Boltzmann equation. For this purpose, one has to introduce a relevant notion of convergence for distributions on the μ-space $\Lambda \times \mathbb{R}^3$. A reasonable choice is to say that an arbitrary sequence of distributions f_n (either proper density functions or in the distributional sense) converges to a distribution f, $f_n \longrightarrow f$, iff the following conditions hold:

For each rectangular parallelepiped $\Delta \subset \Lambda \times \mathbb{R}^3$:

$$\lim_{n \longrightarrow \infty} \int_\Delta f_N d^3\vec{r} d^3\vec{v} = \int_\Delta f d^3\vec{r} d^3\vec{v}, \tag{139}$$

and $$\lim_{n \longrightarrow \infty} \int \vec{v}^2 f_n d^3\vec{r} d^3\vec{v} = \int \vec{v}^2 f d^3\vec{r} d^3\vec{v}, \tag{140}$$

where the second condition is meant to guarantee the convergence of the mean kinetic energy.

It is also convenient to introduce some distance function between (proper or improper) distributions that quantifies the sense in which one distribution is close to another in the above sense. That

---

[67]The condition can be explained by the hand-waving argument that Nd^2/V is proportional to the 'mean free path', i.e. a typical scale for the distance traveled by a particle between collisions, or also by noting that the collision integral in the Boltzmann equation is proportional to Nd^2/V, so that by keeping this combination constant, we keep the Boltzmann equation unchanged.

is to say, one might define some distance $d(f, g)$ between density functions on $\Lambda \times R^3$ such that

$$d(f_n, f) \longrightarrow 0 \Longrightarrow f_n \longrightarrow f. \tag{141}$$

There are many distance functions that could do this job, but I won't go into the question of how to pick out a particular one.

The hope is then, to repeat, that $F^{[x^N]} \longrightarrow f$ in the above sense when $N \xrightarrow{BG} \infty$, where f is sufficiently smooth to serve as an initial condition in the Boltzmann equation, and that with this definition, the Boltzmannian and Hamiltonian evolution become consistent in the sense that the diagram (136) commutes. But clearly this will still be a delicate matter. Indeed, increasing N means a transition from one mechanical system to another with more particles. But there is no obvious algorithm to construct the state x^{N+1} from x^N, and thus no way to enforce convergence on the level of individual states.

Still, one might entertain an optimistic guess, which, if true, would solve the consistency problem between the Boltzmann and the Hamiltonian evolution in an approximate fashion if N is very large.

> OPTIMISTIC GUESS: If $F^{[x_0^N]}$ is near to f then $F^{[x_t^N]}$ is near to f_t for all $t > 0$, and
> where f_t is the solution of the Boltzmann equation with initial condition f.

As Lanford (1976) points out, the optimistic guess cannot be right. This is an immediate consequence of the reversibility objection: Indeed, suppose it were true for all $x \in \Gamma$, and $t > 0$. (Here, we momentarily drop the superscript N from x^N to relieve the notation.) Consider the phase point Rx obtained from x by reversing all momenta: $R(\vec{q_1}, \vec{p_1}; \ldots; \vec{q_N}, \vec{p_N}) = (\vec{q_1}, -\vec{p_1}; \ldots; \vec{q_N}, -\vec{p_N})$. If $F^{[x]}(\vec{r}, \vec{v})$ is near to some distribution $f(\vec{r}, \vec{v})$, then $F^{[Rx]}(\vec{r}, \vec{v})$ is near to $f(\vec{r}, -\vec{v})$. But as x evolves to x_t, Rx_t evolves to $T_t Rx_t = RT_{-t}x_t = Rx$. Hence $F^{[T_t Rx_t]}(\vec{r}, \vec{v}) = F^{[Rx]}(\vec{r}, \vec{v})$ is near to $f(\vec{r}, -\vec{v})$. But the validity of the conjecture for Rx_t would require that $F^{[T_t Rx_t]}(\vec{r}, \vec{v})$ is near to $f_t(\vec{r}, -\vec{v})$ and these two distributions of state are definitely not near to each other, except in some trivial cases.

But even though the optimistic guess is false in general, one might hope that it is 'very likely' to be true, with some overwhelming probability, at least for some finite stretch of time. In order to make such a strategy more explicit, Lanford takes recourse to a probability measure on Γ, or more precisely a sequence of probability measures on the sequence of Γ_N's.

Apart from thus introducing a statistical element into what otherwise would have remained a purely kinetic theory account of the problem, there is a definite advantage to this procedure. As mentioned above, there is no obvious algorithm to construct a sequence of microstates in the Boltzmann-Grad limit. But for measures this is different. The microcanonical measure, for example is not just a measure for the energy hypersurface of one N-particles-system; it defines an algorithmic sequence

of such measures for each N.

In the light of this discussion, we can now state Lanford's theorem as follows (Lanford 1975, 1976):

> LANFORD'S THEOREM: Let $t \mapsto f_t$ be some solution of the Boltzmann equation, say for $t \in [0, a) \subset \mathbb{R}$. For each N, let Δ_N denote the set in the phase space Γ_N of N particles, on which $F^{[x^N]}$ is near to f_0 (the initial condition in the solution of the Boltzmann equation) in the sense that for some chosen distance function d and for tolerance $\epsilon > 0$:
>
> $$\Delta_N = \{x^N \in \Gamma_N \ : \ d(F^{[x^N]}, f_0) < \epsilon\}. \tag{142}$$
>
> Further, for each N, conditionalize the microcanonical measure μ_N on Δ_N:
>
> $$\mu_{\Delta,N}(\cdot) := \mu_N(\cdot|\Delta_N). \tag{143}$$
>
> In other words, $\mu_{\Delta,N}$ is a sequence of measures on the various Γ_N that assign measure 1 to the set of microstates $x^N \in \Gamma_N$ that are close to f_0 in the sense that $d(F^{[x^N]}, f_0) < \epsilon$. Then: $\exists \tau, 0 < \tau < a$ such that for all t with $0 < t < \tau$:
>
> $$\mu_{\Delta,N}(\{x^N \in \Gamma_N \ : \ d(F^{[x_t^N]}, f_t) < \epsilon\}) > 1 - \delta \tag{144}$$
>
> where $\delta \longrightarrow 0$ as both $\epsilon \longrightarrow 0$ and $N \overset{BG}{\longrightarrow} \infty$.

In other words: as judged from the microcanonical measure on Γ_N restricted to those states x^N that have their exact distribution of state close to a given initial function f_0, a very large proportion $(1 - \delta)$ evolve by the Hamiltonian dynamics in such a way that their later exact distribution of state $F^{[x_t^N]}$ remains close to the function f_t, as evolved from f_0 by the Boltzmann equation.

6.4.1 Remarks

Lanford's theorem shows that a statistical and approximate version of the Boltzmann equation can be derived from Hamiltonian mechanics and the choice of an initial condition in the Boltzmann-Grad limit. This is a remarkable achievement, that in a sense vindicates Boltzmann's intuitions. According to Lanford (1976, p. 14), the theorem says that the approximate validity of the Boltzmann equation, and hence the H-theorem, can be obtained from mechanics alone and a consideration of the initial conditions.

Still the result established has several remarkable features, all of which are already acknowledged by Lanford. First, there are some drawbacks that prevent the result from having practical impact for

the project of justifying the validity of the Boltzmann equation in real-life physical applications. The density of the gas behaves like N/d^3, and in the Boltzmann-Grad limit this goes to zero. The result thus holds for extremely rarified gases. Moreover, the length of time for which the result holds, i.e. τ, depends on the constant in (138), which also provides a rough order of magnitude for the mean free path of the gas . It turns out that, by the same order of magnitude considerations, τ is roughly two fifths of the mean duration between collisions. This is a disappointingly short period: in air at room temperature and density, τ is in the order of microseconds. Thus, the theorem does not help to justify the usual applications of the Boltzmann equation to macroscopic phenomena which demand a much longer time-scale.

Yet note that the time scale is not trivially short. It would be a misunderstanding to say that the theorem establishes only the validity of the Boltzmann equation for times so short that the particles have had no chance of colliding: In two fifths of the mean duration between collisions, about 40 % of the particles have performed a collision.

Another issue is that in comparison with Boltzmann's own derivation no explicit mention seems to have been of the *Stoßzahlansatz*. In part this is merely apparent. In a more elaborate presentation (cf. Lanford 1975, 1976), the theorem is not presented in terms of the microcanonical measure, but an arbitrary sequence of measures ν_N on (the sequence of phase spaces) Γ_N. These measures are subject to various assumptions. One is that each ν_N should be absolutely continuous with respect to the microcanonical measure μ_N, i.e. ν_N should have a proper density function

$$d\nu_N(x) = n_N(x_1, \ldots x_N) dx_1 \cdots x_N \tag{145}$$

where $x_i = (\vec{q}_i, \vec{p}_i)$ denotes the canonical coordinates of particle i. Further, one defines, for each N and $m < N$, the reduced density functions by

$$n_N^{(m)}(x_1, \ldots x_m) := \frac{N!}{(N-m)!} \frac{1}{N^m} \int n_N(x_1, \ldots x_N) dx_{m+1} \cdots dx_N \tag{146}$$

i.e. as (slightly renormalized) marginal probability distributions for the first m particles. The crucial assumption is now that

$$\lim_{N \xrightarrow{\text{BG}} \infty} n_N^{(m)}(x_1, \ldots x_m) = n^{(1)}(x_1) \cdots n^{(1)}(x_m) \tag{147}$$

uniformly on compact subsets of $(\Lambda \times \mathbb{R}^3)^m$. This assumption (which can be shown to hold for the microcanonical measures) is easily recognized as a measure-theoretic analogy to the *Stoßzahlansatz*. It demands, in the Boltzmann-Grad limit, statistical independence of the molecular quantities for any pair or m-tuple of particles at time $t = 0$. As Lanford also makes clear, it is assumption (146) that

would fail to hold if we run the construction of the reversibility objection; (i.e. if we follow the states x in Δ_N for some time t, $0t < \tau$, then reverse the momenta, and try to apply the theorem to the set $\Delta'_N = \{Rx_t \ : \ x \in \Delta_N\}$).

But another aspect is more positive. Namely: Lanford's theorem does not need to assume explicitly that the *Stoßzahlansatz* holds *repeatedly*. Indeed a remarkable achievement is that once the factorization condition (146) holds for time $t = 0$ it will also hold for $0 < t < \tau$, albeit in a weaker form (as convergence in measure, rather than uniform convergence). This is sometimes referred to as "propagation of chaos" (Cercignani, Illner &Pulvirenti 1994).

But the main conceptual problem concerning Lanford's theorem is where the apparent irreversibility or time-reversal non-invariance comes from. On this issue, various opinions have been expressed. Lanford (1975, p. 110) argues that irreversibility is the result of passing to the Boltzmann-Grad limit. Instead, Lanford (1976) argues that it is due to condition (146) plus the initial conditions (i.e.: $x_N \in \Delta_N$).

However, I would take a different position. The theorem equally holds for $-\tau < t < 0$, with the proviso that f_t is now a solution of the anti-Boltzmann equation. This means that the theorem is, in fact, invariant under time-reversal.

6.5   The BBGKY approach

The so-called BBGKY-hierarchy (named after Bogolyubov, Born, Green, Kirkwood and Yvon) is a unique amalgam of the description of Gibbs and the approach of Boltzmann. The goal of the approach is to describe the evolution of ensembles by means of reduced probability densities, and to see whether a Boltzmann-like equation can be obtained under suitable conditions —and thereby an approach to statistical equilibrium.

First, consider an arbitrary time-dependent probability density ρ_t. The evolution of ρ is determined via the Liouville-equation by the Hamiltonian:

$$\frac{\partial \rho_t}{\partial t} = \{H, \rho\}. \tag{148}$$

Central in the present approach is the observation that for relevant systems in statistical mechanics, this Hamiltonian will be symmetric under permutation of the particles. Indeed, the Hamiltonian for a system of N indistinguishable particles usually takes the form

$$H(\vec{q_1}, \vec{p_1}; \ldots; \vec{q_N}, \vec{p_N}) = \sum_{=1}^{N} \frac{\vec{p_i}^2}{2m} + \sum_{i}^{N} V(\vec{q_i}) + \sum_{i<j}^{N} \phi(\|q_i - \vec{q_j}\|) \tag{149}$$