Mechanistic slumber vs. statistical insomnia: The early phase of Boltzmann's H-theorem (1868-1877)

Massimiliano Badino*

1 The riddle of the statistical turn

1.1 The received view

In 1872 Ludwig Boltzmann published a lengthy memoir containing two fundamental results: an integro-differential equation describing the time evolution of an ideal gas (Boltzmann equation) and a mathematical argument proving that such evolution must reach, sooner or later, the state of equilibrium described by the Maxwell distribution. The latter achievement, later called the H-theorem, expressed in mechanistic language the essence of the second law of thermodynamics according to which thermodynamic systems tend irreversibly toward a final state. Five years later, in 1877, Boltzmann tackled the equilibrium problem from a completely different angle and emphasized that the evolution toward the Maxwell distribution is a matter of probability: the equilibrium is reached and maintained because is the overwhelmingly most probable state. How Boltzmann moved from the kinetic language of 1872 to the overtly statistical terminology of 1877 has been the subject of a long debate among the specialists.

Following the lead of Paul and Tatiana Ehrenfest,[1] many working physicists and historians have supported a narrative according to which, while in 1872 Boltzmann believed the H-theorem to be free of exceptions, he was abruptly awoken from his 'mechanistic slumber' in 1876 by a brilliant argument of Josef Loschmidt. This criticism made him realize that the H-theorem, as well as the second law, have merely a statistical meaning, which he eventually spelled out in his 1877 paper. Martin Klein has been the first upholder of this narrative. He pointed out that 'it was often the pressure of external criticism that forced Boltzmann to re-examine his position and refine his understanding.'[2] These objections led 'Boltzmann to rethink the very basis of his proof of the second law',[3] that is the H-theorem. Klein's main argument hinges on a literal reading of the 1872 paper: 'I find no indication in his 1872 memoir that Boltzmann conceived of possible exceptions to the H-theorem, as he later called it. His argument made essential use

*Email: mbadino@mpiwg-berlin.mpg.de

[1]Ehrenfest and Ehrenfest (1911).

[2](Klein, 1973, 55); in his book on Ehfenfest (Klein, 1970, 94-112), Klein had taken a more prudent stance although still on the line of the mechanistic slumber narrative.

[3](Klein, 1973, 71).

of the distribution function, to be sure, but his conclusion was presented in absolute form.'[4] Starting from this reading, Boltzmann's extensive use of probabilistic concepts in 1877 must necessarily appear as a radical turn.

This interpretation has been endorsed by the majority of the commentators and it still represents the standard view. H. Brown, W. Myrvold, and J. Uffink, for example, agree with Klein that the importance of Boltzmann's shift from 1872 to 1877 'cannot be overstressed.'[5] Their argument, like Klein's, also relies on the lack of evidence: 'there is no paper of Boltzmann that clearly states that his *H*-theorem should be read as merely a probable decrease rather than a strict one before 1877.'[6]

The number of the critics of the 'mechanistic slumber' narrative is not as much large. Jan von Plato has been one of the first to claim that Boltzmann might have envisioned a statistical meaning of the *H*-theorem from the very beginning,[7] while Michel Janssen has pointed out the role of the authority of the Ehrenfests in making the mechanistic slumber narrative popular among the historians. More particularly, he has suggested that Klein committed a sort of 'creative misreading' by superimposing Ehrenfests's account on his reading of Boltzmann somewhat in the same way as he also superimposed Ehrenfests's account on his reading of Planck.[8]

1.2 Two puzzles

Appealing though the mechanistic slumber narrative may appear, it does not fit too well some historical facts. First, if in 1877 Boltzmann abandoned the strict mechanistic view in favor of the probabilistic one, why did he keep consistently using the 1872 approach throughout his publications until his death? The insistence on Loschmidt's criticism as a watershed leads us to think that the probabilistic approach dominated the late work of a finally converted Boltzmann. However even a cursory examination reveals that he hardly used his probabilistic theory again after 1877.[9] Turning the argument of the mechanistic slumber narrative against itself we can ask: why do not we find any Boltzmann's pronouncement of his own statistical turn? In response to this puzzle historians have disappointingly appealed to an alleged unreliability of Boltzmann's style. Klein stressed that Boltzmann 'changed his point of view without informing the reader',[10] while Brown *et al.* went so far as to claim that a radical change actually occurred 'despite some very misleading remarks by Boltzmann to the contrary.'[11]

[4](Klein, 1973, 73); a similar account can be found in (Kuhn, 1978, 39-54).

[5](Brown et al., 2009, 185); see also (Uffink, 2007, 967-971).

[6](Brown et al., 2009, 187); the same argument can be found in (Uffink, 2007, 967): 'I can find no evidence in the paper [of Boltzmann] that he intended this claim [the strict validity of the *H*-theorem] to be read with a pinch of salt'.

[7](Von Plato, 1994, 77-82).

[8]Janssen (2002). In effect, the historiographical case of Boltzmann resembles closely that of Planck, cf. Darrigol (2001), Badino (2009).

[9]As an effective working tool it appears again in (Boltzmann, 1884, 1909, III, 66-100) where it is used to deal with dissociation theory, a subject particularly suitable for the combinatorial formalism.

[10](Klein, 1973, 83).

[11](Brown et al., 2009, 185).

Obviously this liberality in interpreting post-1877 Boltzmann's own words conflicts with the literal reading that was requested to handle the 1872 theory. This brings me to the second point. The main argument for the mechanistic slumber claims that Boltzmann meant his H-theorem to be strictly true because in 1872 *he did not mention any exception*. But the lack of evidence is seldom evidence of the lack and, in this specific case, we have historical testimony that Boltzmann was aware of exceptions to the second law well before Loschmidt's 1876 fatal argument. To understand this point we have to look deeper into the relation between Boltzmann and Loschmidt.

The issue of the statistical meaning of the second law is usually related to the thought experiment of the Maxwell demon published in 1871: an imaginary being able to select gas particles from either side of a partition in order to turn equilibrium state into non-equilibrium ones.[12] However, as early as 1869, Josef Loschmidt, Boltzmann's colleague at the Physics Department of the University of Vienna and his close personal friend, published a paper in which an 'inanimate' version of the Maxwell demon was presented.[13] Loschmidt argued that the physical descriptions given by kinetic theory deploy average values, for instance of the energy, and this means that at a microscopic level some particles have an energy above the average, some have one under the average. It is therefore possible to induce macroscopic changes in the system by a suitable selection procedure. Loschmidt supposes a volume V separated by a partition from a sub-volume v, a setup quite similar to Maxwell's. The selection procedure works as follows:[14]

> If the initial states of all molecules in V are known, then the molecules hitting on a particular unit surface σ at any ensuing instant are completely determined. Now we assume that at the position of σ an opening in the partition is placed, which is able to open and close at deliberate moments; it is thus possible to arrange this device so that only those molecules will enter [a sub-volume v] whose velocity is higher that the mean value c and will be even possible to increase their number so that also the gas density in v will become higher than that in V. It is therefore theoretically possible to raise a gas from a lower to a higher temperature or to increase its density without expense of work or specific compensations.

We know for certain that Boltzmann knew this argument from the outset. During his stay in Berlin he published a long review paper in *Die Forschritte der Physik*, the official journal of the Berlin Physical Society, in which he discussed the most significant publications on the theory of heat for the years 1869-1870.[15] Loschmidt's paper is awarded of a long and detailed review that focuses chiefly upon some chemical consequences. At the end of the review Boltzmann summarized the argument of the 'demonic' device and pointed out the main consequence:[16]

---

[12](Maxwell, 1871, 308-309).

[13]Loschmidt (1869).

[14](Loschmidt, 1869, 401).

[15]Boltzmann (1870). Note that this article has not been included in Boltzmann's *Collected Papers*.

[16](Boltzmann, 1870, 470).

> For example, if a gas at constant temperature is divided into two parts by a partition with a small hole, it would be possible to place a device before the hole which lets the faster molecules enter in a part, the slower in the other and separates the gas into a warmer and colder part, which would contradict the second law.

We also know that the issue of the second law was lively debated at the Viennese Physics Department. In a speech in memory of Loschmidt held on July 8, 1895, Boltzmann colorfully related that Loschmidt speculated on the violation of the second law before Maxwell and that his argument was superior to that of the English physicist. In fact, Boltzmann did not like the dependence on a foreign intelligence because:[17]

> [I]f all differences in temperature have been equalized, no intelligent entity could any longer exist. In a cellar at uniform temperature, I said, no intelligence can be present. As it were today, I see before me Stefan [the director of the Physics Department], who had remained silent during our lively quarrel, commenting laconically: 'Now I realize why your experiments with the big glass tubes in the cellar have failed so deplorably.'

Thus discussions on the limitations of the second law were order of the day in Stefan's department. Since these discussions started already before 1872, Boltzmann had to conceive his *H*-theorem with the issue of exceptions in mind. More interestingly, in Boltzmann's passionate recollections of Loschmidt's accomplishments the 1876 argument was repeatedly mentioned, but he never alludes to any effect on his own outlook of the problem. There would have been no better occasion than a memorial speech in honor of his lifelong colleague and friend to admit a change in his conception of the second law and *H*-theorem due to the force of Loschmidt's argument. But nothing similar happened.[18] Loschmidt's argument is presented as an intriguing reflection, but not as a decisive step toward a new understanding of the nature of the second law. In sum, there is no single line written by Boltzmann in which he admits a radical conversion toward probability of the sort supposed by the mechanistic slumber narrative.

1.3 Out of the mechanistic slumber

The thesis of the mechanistic slumber provides us with a narrative able to link together Boltzmann's apparent deterministic phraseology in 1872 with the abundance of probabilistic concepts in 1877. But it proves itself untenable after a closer scrutiny. Thus the problem is to find a new narrative able to reconcile the body of evidence at our disposal. Since the material gathered in the previous section seems to rule out the thesis of a sudden conversion, we are led to conclude that it is Boltzmann's pronouncements in 1872 that need to be reinterpreted. In the next sections I will present an alternative reading which relies on the following

[17](Boltzmann, 1905, 231).

[18]The same is true for a second speech Boltzmann held on November 5, 1895 (Boltzmann, 1905, 240-252).

claim: Boltzmann's 1872 deterministic phraseology does not refer to a physical situation, but to a logical relation. This claim calls for a bit of explanation.

Due to the huge complications and to the necessity of using statistical tools, physicists of mid-1800 customarily described gaseous systems in a suitable conceptual space defined by assumptions of asymptotic nature: they supposed numbers *sufficiently* large of molecules, time periods *sufficiently* long, volumes *sufficiently* extended, and so forth. Practically all arguments involving systems with many degrees of freedom such as gases explicitly or implicitly deployed these assumptions. The reason was that asymptotic assumptions provide a natural way to avoid issues of fluctuations and exceptions implicit in the use of averaging procedures. This tendency to create a fictitious conceptual space in which probability theory can be applied rigorously eventually originated the ensemble approach to kinetic theory. I will argue that Boltzmann's apparently deterministic phrasing in 1872 is mostly due to the fact that he framed his theory within the framework of proper asymptotic conditions. Boltzmann understood the *H*-theorem as a rigorous probabilistic law following from suitable hypotheses. According to this reading, Boltzmann did not overlook the probabilistic meaning of the *H*-theorem, he formulated the law within a conceptual space in which exceptions had been excluded from the beginning.

The upholders of the mechanistic slumber narrative might here raise an objection. They can argue that Loschmidt's 1876 argument provides exceptions to the *H*-theorem that are remarkably different from those usually avoided by asymptotic assumptions. The reversibility argument, as it is often called, does not only show that there are microscopical arrangements of molecules that lead to pathological behaviors of the system, it shows that we can contrive a possible violation of the second law *whatever the microstate of the system*. Thus, after all, Loschmidt's argument did change Boltzmann's view, as testified in the 1877 combinatorial paper.

This ameliorated version of the mechanistic slumber, though, suffers from the same shortcomings as the original one. It is incompatible with historical evidence because, as I will show in section 5, the reversibility argument was not Loschmidt's original discover. It had been already used in 1874 by William Thomson. More importantly, this new version superimposes on the original debate later understandings of the problem of irreversibility. If we look carefully into Loschmidt's 1876 paper and into Boltzmann's 1877 response, we realize that their debate was not about the *H*-theorem at all. Instead, it concerned Loschmidt's idiosyncratic polemic against the use of the distribution function.

Finally, part of the problem with the mechanistic slumber narrative is that it considers the 1872 paper as a self-contained and complete theory. In reality, it was the culmination of a research program that Boltzmann had begun back in 1868 and had developed through three important articles published in 1871. The reconstruction of this route, which I carry out in the sections 2 - 3, will illustrate two important points. First, it will clarify the way in which asymptotic assumptions entered Boltzmann's theory and will contextualize the logical structure of the 1872 paper. Second, it will show that the great results of 1872, the Boltzmann equation and the *H*-theorem, grew out of a continuous interplay between mechanical and probabilistic arguments and were by no means understood as mere mechanical laws.

2 Many ways to the same target

In 1868 Boltzmann published a paper devoted to a thorough study of the equilibrium in gases.[19] This paper has been often regarded by the historians as a sort of exercise in mechanics aiming merely at extending Maxwell's distribution to more realistic cases.[20] By contrast, I will argue that in this paper Boltzmann did more than furthering Maxwell's line of thought: he elaborated a general theorem which led him to prove the necessity of the Maxwell distribution.

2.1 Playing with collisions

The first five sections of the paper, amounting to 31 pages out of 47, are occupied by the study of mechanical collisions between material particles. Following a suggestion of ter Haar, I will call it the *kinetic* approach to equilibrium.[21] The main concept is the distribution function of velocity $f(v)dv$ defined as the fraction of time during which a certain particle has velocity v. If N is the average number of particles in the unit volume (or surface), $Nf(v)dv$ is the number of particles in such volume or surface moving with velocity v.[22] Boltzmann alternated two notions of probability. Preferably he treated the probability for a particle to have velocity v as a fraction of time, but occasionally he defined the same probability as the fraction of particles having that velocity. In both cases, the distribution function has a clear probabilistic meaning.

In addition, all positions and directions of motion in the space are assumed equally probable (homogeneity and isotropy). To make the calculation easier Boltzmann introduced some asymptotic conditions: the number of particles, the volume at disposal, and the energy of the system are infinite. Furthermore, the time over which he calculated the averages was supposed very long.[23]

The kinetic argument is simple. Let us consider an arbitrary collision in which molecules enter with velocities v_1, v_2 and come out with v_1', v_2'. The inverse collision has the latter as initial velocities and the former as final ones. Boltzmann calculated the numbers of direct and inverse collisions occurring in a unit volume, which turned out to be proportional to $f(v_1) \cdot f(v_2)\phi dv_1 dv_2$ and $f(v_1') \cdot f(v_2')\phi' dv_1' dv_2'$ respectively, where ϕ, ϕ' are functions expressing the geometrical details of the collision (differential cross section). To arrive at this result Boltzmann made use of the hypothesis of *Stosszahlansatz* (assumption on the number of collisions, SZA for short) according to which the number of collisions involving molecules of velocities v_1, v_2 depends only on the product of the probabilities (and therefore the distribution functions) of each velocity. This hypothesis had been already successfully deployed both by Clausius and by Maxwell and was a basic assumption of kinetic theory.

[19](Boltzmann, 1868, 1909, I, 49-96).

[20]See for example (Klein, 1973, 61-62) and (Brush, 1983, 62).

[21](ter Haar, 1954, 359).

[22]To be sure, Boltzmann worked with cells in the velocity space defined by the values v and $v + dv$. I will speak of exact velocity for brevity's sake.

[23]The consequence of these assumptions is that one can interpret $f(v)dv$ and $Nf(v)dv$ as *actual* fraction of time and density of particles respectively. This is a clear example of how the asymptotic assumptions can turn the probabilistic phraseology in the deterministic one.

If the number of direct collisions is equal to that of inverse collisions, the distribution of velocity will remain stable over time, therefore a sufficient condition of equilibrium reads:

$$f(v_1) \cdot f(v_2) v_1 v_2 = f(v'_1) \cdot f(v'_2) v'_1 v'_2 \tag{1}$$

Equation (1) represents the kinetic condition of equilibrium in case of no constraint on the total energy of the system. The collisions are also subject to the conservation of energy $v_1^2 + v_2^2 = v_1'^2 + v_2'^2$. These two requirements can be fulfilled at the same time if the distribution function has an exponential form $f(v) \propto e^{-hv^2}$, where h is a constant. If the geometrical details of the collision are taken into account it is easy to arrive at the Maxwell distribution.

2.2 The combinatorial procedure

In the first part of the article Boltzmann repeated this argument for more and more complicated systems, but at some point he stopped and stated that there was no need to examine other examples because, in actuality, all cases could be derived from a more general theorem. This theorem constitutes the subject of the second part of the paper. Boltzmann's construction of the theorem moves through the analysis of some specific problems. For the sake of clarity I will turn the order upside down: I will first present Boltzmann's general theorem and then discuss the single cases.[24]

The theorem is formulated in the final section significantly entitled *Allgemeine Lösung des Problems des Gleichgewichtes der lebendigen Kraft* (General solution to the problem of equilibrium of kinetic energy). Boltzmann considered a system made up of n particles moving in a volume and interacting according to an arbitrary potential function. Furthermore, the total energy of the system is now a fixed quantity. This last requirement modifies substantially the condition (1) because the probability for a particle to have a certain velocity is no longer independent of the velocities of the remaining particles:[25]

> [T]he probability that the velocity of one point lies within given limits and, at the same time, the velocity of another lies within other limits will be by no means the product of the two individual probabilities; rather, the second one will depend on the quantity of velocity of the first point.

The equilibrium condition must be modified accordingly: Boltzmann has to shift his focus from the distribution function of a single velocity to the distribution function expressing the state of the system as a whole. By replacing Boltzmann's inconvenient notation with modern phase space terminology,

[24] The second part of Boltzmann (1868) is hardly mentioned in classical studies like Klein (1973) or Brush (1976) especially because of its obscurity. The combinatorial procedure presented in it has been object of some statistical analyses in Bach (1990), Costantini et al. (1996), and Costantini and Garibaldi (1997), but, to the best of my knowledge, the only historical accounts in which it has been (partly) analyzed are Uffink (2007) and Badino (2009).

[25] (Boltzmann, 1909, I, 80-81).

this means that the description of the system is now provided by the function $f(q_1, \ldots, q_n, p_1, \ldots, p_n) d\mathbf{q} d\mathbf{p}$, which gives the fraction of time spent by the system in the $2n$-dimensional phase space region characterized by the values $(q_i, q_i + dq_i)$ of the generalized coordinates and $(p_i, p_i + dp_i)$ of the conjugate momenta $(i = 1, \ldots n)$.[26] This phase region is transformed under the effect of the equations of motion: after a time interval δt, the system will lie in the region $(q'_i, q'_i + dq'_i; p'_i, p'_i + dp'_i)$ and the function becomes $f(q'_i, p'_i) d\mathbf{q}' d\mathbf{p}'$, after $2\delta t$ it becomes $f(q''_i, p''_i) d\mathbf{q}'' d\mathbf{p}''$ and so on. Since the time intervals are equal, Boltzmann argued, the system spends equal amount of time in each region, therefore $f(q_i, p_i) d\mathbf{q} d\mathbf{p} = f(q'_i, p'_i) d\mathbf{q}' d\mathbf{p}' = f(q''_i, p''_i) d\mathbf{q}'' d\mathbf{p}'' = \ldots$ and so on. Next, Boltzmann calculated how the phase volume is transformed by the evolution of the system. The result is a special case of the Liouville theorem according to which the phase space volume singled out by the particles of the system does not change with time, that is $d\mathbf{q} d\mathbf{p} = d\mathbf{q}' d\mathbf{p}'$. It must be noticed that here Boltzmann did not give a rigorous demonstration of this proposition, but confined himself to a very synthetic — and obscure — calculation.[27] Putting together all these results, Boltzmann arrived at the general condition of stationarity for a system under the constraint of the total energy:

$$f(q_i, p_i) = f(q'_i, p'_i) \tag{2}$$

Now, since this condition holds for the entire trajectory, it entails that the stationary distribution function must be a constant of motion, $f(q_i, p_i) = h$. From this theorem Boltzmann drew important consequences.[28] First, the probability for a system to be in a certain state is proportional to the volume $d\mathbf{q} d\mathbf{p}$ of the state: $f(q_i, p_i) d\mathbf{q} d\mathbf{p} \propto d\mathbf{q} d\mathbf{p}$. In other words, equal volume states are equiprobable. Second, the distribution is a function of the total energy only.[29]

This theorem founds a completely new approach to the calculation of the specific form of the equilibrium distribution. Instead of investigating the average behavior of the system during mechanical collisions, one can divide the space of the allowed states into regions of equal volume and evaluate the probability for a single particle to be into a certain region. This probability, according to Boltzmann, is proportional to the number of ways in which the remaining particles can be arranged into the remaining regions. Because of the energy constraint, the fact that a particle lies in a certain state entails that some of the states are no longer reachable for the rest of the system. As an extreme example, if one particle owns the whole energy of the system, only the states with zero energy are allowed to the other particles. Thus, the number of combinations that the rest of the system can assume is a measure of the probability for a particle to lie

[26] I adopt the notation $d\mathbf{q} = dq_1 \cdots dq_n$, and analogously for the momenta p_i, to indicate synthetically the volume of the phase region.

[27] In a paper published few weeks later, Boltzmann came back to the problem, but only for the case of particles under the effect of a force (**?**Boltzmann, 1909, I, 97-105). It is interesting to note, incidentally, that in Liouville (1838) the main stress is on the application of the theorem to a system of differential equations. Only around 1855 Liouville became fully aware of the possible application of his result to analytical mechanics (cf. Lützen (1990)). Boltzmann does not mention Liouville in his early papers, he however refers to the Liouville theorem in the *Gastheorie* (Boltzmann (1898)).

[28] (Boltzmann, 1909, I, 95).

[29] Boltzmann implicitly assumed that the total energy is the only independent constant of motion. As a matter of fact, the theorem is more general because in 1871 Boltzmann would prove that each arbitrary function of the constants of motion is stationary. On this point see also (**?**, 70-71).

in a certain state. This procedure is fully combinatorial and relies on the fact that the distribution is a function of the total energy only.

Boltzmann developed this procedures for a system of particles moving on a surface (two-dimensional case) and in a volume (three-dimensional case). The first example is very instructive because Boltzmann arrived at the final formula by increasing progressively the complexity of the system. He assumed the energy be divided into p elements of magnitude ϵ and calculated the probability for a molecule to have energy between $i\epsilon$ and $(i+1)\epsilon$. He began with two molecules ($n = 2$) and showed how to construct the combinatorial formula for increasing n. The general expression turned out to be a special case of the Polya distribution.[30] To obtain the Maxwell distribution in two dimensions, Boltzmann calculated the limits $n \to \infty$ and $p \to \infty$. For the three-dimensional case the procedure is completely analogous except that Boltzmann considered infinitesimal energy cells and replaced the combinatorial formula with integrals. Again, the idea is that the probability for a particle to lie in a certain state is given by integrating over all possible states accessible to the remaining particles.[31] The result is a complex expression containing Gamma functions which can be reduced to the Maxwell distribution by taking the limit $n \to \infty$.[32]

The theorem of the *Allegemein Lösung* and the combinatorial argument represent a remarkable step forward that 'fills a gap of all others derivations' because they entail 'not only that for [Maxwell's] distribution of velocity the equilibrium takes place, but also that the [equilibrium] is possible in no other way.'[33] This statement of Boltzmann has not been emphasized enough by the commentators. Here the Austrian physicist claims that, contrary to the kinetic approach, the combinatorial argument yields a proof of the *necessity* of the Maxwell distribution. The necessity relies, obviously, on the fact that this distribution is the most probable among the possible ways to distribute the time spent by a single particle in each state.[34] This result would mark Boltzmann's research program in the subsequent years.

There is a broad point that I think we can derive from this second part of the 1868 paper: Boltzmann did not appear to draw a clear-cut line between mechanistic and probability-based approach to equilibrium. Especially in the *Allgemeine Lösung* he seemed to look for the general conditions under which a mechanical system, such as a gas, can be fruitfully described using combinatorial and probabilistic means. As we will see better in the next few sessions, probability and mechanics are neither alternative nor in competition, but rather complementary tools to clarify the nature of equilibrium. Specifically, in 1871 Boltzmann would try to isolate and clarify the various ingredients of the combinatorial approach in

[30] For the details see (Costantini et al., 1996; Badino, 2009, 83-85).

[31] This part of the argument has been analyzed in (Uffink, 2007, 955-956). Incidentally, Boltzmann's procedure is totally equivalent to the technique used today to obtain the canonical distribution. Normally, one calculates the distribution for a system combined with a heat reservoir by integrating out the states of the reservoir. Boltzmann proceeds in the same way, because if n is large enough, the single particle can be considered as immersed in a heat bath.

[32] (Boltzmann, 1909, I, 88-89).

[33] (Boltzmann, 1909, I, 96).

[34] The adoption of a 'time' definition of probability prevents Boltzmann from comparing possible distributions as he would do in 1877. The Maxwell distribution is however the most probable because any other partition of the time would force the remaining particles into improbable configurations.

order to integrate them with the kinetic one and to reach a kinetic proof of the uniqueness of the equilibrium.

A further important point is that the rigorous derivation of the equilibrium distribution requires asymptotic conditions. Only when the number of particles becomes infinite the combinatorial formula turns into the usual Maxwell distribution. On the last page of the paper Boltzmann discussed the limitations of his procedure. The combinatorial calculation relies on the assumption that all states allowed by the total energy condition will be eventually passed through by the system. This assumption will come to be known as 'ergodic hypothesis'[35] and is essential to the combinatorial procedure. If the actual trajectory of the system is constrained into a fraction of the possible states, the combinatorial procedure does not provide any information on the equilibrium. Boltzmann argued that this occurs when the phase variables are not independent, but are interwoven by some constraint. Specifically, there are two cases in which such a phenomenon can happen. First, the system is strictly periodic and the trajectory closes itself 'without having assumed all values compatible with the principle of energy.'[36] This is a case of stable 'false equilibrium' to which the combinatorial procedure does not apply.[37] Alternatively, the system can present an unstable 'false equilibrium'. This occurs when the particles are arranged in a very peculiar way, which forces the system into a weird, pathological behavior. Boltzmann mentioned the example of gas molecules lined on a straight line. If the line is perfectly straight and the molecules are properly directed, they will continue to oscillate along it without reaching the homogeneous and isotropic distribution they are theoretically capable of. Boltzmann clearly considered these microarrangements as improbable. Although he could not provide a formal proof, he appeared to think that an increasing number of particles would make these arrangements even more unstable and therefore improbable. Here we can see a second important role played by the asymptotic conditions: they are able to rule out theoretical exceptions to the rigorous derivation of the equilibrium distribution.

3 Toward the Boltzmann equation

In 1870 and 1871, Boltzmann went to Heidelberg and Berlin to work with G. Kirchhoff and H. Helmholtz chiefly on problems of electrodynamics.[38] He was however still struggling with the problem of understanding the results achieved in 1868. The year 1871 was particularly productive for Boltzmann. I will especially discuss a trilogy of papers in which he explored the interplay between the kinetic approach and the combinatorial procedure. As we have seen, the general theorem of equilibrium had been developed for a very abstract system of particles. In the first two papers Boltzmann elaborated a more specific kinetic model (polyatomic molecule) and unfolded the dynamic characteristics of the ergodic hypothesis. In the last one, he generalized the collision mechanism. These reflections would lead to Boltzmann's theory of irreversibility.

[35] Maxwell (1879); Ehrenfest and Ehrenfest (1911); Von Plato (1991).

[36] (Boltzmann, 1909, I, 96).

[37] These periodic trajectories had been the subject of Boltzmann's first important paper (Boltzmann, 1866, 1909, I, 9-33).

[38] (Hörz and Laass, 1989; Höflechner, 1994, 20-24).

3.1 The polyatomic molecule and the ergodic hypothesis

The general theorem of equilibrium relied on two pillars: the invariance of the phase volume, which is today called the Liouville theorem, and the ergodic hypothesis. In 1868 Boltzmann had given only a cursory discussion of these two ingredients. At the beginning of 1871 he came back to the topic. In the first paper of the trilogy[39] Boltzmann provided a more rigorous and elegant proof of the Liouville theorem for a general kinetic model, the polyatomic molecule. A polyatomic molecule is a system of non-interacting atoms in free motion within the molecule. In turn the molecules collide so that one can apply the usual kinetic formalism. Boltzmann showed that the invariance of the phase volume holds true both for the free motion and for the molecular collisions. To prove the theorem Boltzmann calculated the change of the phase volume under the effect of the equations of motion. The transformation of the phase variables is given by the Jacobian determinant:

$$d\mathbf{q}d\mathbf{p} = \frac{\partial(q'_1, \dots, p'_n)}{\partial(q_1, \dots, p_n)} \cdot d\mathbf{q}'d\mathbf{p}' \qquad (3)$$

Next he proved that the Jacobian must be 1, so that the phase volume remains the same.[40] The extension to collision is possible by understanding two interacting molecules as a new complex system in free motion ruled by suitable equations that fulfill the Liouville theorem. This procedure leads easily to retrieve the kinetic equilibrium condition (1). Thus, by combining the Liouville theorem with a realistic kinetic model, Boltzmann was able to extend the kinetic formalism to systems with a constrained total energy.

The second paper of the trilogy is dedicated to understanding better the features of ergodicity.[41] In the first part Boltzmann generalized the theorem presented in the *Allgemeine Lösung*. Let us study a system of non-interacting polyatomic molecules. The state of each molecule is described by n generalized variables $(s_1, \dots, s_n)$. The number of molecules in a given state $d\mathbf{s} = ds_1 \dots ds_n$ at an instant t is $dN = f(t, \mathbf{s})d\mathbf{s}$. Under these conditions, and applying the Liouville theorem, Boltzmann showed that the stationary distribution depends only on the constants of motion $\phi_1, \dots, \phi_n$:

$$dN = \frac{f(\phi_1, \dots, \phi_n)}{\frac{\partial(\phi_1, \dots, \phi_n)}{\partial(s_1 \dots s_n)}} d\phi_1 \dots d\phi_n \qquad (4)$$

Boltzmann understood that this result represented a bridge between kinetic theory and the general theory of differential equations. He mentioned an important relation between the Liouville theorem and Jacobi's principle of last multiplier

[39](Boltzmann, 1871c, 1909, I, 237-258).

[40]Though, Boltzmann's proof relies on a restrictive assumption on the equations of motion. He supposed that the time evolution of a set of phase variables, say the generalized coordinates, depends on a function the other variables only, that is the momenta, and vice-versa. This assumption allows Boltzmann to simplify the Jacobian matrix. Modern proofs do not deploy the Jacobian, but use the Hamiltonian (which is a function of both sets of variables) and the equation of continuity.

[41](Boltzmann, 1871b, 1909, I, 259-287).

according to which if one knows $n-1$ integrals of a set of n differential equations, it is possible to construct the integrating factor of the remaining equation.[42]

In the second part of the paper Boltzmann investigated the ergodic hypothesis he had already mentioned in passing in 1868. He noticed that, in mechanics, we usually deal with trajectories whereby if one knows a certain number of variables, the remaining ones may be established *via* equations of motion. But one can conceive cases in which the information deriving from some variables is insufficient to fix the others. To substantiate this thought Boltzmann imagined a point orbiting a centre of force by which it is attracted with a force $(a/r)+(b/r^2)$. The resulting motion is a series of ellipses. If the angle formed by the apsidal lines of two consecutive ellipses is an irrational multiple of π, a precession of the elliptical orbit takes place and the resulting trajectory will tend to fill all the circular region between the circumferences described by the major apsis and the one described by the minor apsis. A second example concerns an oscillatory system with equation ax^2+by^2. Also in this case, if a/b is irrational, the system tends to cover densely all the allowed space.

These simple examples display an important feature of the ergodic (in this case 'quasi-ergodic') motion, namely that the coordinates 'are mutually independent (except that they confine each other within given limits)':[43] knowing one variable is not sufficient to establish the other. In an ergodic system the integrals of motion fix some variables, but the remaining are free to assume whatever value.

This brings Boltzmann to a further remark about his extension of the 1868 general theorem. In fact, it is now simple to distinguish between dependent variables (fixed by the constraints) and independent variables (free to change). Let us assume that $s_1,\dots,s_k$ variables are independent and the remaining $n-k$ are fixed by the integrals $\phi_{k+1},\dots,\phi_n$. By applying the same argument as before, Boltzmann demonstrated that one can express the average time spent by the system in a certain state of the independent variables in terms of the integrals of motion:

$$f(s_1,\dots,s_k)ds_1,\dots,ds_k = \frac{Cds_1,\dots,ds_k}{\frac{\partial(\phi_{k+1}\dots\phi_n)}{\partial(s_{k+1}\dots s_n)}} \tag{5}$$

where C is a constant. Since the independent variables $s_1,\dots,s_k$ may assume all values consistent with the general constraints of the problem, one needs only to know the integrals of motion 'without that be necessary to know something on the way in which $s_1,\dots,s_k$ actually change'.[44] Equation (5) is actually another expression of the 1868 general theorem. As Boltzmann pointed out, the

[42]Jacobi (1844); for a modern perspective see Nucci and Leach (2008). Historically, this relation is remarkable. Boltzmann envisioned — and actually used, for instance in (Boltzmann, 1871d, 1909, I, 200-227) — the connection between Jacobi's principle and the concept of invariant integral (cf. Berrone and Giacomini (2003)) that would be formalized by Poincaré many years later in a work that, ironically, was to lead to the recurrence theorem, the basis of Zermelo's objection to Boltzmann's approach in mid-1890s. Both Liouville and Jacobi had seen the same connection (cf. Lützen (1990)), but it seems that Boltzmann was the first to figure out physical applications of the theorem.

[43](Boltzmann, 1909, I, 270).

[44](Boltzmann, 1909, I, 277).

specific form of the distribution might be now determined by integrating out all allowed states of the independent variables, which is precisely the combinatorial procedure developed in 1868. But what is the physical justification of the ergodic hypothesis? On this issue, Boltzmann had to offer little more than a vague appeal to the disorder of the motion:[45]

> The great irregularity of the thermal motion and the multiplicity of the forces acting on the body from outwards make probable that its atoms [. . .] pass through all the possible positions and speeds consistent with the equation of energy.

The ergodic behavior is related to the internal irregularity of the gas and to the large number of degrees of freedom, but a definite proof of the ergodic hypothesis is still a desideratum.

3.2 Generalizing the collisions

The third paper, the *Analytischer Beweis*, is the most widely studied of the 1871 trilogy, especially because the title allegedly suggests a shift of emphasis to the problem of irreversibility.[46] Here, however, I am not interested in the analytical proof of the second law, but rather in the collision theory presented in the first part of the article. This collision theory presents some novelties directly related to the combinatorial procedure and, more importantly, is a crucial ingredient of the Boltzmann equation.

In the 1868 *Studien*, Boltzmann had investigated the collision process between two particles of arbitrary velocities. This investigation had led to the equilibrium relation (1), but it had not given any clue about the evolution toward the equilibrium. In the *Analytischer Beweis* Boltzmann changed tack: he zoomed in on a single molecule and investigated its behavior when colliding with other molecules under all possible conditions. This new approach resembles closely the essence of the combinatorial procedure. In that case Boltzmann had sought for the probability for a single particle to lie in a certain state independently of the state of the remaining particles. To do that, he calculated the probability by integrating out all the allowed states of the rest of the system. This calculation yielded a distribution law that, in the asymptotic condition of infinitely many particles, coincided with Maxwell's. In the *Analytischer Beweis* Boltzmann framed a sort of kinetic analogon of that combinatorial procedure: he followed the behavior of a single molecule in collision over time with the remaining molecules in all possible conditions. For his calculation he made also use of the Liouville theorem and of the ergodic hypothesis.

Let us briefly see the new collision theory. The main idea boils down to evaluate two processes. First, one has to compute the number of collisions that take a molecule[47] from an arbitrary initial state to a certain final state $d\mathbf{q}d\mathbf{p}$. To do this it is sufficient to integrate the usual number of collisions that end up in $d\mathbf{q}d\mathbf{p}$ over all possible initial states. Analogously, the number of collisions that take a

[45] (Boltzmann, 1909, I, 284).
[46] (Boltzmann, 1871a, 1909, I, 288-308).
[47] Boltzmann considered polyatomic molecules of the kind examined in Boltzmann (1871c).

molecule away from $d\mathbf{q}d\mathbf{p}$ is given by the integration of the possible collisions that have $d\mathbf{q}d\mathbf{p}$ as the initial state. The equilibrium state is of course reached when the number of molecules entering the arbitrary state $d\mathbf{q}d\mathbf{p}$ is equal to the number of those getting out of it. Boltzmann showed that this condition yields a Maxwellian form of the distribution function. But this new collision theory does more, because it also gives a characterization of all intermediate steps between an arbitrary state and the equilibrium. In the next section we will see how Boltzmann used this theory to construct the equation named after him. Here I want to point out that also this new collision theory is embedded in the framework of asymptotic conditions. For example Boltzmann stated that there must be 'infinitely many gas molecules' in the system.[48] Even more importantly for what follows, Boltzmann assumed that the time span over which he calculated the collisions was 'very large'. This asymptotic condition on the time will become extremely relevant in the next section to interpret the thorniest passage of the 1872 paper.

4 The Boltzmann Equation and the *H*-theorem

The foregoing sections have established two main points. First, in the wake of the results achieved in 1868, Boltzmann took a pluralistic strategy according to which elements of the successful combinatorial procedure were clarified and integrated with the established kinetic approach. This process led to important ingredients of the Boltzmann equation that we will meet again in this section. Broadly speaking, this process shows that the 1872 theory emerged by the interplay between a kinetic model and a combinatorial procedure which was imbued with probabilistic language. Since, for historical reasons that we will see in a moment, the 1872 paper was completely contiguous to the 1871 trilogy, one has to conclude that Boltzmann could not possibly consider the 1872 theory as an exclusively mechanical theory.

The second point concerns the asymptotic conditions. They are used consistently during the period 1868-1871. They ensure a rigorous derivation of the equilibrium distribution, they make pathological microarrangements improbable, and they represent the framework of the collision theory. Unsurprisingly, they will play a crucial role in our interpretation of Boltzmann's attitude toward the *H*-theorem.

4.1 Putting the mosaic together

We do not know much about the historical background of the famous 1872 paper in which Boltzmann put forward his equation and the *H*-theorem.[49] It is established that the essence of the paper was elaborated during his stay in Berlin. In a famous letter to his mother Katharina on 27 January 1872, Boltzmann wrote about having presented an outline of the paper before the Berliner Physikalische

[48](Boltzmann, 1909, I, 289).

[49](Boltzmann, 1872, 1909, I, 316-402). The title of the article, "Further studies on the equilibrium between the gas molecules", characterizes it as a direct continuation of the 1868 paper.

Gesellschaft. However, the talk attracted no attention, except for some remarks of Helmholtz:[50]

> Yesterday I spoke at the Berlin Physical Society. You can imagine how hard I tried to do my best not to put our homeland in a bad light. Thus, in the previous days, my head was full of integrals. [. . .] Incidentally there was no need for such an effort, because most of the listeners would have not understood my talk anyway. However, Helmholtz was also present and an interesting discussion developed between the two of us.

From another letter to Josef Stefan on 2 February 1872 we are also informed that Boltzmann was bothered about loosing the priority, after all an understandable worry for an ambitious young scholar. For this reason he had worked out the paper in a hurry: he intended to publish the talk in the *Annalen der Physik* and a lengthier, elaborated version in the *Wiener Berichte*.[51] On the insistence of Stefan, Boltzmann eventually resolved to publish the paper only in the *Wiener Berichte* and sent off to *Annalen* an account of his studies on electrodynamics.

These scant pieces of information establish two points relevant for our discussion. First, the 1872 paper was elaborated together with the cluster of reflections that found place in the 1871 series and it should be read together with its 1871 predecessors. Secondly, for contingent reasons — worries about the career, concomitant work on other subjects — the paper was written under pressure. This circumstance should encourage a contextualization of the work and discourage too literal a reading.

Let us now see how the different threads elaborated in 1871 are interwoven in the 1872 paper. In the introduction Boltzmann laid down the essential message of the work. The aim is to find the necessary and sufficient conditions for the Maxwell distribution. Accordingly, the argument has a strong emphasis on the inferential structure: Boltzmann wants to investigate the logical relation between some assumptions and their consequences. Moreover the argument he is about to present is essentially probabilistic in nature and this fact deserves a specific discussion:[52]

> [T]he problems of the mechanical theory of heat are also problems of probability theory. It would, however, be erroneous to believe that the mechanical theory of heat is therefore afflicted with some uncertainty because the principles of probability theory are used. One must not confuse an incompletely known law, whose validity is therefore in doubt, with a completely known law of the calculus of probabilities; the latter, like the result of any other calculus, is a necessary consequence of definite premises, and it is confirmed, insofar as these are correct, by experiment, provided sufficiently many observations have been made, which is always the case in the mechanical theory of heat because of the enormous number of molecules involved.

[50](Höflechner, 1994, II, 9), (Cercignani, 1998, 10).
[51](Höflechner, 1994, II, 10-11).
[52](Boltzmann, 1909, I, 316-317).

This passage shows very explicitly the tight relation between theory of heat and probability and the inferential structure of the argument. Boltzmann saw a probabilistic theory as a perfectly exact theory in the sense that *the consequences follow logically from the premises*, therefore the conclusions that one can draw from the initial assumptions are not undermined by the fact that one is using probabilistic arguments. Another point mentioned in the quotation concerns the physical plausibility of these assumptions. The alleged uncertainty that might plague a probability-based theory of heat lies obviously in the presence of exceptions. The fact that something is merely 'probable' means that there are exceptions to its occurrence. After all, it was well known that, in statistical reasoning, pathological microscopic arrangements exist that can possibly lead to odd behaviors. Boltzmann himself had discussed examples of these pathological arrangements at the end of Boltzmann (1868) (section 2.2). However Boltzmann here explicitly excludes the exceptions on the ground of the asymptotically large number of molecules. Gases are constituted of a huge number of molecules, we can take volumes and observation time as large as we please. These conditions, as far as all experimental outcomes are concerned, ensure that exceptions will not occur, hence we can consider the theory as strictly valid. Cast in the framework of asymptotic conditions, Boltzmann's argument becomes a purely logical and mathematical argument. The asymptotic conditions are supposed to dispose of these arrangements or, at least, to confine them to the realm of purely theoretical possibilities. Thus, the previous quotation clarifies that when Boltzmann ascribes 'necessity' to his conclusions he means *the internal necessity of his inferential argument*, namely the logical necessity and the physical necessity insofar the asymptotic conditions are fulfilled.

As an illustration of how Boltzmann constructed his 1872 theory we can take the case of monoatomic gas. He insisted that there are only two general assumptions. First, all directions of motion are equally probable (isotropy), second, the molecules are uniformly distributed over all possible positions (homogeneity). Boltzmann pointed out that, even though at the beginning such conditions may not hold true, the effect of the collisions is precisely to equalize directions and positions. This means that the equiprobability should be understood as an equal tendency of the molecules to assume directions and positions as a consequence of the collisions. This point is important because in the second part of the paper Boltzmann planned to deal with transport phenomena and the equal tendency based on collisions allowed him to introduce a non-homogeneous distribution function without major changes in the mechanism of equilibration. Hence, the inferential structure of Boltzmann's argument is the following: he showed that *the equilibration of velocity follows from the assumed equilibration of positions and directions of motion*.

The derivation of the Boltzmann equation is well known. Let us focus on the function $f(\mathbf{v}_1, t)d\mathbf{v}_1$ which gives the number of molecules that have velocity $\mathbf{v}_1$ at the time t. Boltzmann evaluated the time variation of $f(\mathbf{v}_1, t)d\mathbf{v}_1$ by applying a version of the collision mechanism worked out in the *Analytischer Beweis*. The change of the distribution depends on the balancing between the molecules gaining velocity $\mathbf{v}_1$ and those loosing it as an effect of collisions. Therefore Boltzmann calculated the total number of collisions during the infinitesimal time interval τ that have $\mathbf{v}_1$ as starting velocity (that is, molecules that leave the corresponding velocity cell) and compared it with the number of collisions that have $\mathbf{v}_1$ as final velocity (that is, molecules that enter the corresponding velocity cell). As in the *Analytischer Beweis*, this calculation requires to integrate out the other possible velocities occurring in the collision. The sought-for numbers are, respectively:

$$\int dn_1 = \tau d\mathbf{v}_1 \int_0^{\infty} \int_0^{\mathbf{v}_1+\mathbf{v}_2} d\mathbf{v}_2 d\mathbf{v}_1' \psi(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1') f(\mathbf{v}_1, t) f(\mathbf{v}_2, t) \tag{6}$$

$$\int dn_2 = \tau d\mathbf{v}_1 \int_0^{\infty} \int_0^{\mathbf{v}_1+\mathbf{v}_2} d\mathbf{v}_2 d\mathbf{v}_1' \psi(\mathbf{v}_1', \mathbf{v}_2', \mathbf{v}_2) f(\mathbf{v}_1', t) f(\mathbf{v}_2', t) \tag{7}$$

where $\mathbf{v}_2$ is the velocity of the second colliding molecule, $\mathbf{v}_1', \mathbf{v}_2'$ are the velocities of the molecules after the collision and ψ is the differential cross-section.[53] By series expanding $f(\mathbf{v}_1, t)$ at the point $\mathbf{v}_1$ of the velocity space and neglecting higher order differentials, it can be shown that:

$$\frac{\partial f(\mathbf{v}_1, t)}{\partial t} = \int dn_2 - \int dn_1 \tag{8}$$

therefore a time equation for $f(\mathbf{v}_1, t)$ may be obtained by directly comparing the two numbers of collisions. Here Boltzmann deployed the results arrived at in Boltzmann (1871b) and Boltzmann (1871c) and noticed that the function ψ must fulfill the Liouville theorem. This means that ψ remains the same for both types of collisions. The Liouville theorem allows a remarkable simplification of the integrals. Now the time variation of the distribution function can be written as a unique integro-differential equation of balancing between the distributions of the velocities involved in the collision:

$$\begin{aligned} \frac{\partial f(\mathbf{v}_1, t)}{\partial t} = \int_0^{\infty} \int_0^{\mathbf{v}_1+\mathbf{v}_2} & d\mathbf{v}_2 d\mathbf{v}_1' \psi(\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_1') [f(\mathbf{v}_1', t) f(\mathbf{v}_2', t) + \\ - & \ f(\mathbf{v}_1, t) f(\mathbf{v}_2, t)] \end{aligned}$$

This is the famous Boltzmann equation. It is easy to prove that the Maxwell distribution fulfills the equilibrium requirement $\partial f(\mathbf{v}_1, t)/\partial t = 0$, therefore the initial conditions assumed by Boltzmann are sufficient. To prove that they are also necessary, Boltzmann must show that the Maxwell distribution is the only equilibrium distribution. To accomplish that, he stated the *H*-theorem. He introduced a special function of f, namely $H(f, t) = \int d\mathbf{v} f(\mathbf{v}, t) \log f(\mathbf{v}, t)$ and showed that its time derivative must decrease monotonically to a minimum value, corresponding to the Maxwell distribution.[54] I will say more on the H function in the next section, but now I would like to discuss Boltzmann's comments on this outstanding result.

[53]The second colliding molecule can assume whatever value, therefore its integration ranges from zero to infinity. On the contrary, the exit velocities are constrained by the condition on the total energy. In effect, one of these velocities, in this case $\mathbf{v}_2'$, is uniquely determined by the others.

[54]Famously in 1872 Boltzmann proposed the letter E whereas the letter H we used today was first introduced by Burbury. Moreover, in Boltzmann's paper the function reads $H =$

After pointing out that '[H] must necessarily decrease', Boltzmann stated:[55]

> It has been rigorously proved that, whatever may be the initial distribution of kinetic energy, in the course of a very long time, it must always necessarily approach the one found by Maxwell.

I have highlighted this sentence because most of the strength of the mechanistic slumber narrative relies on it. Here Boltzmann does not mention any exception to his theorem. On the contrary he emphasizes the necessity of the conclusion. Is thus Boltzmann denying the *existence* of exceptions? The foregoing analysis suggests that this is not the case. In reality, the issue of physical exceptions has no room whatever here: Boltzmann is talking about the logical necessity of the relation between the initial assumptions and the conclusions of the theorem. In fact, he states that the theorem holds *for any arbitrary distribution of velocity, not for any arbitrary microscopic arrangement*. The hasty inference that, if it holds for each distribution, then it must hold for each microarrangement (because each distribution corresponds to many arrangements, included the pathological ones) misses the important point that the problem of the pathological arrangements has already been left out of the inferential argument: Boltzmann focuses only on the transformation of the distribution and does not mention exceptions because they are excluded from the start. Put in other words: Boltzmann knew that exceptions to the second law are possible. The evidence displayed in section 1.2 demonstrates this point beyond doubt, I think. For this reason, he formulated his H-theorem under conditions apt to immunize it from the exceptional cases.

Now, one of the consequences of the mechanistic slumber narrative, on the basis of the previous quotation, is that Boltzmann realized a clear distinction between distribution and microarrangements only around 1877, thus his statements about distribution in 1872 should be read as statements about microarrangements. However the previous sections have presented several reasons to dispute this reading. For one, the combinatorial procedure relies precisely on picturing the distribution as a combination of microscopic arrangements. Further, Boltzmann's explicit mention of exceptions as early as 1868 suggests that he knew well that some of them might be exceptional. So much so that he repeated this point in 1872 when he turned to the discrete case. After the derivation we have examined before, Boltzmann analyzed the system under the assumption that the molecules posses only discrete amounts of energy. In this case he pointed out that the distribution tends to reach the equilibrium after a very long time 'with the exception of very special cases', namely pathological microarrangements.[56] In addition, we have seen that Boltzmann's way to dispose of the exceptions was to cast the argument in the framework of asymptotic conditions. The use of these conditions in 1872 signalizes Boltzmann's awareness that his reasoning does not concern the transformation of individual microscopic arrangements, but rather the transformation of the distribution function as a whole.

$\int d\mathbf{v}f(\mathbf{v},t)[\log f(\mathbf{v},t) - 1]$ because the term -1 simplifies a bit the calculation. The term, however, is not strictly necessary because the additional integral $\int \partial f(\mathbf{v},t)/\partial t$ coming out from the time derivative of H can be eliminated by appealing to the conservation of the total number of molecules. Boltzmann mentioned this point in a footnote.

[55] (Boltzmann, 1909, I, 345).

[56] (Boltzmann, 1909, I, 357-358).

Finally, Boltzmann himself cautioned the reader against too a literal understanding of the theorem for the monoatomic gas. He stressed that, in this case, 'the procedure used [...] is of course nothing more than a mathematical artifice.' A physical meaning of this argument, Boltzmann stated, can be gained by applying it to the case of polyatomic molecule, his favored kinetic model.

One further important clue for my reading comes from the apparently mysterious proviso that the monotonic decrease of H takes place 'in the course of a very long time.' Why is Boltzmann stating this condition? Nothing in his theory justifies such a scale evaluation: Boltzmann has given no reason to conclude that a very long time is necessary to reach the equilibrium. Even more strikingly, if taken as a physical pronouncement, this condition is wrong! It is well known that the collision term of the generalized Boltzmann equation changes very rapidly and the system reaches the equilibrium in a fraction of second.[57] The conclusion is that we should not read Boltzmann's remark on the long time as referring to a physical situation at all. Instead, it is another example of the same asymptotic condition we have already met in the *Analytischer Beweis*. As in that case, Boltzmann is assuming an asymptotically long time to transform an uncertain statement in a purely logical proposition of probability calculus.

In the second and third part of the paper Boltzmann made extensive use of the analytical machinery developed in 1871. To treat the transport phenomena, for instance, he exploited concepts introduced for the polyatomic molecule. He introduced the generalized distribution function $f(\mathbf{q}, \mathbf{v}, t)d\mathbf{q}d\mathbf{v}$ depending on the velocity as well as on the position because of the local differences in density. Then he obtained a generalized Boltzmann equation by combining the free evolution of the local volumes (streaming term) with the collision term and applied the analytic arsenal developed for the polyatomic molecule to the problem of transport phenomena. The same arsenal came explicitly to the fore in the third part where Boltzmann generalized the H-theorem to the polyatomic molecule. However also in these cases, he showed that the argument retains its inferential structure.

4.2 The origin of the H-function

The emergence of the H-theorem is a long-standing mystery. As it stands in 1872, the theorem consists merely in stating the suitable H-function and in using the Boltzmann equation to show that it has the desired properties. In other words, given the correct H-function the theorem becomes an analytical statement whose validity is constrained by the conditions of validity of the Boltzmann equation. It is clear that Boltzmann had no argument to derive that function on the ground of the dynamics of the system. Therefore, the question: how did he find out the correct H-function? The usual answer resorts to a pure flash of genius. Stephen Brush has ventured the hypothesis that the H-function was 'a brilliant inspiration' and suggested that it was 'probably the result of educated guesses on his previous work with entropy formulae, combined with some trial-and-error work.'[58] Most of other historians have simply restrained from advancing hypotheses and have taken the H-function as a miraculous gift of Boltzmann's ingenuity.

[57] See for example Cohen (1962).
[58] (Brush, 1976, 600).

The mechanistic slumber narrative is a serious obtacle to any plausible guess on the birth of the *H*-function. Indeed, it pictures Boltzmann's pre-1877 work as purely mechanical and, since there is no clear mechanical way to arrive at the function, we are only left with the option of the flash of genius. But as soon as we release this constraint and concede the possibility that Boltzmann attained the *H*-function through statistical arguments, new possibilities become available. I will put forward here a conjecture on the genesis of the *H*-function that draws upon Boltzmann's wholesale use of probabilistic arguments before 1877. Admittedly, it has no direct support, but it borrows plausibility from my general argument. The conjecture runs as follows.

We have seen that Boltzmann had the solution to the uniqueness of the Maxwell distribution as early as 1868. He knew that such uniqueness followed from the fact that the Maxwell distribution is the most probable one, to wit compatible with the largest number of microscopic configurations. If we look at the way in which Boltzmann entwined mechanical and probabilistic arguments, it becomes reasonable that he had elaborated the *H*-function already around 1870. Looking for a solution to the problem of uniqueness, Boltzmann might have resorted again to combinatorial arguments as a heuristic tool. In particular, he might have asked for a function of the distribution which becomes an extreme (a maximum or a minimum) for the Maxwell distribution only. This problem is very difficult in the kinetic approach, but rather accessible in the combinatorial one: Boltzmann could have solved this problem by a procedure very similar to that used in 1877, that is by calculating the state probability and by maximizing it. He could have realized very soon that the state probability can be written:

$$W = \frac{N!}{\prod n_i!} \tag{9}$$

where N is the total number of molecules and n_i is the number of molecules in the i-th energy cell. If we now pass to the logarithm, it turns out that $\log W = \log N! - \sum \log n!$. Using the Stirling approximation that transforms $n_i!$ into $n_i^{n_i}$, it is immediate to conclude that $\log W$ is a maximum when the function $\sum n_i \log n_i$ is a minimum (the term $N \log N$ is a constant). But this function is precisely the *H*-function.

It is important to notice that this procedure does not require conceptual resources unknown to Boltzmann in early 1870s. Moreover, this argument is not essentially different from the one put forward in 1868, which also relied on counting the number of microscopic configurations compatible with some state.[59] If we admit, contrary to the mechanistic slumber narrative, that Boltzmann took seriously the combinatorial procedure as a heuristic tool, then the conjecture sketched above becomes plausible. Furthermore, we do not need to assume that Boltzmann had a full-fledged version of the 1877 theory already at his disposal before 1872, but only the fundamental idea. If this is the case, Boltzmann might have realized, *via* combinatorial reasoning, that the minimum of the function $H = \int d\mathbf{v} f(\mathbf{v}, t) \log f(\mathbf{v}, t)$ corresponds to the maximum of probability and therefore to the Maxwell distribution.

[59] Indeed, in an interesting paper published in 1910 Wilhelm Lenz showed the amazingly close connection between the combinatorial arguments of 1868 and 1877 (Lenz (1910)). For a recent comparison see (Badino, 2009, 85-87).

This conjecture allows us to dispose of some puzzles. First, as already reminded, there is no account for the genesis of the *H*-theorem apart from those based on the stroke of genius or trial-and-error. Secondly, in the 1877 paper Boltzmann used equation (9) to define the probability of state, but, amazingly, there is no comment whatever on the identity between the denominator of the state probability and the *H*-function. It is rather baffling that Boltzmann did not comment on the formal similarity, even more so if he had discovered it for the first time in 1877 through a combinatorial way. Instead, according to my conjecture he had no reason to comment because the *H*-function had originally been derived along the same line.

Finally, another support for this conjecture comes from the historical analysis of Boltzmann's work in 1871 and his pluralistic strategy. In the foregoing sections we have seen that time and again Boltzmann fell back to the combinatorial procedure to improve the kinetic approach. Combinatorially, the proof of uniqueness was rather straightforward, it is therefore reasonable that Boltzmann looked at a combinatorial path to solve the problem and successively tried to integrate this path in the kinetic approach. This conjecture has thus the advantage of making the genesis of the *H*-function consistent with Boltzmann's overall strategy.

5 Of drawing and counting

According to the mechanistic slumber narrative Boltzmann changed radically his views after Loschmidt's 1876 reversibility argument. To be sure, 1876 was not likely the first time Boltzmann heard of the reversibility argument, though. For it had made its appearance already in an interesting paper that William Thomson published in *Nature* on April 9, 1874.[60] In this paper Thomson discussed minutely the argument of Maxwell's demon and its meaning for the second law.[61] At a certain point he even presented an early version of the reversibility argument:[62]

> Suppose now the temperature to have become thus very approximately equalized at a certain time from the beginning, and let the motion of every particle become instantaneously reversed. Each molecule will retrace its former path, and at the end of a second interval of time, equal to the former, every molecule will be in the same position, and moving with the same velocity, as at the beginning; so that the given initial unequal distribution of temperature will again be found, with only the difference that each particle is moving in the direction reverse to that of its initial motion. This difference will not prevent an instantaneous subsequent commencement of equalization, which, with entirely different paths for the individual molecules, will go on in the average according to the same law as that which took place immediately after the system was first left to itself.

[60]Thomson (1874).

[61]He hypothesized an 'army' of demons to select suitable molecules and to change artificially the distribution of temperature in a system.

[62](Thomson, 1874, 442).

This passage hints at many intriguing points. First, Thomson alludes to the fact that, after having reached the initial state, the system will go on to a new uniform distribution of temperature albeit through a different path. This means that reversibility undermines equilibrium only locally. Boltzmann would claim the same some years later.[63] Second, Thomson somewhat parallels the reversibility argument with the demon argument in considering both as external interventions and 'very special' cases. The behavior of the system 'left to itself' is entirely different and Thomson claimed that if the number of molecules approaches infinity, the physical plausibility of the argument become negligible: 'the greater the number of molecules, the shorter will be the time during which the disequalizing will continue; and it is only when we regard the number of molecules as practically infinite that we can regard spontaneous disequalization as practically impossible.' Thus Thomson admitted that the demon argument and the reversibility argument indeed prove the theoretical possibility of violation of the second law. To be sure, he even evaluated — qualitatively and quantitatively in a very particular case — the probability of such violation. In general, however, he relied on the asymptotic conditions to discard these conclusions from the realm of the physical possibilities. When molecules grow to infinity, such a coherent behavior becomes implausible.

Before moving on to Loschmidt's use of the reversibility argument, I would like to discuss more thoroughly the objection I mentioned in section 1.3. One might concede that Boltzmann was aware of a certain statistical meaning of the second law and that his deterministic language in 1872 only referred to the inferential structure of the theory, but these concessions notwithstanding, the reversibility argument raises genuine problems for the *H*-theorem. In fact, it provides a powerful algorithm to construct counterexamples to the *H*-theorem even when the two fundamental assumptions of isotropy and homogeneity (and the asymptotic conditions as well) hold true. So, an inferential reading of the *H*-theorem would not save Boltzmann from facing the limitations of his theory. More importantly, the reversibility argument forces us to reconsider the application of the SZA, the assumption at the root of the Boltzmann equation. Brown, Myrvold, and Uffink have efficaciously summarized the challenges implicit in the reversibility argument:[64]

> What the Loschdmidt objection does is to demonstrate that Boltzmann's use of the *H*-theorem is seriously *incomplete*. First, there is no reason given as to why the SZA holds for pre-collision velocities rather than post-collisions ones. But secondly, and more to the point, so far there is no categorical reason to think that it could not be a contingent fact (unexplained for sure) that the SZA in its standard form holds at all times.

Thus, there may be a bottom line of truth in the mechanistic slumber, after all. Loschmidt's argument convinced Boltzmann that his interpretation of the *H*-theorem was insufficient and that the statistical illness affected the SZA as well. The objection keeps alive the claim that in 1877 Boltzmann's interpretation of the

[63](Boltzmann, 1898, sect. 6).
[64](Brown et al., 2009, 181).

H-theorem underwent a drastic conceptual modification because of Loschmidt's criticism, although it fine-tunes the extent of such modification.

However I do not think that this revised version of the mechanistic slumber narrative is tenable either. It hinges on a relation between the reversibility argument and the *H*-theorem that is largely an outcome of later discussions. On the contrary, *Loschmidt's original use of the reversibility argument is not directed against the H-theorem and, unsurprisingly, Boltzmann's response to it does not concern the theorem either*. To understand this point we must look carefully into Loschmidt's papers.

The main target of the series of four papers[65] published between 1876 and 1877 is a prime consequence of the Maxwell distribution: the equipartition theorem. Loschmidt found extremely difficult to believe that, in real physical systems, the average energy of each degree of freedom is the same. In particular, he did not accept the barometric law, the statement that the average energy ascribed to each vertical level is a constant when the system, for instance a column of gas, is subject to a gravitational field. Kinetic theory and the machinery of the distribution function, Loschmidt thought, work well for properly idealized systems, but the introduction of external forces changes everything. In his opinion, the effect of external forces cannot possibly be described by the simple distribution law, but must affect the selection of microscopic states.

In the first paper Loschmidt conceived counterexamples to the barometric law based on peculiar vertical arrangements that are possible for solids but, he conceded, extremely improbable in the case of freely moving molecules. Therefore he translated the same idea into another argument. Let us imagine a vessel in which one single atom is placed at the top and the others are at rest at the bottom. The atom falls and hits the remaining so that, after a while, its potential energy is turned into kinetic energy and distributed among all atoms of the gas.

Let us now suppose to divide the total volume into horizontal layers piled one on the other. The barometric law states that the mean energy of each layer — the total energy of the layer divided by the number of atoms in the layer — must be the same. However, Loschmidt argued, no atom can be at the top of the vessel because such a condition is compatible only with the state in which one atom is at the top and all remaining are at rest at the bottom. Interestingly, Loschmidt wrote that 'it is very probable that, insofar [the number of atoms] is considerably larger than one, an atom will never come at the top.'[66] Thus, Loschmidt shared the contention that if the number of atoms is large the probability of this peculiar arrangement is very low. But he wanted to explore the limits of this contention.

It is here that the argument of reversibility comes to the fore.[67] Given the equilibrium described by the Maxwell distribution, Loschmidt argued that it is possible to conceive a new microscopical state which is still described by the same distribution, but gives rise to a completely different course of events:[68]

> If, after a sufficient time τ is elapsed from the establishment of the

[65] Loschmidt (1876).

[66] (Loschmidt, 1876, 138-139).

[67] Note that the reversibility argument, that now we regard as the essence of Loschmidt's work in 1876, only occupies half a page: it is mentioned in the first paper only.

[68] (Loschmidt, 1876, 139).

> stationary state, we suddenly turn the velocities of all atoms in the opposite directions, then we will find ourselves at the beginning of a state to which the character of stationarity can apparently be ascribed. This would last for a certain time, but then the stationary state would start gradually to deteriorate and after a time τ we would arrive unavoidably again to our initial state.

The gist of the argument is that *the distribution function is compatible with microscopical states that lead to completely different evolutions*. This means that the 'character of stationarity' embodied by the distribution is insufficient. The argument hits the heart of the equipartition theorem and the barometric law because it undermines the reliability of the distribution function as a description of thermodynamic phenomena and, as a consequence, entails that the barometric law is a pure artifact. Note, however, that this use of the reversibility argument does not concern the *H*-theorem.[69] Instead, Loschmidt's emphasis is on the fact that the distribution function, so common in kinetic theory, misses important bits of information, therefore it comes as no surprise that it produces outcomes as strange as the barometric law.

Having clarified the real point made by Loschmidt helps us to understand Boltzmann's reply. In fact, Boltzmann's 1877 combinatorial theory[70] is a very disappointing answer to the reversibility argument as we now understand it. For us, the essence of the problem is that, given a distribution describing the equilibrium state, half of the microstates covered by that distribution are the final states of a *H*-decreasing process and half of them are the starting states of a *H*-increasing process, therefore some modification in the SZA is required to warrant an asymmetry in the time evolution of *H*. However, Boltzmann's 1877 argument only shows that there are overwhelmingly more microstates corresponding to the equilibrium distribution than to any other distribution.[71] It is immediate to realize that this argument can not possibly offer a satisfactory answer to the modern reading of the reversibility challenge: the fact that a randomly selected microstate most probably corresponds to the equilibrium distribution does not exclude the possibility of reversing the evolution of the system. It merely tells us one part of the story, without explaining why the underlying dynamics should be sensitive to probability differences. But if we set aside our modern perspective and bring Loschmidt's original point into play, the situation becomes less puzzling. Boltzmann was responding directly to Loschmidt's statement on the reliability of the distribution function: *analogously to Loschmidt's original argument, Boltzmann's 1877 paper does not concern the H-theorem at all*.

Initially, Boltzmann replied with a paper which was only partially devoted to the issue.[72] There the reversibility argument is labelled as a 'sophism' based on a 'fallacy', which, however, leads to 'the correct understanding of the sec-

[69] To be sure, it barely concerns the second law. In the fourth paper, for instance, Loschmidt pointed out that he was arguing against the odd consequences of the formalism of kinetic theory, not against the second law itself (Loschmidt, 1876, 213).

[70] (Boltzmann, 1877b, 1909, II, 164-223).

[71] A detailed description of the argument can be found in (Klein, 1973, 77-82) and (Uffink, 2007, 975-976). Basically, it boils down to using equation (9) above to calculate the probability of a state in terms of the number of microscopic configurations ('complexions') corresponding to it and to maximizing this probability by the method of the Lagrange multipliers.

[72] (Boltzmann, 1877a, 1909, II, 112-148).

ond law.'[73] Thus it contains a truth disguised as a paradox. The truth is that the distribution function leaves out bits of information about the system. The paradox, Boltzmann argued, consists in drawing the conclusion that the distribution function is therefore an unreliable tool. To unmask the sophistical nature of Loschmidt's conclusion it is necessary to calculate the numerical relation between distribution function and microstates. Through this precise numerical relation it is easy to show that the distribution function is indeed an effective tool and its manipulation tells us something on the evolution of the system: it tells us that the equilibrium state is the most probable one, a conclusion Boltzmann had already arrived at in 1868.

It is hard for our modern eye to see Loschmidt's real point. We automatically consider the reversibility argument as an objection against the *H*-theorem. But taking into account the real object of the dispute the *H*-theorem disappears from the scene. Michel Janssen has first suggested that the 1877 paper should not be regarded as a new formulation of the *H*-theorem, but rather as a rephrasing of the second law in terms already, though confusingly, present in 1872.[74]

My reading goes along a similar direction. Commentators have been impressed by Boltzmann's statement that the reversibility argument provides the correct understanding of the second law and by the fact that the 1877 approach looks very different from the 1872 theory. These circumstances have contributed to the mechanistic slumber narrative. According to my reading, the 1877 theory emerged in the context of the debate with Loschmidt about the significance of the distribution function. Neither the *H*-theorem, nor the reinterpretation in a combinatorial key of the *H*-function are mentioned or commented in 1877. Instead, the theory illustrates an alternative approach to the equilibrium, that combinatorial approach Boltzmann had already developed in the *Allgemeine Lösung* and that was the background of the 1872 theory. There is no intention to break with the previous work: there is, instead, the attempt at clarifying and reformulating elements that Boltzmann had been investigating for almost ten years then.

A second argument in support of the thesis that Boltzmann did not perceived Loschmidt's 1876 paper as a threat against the *H*-theorem is that, as I mentioned in section 1.2, he kept relying on the 1872 theory even after 1877. An example of Boltzmann's own consideration of the 1872 theory is a letter to Lord Rayleigh on December 11, 1892.[75] In the attempt at getting Rayleigh interested in his work, Boltzmann sent to him a copy of the 1872 paper that, in his own words, concerned 'a partial differential equation for the variation of the law of distribution.' As late as 1892, thus, Boltzmann viewed that theory as the apex of his contribution. Why did not Boltzmann use more extensively the combinatorial theory? An explanation is that Boltzmann probably saw in the 'reversed' microstates the same character of artificiality he had denounced in the pathological micro-arrangements as early as 1868. For example, on 14 December 1876 Boltzmann presented a paper that replied to all Loschmidt's weird special cases, which supposedly proved that probability theory is not applicable in presence of external forces. He conceded that arrangements can be figured out in which the state probability depends on the initial states and he referred to 'a sim-

[73](Boltzmann, 1909, II, 119).
[74]Janssen (2002).
[75](Höflechner, 1994, II, 187).

ilar example I have introduced at the conclusion of my paper *Studies über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten*',[76] namely the case of molecules lined on a straight line. In short, Boltzmann knew that, using demons, exotic devices or simply our imagination we can force the system as a whole to behave weirdly. Even though these issues suggested caution in the conclusions of kinetic theory, Boltzmann could not help deeming these purpose-made cases as somewhat different from the 'typical' behavior.

This attitude was very persistent. As late as 1895 in his first reply to E. P. Culverwell, who was rising doubts on the *H*-theorem, Boltzmann wrote:[77]

> It can never be proved from the equations of motion alone, that the minimum function *H* must always decrease. It can only be deduced from the laws of probability, that *if the initial state is not specially arranged for a certain purpose*, but haphazard governs freely, the probability that *H* decreases is always greater than it decreases.

What Boltzmann realized only gradually after 1877 is that his intuitive dismissal of the pathological arrangements, largely based on the use of asymptotic conditions, at some point became a sidestepping of the question. A clarification of the way in which 'haphazard' elements enter the laws of dynamical systems was required and this progressively convinced him that the basic mechanical assumptions, particularly the SZA, must have been reinterpreted. Thus I do not dispute that Boltzmann's position in the 1890s differs from his position in 1872. However, I claim, *contra* the mechanistic slumber narrative, that this evolution was much more complex and took much more time than it is usually pictured.

6 Conclusion

The historiography on Boltzmann has been often afflicted by the temptation of using the modern understanding of the subtle problems of statistical mechanics as a key to read Boltzmann's original theory. This tendency is only natural and, to a certain degree, even recommendable, but it camouflages the pitfall of a 'creative misreading' as Janssen (2002) has emphasized. A second dangerous leaning that can be found in the literature is to look at the 1872 theory as popping fully armed out of Boltzmann's head like Athena out of Zeus's. In this paper I have tried to avoid these drawbacks. Instead of trusting the useful but debatable guide of Ehrenfests's retrospective outlook, I have focused upon the crucial years that preceded irreversibility theory to unfold the conceptual elements whereby the Boltzmann equation and the *H*-theorem were constructed.

What kind of picture comes out of this analysis? From a broader perspective, we should always remember that investigations on complex systems in the second half of nineteenth century made use of asymptotic conditions. In most cases, these conditions were the justification for introducing and manipulating averages and distributions. Under these ideal circumstances many problems related to exceptions became immediately dismissible. Moreover, there were

[76](Boltzmann, 1876, 1909, II, 55-102), quotation on page 71.
[77](Boltzmann, 1895, 1909, III, 535-544), quotation on page 540, italic added.

very good reasons to adopt this attitude. First, statistical mechanics was a relatively young discipline in need to establish its own scientific status. In this phase foundational problems are usually swept under the rug. Secondly, no adequate conceptual tools for dealing with exceptions were at disposal. Set-theoretical or measure-theoretical approaches emerged only at the turn of the century.

These points must be born in mind to justly place Boltzmann's interpretation of his own mathematical results. He worked with averages and knew very well that averages experience fluctuations and their evolution calls for some microscopic disorder. So much so, that he used probabilistic arguments to handle them. Far from falling into a mechanistic slumber, from the very beginning Boltzmann lived in a state of 'statistical insomnia': he could not avoid statistical arguments and probability with the consequence that they are always present at different levels of his theory.

The specific method adopted by Boltzmann constitutes another important part of the picture. From 1868 on he developed a pluralistic strategy in which kinetic and combinatorial procedure were treated on the same footing. On the one hand the kinetic approach was better established in the community, on the other hand the combinatorial approach was more general. More importantly, in the combinatorial approach the issue of the uniqueness of the Maxwell distribution could be immediately solved. Therefore, Boltzmann exploited considerations drawn from the combinatorial procedure and from the general theorem to improve the kinetic technique. At the same time, he was aware of the presence of limitations and exceptions (pathological arrangements, fluctuations) in the combinatorial procedure. Thus, he was aware of analogous limitations in the kinetic approach, but he considered them as the artificial constructions of a playful demon or of a fervent imagination.

Acknowledgments

This paper has been written as part of the Project of the History of Quantum Physics of the Max-Planck-Institut für Wissenschaftsgeschichte and of the Fritz-Haber-Institut of the Max Planck Society. My discussions with Michel Janssen have largely contributed to the final shape of the paper. I am also indebted to two anonymous referees and to my colleagues at the History of Quantum Physics Group Marta Jordi, Shaul Katzir, and Christoph Lehner for their careful reading of the manuscript and for their suggestions.

References

Bach, A. (1990). Boltzmann's probability distribution of 1877. *Archive for History of Exact Sciences 41*, 1–40.

Badino, M. (2009). The odd couple: Boltzmann, Planck and the application of statistics to physics (1900-1913). *Annalen der Physik 18*(2-3), 81–101.

Berrone, L. R. and H. Giacomini (2003). Inverse Jacobi multipliers. *Rendiconti del Circolo Matematico di Palermo 52*, 77–130.

Boltzmann, L. (1866). Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 53*, 195–220.

Boltzmann, L. (1868). Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 58*, 517–560.

Boltzmann, L. (1870). Theorie der Wärme. *Die Forschritte der Physik 26*, 441–504.

Boltzmann, L. (1871a). Analytischer Beweis des zweiten Hauptsatzes der mechanischen Wärmetheorie aus den Sätzen über das Gleichgewicht der lebendigen Kraft. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 63*, 712–732.

Boltzmann, L. (1871b). Einige allgemeine Sätze über Wärmegleichgewicht. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 63*, 679–711.

Boltzmann, L. (1871c). Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 63*, 397–418.

Boltzmann, L. (1871d). Über die Druckkräfte, welche auf Ringe wirksam sind, die in bewegte Flüssigkeit tauchen. *Journal für die reine und angewandte Mathematik 73*, 111–134.

Boltzmann, L. (1872). Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 66*, 275–370.

Boltzmann, L. (1876). Über die Aufstellung und Integration der Gleichungen, welche die Molekularbewegung in Gasen bestimmen. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 74*, 503–552.

Boltzmann, L. (1877a). Bemerkungen über einige Probleme der mechanischen Wärmetheorie. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 75*, 62–100.

Boltzmann, L. (1877b). Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Sätzen über das Wärmegleichgewicht. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 76*, 373–435.

Boltzmann, L. (1884). Über das Arbeitsquantum, welches bei chemischen Verbindungen gewonnen werden kann. *Annalen der Physik 22*, 39–72.

Boltzmann, L. (1895). On certain questions of the theory of gases. *Nature 51*(1322), 413–415.

Boltzmann, L. (1898). *Vorlesungen über Gastheorie*. Leipzig: Barth.

Boltzmann, L. (1905). Zur Erinnerung an Josef Loschmidt. In *Populäre Schriften*, pp. 228–252. Leipzig: Barth.

Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen*. Leipzig: Barth.

Brown, H. R., W. Myrvold, and J. Uffink (2009). Boltzmann's H-theorem, its discontents, and the birth of statistical mechanics. *Studies in History and Philosophy of Modern Physics 40*, 174–191.

Brush, S. G. (1976). *The Kind of Motion We Call Heat*. Amsterdam: North Holland.

Brush, S. G. (1983). *Statistical Physics and the Atomic Theory of Matter*. Princeton: Princeton University Press.

Cercignani, C. (1998). *Ludwig Boltzmann. The Man Who Trusted Atoms*. Oxford: Oxford University Press.

Clausius, R. (1859). On the mean lenght of paths described by separate molecules of gaseous bodies. *Philosophical Magazine 18*, 81–91.

Clausius, R. (1870). On a mechanical theorem applicable to heat. *Philosophical Magazine 40*, 122–127.

Clausius, R. (1871). On the reduction of the second axiom of the mechanical theory of heat to general mechanical principles. *Philosophical Magazine 42*, 162–181.

Cohen, E. G. D. (1962). The Boltzmann equation and its generalization to higher densities. In E. G. D. Cohen (Ed.), *Fundamental Problems in Statistical Mechanics*, pp. 110–156. Amsterdam: North Holland.

Costantini, D. and U. Garibaldi (1997). A probabilistic foundation of elementary particle statistics. Part I. *Studies in History and Philosophy of Modern Physics 28*, 483–506.

Costantini, D., U. Garibaldi, and M. A. Penco (1996). Ludwig Boltzmann alla nascita della meccanica statistica. *Statistica 3*, 279–300.

Darrigol, O. (2001). The historians' disagreements over the meaning of Planck's quantum. *Centaurus 43*, 219–239.

Dias, P. M. C. (1994). "Will someone say exactly what the H-theorem proves?" A Study of Burbury's Condition A and Maxwell's Proposition II. *Archive for History of Exact Sciences 46*(4), 341–366.

Ehrenfest, P. and T. Ehrenfest (1911). *The Conceptual Foundations of the Statistical Approach in Mechanics*. New York: Dover.

Höflechner, W. (1994). *Ludwig Boltzmann. Leben und Briefe*. Graz: Akademische Druck und Verlaganstalt.

Hörz, H. and A. Laass (1989). *Ludwig Boltzmanns Wege nach Berlin. Ein Kapitel österreichisch-deutscher Wissenschafts-beziehungen*. Berlin: Akademie-Verlag.

Jacobi, C. G. (1844). Theoria novi multiplicatoris systemati aequationum differentialium vulgarium applicandi. *Journal für die reine und angewandte Mathematik*, 333–376.

Janssen, M. (2002). Dogs, fleas, and tree trunks: The Ehrenfests marking the territory of Boltzmann's H-theorem. Unpublished manuscript.

Klein, M. J. (1970). *Paul Ehrenfest. Volume 1: The Making of a theoretical Physicist*. Amsterdam: North Holland.

Klein, M. J. (1973). The development of Boltzmann's statistical ideas. *Acta Physica Austriaca, Supplementum 10*, 53–106.

Kuhn, T. (1978). *Black-Body Theory and the Quantum Discontinuity, 1894-1912*. Oxford: Oxford University Press.

Lenz, W. (1910). Zum Maxwellschen Verteilungsgesetz. *Physikalische Zeitschrift 11*, 1175–1177.

Liouville, J. (1838). Note sur la théorie des équations différentielles. *Journal de Mathématiques pures et appliquées 3*, 342–349.

Loschmidt, J. (1869). Der zweite Satz der mechanischen Wärmetheorie. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 59*, 395–418.

Loschmidt, J. (1876). Über den Zustand des Wärmegleichgewichtes eines Systemes von Körpern mit Rücksicht auf die Schwerkraft I-IV. *Sitzungsberichte der Akademie der Wissenschaften zu Wien 73-74*, 128–139; 366–372; 287–298; 209–225.

Lützen, J. (1990). *Joseph Liouville 1809-1882: Master of Pure and Applied Mathematics*. New York: Springer.

Maxwell, J. (1871). *Theory of Heat*. London: Longmans, Green and Co.

Maxwell, J. (1879). On Boltzmann's theorem on the average distribution of energy in a system of material points. *Cambridge Philosophical Society Transactions 12*, 547–574.

Nucci, M. C. and P. G. L. Leach (2008). The Jacobi Last Multiplier and its applications in mechanics. *Physica Scripta 78*(6), 1–6.

Penrose, O. (1979). Foundations of statistical mechanics. *Reports on Progress in Physics 42*, 1937–2006.

ter Haar, D. (1954). *Elements of Statistical Mechanics*. New York: Rinehart and Co.

Thomson, W. (1874, April 9). The kinetic theory of the dissipation of energy. *Nature 9*, 441–444.

Uffink, J. (2007). Compendium of the foundations of classical statistical mechanics. In J. Butterfield and J. Earman (Eds.), *Philosophy of Physics*, Volume 2, pp. 923–1074. Amsterdam: North Holland.

Von Plato, J. (1991). Boltzmann's ergodic hypothesis. *Archive for History of Exact Sciences 42*, 71–89.

Von Plato, J. (1994). *Creating Modern Probability*. Cambridge: Cambridge University Press.