



Boltzmann's reply to the Loschmidt paradox: a commented translation

Olivier Darrigol

► To cite this version:

Olivier Darrigol. Boltzmann's reply to the Loschmidt paradox: a commented translation. European Physical Journal H, 2021, 46 (1), pp.29. 10.1140/epjh/s13129-021-00029-2 . hal-03467467

HAL Id: hal-03467467

<https://hal.science/hal-03467467>

Submitted on 7 Dec 2021

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

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

BOLTZMANN'S REPLY TO THE LOSCHMIDT PARADOX:

A COMMENTED TRANSLATION

**[This is an early draft of a paper published in EPJH : *The European physical journal H* (2021), <https://doi.org/10.1140/epjh/s13129-021-00029-2>
Please use only the published version for citation,]**

Abstract

Boltzmann's reply to Loschmidt's reversibility paradox (1877) has baffled many readers, owing to imprecise language and unproven assumptions. Based on a new translation and detailed commentary, it will be shown that this text nevertheless contains the essentials of a correct, insightful interpretation of thermodynamic irreversibility in statistico-mechanical context.

Keywords: reversibility paradox, Loschmidt, Boltzmann, thermodynamic irreversibility, entropy and probability.

In 1876, Ludwig Boltzmann's old friend Joseph Loschmidt published his version of the reversibility paradox, which opposes the reversibility of the equations of molecular dynamics to the irreversibility of thermodynamic evolution. Boltzmann published his reaction to this argument a few months later, as the second section of a larger memoir that covered several topics in the kinetic-molecular theory (see Fig. 1). In retrospect this text contains the essentials of Boltzmann's probabilistic analysis of thermodynamic irreversibility. There are, however, a few obstacles to penetrating Boltzmann's ingenious considerations. The first is the lack of precise terminology for basic concepts: for instance, Boltzmann's "states" sometimes refer to microstates, sometimes to macrostates, although the distinction between the two kinds of states is needed for a proper definition

of thermodynamic probability. Secondly, the relevant concept of probability is not precisely defined; it is only hinted at. Thirdly, Boltzmann mixes intuitions, assumptions, and theorems without making the categorical difference clear. Lastly, Boltzmann sometimes writes in a convolute style, at least from a non-native German reader's perspective.¹

II.

Über die Beziehung eines allgemeinen mechanischen Satzes zum zweiten Hauptsatze der Wärmetheorie.

In seiner Abhandlung über den Zustand des Wärmegleichgewichtes eines Systems von Körpern, mit Rücksicht auf die Schwerkraft, hat Loschmidt einen Satz ausgesprochen, welcher ein Bedenken gegen die Möglichkeit eines rein mechanischen Beweises des zweiten Hauptsatzes involvirt. Da dasselbe äusserst scharfsinnig erdacht ist und mir für das richtige Verständniss des zweiten Hauptsatzes von hoher Bedeutung zu sein scheint, an der bezeichneten Stelle aber eine mehr philosophische Einkleidung gefunden hat, in der es vielleicht manchem Physiker nur schwer verständlich sein wird, so will ich zunächst versuchen, es mit anderen Worten wiederzugeben. Gesetzt wir wollten auf rein mechanischem Wege beweisen, dass alle Naturprocesse immer so vor sich gehen, dass dabei

$$\int \frac{dQ}{T} \leq 0$$

ist, so fassen wir dabei die Körper als Aggregate materieller Punkte auf. Die zwischen diesen materiellen Punkten thätigen Kräfte fassen wir als Functionen der relativen Lage der materiellen Punkte auf. Wenn sie als Function dieser relativen Lage bekannt wären, so würden wir sagen das Wirkungsgesetz der

Fig. 1: The first page of Boltzmann's reply to Loschmidt in the *Wiener Berichte*.

¹ Loschmidt 1876; Boltzmann 1877a, pp. 116–122.

The purport of this article is to make Boltzmann's important text understandable to a variety of readers through a line-by-line commentary. Section 1 recalls the context of Boltzmann's intervention, with a full citation of Loschmidt's argument. Section 2 offers a new translation of Boltzmann's text. Section 3 is a detailed commentary, sentence by sentence. Section 4 is a condensed, clarified version of Boltzmann's main points. Section 5 is a criticism mostly inspired by Boltzmann's later writings on the same topic.²

The proposed translation is as literal as the difference between English and German grammar permits (I have sometimes divided long sentences into smaller ones). Significant departures from Brush's old translation are explained in footnotes. My commentary is based on internal textual evidence and on knowledge of Boltzmann's contemporary writings. One would otherwise run the risk of retrospective overinterpretation and could not establish the extent to which Boltzmann then anticipated the views he more clearly expounded in the 1890s. That said, my commentary employs vocabulary ('microstates', 'macrostates', 'Liouville's theorem', 'microcanonical distribution'...) that Boltzmann did not use, and it sometimes explicates things that Boltzmann left implicit (for instance how to define the number of states of the kinetic-molecular model compatible with a given macroscopic appearance of the system). This will ease understanding by modern readers, and this can be done, I believe, without making Boltzmann tell more than he intended.

A frequently debated historical question regarding Boltzmann's reply to Loschmidt is how much he thereby departed from his earlier understanding of the role of

² There already are a nearly complete French translation (Dugas 1959, pp. 188–191) and two complete English translations: Brush 1976, pp. 188–193; Gallavotti 2014, pp. 171–174. Translated extracts and critical comments are found in Klein 1973, pp. 71–74; Brush 1976, pp. 239, 605–607; Sklar 1993, p. 39; von Plato 1994, pp. 87–89; Cercignani 1998, §5.2; Uffink 2007, §4.3; Badino 2011, pp. 373–374.

probabilities in the kinetic molecular theory.³ The present article does not address this question, its focus being on the contents of Boltzmann's text of 1877. However, Section 5 briefly describes the arguments through which Boltzmann later clarified and modified his reasoning of 1877.

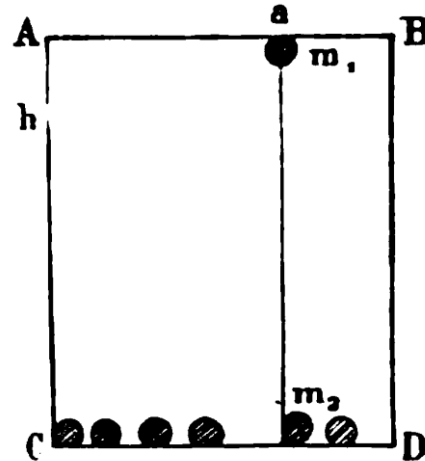
1. The Loschmidt paradox

Joseph Loschmidt, whom Boltzmann had befriended in his early years at the Physics Institute in Vienna, was an ardent defender of the kinetic-molecular theory of heat and he was among those who believed that this theory might save us from the "heat death" implied by the second law of thermodynamics. In particular, he believed gravity could be used to counteract the temperature uniformity implied by the second law: intuitively, the kinetic energy of the molecules of a gas should be lower at higher elevation in compensation for the higher potential energy. Fig. 2 represents one of the devices imagined by Loschmidt: at time zero, all molecules are at rest in the container ABCD. All molecules except one, m_1 , are on the bottom of the container. This molecule is released from the top of the container, and gradually communicates kinetic energy to all the other molecules through a succession of collisions. In the end the gas reaches a stationary state

³ The chief protagonists of this debate are René Dugas, Martin Klein, and Jos Uffink (in favor of radical departure); Jan von Plato, Michel Janssen, and Massimo Badino (in favor of continuity). For references and further discussion, see Darrigol 2018, pp. 542–544.

in which all molecules are moving, with a kinetic energy depending on their height (so thinks Loschmidt).⁴

Fig. 2: Loschmidt's device for producing a vertical temperature gradient in a gas in equilibrium. From Loschmidt 1876, p. 137.



At this point, Loschmidt makes the following remark:

One should be careful with the following assertion: in a system that has evolved from a given initial state to the so-called stationary state, the latter average state will persist for all times. One can predict so much with sufficient confidence only for a short time, I believe.

Indeed, if in the above-given case, after a time τ sufficiently long for the stationary state to obtain, we suddenly reversed the velocities of all atoms, we would thus be at the beginning of a state that would look like the stationary state. This would hold for some time, but the stationary state would deteriorate in the long run and after the time τ we would inevitably return to the initial state, namely, a single atom m_1 would have absorbed the total live force of the system and would have converted it into potential energy while all other atoms would have returned to their static position at the bottom of the container.

Evidently, in full generality in any system one can revert the entire course of events by suddenly inverting the velocities of all the elements of the system. To the egregious problem of undoing what has happened this does not give a solution but a simple formulation, which consists in the following simple

⁴ Loschmidt 1876. Cf. Dugas 1959, pp. 158–184; Darrigol 2018, pp. 181–188. William Thomson anticipated the Loschmidt paradox in 1871 (in private) and in 1874 (in print): see Klein 1970, p. 93; Garber, Brush, and Everitt 1995, pp. 178, 187–188, 192–193, 205.

stipulation: to suddenly revert the instantaneous velocities of all atoms of the universe.

This argument is twofold. Firstly, it involves a simple mechanical theorem, according to which (for the class of mechanical systems usually considered in the kinetic theory of gases) the motion from time $-\tau$ to time 0 under given conditions and the motion from time 0 to time τ after sudden reversal of all velocities at time 0 are the exact reverse of each other. Secondly, it exhibits initial states of the kinetic-molecular model for which the usual tendency toward uniformity does not apply: one just has to take a state that has evolved from a non-uniform state to a uniform state from time $-\tau$ to time 0, and then suddenly inverse all velocities (this operation obviously preserves the uniformity of the initial state).⁵

Loschmidt does not regard the latter result as a paradox. On the contrary, he welcomes this additional blow to the alleged universality of the second law of thermodynamics. More broadly he remarks that in the light of his theorem, the undoing of history becomes conceivable though still not practically feasible.⁶

2. "On the relation between a general mechanical theorem and the second principle of the theory of heat"

⁵ Loschmidt 1876, p. 139.

⁶ Today's physicists know that for some systems the Loschmidt return (better: return to the initial microstate after a double velocity-reversal) can be concretely achieve in some systems. See Hahn 1950; Waugh, Rhim, and Pines 1972.

This is a translation of Boltzmann's text. The notes labelled by letters are Boltzmann's. The numbered notes are mine, and they give extracts of Brush's translation whenever it notably departs from mine.

In his memoir on the state of equilibrium of a system of bodies under gravity, Loschmidt enunciated a theorem [*Satz*] involving a reservation [*Bedenken*] about the possibility of a purely mechanical derivation of the second principle. This reservation⁷ is devised in an extremely astute manner and I think it is highly significant for the correct understanding of the second principle. However, in the given reference it appears in a rather philosophical form that could make it difficult to grasp for many physicists. This is why I will first try to restate it in different words. Suppose we want to prove, in a purely mechanical way, that all natural processes always satisfy

$$\int \frac{dQ}{T} \leq 0.$$

For this purpose, we conceive⁸ the implied bodies as aggregates of material points. We assume the forces acting between these points to be functions of their relative positions. If these functions were known, we would say that the law of force [*Wirkungsgesetz der Kräfte*] is known. In order to compute the actual motion of the material points and the resulting changes in the states of the bodies, we would still have to know the initial positions and the initial velocities of all the material points. That is to say, the initial conditions should be given. When one wants to prove the second principle by mechanics, he always seeks to derive it from the nature of the force law without any reference to the

⁷ Brush: this *theorem*.

⁸ Brush: we *must* conceive.

initial conditions, about which we know nothing. Thus, he tries to prove that—whatever be the initial conditions—the states of the bodies evolve in accordance with

$$\int \frac{dQ}{T} \leq 0.$$

Now suppose certain bodies⁹ to be given as collections of certain material points. The initial conditions at time zero are assumed to be such that the bodies change their state in conformity with

$$\int \frac{dQ}{T} \leq 0.$$

We will show that with the same law of force, one can always find other initial conditions for which conversely

$$\int \frac{dQ}{T} \geq 0.$$

Indeed, consider the positions and the velocities of all the material points after the arbitrary time t_1 has elapsed. In place of the former initial conditions, we now take the following ones: at the initial time, that is, at time zero, all the material points^a have the same positions as the ones they reached at time t_1 under the former initial conditions, and they have equal and opposite velocities. For brevity, we will henceforth call this state the state opposite [*entgegengesetzt*] to the state that formerly occurred at time t_1 .

Clearly, the material points will now pass in reverse order [*in der verkehrten Weise*] the states through which they passed in direct order under the former initial

⁹ Brush: *a certain body*.

^a We thereby include all material points of all bodies that interact with the bodies under consideration, directly or indirectly. Strictly speaking, all bodies of the universe ought to be included, because we cannot build a complex of bodies that has strictly no relation to the other bodies of the universe, we can only imagine it.

conditions.^b The initial state that they formerly had at time zero, is now reached only after the time t_1 has elapsed. Consequently, whereas we formerly had

$$\int \frac{dQ}{T} \leq 0,$$

this integral now is ≥ 0 . That is to say, the sign of this integral cannot be inferred from the law of force;¹⁰ it can be inferred only from the initial conditions. That this integral, as experience teaches us, is always ≤ 0 for processes in the world in which we live, is not founded on the interaction law for the implied forces but only on the initial conditions. If at time zero the state of all the material points of the universe were exactly the state opposite to the one it otherwise reaches at the much later time t_1 , then the course of all events between the times t_1 and zero would be exactly the reverse course, and thus would always be such that

$$\int \frac{dQ}{T} \geq 0.$$

Any attempt to deduce

$$\int \frac{dQ}{T} \leq 0$$

from the nature of the bodies and from the law for the forces acting among them, without invoking the initial conditions, must therefore be in vain. We see that this conclusion has much to entice us and that we should rightly¹¹ call it an interesting sophism. In order to

^b Needless to say, for a picture [*Anschauung*] of the mode of action of natural forces in which this [reversal] is not true (as could happen for instance in a dynamic picture), the following considerations would also fail to apply. [The reference to this footnote is missing in the original *WB* publication; Hasenöhrle reintroduced it in *BWA*, at a location different from the one assumed in this translation.] [By "dynamic picture," Boltzmann probably means a view, such as Lesage's, in which the forces between atoms are traced to motion of intermediate matter.]

¹⁰ Brush: *the sign of this integral ... does not depend on the force law.*

¹¹ Brush omits "rightly" (*geradezu*).

track down the error in the former reasoning, we will now imagine a system made of a finite number of material points that have strictly no interaction with the rest of the universe.

Let a large though not infinite number of absolutely elastic balls move in a closed container whose walls are strictly immobile and likewise absolutely elastic. No external force is allowed to act on the balls. Suppose that at time zero the distribution of the balls in the container is non-uniform. For instance, the density of the balls could be much higher on the right side than on the left side of the container, or¹² they could move faster on the upper part than on the lower part, or anything of the sort. The sophism now reads: without bringing in the initial conditions one cannot prove that the balls will mix uniformly in the long run. Under the originally assumed initial conditions, the balls might for instance be almost uniformly mixed at time t_1 . Now we may replace these initial conditions with the state-distribution [*Zustandsvertheilung*] that is the exact opposite of the state-distribution that would have occurred after time t_1 under the originally assumed initial conditions. Then the elastic balls will sort themselves out more and more in the course of time. In the end at time t_1 there will be a highly non-uniform distribution, even though the initial distribution was nearly uniform. The following remark is in order: a proof that after a sufficient time t_1 has elapsed the mixture of spheres will become uniform with absolute necessity whatever be the state-distribution at the beginning of time, cannot be delivered. This we already know from probability calculus: a non-uniform state-distribution (no matter how much it departs from uniformity) is not absolutely impossible, although it is extremely improbable. Indeed, it is clear that every

¹² Brush: *and*.

individual state-distribution that arises from a given initial state after a given time has elapsed, is just as improbable as any individual state-distribution of arbitrarily high non-uniformity, just as in the game of lotto every individual sequence of five numbers [Quinterne] is equally improbable as the sequence 1, 2, 3, 4, 5. It is only because there are many more uniform than non-uniform state-distributions that the state-distribution is more likely¹³ to become uniform in the course of time. One therefore cannot prove that for every initial value of the positions and velocities of the balls the distribution will always be uniform after a very long time.¹⁴ What one can prove is that after a given long time there are infinitely many more initial states leading to a uniform state-distribution than there are initial states leading to a non-uniform state-distribution; and that even in the latter case the distribution will again become uniform after an even longer time has elapsed.¹⁵ So, Loschmidt's theorem teaches only how to identify initial states that truly lead to a highly non-uniform state-distribution, but it does not give a proof that there are not infinitely many more initial states that would lead to a uniform state-distribution after the same time t_1 has elapsed. On the contrary, this [the dominance of the initial states leading to uniformity] follows from the theorem itself, because, as there are infinitely many more uniform state-distributions than non-uniform ones, the number of states that ensue from uniform states after a given time t_1 has elapsed is much larger than the number of states that ensue from non-uniform states, and the latter states are those which

¹³ Brush omits the "more likely": *will become uniform*.

¹⁴ Brush: *a long time*.

¹⁵ Brush omits this last sentence (after the semicolon).

according to Loschmidt should be taken as initial conditions in order that a non-uniform state-distribution occurs after the time t_1 .¹⁶

One could even compute the probability of the various state-distributions from their relative numbers, which might perhaps lead to an interesting method for calculating thermal equilibrium. Things are quite analogous for the second principle. In a few special cases at least,¹⁷ it has been possible to prove that when a system goes over from a non-uniform to a uniform state-distribution, then $\int dQ / T$ will be negative for this system, and positive in the reverse case. Now, as there are infinitely many more uniform than non-uniform state-distributions, the latter case is extremely improbable and in practice it can be regarded as impossible: just as we can regard it as impossible that oxygen and nitrogen be initially mixed in such a manner that¹⁸ a month later the oxygen be found chemically pure in the lower half of the container and the nitrogen in the upper half, which is extremely improbable but not impossible according to probability calculus. Nevertheless, Loschmidt's theorem seems highly important to me, for it shows how intimately the second principle is connected to probability calculus, whereas the first principle is completely independent from this calculus. In all cases where $\int dQ / T$ can be negative, there still are particular, extremely improbable initial conditions¹⁹ for which it is positive; and the proof that it will be almost always negative can be carried out only through probability calculus. It seems to me that for closed trajectories of the atoms²⁰ $\int dQ / T$

¹⁶ Brush gives a completely different translation of this sentence. In particular, he renders *folgen auf* as "result in" instead of "ensue from." So does Dugas, in French.

¹⁷ Brush: *It is only in special cases that...*

¹⁸ Brush: *if one starts with oxygen and hydrogen mixed in a container...*

¹⁹ Brush: *there is also an individual very improbably initial condition...*

²⁰ Brush: *atom.*

must always be zero, which can thus be proved independently of probability calculus. In contrast, for unclosed trajectories this quantity can also be negative. I will still mention a peculiar consequence of Loschmidt's theorem: whether we follow the states of the universe into an infinitely remote past or into the remotest future, we have fundamentally equal reasons to regard it as very probable that we will reach a state in which all temperature differences have vanished. This is analogous to the following case: if we know that in a gas at a certain time there is a non-uniform state-distribution and that the gas has been present in the same container for a very long time without any external influence, we ought to conclude that much earlier the state-distribution was uniform and that the rare case has occurred in which the distribution gradually becomes non-uniform. In other words: any non-uniform state-distribution leads after a long time t_1 to a nearly uniform distribution. The distribution opposite to the latter distribution leads after the same time t_1 back to the initial non-uniform distribution (more exactly to the opposite of this distribution). However, the distribution opposite to the initial distribution would, when chosen as initial condition, likewise lead to a nearly²¹ uniform distribution after the time t_1 .

However, this transfer [*Verweisung*] of the second principle to the domain of probability calculus may perhaps raise strong doubts about applying this principle to the entire universe, no matter how certainly the laws of probability calculus will be confirmed in laboratory experiments.²²

²¹ Brush omits "nearly".

²² Brush: *If perhaps this reduction of the second law to the realm of probability makes its application to the entire universe appear dubious, yet the laws of probability theory are confirmed by all experiments carried out in the laboratory.*

3. Comments

Boltzmann's successive assertions are given in italics, my comments are sandwiched in Roman.

In his memoir on the state of equilibrium of a system of bodies under gravity, Loschmidt enunciated a theorem involving a reservation about the possibility of a purely mechanical derivation of the second principle.

As will appear in a moment, what Boltzmann calls "Loschmidt's theorem" is the possibility of reversing the evolution of a purely mechanical system to a given state by inverting the velocities of all the material points of the system in this state. The "reservation" refers to Loschmidt's remark that, as a consequence of this theorem, it is possible to choose the initial positions and velocities of the molecules of the gas so that the distribution of the molecules is initially uniform and yet evolves toward a non-uniform distribution. This contradicts the second law of thermodynamics. Therefore, it should not be possible to derive this law by purely mechanical considerations.

Boltzmann had earlier asserted this impossibility for different reasons. In 1866, he gave a purely mechanical expression of the entropy for periodic systems but later convinced himself of the impossibility of generalizing this result to non-periodic systems. In 1871, he obtained the canonical entropy formula

$$S = -k \int \rho \ln \rho d\sigma ,$$

in which $d\sigma$ is the standard measure in phase space and $\rho \propto e^{-\beta H}$ is the canonical distribution of the phases of the global system. He interpreted this distribution as a temporal *probability*, namely: he defined $\rho d\sigma$ as the fraction of time that the phase of the system spends in the element $d\sigma$ of phase-space in the long run. He then insisted that the derivation of an entropy for the kinetic-molecular model "required the consideration of this probability."²³

[Loschmidt's] reservation is devised in an extremely astute manner and I think it is highly significant for the correct understanding of the second principle. However, in the given reference it appears in a rather philosophical form that could make it difficult to grasp for many physicists. This is why I will first try to restate it in different words.

Clearly, Boltzmann does not treat Loschmidt's remark as an obvious point, and he promptly acknowledges its importance, even though he was already somehow convinced that probabilities should play an important role in deducing entropy from a mechanical (kinetic-molecular) model. His perception of Loschmidt's phrasing as "philosophical" may refer to Loschmidt's brief allusion to the "problem of undoing what has happened" and more broadly to Loschmidt's concern with the heat death of the universe. As we are about to see, the main originality of Boltzmann's rephrasing is its recourse to Clausius's entropy and its reliance on a generic model of interacting material points.

²³ Boltzmann 1866; 1871, p. 295. Cf. Darrigol 2018, pp. 70–79; 128–133.

Suppose we want to prove, in a purely mechanical way, that all natural processes always satisfy

$$\int \frac{dQ}{T} \leq 0.$$

Here and below, Boltzmann uses this inequality as a symbolic expression for the entropy increase of a closed system when left to evolve by itself. He had already done so in 1872 while discussing the H theorem.²⁴ This is of course improper, since for a closed system there is no exchanged heat. Clausius's inequality concerns the evolution of a system exchanging the heat dQ with sources at the temperature T . It implies entropy increase for the global, closed system comprehending the former system and the surrounding sources. Boltzmann knew all of that, but he was also aware of his contemporaries' difficulties with the abstract concept of entropy, and he preferred a loose expression involving the familiar notions of heat and temperature.

For this purpose, we conceive the implied bodies as aggregates of material points. We assume the forces acting between these points to be functions of their relative positions. If these functions were known, we would say that the law of force is known.

For instance, we may assume that all forces derive from a potential V that depends only on the differences $\mathbf{r}_i - \mathbf{r}_j$ between the positions of two material points. In modern terms, the resulting equations of motion, $m_i \ddot{\mathbf{r}}_i = -\nabla_i V$, are invariant under time reversal.

²⁴ Boltzmann 1872, pp. 345–346.

In order to compute the actual motion of the material points and the resulting changes in the states of the bodies, we would still have to know the initial positions and the initial velocities of all the material points. That is to say, the initial conditions should be given.

This follows immediately from the fact that the equations of motion are of second differential order.

When one wants to prove the second principle by mechanics, he seeks to derive it from the nature of the force law without any reference to the initial conditions, about which we know nothing. Thus, he tries to prove that—whatever be the initial conditions—the states of the bodies evolve in accordance with

$$\int \frac{dQ}{T} \leq 0.$$

That we know nothing about the initial conditions is an exaggeration: we know that the initial microstates should be compatible with the initial macrostate. But since this is the only information we have, it seems natural to expect that the entropy law should hold as a consequence of the laws of motion, for any choice of the initial microstate compatible with this information.

Now suppose the bodies to be given as collections of material points. The initial conditions at time zero are assumed to be such that the bodies change their state in conformity with

$$\int \frac{dQ}{T} \leq 0.$$

No one would deny the existence of initial microstates for which the entropy increases, since in nature we constantly witness entropy-increasing processes.

We will show that for the same law of force, one can always find other initial conditions for which

$$\int \frac{dQ}{T} \geq 0.$$

Consider the positions and the velocities of all material points after the arbitrary time t_1 has elapsed. In place of the former initial conditions, we now take the following ones: at the initial time, that is, at time zero, all the material points have the same positions as the ones they reached at time t_1 under the former initial conditions, and they have equal and opposite velocities. For the sake of concision, we will henceforth call this state the one opposite to the state that formerly occurred at time t_1 .

In symbols, the original motion $\mathbf{r}_i(t)$ has the initial conditions $\mathbf{r}_i(0) = \mathbf{r}_i^0$, $\dot{\mathbf{r}}_i(0) = \mathbf{v}_i^0$, under which $\mathbf{r}_i(t_1) = \mathbf{r}_i^1$, $\dot{\mathbf{r}}_i(t_1) = \mathbf{v}_i^1$. The new motion $\bar{\mathbf{r}}_i(t)$ has the initial conditions $\bar{\mathbf{r}}_i(0) = \mathbf{r}_i^1$, $\dot{\bar{\mathbf{r}}}_i(0) = -\mathbf{v}_i^1$.

Clearly, the material points will now pass in reverse order the states through which they passed in direct order under the former initial conditions. The initial state that they formerly had at time zero, is now reached only after the time t_1 has elapsed.

In symbols, we have $\bar{\mathbf{r}}_i(t) = \mathbf{r}_i(t_1 - t)$ and $\dot{\bar{\mathbf{r}}}_i(t) = -\dot{\mathbf{r}}_i(t_1 - t)$. Strictly speaking (as Boltzmann will note in a moment), the state reached by the new motion at time t_1 is the *velocity-reverse* of the state it formerly had at time zero.

Consequently, whereas we formerly had

$$\int \frac{dQ}{T} \leq 0,$$

this integral now is ≥ 0 . That is to say, the sign of this integral cannot be inferred from the law of force; it can be inferred only from the initial conditions. That this integral, as experience teaches us, is always ≤ 0 for processes in the world in which we live, is not founded on the interaction law for the implied forces but only on the initial conditions.

For any given law of force, there are initial conditions for which the entropy of the system is decreasing. Then it seems obvious that the systematic entropy increase observed in actual experiments (on closed systems) cannot be a consequence of the force law by itself and should be traced to the choice of initial conditions. Boltzmann will soon denounce the latter assertion as a sophism.

If at time zero the state of all the material points of the universe was exactly the state opposite to the one it otherwise reaches at the much later time t_1 , then the course of all events between the times t_1 and zero would be exactly the reverse course, and thus would always be such that

$$\int \frac{dQ}{T} \geq 0.$$

Any attempt to deduce

$$\int \frac{dQ}{T} \leq 0$$

from the nature of the bodies and from the law for the forces acting among them, without invoking the initial conditions, must therefore be in vain.

This is just a repeat of what has been just said for a closed system, now applied to the entire universe regarded as a closed system (as Boltzmann mentions in a footnote, the universe might be the only strictly closed system).

We see that this conclusion has much to entice us and that we should rightly call it an interesting sophism. In order to track down the error in the former reasoning, we will now imagine a system made of a finite number of material points that have strictly no interaction with the rest of the universe.

What Boltzmann calls a sophism is not Loschmidt's theorem, which is a mere consequence of the laws of mechanics. Nor is it the assertion that we can imagine choices

of the microscopic initial conditions for which the entropy decreases.²⁵ Nor else is it the assertion that the laws of mechanics are by themselves insufficient to derive the observed entropy increase in natural processes. It is the assertion that the sign of the observed entropy variation is purely contingent on the choice of the initial conditions, or that for some reason the only initial conditions realized in nature are those for which entropy increases.

In order to identify the source of this sophism, Boltzmann returns to a closed system of material points. The number of these points is finite, and in the following they are concretely represented as the centers of spherical balls that bounce on each other or on the walls of the container in a perfectly elastic manner.

Let a large though not infinite number of absolutely elastic balls move in a closed container whose walls are strictly immobile and likewise absolutely elastic. No external force is allowed to act on the balls. Suppose that at time zero the distribution of the balls in the container is non-uniform. For instance, the density of the balls could be much higher on the right side than on the left side of the container, or they could move faster on the upper part than on the lower part, or anything of the sort. The sophism now reads: without bringing in the initial conditions one cannot prove that the balls will in the long run mix uniformly. Under the originally assumed initial conditions, the balls might for instance be almost uniformly mixed at time t_1 . Now we may replace these initial conditions with the state-distribution that is the exact opposite of the state-distribution

²⁵ Jos Uffink (2007, § 4.3.1, Point 1) gives this interpretation of the "sophism" and therefore sees a contradiction with Boltzmann's subsequent admission of initial states for which the entropy law is violated. Without using the word sophism, Lawrence Sklar (1993, p. 39) correctly remarks that in Boltzmann's opinion, Loschmidt's argument does not force us "to posit the existence of specific initial conditions" and that we can instead "take the statistical viewpoint."

that would have occurred after time t_1 under the originally assumed initial conditions.

Then the elastic balls will be more and more unequally distributed. At time t_1 there will be a highly non-uniform distribution, even though the initial distribution was nearly uniform.

This is just the concrete illustration of Loschmidt's theorem, in the same spirit as the gas model used by Loschmidt himself, and without explicit recourse to the entropy.

The following remark is in order: a proof that after a given time t_1 the mixture of spheres will become uniform with absolute necessity whatever be the state-distribution at the beginning of time, cannot be delivered.

The words "mixture" and "state-distribution" call for some explanation. Boltzmann's talk of uniform mixing (*Mischung*) is easily understood if he has in mind the uniformity of temperature or of chemical composition of the gas (when assumed to be made of two kinds of molecules) rather than uniformity of density. From the context, "state-distribution" (*Zustandsvertheilung*) clearly refers to what we would now call a microstate: the list of the positions and velocities of all the balls.²⁶ Then Boltzmann's "remark" is just a consequence of Loschmidt's theorem: even when the spheres are uniformly distributed in the container at time zero, there are choices of their initial positions and velocities such that their distribution at time t_1 will be non-uniform. Yet,

²⁶ This is properly noted in Uffink 2007, §4.3.1, Point 2.

Boltzmann is about to give another justification of possible anti-thermodynamic evolutions:

This we already know from probability calculus: a non-uniform state-distribution (no matter how much it departs from uniformity) is not absolutely impossible, though it is extremely improbable.

At first glance, this assertion looks strange because it seems to concern only the relative probability of different kinds of states of the system at a given time, and not the evolution of the system. From the preceding and succeeding sentences, we can guess Boltzmann's true meaning: for a randomly chosen initial microstate, the final microstate has a non-zero probability of being non-uniform. But there is another difficulty: what is the intended meaning of probability in the latter assertion? Boltzmann's gives us a clue in his next sentence:

Indeed it is clear that every individual state-distribution that arises from a given initial state after a given time has elapsed, is exactly as improbable as any individual state-distribution of arbitrarily high non-uniformity, just as in the game of lotto every individual sequence of five numbers is equally improbable as the sequence 1, 2, 3, 4, 5. It is only because there are many more uniform than non-uniform state-distributions that the state-distribution is more likely to become uniform in the course of time.

This analogy indicates that Boltzmann has a discrete model in mind, in which the system can have only a finite number of discrete microstates. The evolution from time zero to time t_1 is reversible, and there is a one-to-one correspondence between initial and final microstates. Consequently, if the initial microstate is chosen randomly, any microstate can be reached at time t_1 , including those compatible with a non-uniform macroscopic appearance (I will call them non-uniform microstates). If, in addition, the number of non-uniform microstates is much smaller than the number of uniform ones, then there is a very small probability to reach non-uniformity at time t_1 . Boltzmann does not explain why the uniform microstates are so much more numerous than the non-uniform one. This is easily seen on a simple discrete model, for instance by evaluating the probability of a given proportion of heads and tails after flipping a coin many times. Evidently, the probability of a proportion close to one (meaning uniformity) is very close to certainty.

One therefore cannot prove that for every initial value of the positions and velocities of the balls the distribution will always be uniform after a very long time. What one can prove is that after a given long time there are infinitely many more initial states leading to a uniform state-distribution than there are initial states leading to a non-uniform state-distribution; and that even in the latter case the distribution will again become uniform after an even longer time has elapsed.

This is in part a fuller statement of what has been announced before, with a discrete model in mind: for a randomly chosen initial microstate, the system almost certainly

reaches a uniform state after a given time. Here and elsewhere, Boltzmann's "infinitely many more" should be read as "vastly many more."

But there is a first novelty: Boltzmann now writes "after a given long time" instead of "after a given time." Presumably, he has in mind not only the Loschmidt case in which the initial distribution is uniform and is most likely to remain uniform after *any* given time, but also a different case in which the initial distribution is non-uniform and most likely leads to uniformity after a *sufficiently long* time. In the latter case, the explanation would be that most of the microstates compatible with the initial non-uniform macro-distribution of the balls lead to a uniform distribution. This is far less obvious than the former statement regarding the improbability of a fluctuation out of equilibrium.

A second, interesting novelty is the remark that in the rare event in which the end state is non-uniform, the system is extremely likely to return to uniformity after some time. Again, this is not as obvious as the improbability of a Loschmidt fluctuation, because the more we restrict the initial microstates (here to the ones that lead to a non-uniform state at t_1), the less we can be sure that the end state will be uniform. Boltzmann implicitly relies on an additional principle, according to which the uniformity of the state of the system is always extremely likely to increase over a given fixed time, no matter what the previous (macro-) history may be.

So, Loschmidt's theorem teaches only how to identify initial states that truly lead to a highly non-uniform state-distribution, but it does not give a proof that there are not infinitely many more initial states that would lead to a uniform state-distribution after the same time t_1 has elapsed.

By Boltzmann's earlier argument, there surely are initial microstates for which the final state of the system is non-uniform. Loschmidt tells us how to find such microstates. He does not tell us that the number of these initial states is comparable to the number of those for which the final state is uniform (although naive common sense might suggest so much, based on the symmetry of the laws of mechanics by time reversal).

On the contrary, this [the dominance of initial microstates leading to uniformity] follows from the theorem itself, because, as there are infinitely many more uniform state-distributions than non-uniform ones, the number of states that ensue from uniform states after a given time t_1 has elapsed is much larger than the number of states that ensue from non-uniform states, and the latter states are those which according to Loschmidt should be taken as initial conditions in order that a non-uniform state-distribution occurs after the time t_1 .

Call Ω the set of uniform states, ω the set of non-uniform microstates, U the evolution over the given time lapse, and μ the uniform measure on the energy shell in phase space. In symbols, Boltzmann means $\mu(\Omega) \gg \mu(\omega)$, which implies $\mu(U\Omega) \gg \mu(U\omega)$, because by Liouville's theorem the measure μ is preserved by the evolution U . The initial microstates that lead to uniform and non-uniform states after the time t_1 are the velocity-reverse of the microstates belonging to $U\Omega$ and $U\omega$ retrospectively (the velocity reversal is only implicit in Boltzmann's "the latter states are those which *according to*

Loschmidt..."). As the velocity-reversal preserves the measure μ , the former states are immensely more probable than the latter.²⁷

In a simpler, equivalent reasoning, the dominance of microstates leading to uniformity may be expressed as $\mu(U^{-1}\Omega) \gg \mu(U^{-1}\omega)$. This inequality results from $\mu(U^{-1}\Omega) = \mu(\Omega)$, $\mu(U^{-1}\omega) = \mu(\omega)$, and $\mu(\Omega) \gg \mu(\omega)$.²⁸

One could even compute the probability of the various state-distributions from their relative numbers, which might perhaps lead to an interesting method for calculating thermal equilibrium.

In this sentence, Boltzmann has silently changed the meaning of "state-distribution." It cannot refer, as it did so far, to a microstate of the system, since by Boltzmann's lotto analogy all microstates are equiprobable. It can only refer to a macrostate, for instance defined by a macroscopic partition of the molecules. The probability of a macrostate can then be defined as proportional to the number of compatible microstates. This is exactly what Boltzmann will do a few months later in his famous memoir "On the relation between the second law of the mechanical theory of heat and probability calculus with regard to the laws of thermal equilibrium." There the microstates are discrete "complexions" and the "state-distribution" is defined by cutting off the phase space of a molecule into equal cells and listing the number of molecules in each cell. The most

²⁷ Boltzmann had used special cases of Liouville's theorem since 1868, and he had recently read Watson's treatise (Watson 1876) in which the theorem was derived and used in broad Hamiltonian context: see Darrigol 2018, pp. 472–475.

²⁸ Cf. Uffink 2007, §4.3.1, Point 4.

probable velocity distribution in this sense is a discrete version of Maxwell's law, and it therefore represents the state of equilibrium of the gas.²⁹

Things are quite analogous for the second principle.

Boltzmann is indeed going to argue the following: just as we may compute equilibrium by probability considerations (computing the state of maximal probability), we may justify the second law by probability considerations (requiring the system to evolve toward states of increasing probability).

In a few special cases at least, it has been possible to prove that when a system goes over from a non-uniform to a uniform state-distribution, then $\int dQ / T$ will be negative for this system, and positive in the inverse case.

Boltzmann means that the entropy variation calculated according to thermodynamics³⁰ is positive when the system goes over from a non-uniform to a uniform state, for instance when a gas expands into a vacuum (by removing a wall in a double container) or when two different gases are allowed to diffuse into each other. The entropy variation would therefore be negative in the inverse transition.

²⁹ Boltzmann 1877b. The alternative meaning of "state-distribution" already occurred in Boltzmann's memoirs of 1871 and 1872.

³⁰ As is well known, this calculation is done by imagining a reversible transformation connecting the initial state to the end state and integrating the ratio dQ / T .

Now, as there are infinitely many more uniform than non-uniform state-distributions, the latter case is extremely improbable and in practice it can be regarded as impossible: just as we can regard it as impossible that oxygen and nitrogen be initially mixed in such a manner that a month later the oxygen be found chemically pure in the lower half of the container and the nitrogen in the upper half, which is extremely improbable but not impossible according to probability calculus.

Here is again the remark that a system is extremely unlikely to evolve from a uniform to a non-uniform state, now illustrated by the unmixing of two gases. Evolution from the mixed state to the unmixed state is not impossible, because the relative probability of the unmixed state is not strictly zero; but it is highly improbable because the relative probability is extremely small. This observation does not quite justify Boltzmann's earlier suggestion that the second law can be derived through considerations of probability. To prove so much, he would have to show that the possibility and probability of the evolution from a given macrostate to another are determined by the relative probability of these two states not only when the initial microstate is uniform but also when it is non-uniform.

Nevertheless, Loschmidt's theorem seems highly important to me, for it shows how intimately the second principle is connected to probability calculus, while the first principle is completely independent from this calculus.

Boltzmann has earlier announced the significance of Loschmidt's theorem for a proper understanding of the second law. He now explains that the theorem is important in showing that the second law, unlike the first law of thermodynamics, cannot be derived from the laws of mechanical evolution only: the derivation must involve considerations of probability. Not so easy to understand is Boltzmann's "nevertheless." Plausibly, he means that the oxygen-nitrogen mixing illustration and the concomitant intuition that the probability of an evolution is determined by the relative probability of the initial and final macrostates sufficiently indicate, without Loschmidt's theorem, that anti-thermodynamic evolution is possible. But Loschmidt's theorem more rigorously shows the impossibility of ignoring probability considerations in relating the second law to the laws of mechanics.

In all cases for which $\int dQ / T$ can be negative, there still are special, extremely improbable initial conditions for which it is positive; and the proof that it will be almost always negative can be given only through probability calculus.

The first of these two statements is a consequence of Loschmidt's theorem. It excludes a derivation of the second law based only on the law of mechanics. Yet it does not force us to rule out the anomalous initial conditions (the contrary belief is what Boltzmann earlier called a sophism). What is required is a probability calculus through which these conditions are shown to be extremely improbable. Here Boltzmann does not mean that an individual initial microstate for which the entropy increases is much more probable than an individual initial microstate for which the entropy decreases (he has already asserted

that all microstates are a priori equiprobable). What he means is that for a given initial macrostate, the entropy increases for almost every microstate compatible with this macrostate (and this is true in both directions of time!).

It seems to me that for closed trajectories of the atoms, $\int dQ / T$ must always be zero, which can thus be proved independently of probability calculus. In contrast, for unclosed trajectories it can be negative.

This is an allusion to Boltzmann's memoir of 1866, in which he gave a mechanical entropy formula for a set of atoms that all (very nearly) return to their original position after some time. In this case, no probability consideration is needed to prove that the entropy does not vary during a spontaneous evolution of the system. As was already mentioned, by 1871 Boltzmann had become convinced that the case of non-periodic motions, for which the entropy may vary, requires probability considerations. The next section of Boltzmann's present memoir indeed gives a new proof that Boltzmann's theorem of 1866 cannot be extended to non-periodic motion.³¹ Loschmidt's theorem further shows that the sign of the entropy variation cannot result from the laws of mechanics alone.

I will still mention a peculiar consequence of Loschmidt's theorem: whether we trace back the states of the universe into an infinitely remote past or into the remotest future,

³¹ Boltzmann 1866; 1877a, pp. 122–148. On the latter text, see Darrigol 2018, pp. 199–203.

we have fundamentally equal reasons to regard as very probable that we will reach a state in which all temperature differences have vanished.

Boltzmann again ventures to apply his probabilistic considerations to the entire universe. He implicitly assumes that all microstates compatible with the present ($t = 0$) macrostate of the universe are equiprobable. The set of these microstates is globally invariant by velocity-reversal, so that it leads to the same relative probability of uniform and non-uniform states at times t and $-t$. If the uniform state is most probable in a remote future, so is it too in the remote past. Somewhat paradoxically, Boltzmann's probabilistic concept of irreversibility is time-symmetric!

This is analogous to the following case: when we know that in a gas at a certain time there is a non-uniform state-distribution and that the gas has been present in the same container for a very long time without any external influence, we ought to conclude that much earlier the state-distribution was uniform and that the rare case has occurred in which the distribution gradually becomes non-uniform [Assertion 1]. In other words: any non-uniform state-distribution leads after a long time t_1 to a nearly uniform distribution [Assertion 2]. The distribution opposite to the latter distribution leads after the same time t_1 back to the initial non-uniform distribution (more exactly to the opposite distribution) [Assertion 3]. However, the distribution opposite to the initial distribution would, when chosen as initial condition, all the same leads to a uniform distribution after the time t_1 [Assertion 4].

Boltzmann here returns to safer grounds: the system is simply a gas that has been for a long time in the same container. In this case, it seems natural to assume that in the present non-equilibrium macrostate of the gas, the compatible microstates are all equiprobable. Assertions 2 and 4 follow from the additional assumption that almost every microstate compatible with a given non-uniform macrostate leads to a uniform state after the sufficiently long time t_1 . Assertion 3 is just an expression of Loschmidt's theorem, and it is equivalent to Assertion 1 since the backward evolution is the same as the forward evolution after velocity reversal of the initial state. Assertions 3 and 4 together confirm Boltzmann's earlier remark that "even in the [Loschmidt] case the distribution will again become uniform after an even longer time has elapsed."

However, this transfer of the second principle to the domain of probability calculus may perhaps raise strong doubts about the application of this principle to the entire universe, no matter how certainly the laws of probability calculus will be confirmed in laboratory experiments.

Boltzmann now returns to the case of the complete universe and questions the analogy he has just used with the gas case. He probably has in mind that the universe exists only in one global microstate, and that we cannot repeat observations of the universe in which the microstate compatible with its present macrostate would vary. Then the evolution of the universe is not a matter of probability. In contrast, the law of entropy increase is a matter of probability. Its applicability to the universe is therefore questionable.

4. A condensed version of Boltzmann's reply

Based on the former analysis we may extract the essence of Boltzmann's reply to Loschmidt as follows. Firstly, Boltzmann readily admits Loschmidt's theorem (T1) according to which the evolution of a (purely) mechanical system can be reversed by sign-reversing the velocities in the final state, as well as the disturbing consequence that in a molecular-mechanical model of thermodynamic systems, there are initial states of this model for which the entropy decreases. From this anomaly he does not infer that an ad hoc selection of the mechanically possible states is needed to conciliate the micro-model with the entropy law. This inference he regards as a sophism.

His solution is to remark that the initial microstates for which the entropy law is violated are extremely few, and then to downgrade the entropy law to a statement about the most probable evolution of a system. Whatever the examples taken by Boltzmann for the micromodel, it corresponds to a Hamiltonian system for which the motion is described by Hamilton's equations. The system is isolated, and all forces are internal.

In this context, a first component of Boltzmann's probabilistic reinterpretation of the entropy law is the rigorous theorem:

T2: The subsets of initial microstates that lead, after a given time has elapsed, to microstates compatible with the equilibrium macrostate and with a nonequilibrium macrostate respectively, have the same relative measure as the subsets of microstates compatible with these two macrostates.

This theorem is a straightforward consequence of Liouville's theorem, when applied to the standard measure on the energy shell of the Hamiltonian system.

By analogy with simple discrete models, Boltzmann knows

K1: The relative measure of the subsets of non-equilibrium and equilibrium microstates is extremely small when the number of components (atoms) of the micromodel is huge.

In addition, he assumes A1: in actual laboratory experiments, all microstates compatible with the initial macrostate are equiprobable. Then, if the initial macrostate is a state of equilibrium, K1 implies

P1: the probability of a fluctuation in which the system would be found out of equilibrium at a given later time is vanishingly small.

This is Boltzmann's first, best-founded step toward conciliating the entropy law with Loschmidt's theorem. P1 may be rephrased as follows: a closed system is very unlikely to evolve from the macrostate of maximal probability (the equilibrium state) to a state of much lower probability (a non-equilibrium state).

Without warning, Boltzmann moves from P1 to

P2: when a system is originally out of equilibrium, the probability that after a sufficiently long time it will reach equilibrium is extremely large.

By A1 and K1, the equilibrium macrostate is immensely more probable than any non-equilibrium macrostate. Boltzmann seems to regard it as intuitively obvious that

A2: A closed system made of an enormous number of molecules should almost certainly evolve toward states of increasingly high probability, until the state of highest probability (the equilibrium state) is reached.

Or he may just be unaware of the difference between P1 and P2.³² At any rate, his reasoning is supported by simple models in which the tendency toward states of higher probability is intuitively obvious. For instance, if a great number of rapidly moving elastic balls are originally confined in the left half of a rigid container through a sliding wall, we easily imagine that after removing the sliding wall the balls will quickly spread

³² This confusion is unfortunately reflected in Darrigol 2018, pp. 197, 560. It will be corrected in the paperback edition (2021).

through the new available volume until approximate uniformity is reached. In contrast, we do not imagine that an originally uniform distribution of the balls in the volume of the container could lead after some time to a significantly non-uniform distribution. Things are different if the number of balls is small: it becomes quite conceivable that after some not too large time all the balls are found in the left half of the container. But this becomes extremely unlikely if the number of balls is of the order of the Avogadro number.

Although A2 (in full generality) is only implicit in Boltzmann's reply to Loschmidt, it appears a few months later in his famous memoir relating entropy to combinatorial probability, as "A [closed] system of bodies constantly evolves from a less probable to a more probable state." Similar statements, sometimes more carefully expressed as concerning the *most probable* evolution, frequently appear in Boltzmann's later writings. I call them *the Boltzmann principle*, because there is much evidence that Boltzmann used this principle to justify the entropy law (intuitively at least), not vice-versa.³³

As Boltzmann realizes, A1 and A2 together lead to the conclusion that a non-equilibrium macrostate at time zero almost certainly leads to states increasingly close to the equilibrium state for positive times *and also for negative times*. This is counterintuitive, because we can easily imagine a situation in which at time zero the system is in the middle of a transition from a non-equilibrium to an equilibrium state. Then the system should be further from equilibrium for negative times, just as a Loschmidt state would do for positive times. Boltzmann, however, does not mean his probabilistic prediction to apply in such transitory situations. He assumes either that the system exists only for positive time (as is the case in the sliding-wall experiment) or that

³³ Boltzmann 1877b, p. 166. See Darrigol 2021, pp. 560–563.

the system has long been evolving by itself before time zero. In these cases, he believes he can safely assume that all microstates compatible with the initial macrostate are equally probable (A1).³⁴

5. Criticism

In retrospect, we know that most of Boltzmann's statements regarding the evolution of a suitable mechanical micro-model are correct and that the probabilistic kind of irreversibility he sees in this evolution is genuine.³⁵ This can be verified, for instance, by numerical simulation of one of Boltzmann's model: a large number of elastic disks moving in a plane between rigid walls.³⁶ However, his probabilistic principles of evolution (P) result from a heterogeneous mixture of theorems (T), unproven assumptions (A), and knowledge inferred from simple models (K). While the T's and K's are sound enough, the A's raised the suspicion of Boltzmann's critics. Why and in what sense should the various microstates compatible with a given macrostate be declared equiprobable? Why should a system evolve toward states of higher and higher probability? I will briefly indicate Boltzmann's later reply to these questions, based on the so-called H curve.³⁷

³⁴ In contrast, when we already know that the system has been evolving from a state out of equilibrium, there is only a very small fraction of the microstates compatible with the present macrostate that are compatible with this knowledge, and A1 cannot be true.

³⁵ See Cercignani 1998; Gallavotti 2014.

³⁶ See Alder and Wainwright 1960; Orban and Bellemans 1967.

³⁷ Ernst Zermelo (1896, pp. 795–796) strongly attacked A2 in 1896, arguing that probabilities by themselves could say nothing on the evolution of a system. Jos Uffink (2007, §4.3.1, Points 2 and 3) deplores the lack of justification of both A1 and A2, and his overall judgment of Boltzmann's reply is

Let a plausible mechanical model of a monatomic gas evolve during a time much longer than the relaxation time (which is the time that the gas takes to return to equilibrium when starting from a non-equilibrium state). Divide the phase-space of a molecule into cells of uniform size ε , and call N_i the number of molecules in the cell i at a given time. Define the macrostate of the gas as the distribution $(N_1, N_2, \dots, N_i, \dots)$. The volume of the portion of the global space associated with a given macrostate is

$$\Omega = \varepsilon^N \frac{N!}{N_1! N_2! \dots N_i! \dots}.$$

By Boltzmann's A1, this quantity should be proportional to the probability of the macrostate. For a very large value of the total number N of molecules, we very nearly have

$$\ln \Omega = -\sum_i N_i \ln N_i + \text{constant}.$$

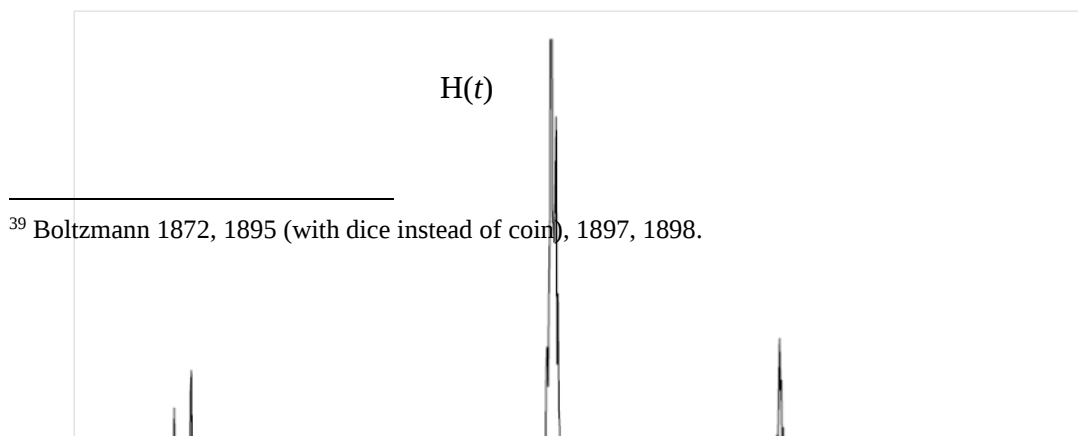
By definition, the H function is the sum $\sum_i N_i \ln N_i$. As Boltzmann proved in 1877, the distribution for which the probability is maximal or, equivalently, the one for which H is minimal under the constraints of constant N and constant total energy is a discrete version of Maxwell's distribution (it is spatially uniform and it is Gaussian for the velocities). Moreover, the value of H for this distribution agrees with the classical entropy of a monatomic gas, and its value out of equilibrium agrees with the function which forever increases according to Boltzmann's H theorem of 1872. Boltzmann therefore regards $-H$ as an extension of the entropy for states out of equilibrium.³⁸

unflattering: "One may question whether his considerations of the probability of the initial state hit the nail on the head. Probability theory is equally neutral to the direction of time as is mechanics."

³⁸ Boltzmann 1877b, 1881a.

The Boltzmann equation of 1872 for the evolution of the distribution of the molecules is based on Maxwell's intuitive expression for the number of collisions of various kinds occurring in the gas per unit time. It leads to a strict decrease of the H function in time, as was just said. But this is only an approximation, in a sense Boltzmann did not clarify in 1872. The exact evolution of H follows from the equations of motion of the molecules, since these equations determine which molecules enter or leave a given cell i in the course of time. Boltzmann imagines a coin being flipped an indefinite number of times and compares the evolution of H with the variation of the proportion p of heads and tails for the N flips of ranks $n, n+1, \dots, n+N$ when n increases (n here plays the role of a discrete time). He probably has in mind some similarity between the random flips and the molecular collisions, which cause random jumps of the velocity values. The resulting curve is represented in Fig. 2.³⁹

In the course of time, the value of H is most likely to be very close to the equilibrium value $H_{\min}$, just as the value of p is most likely to be equal to $1/2$. The stochastic character of the dynamics implies that the evolution of H is highly irregular, with occasional fluctuations, fluctuations within the fluctuations, and so forth. The frequency of the fluctuations of H is a very rapidly decreasing function of their amplitude, as is the case for the departure $|p - \frac{1}{2}|$ of p from the equilibrium value $1/2$.



³⁹ Boltzmann 1872, 1895 (with dice instead of coin), 1897, 1898.

Fig. 2: Boltzmann's H curve. From Uffink 2007, §4.5.3.

Suppose that at a given time the gas is seen out of equilibrium. On the graph of the H function, draw the horizontal line corresponding to the value of H for this non-equilibrium state. This line occasionally intersects a peak of the H curve, and it most probably does so very near the summit of this peak since the frequency of the fluctuations is a very rapidly decreasing function of their amplitude. If the gas system is on the right side of the peak, then H will decrease in subsequent times. If it is on the left side, it will increase for a short time until it passes the summit and starts decreasing. This is how Boltzmann justifies his statements of 1877 regarding the most probable increase of entropy. This picture is evidently time-symmetric: when starting from a state out of equilibrium, the H function is most likely to decrease, both in the positive and in the negative direction of time.

There are two difficulties with this picture. Firstly, Boltzmann does not give any rigorous proof that the H curve truly has the shape he imagines for a gas system. Secondly, it is not clear why in actual experiments the probability of the outcome is

dictated by the probability of being at a given height of the H curve *in the course of time*. Nowadays, we might alleviate the first difficulty by noting that numerical simulations of the evolution of gas models do yield the desired behavior of H. In a more mathematical vein, one could assume the ergodic hypothesis according to which the global phase of the gas system densely fills the energy shell in phase space in the long run. The fraction of time that the system spends in a given portion of this shell in the long run is then proportional to the measure of this portion. In particular, the temporal probability of a given distribution of the gas molecules is proportional to e^{-H} , which measures the volume of phase space associated with the distribution. This explains why H, in the course of time, is very likely to be very close to the equilibrium value, and why significant departures from this value are very rare, with a frequency rapidly diminishing with the amplitude of the departure.

Boltzmann briefly assumed ergodicity in 1868, but soon judged it implausible. Moreover, he understood that the time scale for the ergodic filling of the energy shell would be enormously larger than any physically relevant time. For these reasons, he replaced ergodicity with a weaker hypothesis, according to which for almost every choice of the initial phase, the temporal average of the observable properties of the system reaches a very nearly constant limit after a moderately long time (of the order of the relaxation time). This is what I called *the Boltzmann hypothesis*. Boltzmann had of course no way to derive this hypothesis from the dynamics of the micromodel, but he could allege conformity with the observed existence and uniqueness of thermodynamic

equilibrium.⁴⁰ Under the ergodic hypothesis, the temporal average $\bar{Q}$ of a physical quantity Q over an infinite time can be replaced with its average $\langle Q \rangle$ over the energy shell (microcanonical average), because we have

$$\bar{Q} = \langle \bar{Q} \rangle = \overline{\langle Q \rangle} = \langle Q \rangle.$$

This is still true under the Boltzmann hypothesis, except that the first equality is no longer exact, and that the temporal average is now performed over a physically accessible long time. In particular, the temporal average of $H, H^2, H^3, \dots$ can be replaced with the microcanonical average, so that the temporal distribution of the values of H is very nearly given by their microcanonical weight e^{-H} , as was exactly the case under the ergodic hypothesis.⁴¹

We are left with the second difficulty, regarding Boltzmann's way of relating the H curve with the probable evolution of a single gas system from a given initial state. In a laboratory experiment, we may artificially prepare a gas in a state in which half of the molecules are in the left half of the container, by suddenly removing a sliding wall. This implies that at time zero, the system is either on the left side or the right side of a peak of the H curve. The choice of the starting point on the H curve being otherwise arbitrary, the system is most likely to be very close to the summit of the peak. Consequently, after a finite time has elapsed, the system is most likely to be found in a state of lower H ,

⁴⁰ This justification is fragile, because there is no warranty that in actual relaxation experiments all possible initial microstates may occur.

⁴¹ Boltzmann 1868, pp. 95–96; 1881b, pp. 592–593. See Darrigol 2018, pp. 558–560. Boltzmann's analysis does not apply to systems for which the microcanonical distribution does not yield the observed time-averages. This the case for systems involving long-range coupling, encountered for instance in plasma physics. For an extension of statistical mechanics to such systems, cf. Tsallis 2009.

although H may very briefly increase in the first instants.⁴² If we wait long enough after equilibrium has been reached and then observe the gas at a randomly chosen time, there is a tiny chance that it will be found out equilibrium. For most of these occurrences, the system will be very near a summit of the H curve, and the value of H should be lower both before and after the chosen time, just as Boltzmann predicted.

These predictions are based on the temporal probability of the various values of H . Under the ergodic hypothesis or under the Boltzmann hypothesis, this temporal probability coincides with the combinatorial probability e^{-H} based on the assumption A1 for the equiprobability of all microstates compatible with a given macrostate. This explains why Boltzmann's predictions of 1877 based on combinatorial probability coincide with predictions based on the shape of the H curve. Boltzmann's assumption A1 is thereby justified.

The H curve also sheds light on the frequent case in which we know the system to be in the process of transiting from a non-equilibrium state to the equilibrium state. For instance, at time zero we observe a not yet completely mixed state of two gases that have been diffusing into each other for a while. In this case, we know that the system must originally be somewhere on the right side of a peak of the H curve, that it proceeded from a higher value of H , and that for a given positive time it is most likely to take a lower value of H . Clearly, we cannot account for this behavior by assuming the equiprobability of all microstates compatible with the macrostate at time zero. This would lead to the false prediction that the system most likely is on top of a peak of the H curve. Indeed, the

⁴² In such experiments there is a common arrow of time, from the unlifted to the lifted sliding wall. In other words, the temporal asymmetry, which is absent in the complete H curve, is generated by the preparation of the initial state.

past of the system cannot be judged by probabilistic considerations since it is already known. At best, such considerations apply to the future of the system.

There is a residual difficulty. On the H curve, there are as many points at which H tends to decrease as there are points at which H tends to increase. This perfect symmetry, which is a straightforward consequence of the temporal symmetry of the microdynamics, seems to be incompatible with the dissymmetry between past and future in our universe: whenever we observe a quasi-closed subsystem in our universe the entropy increases for any of these subsystems. In Eddington's words, there is a common "arrow of time."⁴³ As Boltzmann explains to the author of this objection (Ernst Zermelo), the error is to consider all the subsystems as existing *in abstracto* from an indefinite past to an indefinite future. In reality, all the subsystems are extracted from a universe which we know to have started from a globally improbable state (this is the so-called *past hypothesis*). There is a global arrow of time defined by the global entropy increase of the universe. When a closed subsystem of this universe is formed either by artificial means or by natural means, the sign of the entropy increase for this subsystem is determined by the fact that the formation of the subsystem occurs in the positive direction of time. As judged from the entropy variation, the direction of time flow is therefore the same in each of the subsystems as in the global universe.⁴⁴

It is worth noting that this reasoning contradicts Boltzmann's earlier statement (in 1877) that the entropy law, being reduced to a probabilistic statement, might not apply to the universe (of which we have only one instantiation). The entropy law, as he now conceives it, is not the statement that for a system out of equilibrium at time zero the

⁴³ Eddington 1929, p. 68.

⁴⁴ Boltzmann 1897. Cf. Schrödinger 1951, pp. 193–195.

entropy tends to decrease both in the positive and in the negative direction of time, it is the statement that there is a choice of the direction of time for which the entropy tends to decrease. The universe, as far as it may be regarded as a closed system comparable to a gas system, has its own H curve. Since it is notoriously out of equilibrium and since it has been so for a long time, it must be somewhere on an upward or downward slope of a peak of this curve, not too close to the summit nor to the equilibrium baseline. *By definition*, the positive direction of time is that for which H is decreasing.⁴⁵

Lastly, one may wonder, as Boltzmann's British readers did, how the smooth, monotonic evolution of H given by the Boltzmann equation relates to the exact H curve given by the molecular dynamics. Although Boltzmann never gave a precise answer to this question, a few clues may be found in his reply to British criticism.⁴⁶ For the most, he reasoned at the molecular level, introducing the assumption of "molecular chaos" under which Maxwell's collision formula (which leads to the Boltzmann equation) is *statistically* valid. Since the Boltzmann equation is a consequence of this formula, it also has a limited statistical validity. The evolution of H given by the Boltzmann equation corresponds to the most likely shape of the relevant peaks of the H curve. Starting from a given non-equilibrium distribution of the molecules of the gas, we are most likely on top of a peak, and although for each possible peak (compatible with the initial distribution) there are fluctuations on the slopes of the peak, these fluctuations average out from peak

⁴⁵ Boltzmann 1898, p. 637. See Darrigol 2018, pp. 552–553 (also for the related Boltzmann-Schütz argument).

⁴⁶ Boltzmann 1895, p. 541. On the British debate, cf. Brown, Myrvold, and Uffink 2009, pp. 182–284. Loschmidt and Boltzmann did not discuss the Boltzmann equation in their first exchange; they did so a few months later (see Darrigol 2018, pp. 228–231).

to peak and the evolution is smoothly monotonous on average.⁴⁷ The end result is a symmetric cusp with the horizontal asymptote $H = H_{\min}$. As Boltzmann briefly indicated, the Boltzmann-Grad limit of an infinite number of molecules with vanishingly small collision probability (keeping the mean free path constant during this double limit), would yield the same behavior.⁴⁸ Indeed, in this limit the probability of fluctuations becomes smaller and smaller, whatever their size. The successive peaks on the H curve are more and more distant and they get smoother and smoother. In the end, for a given initial value of H , we are left with a single symmetric cusp. The other peaks are thrown to infinity.⁴⁹

Bibliography

The following abbreviations are used:

*BW*Ai *Wissenschaftliche Abhandlungen von Ludwig Boltzmann* (Leipzig: Barth, 1909), vol. i.

WB (*Wiener Berichte*) Kaiserliche Akademie der Wissenschaften zu Wien, mathematisch-naturwissenschaftliche Classe, *Sitzungsberichte*.

Alder, Berni, and Thomas Wainwright
1960 Studies in molecular dynamics. II. Behavior of a small number of elastic spheres. *Journal of chemical physics*, 33: 1439–1451.

Ardourel, Vincent
2017 Irreversibility in the derivation of the Boltzmann Equation. *Foundations of physics*, 47: 471–489.

⁴⁷ The Ehrenfests (1911, pp. 44–45) gave a similar interpretation of some of Boltzmann's remarks, based on the "concentration curve" (*Verdichtungskurve*) for the bundle of H -curves starting from the microstates compatible with the given initial distribution. See Sklar 1993, pp. 63–67; Uffink 2007, pp. 60–61.

⁴⁸ Boltzmann 1896–1898, vol. 2, p. 112. Cf. Darrigol 2018, p. 443.

⁴⁹ These considerations may shed light on the (in)compatibility of Lanford's theorem with the reversibility of the micro-dynamics, recently debated in Uffink and Valente 2015 and in Ardourel 2017.

Badino, Massimiliano

2011 Mechanistic slumber vs. statistical insomnia: The early history of Boltzmann's H-theorem (1868–1877). *The European physical journal H*, 36: 353–378.

Boltzmann, Ludwig

1866 Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie. *WB*, 53: 195–220. Also in *BWA1*, 9–33.

1868 Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *WB*, 58: 517–560. Also in *BWA1*, 49–96.

1871 Analytischer Beweis des zweiten Hauptsatzes der mechanischen Wärmetheorie aus den Sätzen über das Gleichgewicht der lebendigen Kraft. *WB*, 63: 712–732. Also in *BWA1*, 288–308.

1872 Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *WB*, 66: 275–370. Also in *BWA1*, 316–402.

1877a Bemerkungen über einige Probleme der mechanischen Wärmetheorie. *WB*, 7: 62–100. Also in *BWA2*, 112–148.

1877b Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *WB*, 76: 373–435. Also in *BWA2*, 164–223.

1881a Über einige das Wärmegleichgewicht betreffende Sätze. *WB*, 84: 136–145. Also in *BWA2*, 572–581.

1881b [Review of Maxwell 1879]. In *Beiblätter zu Annalen der Physik*, 5: 403–417. English in *Philosophical magazine*, 14 (1882), 299–313. Also in *BWA2*, 582–595.

1895 On certain questions of the theory of gases. *Nature*, 51: 413–415. Also in *BWA3*, 535–544.

1896–1898 *Vorlesungen über die Gatheorie*. 2 vols. Leipzig: Barth.

1897 Zu Hrn. Zermelos Abhandlung 'über die mechanische Erklärung irreversibler Vorgänge.' *Annalen der Physik*, 60: 392–398. Also in *BWA3*, 579–586.

1898 Über die sogenannte H-Kurve, *Mathematische Annalen*, 50: 325–332. Also in *BWA3*, 629–637.

Brown, Harvey, Wayne Myrvold, and Jos Uffink

2009 Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics. *Studies in the history and philosophy of modern physics*, 40: 174–191.

Brush, Stephen

1966 *Kinetic theory*. Vol. 2: *Irreversible processes*. New York: Pergamon Press.

1976 *The kind of motion we call heat: A history of the kinetic theory of gases in the 19th century*, 2 vols. Amsterdam.

Cercignani, Carlo

1998 *Ludwig Boltzmann: The man who trusted atoms*. Oxford: Oxford University Press.

Darrigol, Olivier

2018 *Atoms, mechanics, and probability: Ludwig Boltzmann's statistico-mechanical writings—an exegesis*. Oxford: Oxford University Press.

2021 Paperback edition of the former, including a few corrections.

Dugas, René

1959 *La théorie physique au sens de Boltzmann*. Neuchâtel: Griffon.

Eddington, Arthur Holly

1929 *The nature of the physical world* [The Gifford lectures of 1927, Edinburgh]. New York: MacMillan.

Ehrenfest, Paul, and Tatiana Ehrenfest

1911 *Begriffliche Grundlagen der statistischen Auffassung in der Mechanik*.

Encyklopädie der mathematischen Wissenschaften, IV/2.II: 1–90.

Gallavotti, Giovanni

2014 *Nonequilibrium and irreversibility*. Heidelberg: Springer.

Garber, Elisabeth, Stephen Brush, and C.W.F. Everitt

1995 *Maxwell on heat and statistical mechanics*. Bethlehem: Lehigh University Press.

Hahn, Erwin

1950 Spin echoes. *Physical review*, 80: 580–594.

Klein, Martin

1970 Maxwell, his demon, and the second law of thermodynamics. *American scientist*, 58: 84–97.

Loschmidt, Josef

1876 Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft. I. *WB*, 73: 128–142.

Orban, J, and A. Bellemans

1967 Velocity inversion and irreversibility in a dilute gas of hard disks. *Physics letters*, 24A: 620–621.

Schrödinger, Erwin

1951 Irreversibility. *Proceedings of the Royal Irish Academy*, 53: 189–195.

Sklar, Lawrence

1993 *Physics and chance: Philosophical issues in the foundations of statistical mechanics*. Cambridge: Cambridge University Press.

Tsallis, Constantino

2009 *Introduction to nonextensive statistical mechanics: Approaching a complex world*. New York: Springer.

Uffink, Jos

2007 Compendium to the foundations of classical statistical mechanics. In Jeremy Butterfield and John Earman (eds.) *Handbook for the philosophy of physics* (Amsterdam: Elsevier), 924–1074.

Uffink, Jos, and Giovanni Valente

2015 Lanford's Theorem and the Emergence of Irreversibility. *Foundations of physics*, 45: 404–438.

von Plato, Jan

1994 *Creating modern probability: Its mathematics, physics and philosophy in historical perspective*. Cambridge: Cambridge University Press.

Watson, Henry William

1876 *A treatise on the kinetic theory of gases*. Oxford: Clarendon Press.

Waugh, John, Won-Kyu Rhim, and Alexander Pines

1972 Spin echoes and Loschmidt's paradox. *Journal of pure and applied chemistry*, 32: 317–324.

Zermelo, Ernst

1896 Über mechanische Erklärungen irreversibler Vorgänge. Eine Antwort auf Hrn. Boltzmann's 'Entgegnung.' *Annalen der Physik*, 59: 793–801.