WAS THERE A STATISTICAL TURN ? THE INTERACTIONS BETWEEN MECHANICS AND PROBABILITY IN BOLTZMANN'S THEORY OF NON EQUILIBRIUM (1872-1877)

1. OVERVIEW

Like many fundamental episodes of the history of science, also Boltzmann's route to the statistical version of thermodynamics has progressively crystalized in a standard historiographical reconstruction, a so-called "received view". According to it, the initial purely mechanical standpoint adopted by Boltzmann was proved unteneable by means of cogent objections which obliged him to change his mind and to framed the Second Principle of thermodynamcs in a statistical way. For seek of simplicity henceforth I refer to this view as the Klein thesis", because it was clearly developed for the first time in Martin J. Klein's important paper "The Development of Boltzmann's Statistical Ideas". Furthermore, for seek of completeness, I divide the Klein thesis in five sub-theses:

(K1) Boltzmann's view of the role and the meaning of probability deeply changed from 1868 to 1877.[1]

(K2) This change is due more to external criticisms than to a natural internal evolution.[2]

(K3) The theory framed in 1872 is a purely mechanical one. Probabilistic elements emerge only later and as a means to reconcile the reversibility of mechanical laws with the irreversibility of thermodynamic phenomena.[3]

(K4) In 1872 Boltzmann maintains a deterministic version of *H*-theorem, namely a version according to which the *H*-function monotonically decreases as a consequence of molecular collisions until it reaches a minimum value. Loschmidt's criticisms induced Boltzmann to change his opinion and to claim a statistical version of the same theory: the temporal decrease of the *H*-function is only a very probable event.[4]

(K5) The combinatorial theory of 1877 merely is an answer to Loschmdt's objections.[5]

This reconstruction has more or less generally been accepted by the scholars. Of course, there are nuances of interpretations, but some aspects, particularly theses (K3) e (K4), gained a wide agreement in literature.[6]

---

[1] Klein 1973, 55.
[2] Klein 1973, 55.
[3] Klein 1973, 56.
[4] Klein 1973, 73.
[5] Klein 1973, 77.
[6] Even if in Brush 1983, 59-68 e 268-270 the general position seems rather neutral, in Brush 1976, 615-616 and in Brush 1986, 607 the agreement with the Klein thesis becomes explicit. Cf. also Elkana 1974, 261-265, Krüger 1981, 100-101; 1987, Porter 1986, 211-213, 375, Lindley 2001, 84-85.

Occasionally some criticisms were arisen. For instance, Jan von Plato sometimes argued against (K3) claiming that the theory of 1872 cannot be considered purely (or completely) mechanical, and pointing out that the elements of continuity in the development of Boltzmann's thought are as important as the elements of discontinuity.[7] Nonetheless, in the literature no systematic discussions of the Klein thesis and no detailed historiographic analysis of all its parts can be found. The question whether the general route it pictures actually fits with the historical facts or rather it has "mithologized" this route itself is not yet settled down. This paper is meant to be a contribution towards this direction.

The real conceptual core of the Klein thesis can be reduced to the two following statements:

(1) Mechanical approach and probabilistic approach are incompatible.

(2) Boltzmann's theory trasforms itself from a mechanical to a probabilistic one.

The statement (1) – which is a very general presupposition – is the foundation of the statement (2), because if mechanical approach and probabilistic approach are incompatible, then they cannot be adopted at the same time, but one must follow the other. In other words, the thesis of a deep caesura in Boltzmann's views on thermodynamics relies on the thesis that the probabilistic tools are completely foreign, and unsuitable, to the mechanical point of view. Indeed, if a standpoint according to which the behavior of the physical world is ruled by deterministic mechanical laws is adopted, then the use of probabilistic concepts and tecniques can be justified only appealing to demands of approximation or to the ignorance of a detailed description of the physical systems. As Boltzmann in 1877 and in his late works often and manifestly ascribes a clear objective meaning to the statistical fluctuations, it follows that, at that time, he has given up the dream of a genuine mechanical foundation of thermodynamics.

Of course, this line of reasoning is unexceptionable, but the question is: did Boltzmann really agree with the statements (1) and (2)? In particular, does the statement (1) actually represent Boltzmann's general position about the relationship between mechanics and probability? An answer to these questions necessarly involves an investigation about Boltzmann's understanding of the role of mechanics in founding thermodynamics, a topic that has often been solved in a form of, almost naive, reductionism. And it also involves a deep analysis of the conceptual relationships linking the theory of the equilibrium state developed by Boltzmann in the period 1868-1871 and the theory of the non equilibrium state developed in 1872-1877. As it will be clearer in the following sections, this analysis reveals a state of affairs essentially different from that pictured in the Klein thesis. It shows to us that the development of Boltzmann's thought is chacterized by continuity, that the equilibrium theory relies on foundations which are, at the same time, mechanical and probabilistic and that the non equilibrium theory stems directly as a development of these foundations. From this point of view the combinatorial theory of 1877 represents the emergence of conceptual elements previously in the background rather than a real methodological and philosophical turn.

[7] Cf. von Plato 1987; 1994, 81.

In order to support these claims I will divide my analysis in three different steps. The first step consists in arguing for the following thesis:

(1*) According to Boltzmann there does not exist any incompatibility between mechanical and probabilistic approach. Using the deterministic laws of mechanics does not compel us to interpret the probabilistic arguments merely as a mean to support our lack of information. In other words, Boltzmann regards probabilistic tecniques as compatible with mechanics. In particular, there is no contraddiction between probabilistic tools and the formal tools of the analytic mechanics developed by Lagrange, Hamilton and Jacobi when they are applied to a certain type of physical systems.

The second step concerns the following thesis:

(2*) Boltzmann's theory is characterized by a – not always consistent – mixture of mechanical and probabilistic elements.

Again, the thesis (1*) is a presupposition for the thesis (2*) because the mixture of mechanical and probabilistic elements follows from the possibility to provide a genuine probabilistic description of mechanical systems.

The third step involves a very detailed investigation of the genesis of the combinatorial theory of 1877 in order to prove that it is not the result of a statistical turn, but rather the explication of ideas already – even if implicitly – present in the theory of 1872 and which can be found in the equilibrium theory of 1868 as well. An important role for this goal is played by the analysis of analogies and differences between the combinatorial arguments in 1868 and in 1877.

2. The General Solution to the problem of equilibrium.

The key to understand the relationship between mechanics and probability in Boltzmann's thought is the concept of "diffuse motion". This concept makes its first official appearance in the 1868 article entitled "Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten" dedicated to a broad discussion of the Maxwell distribution of equilibrium state. In this paper Boltzmann generalizes the equilibrium distribution taking into account the two- and three-dimensional cases, the effect of external forces and the constraint of the conservation of the total energy. The finishing touch of this remarkable analytical effort is the *allgemeine Lösung* (General Solution) to the problem of equilibrium, a section of Boltzmann's article which has not received in the literature the attention it deserves.

Boltzmann considers the gas as a system of n material points of mass m_i ($i = 1, \ldots, n$) whose physical coordinates find themselves within the volumes $ds_i = dx_i dy_i dz_i$ of the position space and $d\sigma_i = du_i dv_i dw_i$ of the velocity space. If the total energy is fixed, one coordinate of the velocity – let us suppose the last – is automatically determined, then the distribution function becomes $f(ds_i, d\sigma_{i-1}, d\omega_n)$ where $d\omega_n$ is an elementary surface defined by:

$$d\omega = \frac{c_n}{w_n} du_n dv_n.$$

Boltzmann proves with purely analitical methods that, if an infinitesimal variation transforms the state $ds_i d\sigma_i d\omega_n$ into the state $ds'_i d\sigma'_i d\omega'_n$ then:

$$(1) \qquad ds'_1 ... d\sigma'_{n-1} d\omega'_n = \left(1 + \frac{\delta w_n}{w_n}\right) ds_1 ... d\sigma_{n-1} d\omega_n.$$

Equation (1) is a particular case of the Liouville theorem and concerns the temporal conservation of the phase volume. Accordingly, for k consecutive trasformations Boltzmann obtains:

$$(2) \qquad c_n f(x_1,...,v_n) = c_n^{(k)} f(x_1^{(k)},...,v_n^{(k)}),$$

i.e. the distribution function is a constant as well. The results (1) and (2) are not usual in the kinetic theories of gases in 1860s. Although kinetic theory essentially was a mechanical theory, the formal tools used were mainly those of elementary mechanics. Boltzmann was among the first using the formal machinery and the concepts of analytical mechanics.[8] However, in the General Solution Boltzmann goes beyond the limits of analytical mechanics itself because his fundamental move is the generalization of (1) and (2) – which are theoretically limited to the actual phase trajectory of the system – *to the whole phase space allowed by the constraint on the total energy*.[9] This move attracted Maxwell's criticisms more than ten years later. Although Maxwell basically agreed with Boltzmann's formal argument, he pointed out that the generalization to the whole phase space allowed by the total energy necessarily requires an ergodic hypothesis.[10] For brevity I will call henceforth the generalization of (1) and (2) to the whole hypersurface of constant energy the "Boltzmann-Liouville theorem".

If the Boltzmann-Liouville theorem is the General Solution to the problem of equilibrium, what is Boltzmann's view of the equilibrium state? In order to comprehend this point let us analyze three important consequences of the General Solution. The first consequence is the following:[11]

> The probability that the point m_1 finds itself within the space element ds_1, the point m_2 within ds_2, …, the end-point of the [velocity vector] c_1 of the first point within $d\sigma_1$, the end-point of the [velocity vector] c_2 of the second point within $d\sigma_2$, …, and finally the end-point of the [velocity vecton] c_n of the last point within the surface element $d\omega_n$ is proportional to the product:

$$\frac{1}{c_n} ds_1 ds_2 ... ds_n d\sigma_1 d\sigma_2 ... d\sigma_{n-1} d\omega_n.$$

In other words, the probability that the representative point of the system finds itself within a certain phase region depends on the measure of this region.

---

[8] Likely, the first systematic expositions of these results can be found in Watson 1876 (see expecially 21-24).

[9] Boltzmann 1868, 95.

[10] Maxwell 1879, 714, 722.

[11] Boltzmann 1868, 95.

Accordingly, elementary regions of equal measure are equiprobable.[12] The second consequence is that the Boltzmann-Liouville theorem justifies the combinatory derivation of the equilibrium distribution developed in the second part of the 1868 paper.[13] The third consequence is that, as Boltzmann always understands the distribution function and the probability as sojourn time in a certain state,[14] the General Solution implies that the system will pass through all the physical states compatible with its total energy, i.e. its motion will be "diffuse" in the entire phase space. The last two consequences, the capability of justifying the combinatorial derivation and the concept of diffuse motion, are the most problematic and represent the real core of the General Solution. Thus, I will focus on these two points and try to answer to the following question: how does, according to Boltzmann, the concept of diffuse motion justify the combinatorial analysis and the use of combinatorial arguments? I claim that the answer lies in two fundamental roles played by this concept.

In the first place, it divides the phase space in "elementary events", an indispensable presupposition for the combinatorial analysis. Indeed, Boltzmann understood the equilibrium as a complex phenomenon which stems from the combination of a lot of elementary equiprobable events. This view was common among the advocates of the so-called "social statistics",[15] but it was absolutely not usual in kinetic theory before Boltzmann. Neither Maxwell, nor Clausius made clear attempts to characterize combinatorially the equilibrium state. In any case, if the motion is diffuse, then the position-space configurations in formula (1) immediately become the sought elementary events which the combinatorial derivation of the equilibrium distribution can build on.

In the second place, diffuse motion justifies the combinatorial analysis because under this condition the details of motion, i.e. the precise specification of the micro-state of the system, are irrelevant. In fact, if a system passes through all the states of the allowed phase space, there is only a phase trajectory covering the whole space, so no detailed specification of initial and final states is required in order to describe the trajectory itself. Indeed, its description coincides with the combinatorial analysis of the elementary states forming the space.

This aspect of the concept of diffuse motion is close connected with Boltzmann's previous application of analytical mechanics to the problems of thermodynamics. In 1866, looking for a mechanical analogy of the Second Principle, Boltzmann investigated the possibility to apply the principle of the Least Action to the study of the motion of a system with many degrees of freedom. In that paper Boltzmann was also concerned with the problem of neglecting the details of motion, in particular by means of special periodicity assumptions that allow to erase the initial and final condition from the action integral.[16] These researches are the real premises of the concept of diffuse motion and of the Boltzmann-Liouville theorem. In fact, Boltzmann shows that in a

---

[12] Of course, this conclusion requires that the probability measure is absolutely continous with the Lebesgue measure of the phase space. Boltzmann, however, could not take into account this aspect which will be come clear only in the next century (cf. von Plato 1994, 27-70).

[13] Boltzmann 1868, 95.

[14] Boltzmann 1868, 50, 85; cf. anche Boltzmann 1881, 582.

[15] Cf. Porter 1986.

[16] Boltzmann 1866, 24, 28-29, Biehrhalter 1981, von Plato 1994, 76-77. On the historical development of this research programme see Bierhalter 1992.

infinitely slow mechanical transformation (equivalent to a thermodynamic adiabatic transformation) the action is an invariant. The connection between this result and the Boltzmann-Lioville theorem will become apparent in 1916 when Paul Ehrenfest will show that the geometrical meaning of this result is just the invariance of the elementary phase volume.[17] Thus, there is a clear connection in Boltzmann's thought between the mechanical analysis of systems of material points with many degrees of freedom and the presuppostions of the combinatorial calculus.

I add two further considerations. First, Jan von Plato suggested that the view of probability as a temporal average was developed by Boltzmann in close analogy with the kinetic interpretation of temperature as a temporal average of the energy.[18] Scarce attention was paid to the possible mechanical origin of this concept and in particular to the relationship between the diffuse motion and the temporal average meaning of probability. Actually, if the motion of the system is diffuse, the time of sojourn in a certain phase region is a natural tools to describe the motion itself, so it seems very plausible that Boltzmann framed this meaning of probability from his studies in analytical mechanics applied to complex systems.

Second, it is apparent that the concept of diffuse motion bears on the modern notion of ergodicity. Nonetheless, although Boltzmann has the general concept of ergodicity and he is aware of its functions and meaning, he never frames, at this stage, an explicit ergodic hypothesis. I think that this aspect cannot be neglected. Whereas Maxwell openly states a formal ergodic hypothesis as a fundamental requirement of the equilibrium theory, in Boltzmann this concept always remains in the background as a general assumption on the dynamic systems. For this reason I prefer to separate the "hard" notion of ergodicity from the "soft" notion of diffuse motion. Indeed, Boltzmann understands the diffuse motion as a sort of presupposition which is weaker than a formal physical hypothesis, which is somehow intuitively clear and which does not require any strictly formal expression.[19] Instead of formulating the diffuse motion formally, Boltzmann provides mechanical models able to represent it, as we will see in the next section.

To be sure, the system of concepts used by Boltzmann to bridge the gap between mechanical approach and probabilistic tools (diffuse motion, equiprobability, independence of phase coordinates) has a very peculiar epistemological *status*: these ideas form a sort of conceptual network without a rigid hierarchy. In different physical contexts each of them can assume a foundational role from which the others can be derived. In other words, it seems that Boltzmann's approach to the relationships between mechanics and probability is not axiomatic, rather analogic: conceptual links and formal similarity are more important than a strict division between axioms and consequences.

[17] Ehrenfest 1916, 333-334.
[18] Von Plato 1991; 1994.
[19] On Boltzmann's and Maxwell's ergodic hypothesis see von Plato 1991.

3. MECHANICAL MODELS OF PROBABILISTIC PROCESSES.

The diffuse motion is a fundamental requirement to apply probabilistic and combinatorial tecniques to the study of mechanical systems. Instead of giving a formal definition of this concept, in 1871 Boltzmann proposes a model of a mechanical system whose motion fills the whole phase space: a mechanical model of the diffuse motion. Boltzmann restricts his argument to the bidimensional case and starts with the following remark: if the equations of motion are able to limit the trajectory of the system to a subset of the phase space and one coordinate is known, then, in general, the other coordinate is automatically determined. A clear example is the motion of a material point around a center of force attracting it with a Newtonian potential. In his 1870 article on the virial theorem, Rudolf Clausius called this kind of motion "stationary".[20] The stationary motions include, like Clausius himself said, periodical motions like the planets' ones.

However, Boltzmann adds, we can imagine motions which do not meet this condition. For this kind of motions, knowing a coordinate allows us only to fix a certain set of possible values for the other.[21] In order to make the problem clearer, Boltzmann supposes that the force attracting the material point is $(a/r) + (b/r^2)$. In this case the trajectory is a sequence of ellipses whose form depends on the constants a and b. If the angle formed by the absidal lines of two ellipses is a rational multiple of π, the motion will be strictly periodic, but if this is not the case, the motion will pass, in the long run, through all the points of the circular region included in the circumferences whose radii are the major and the minor axes of the ellipses. In other words, such a motion will be diffuse into the allowed space. This sort of stationary-diffuse motion has two important consequences.

First, for such a motion a concept of probability as sojourn time in an arbitrary phase region can be defined.[22] This consideration substantiates the thesis according to which the concept of probability as sojourn time stems from a mechanical context.

Second, for a diffuse motion the phase coordinates «are mutually independent (only they limit each other within given limits)».[23] The independence of phase coordinates belongs to the conceptual network used by Boltzmann in facing the problem of the relationship between mechanics and probability. It is an essential condition because if the phase coordinates are mutually independent, the behavior of a mechanical system is completely analogous to the behavior of a stochastic model such as drawing from a lottery or throwing dice.

Boltzmann deals with this problem also in his article of 1877 entitled "Bemerkungen über einige Probleme der mechanischen Wärmetheorie" where the first answer to Loschmidt's objection can be found. In the third part of this article[24] he develops a detailed analysis of the possible motions of a material point which ends with the following remarks. If a material point of mass m moves around a center of force O attracting it with force $f(r)$, where r is the distance from O, the trajectory of the point can (a) diverge towards infinite or (b) remain limited

---

[20] Clausius 1870, 123.
[21] Boltzmann 1871b, 269.
[22] Boltzmann 1871b, 270.
[23] Boltzmann 1871b, 270.
[24] Boltzmann 1877a, 122-148.

within a finite region of the space. Motion of type (a) can be divided into two sub-classes: (a1) trajectories which come near *O* only one time and then diverge and (a2) trajectories which orbit many times around *O* before diverging.

The trajectories (b) can be divided into two sub-classes as well: (b1) spiral-like trajectories which come nearer and nearer *O*, or (b2) trajectories characterized by many approaches to, and removal from, the center of force. Trajectories of type (b2) can assume many different forms but have a sort of periodicity because they remain in a intermediate state between diverging and falling on the center of force.[25] However, Boltzmann notes, this does not mean that a trajectory of the type (b2) is necessarly closed, i.e. it passes only through a sub-region of the allowed space:

> Even if it is completely contained in a finite region, the trajectory is not in general a closed one. Only if the angle between two successive absidal lines has a rational ratio with π, the trajectory is closed. If this angle is π, then *r* […] has only one maximum and one minimum like the motion of the planets. If it is π/2, *r* has two equal maxima and two equal minima, like in the infinitesimal oscillation of a conic pendulum. If this angle is π/4, the trajectory has the form of a four-points star, if it is π/6 a six-points star.[26]

Accordingly, if the ratio of the angle of two absidal lines and π is irrational, the trajectory will pass through all the points of the allowed space. This is exactly the result Boltzmann had already obtained in 1871 and it represents the mechanical analogy of a diffuse motion.

To sum up, according to Boltzmann a real contraddiction between mechanical approach and probability does not exist, provided that some general conditions are respected. Generally (and a bit roughly) speaking, his argument can be summarized as follows. The classical approach to the study of a mechanical system requires the knowledge of the equation of motion (or a complete set of integrals of motion) and the specification of the initial state (or the initial and final state of the trajectory). However, in case of a system with many degrees of freedom this method is unsuitable because the system is far too complex. But, to be sure, this method is really indespensable only if the system passes through a sub-region of the allowed phase space because, if this is the case, we need to distinguish between the states *effectively belonging to the history of the system* from the states which are merely *physically possible*. Now, if the system passes through *all the physically possible states*, the classical method provides us with a lot of *irrelevant information* and it can be replaced by a combinatorial analysis of the phase space.

As far as this Boltzmann's position is concerned, I add two remarks. First, it belongs to the train of thought which endeavours to find a reconciliation between deterministic mechanical laws and an objective view of probability. Indeed, Boltzmann's cases are quite analogous to Nicola Oresme's studies on the "ergodicity of rotation".[27] Jan von Plato pointed out that these attempts use a conceptual system, consisting of notions like "independence", "instability of the trajectories", "annihilation of dynamical correlations" and so on, which is similar

---

[25] Boltzmann 1877a, 146-147.
[26] Boltzmann 1877a, 147.
[27] Von Plato, 1994, 279-287.

to Boltzmann's.[28] It is worth noting that, although Boltzmann occasionally uses the concept of trajectory instability,[29] the main role in his approach to the problem is played by the concept of diffuse motion. Scarce attention was paid by the scholars to the function of this concept in Boltzmann's approach and, more generally, in the train of thought above mentioned.

Second, according to Boltzmann, using probability arguments in the study of mechanical systems is not due to our ignorance of the specific conditions, but to the fact that the information provided us from the details of motion is *irrelevant*. Thus, contrary to Maxwell, the probabilistic analysis completely replaces the classical approach as a logically equivalent, and formally more efficient, alternative. The autonomous *status* of the probabilistic arguments will be defended by Boltzmann also in the introduction of his article of 1872.[30]

4. MECHANICS AND PROBABILITY IN NON EQUILIBRIUM THEORY.

The main consequence of the compatibility between mechanical approach and probabilistic arguments is that a mechanical system with many particles can be compared to «as many individuals in the most different conditions of motion»[31] or to a stochastic model pictured in close analogy with the games of chance.[32] This analogy plays an essential role in the theory of 1877, however, the mixture of probabilistic and mechanical elements can be found also in the "mechanical" version of the non equilibrium theory developed by Boltzmann in 1872. In this section I will particularly focus on the analysis of the collision mechanism whose probabilistic character has not yet been discussed in the literature.

The theory of 1872 relies on two assumptions: (1) the equiprobability of the directions of motion and (2) the equiprobability of the positions in space.[33] Boltzmann hypothesizes that each molecule, whatever the kinetic energy, can assume every direction and every position with the same probability. The argument with which Boltzmann justifies such assumptions is peculiar: he says that in a system without external influences a progressive mixture will take place which will cancel any "non uniform" initial condition, e.g. conditions in which fast and slow molecules occupy different regions of space. It is a natural tendency of the system to leave a non uniform initial condition to reach a sort of uniform distribution as far as positions and directions of motion are concerned. Boltzmann adds that, as this process will take place in the long run, we can assume the equiprobability of positions and directions of motion *from the start*.[34]

These two uniformity assumptions remarkably weaken the argomentative structure of the theory, because they deal with a process of mixing (of positions and directions) completely analogous to that which Boltzmann is willing to derive

---

[28] Von Plato 1983; 1987; 1994.
[29] Cf. e.g. Boltzmann 1868, 96.
[30] Boltzmann 1872, 316-317.
[31] Boltzmann 1872, 317.
[32] For example, Boltzmann explitcitly refers to a comparison between the evolution of the gas and the game of dice in Boltzmann 1895, 540.
[33] Boltzmann 1872, 321-322.
[34] Boltzmann 1872, 322; already in 1871 Boltzmann introduced the mixing process in an intuitive way, without a clear formal justification, cf. Boltzmann 1871a, 240.

for kinetic energy. Accordingly, Boltzmann can prove the uniforming of energy, summarized in the equilibrium distribution, *only assuming that a similar uniforming process takes place for positions and directions of motion*. Why Boltzmann, who was well aware of the role played by the uniformity hypotheses,[35] tolerated such a weakening of his conclusions? There are two main reasons.

First, as positions and directions of motions are the collision parameters characterizing the mechanical process of molecular collision, the *equiprobability* of these parameters justifies, from Boltzmann's standpoint, the possibility of representing the collision process as a stochastic model because, as we have seen, under these conditions the mechanical analysis can be replaced by a probabilistic analysis. In a molecular collision a pair of molecules enters in the process with some kinetic energies (enter energies) and exits from the process with some other energies (exit energies). In the long run this process succeeds in mixing the energy distribution towards the uniformity. But Boltzmann makes clear that the mixing effect ascribed to the collisions is derived from the variety of the impact conditions which make *every pair of exit energies equiprobable independently of the pair of enter energies*.[36] No attempts are made to characterized this transformation from a purely mechanical point of view. Instead, the real core of the uniforming effect of the collision process is understood as a sort of game of chance regarding the two pairs of energies, whose details I will show later on.

Second, the two uniformity assumptions imply that the function f expressing the energy distribution is *representative* of the state of the system. Indeed, if the state of the system is non uniform, the value of the distribution function depends on the region where it is computed. On the contrary, if the state is uniform, the distribution is constant in every region of the space (provided that it is large enough to contain many molecules) and then it describes the system as a whole.

In order to substantiate this interpretation, I will shortly analize the stochastic model implicit in Boltzmann's collision mechanism. To compute the temporal variation of the distribution function, Boltzmann fixes an arbitrary energy and reckons the difference between the numbers of molecules acquiring and losing that energy in an infinitesimal time τ. These two numbers are given by the following formulae:

$$\int dn = \tau f(x,t)dx \int_0^\infty \int_0^{x+x'} f(x',t)\psi(x,x',\xi)dx'd\xi \tag{3}$$

$$\int d\nu = \tau dx \int_0^\infty \int_0^{x+x'} f(\xi,t) f(x+x'-\xi,t)\psi(x+x'-\xi)dx'd\xi \tag{4}$$

---

[35] Cf. Boltzmann 1872, 322: «[the equiprobability of the positions] and the equiprobability of the directions of motions at the beginning of time are the two *limiting* assumptions under which the problem will be dealt with » (italics added).

[36] Boltzmann will try to ascribe the uniforming effect to the inner motion of the atoms when he will generalize his theory to the polyatomic molecule. To accomplish this task, he will assume, like in 1871, that the phases of the molecules are mutually incommensurable (cf. Boltzmann 1872, 396).

where x, x' are the enter energies and ξ, ξ' the exit energies in the collision process.

Let fix our attention on the equation (3). The function ψ expresses the mechanical characteristics of the collision, and Boltzmann will prove that it fulfils the Liouville theorem.[37] Furthermore, the rest of the equation shows an apparent similarity with an assumption already used by Maxwell in 1867 and closely connected with the indipendence hypothesis used to describe the behavior of a game of chance.[38] Actually, in the equation (3) the only variables are the energies x' and ξ because x is fixed and ξ' can be derived from the constraint on the total energy. Now, in calculating the number of collision decreasing the number of molecules with energy x, Boltzmann's independence assumption compares the number of molecules with energy x' calculated on the actual distribution f and the number of molecules with energy ξ merely considered as an infinitesimal element in an integration over all the possibile values. Thus, the collision mechanism Boltzmann frames it is equivalent to a stochastic model of drawing from two different urns:

(A1) Drawing of the energy x' from an urn where the distribution of energy is defined by the initial function f. This drawing also fixes the limits for the total energy in the collision process.

(A2) Drawing of the energy ξ from an urn with uniform distribution on all the possible values in the new total energy interval fixed in (A1).

As in this model the enter energies have a probability depending on the initial distribution and the exit energies have a probability depending on the uniform distribution, this model is able, in the long run, to transform the initial distribution in Maxwell's distribution, uniform on the elementary configurations in the phase space. In other words, Boltzmann's collision mechanism works because it combines an independent assumption with drawing from urns, so that the behavior of a mechanical system is considered completely analogous to that of a game of chance. Similar remarks hold for the equation (4), even though in this case the model is a little more complex because it must fulfil some other analytical requirements.

It may be objected that the numbers of collisions (3) and (4) do not represent a stochastic model, but they are the result of a simple marginalization by means of which Boltzmann computes the number of collisions decreasing or increasing the distribution function on a certain energy *without taking into account the exit energies*. Although this remark is formally correct, it can be questioned from a historical point of view. Indeed, some arguments can be presented in order to claim that Boltzmann's aim is not merely to calculate a marginal number of collisions, but to build a particular stochastic model of collision.

First, Boltzmann does not consider the marginalization procedure a suitable technique to deal with the problem of non equilibrium. I will clarify this point in the section 6.

---

[37] Boltzmann 1872, 332.
[38] Cf. Maxwell 1867; see also Ehrenfest 1911, 5-6.

Second, Boltzmann criticizes Maxwell's theory of collisions because it is far too limited. In fact, it does not take into account the possibility of a complete variety of exit energies.[39] Such a criticism and Boltzmann's new – and more complex – theory of collisions are not understandable if the goal is simply to marginalize the exit energies. To this aim Maxwell's view is complete enough.

Third, although the integration over all the possible values may be formally read as a marginalization, Boltzmann's procedure has another important ingredient: it considers the exit energies uniformally distributed. This uniformity is the real core of Boltzmann's collision mechanism, and its justification lies in the General Solution of 1868 where Boltzmann proved that equilibrium requires uniformity of elementary configurations.

So, it seems that in Boltzmann's collision mechanism something deeper is hidden. It cannot be reduced to a standard formal procedure, but it involves a general view about equilibrium and its conditions, the kind of view that Boltzmann developed in 1868. Furthermore, it shows how closely mechanical approach and probabilistic techniques interact: Boltzmann viewed molecular collisions as a mechanical process which can be described by a stochastic model. Already in 1872, Boltzmann uses, even if only implicitly, the stochastic model of the urn which will become the main conceptual tool in the combinatorial theory of 1877.

5. LOSCHMIDT'S OBJECTION.

The non equilibrium theory of 1872 is unsatisfactory and incomplete. It shows that the system tends to reach a uniform distribution of the energies *provided that* it is already uniform as far as positions and directions of motion are concerned. In other words, it is proved that a certain kind of mixing can happen only if another kind of mixing holds from the beginning. A general theory of the mixing process and an explication of this concept are totally absent. Boltzmann was perfectly aware of these lacks and of the role played by the preliminar assumptions of uniformity, as the long discussion about them in 1872 confirms.

This awareness had a key role in his interpretation of Loschmidt's objection. Indeed, from Boltzmann's point of view, the core of the problem arisen by Loschmidt in 1876 did not consist in the possibility of an occasional failure of the H-theorem or of the uniforming process. These possibilities were openly admitted by Boltzmann already in 1868 and the proof of the H-theorem in 1872 was obtained by means of an average.[40] The real problematic aspects of Loschmidt's objection were localized by Boltzmann elsewhere.

To be sure, Loschmidt's argument stresses that the fundamental presupposition of Boltzmann's non equilibrium theory (the mixing of certain parameters of motion) is not more justified than its contrary *from the point of view of the mechanical laws*.[41] The italicized statement is the essential point of

---

[39] Boltzmann 1872, 319.
[40] Cf. Boltzmann 1868, 96 and Boltzmann 1872, 343-345 for the proof of the H-theorem.
[41] Loschmidt 1876, 139.

Loschmidt's thesis. It suggests that no proof of a preference of the mixing process can be compatible with the laws of mechanics.

But we have seen that, according to Boltzmann, the compatibility between probability and mechanics was a point of departure, a presupposition beyond all dispute. For this reason he simply neglects this aspect of Loschmidt's argument and focusses on the first part: to provide a general theory of the mixing process applicable to every molecular parameter and able to prove that the mixing process (from non-uniform to uniform distribution) is privileged in comparison to the non-mixing process (from uniform to non-uniform distribution). Boltzmann's interpretation of Loschmidt's argument is the result of an intense dialogue and of a personal exchange of ideas, but its general lines are understandable from Boltzmann's first answer in 1877. In this article, which I have already mentioned above, he states that Loschmidt's argument points against the general possibility of a transformation from a non-uniform to a uniform distribution: «the sophism consists in the fact that without taking into account initial conditions, it is not possible to prove that [the molecules] mix uniformely during the time».[42]

Boltzmann's answer to Loschmidt will be developed in detail in the combinatorial theory, but in this first article some interesting elements can be already found. In particular, Boltzmann's first remarks on the problem are divided into two steps corresponding to two consequences of Loschmidt's objection.

First, Boltzmann notes that:[43]

> Loschmidt's law simply allows us to know of initial conditions that, after a certain time t_1, lead to a state distribution which is highly non-uniform.[44]

This possibility was foreseen already in 1868 and it does not represent a real difficult for Boltzmann. Loschmidt's special initial conditions do not fulfil the requirement of diffusion because the temporal reversal of the molecular trajectories leads the molecules to certain collisions, while other (physically) possibile collisions are ruled out *in principle*. In other words, in Loschmidt's objection the initial state shows, to a certain extent, a sort of *syncronization* and this state cannot origin a diffuse motion.[45] Although it is a genuine mechanical state, it is an exception, a case explicitly left out from the presuppositions of the theory. Another important point is that Boltzmann understands Loschmidt's objection as a particular case of the problem he already faced in 1866 and 1868, namely to study the behavior of a system neglecting the details of motion. Such a behavior has to depend on the general conditions only (e.g. the total energy), but not on the initial state of the system. This point clearly emerges in a long paper of 1876 entitled "Über die Austellung und Integration der Gleichungen, welche die

---

[42] Boltzmann 1877a, 119.

[43] Boltzmann 1877a, 120,

[44] In this article the terminology is not precise: only in Boltzmann 1877b we have a clear linguistic distinction between "state distribution" and "complexion". In Boltzmann 1877a the two terms are often confused.

[45] This argument will clearly emerge in 1890s when Boltzmann will frame the notion of "molecular chaos" as a condition of the diffuse motion. See, e.g. Boltzmann 1896, 58-59: «the assumption made there that the state distribution is molecular-disordered is not fulfilled [in Loschmidt's temporal reversal], since after exact reversal of all velocities each molecule does not collide with others according to the laws of probability, but rather it must collide in the previously calculated manner».

Molekularbewegung in Gasen bestimmen”,[46] where Boltzmann argues that irrational ratios in the initial energy distribution make «possible the most different molecular distributions of one [total kinetic energy] (state distribution)».[47] The first step of Boltzmann’s answer is thus to confirm that the theory of non equilibrium requires certain mechanical conditions. This task will be accomplished in the third part of the paper where Boltzmann faces again the problem of the mechanical analogy of the Second Principle of thermodynamics and pursues the detailed analysis of the trajectories of a system of material points I have discussed in section 3.

Second, another consequence of Loschmidt’s argument is the following statement:[48]

> If we know that in a gas, at a certain time, a non uniform distribution is present and that the gas since a long time was in the container without an external influence, then we have to conclude that, much time before, the state distribution was uniform and that the unusual case occurred that it is gradually become non uniform.

Thus, Boltzmann does not question the case presented by Loschmidt, but rather he proves that it is a consequence of his general view on equilibrium: a non uniforming evolution is not a possibility totally excluded, but only «unusual». To substantiate this view an important step is required: proving that high entropy states are also high-probability states.

To sum up, the outlines of Boltzmann’s answer show all the elements already developed in the period 1868-1872: the conditions a mechanical system must fulfil to be probabilistically described, the probabilistic meaning of the molecular distribution and the equiprobability of the elementary events as a condition of the combinatorial analysis.[49] However, in the period 1868-1872 these points were strictly interlaced and the close mixture of mechanical and probabilistic elements made difficult an explication of Boltzmann’s concept of mixing. In particular, the probabilistic side remained in background in the theory of equilibrium and in the 1872 version of the theory of non equilibrium.

In 1877 Boltzmann faces the task to divide more clearly the probabilistic from the mechanical aspect, and proposes a theory where the physical hypotheses do not play a relevant role and which can be easily extended to any physical quantity: a really general theory of the mixing process. These goals are apparently outlined by Boltzmann when he states that his aim concerns «to investigate the relationships of the laws of probability with the Second Principle of the mechanical theory of heat»,[50] moving in the background specific physical hypotheses and pointing out the connections between high entropy states and high probability. Furthermore, he adds (italics added):[51]

> However, I stress that […] I want to derive the probability of a certain state distribution *completely apart from if and how it emerges*, or, more exactly expressed, I want to investigate all the combinations which are possible in the distribution of the $p + 1$ [kinetic energies] among the n

[46] See in particular Boltzmann 1876, 71-73.
[47] Boltzmann 1876, 72.
[48] Boltzmann 1877a, 122.
[49] On this point see Boltzmann 1877a, 120.
[50] Boltzmann 1877b, 166.
[51] Boltzmann 1877b, 168.

molecules and, as a consequence, to calculate how many combinations correspond to each state distribution.

Boltzmann's aim is to build an analysis which leaves out of consideration the problems linked to physical details of the system and thus which can be extended to every systems provided some general conditions, like the diffuse motion.

In short, Loschmidt's argument contained nothing of really new from Boltzmann's point of view, but rather it made present and incited the requirement of further explanation of conceptual elements already implicit in 1868 and in 1872. The probabilistic nature of the mixing process was a presupposition which was not adequately developed in the 1872 theory, because it was overshadowed by the project of building a physical theory of the thermodynamic systems and of the transport phenomena in "Maxwell style". So, the real problem was not the probabilistic nature of the mixing process in itself, but its explanation in general terms.

6. Combinatorial theories of 1868 and of 1877.

The combinatorial character of the theory of 1877 is, without doubt, its main peculiarity. There is no other cases in which Boltzmann makes a use as explicitly of stochastic models and probabilistic tools as in his article "Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Sätzen über das Wärmegleichgewicht". As we have seen, a stochastic model is implicit also in the collision mechanism framed by Boltzmann in 1872, but in that theory the purely kinetic context prevails. In this section I will show that, also in the combinatorial derivation of Maxwell distribution of 1868, Boltzmann makes use implicitly of a stochastic model which is completely analogous to the model of 1877. In this way a deep methodological and conceptual continuity connecting equilibrium and non equilibrium theory will be pointed out. This aim requires the discussion of analogies and differences between the combinatorial approaches of 1868 and of 1877.

In both theories two concepts of probability appear: as ratio between favourable cases and possible cases and as sojourn time in a certain phase region. In 1877 the latter is introduced by means of a famous stochastic model.[52] Boltzmann imagines an urn with many tickets; on each ticket a number of "energy elements" is reported. Each ticket (and then each "packet" of these elements) is equiprobable, and so are all the possible sequences drawn from the urn. A lot of drawings are performed, and, at the end, all the sequences which do not fulfil the condition of the total energy are neglected. But many sequences still remain, each of them represents a possible "complexion", i.e. a micro-configuration of the state of the system. The relative frequency of a state distribution derives from the number of complexions corresponding to the distribution itself assuming that the complexions are equiprobable, namely that each complexion appears «as much often» as any other in the drawing process. Thus, the relative frequency of a state

[52] Boltzmann 1877b, 171-172.

distribution derives from the "equifrequencies" of its complexions.[53] Again, the diffuse motion, translated in a relative-frequencies language, plays the central role.

But there are differences too. Indeed, the combinatorial derivations of 1868 and 1877 have two different argumentative structures. I will shortly discuss both the arguments and I will give an interpretation of their dissimilarity.

In 1868 Boltzmann aims to derive the two- and three-dimensional Maxwell distribution for a system of n particles with total energy $n\chi$ where χ is the average energy. The starting assumption, founded on the Boltzmann-Liouville theorem, is that all the possible individual dispositions of energy among the particles (i.e. all the ways to partition the total energy $n\chi$ among the n individual particles) are equiprobable. Accordingly, the problem simply becomes a combinatorial calculation developed considering the total energy as portioned out in a very large number p of very small "elements of energy" equal to each other.[54]

Boltzmann's argument to obtain the equilibrium distribution relies on the calculation of the marginal distribution of energy, i.e. on evaluing the probability that a particle has a certain energy independently from the energy of the other particles. Boltzmann focusses on the particle with kinetic energy k_1. As the total energy has been portioned out in p elements of dimension $n\chi/p$, p different energetic intervals can be defined:

$$\left[0,\frac{n\chi}{p}\right],\left[\frac{n\chi}{p},2\frac{n\chi}{p}\right],\ldots,\left[(p-1)\frac{n\chi}{p},p\frac{n\chi}{p}\right]. \tag{5}$$

As a first step, Boltzmann discusses the cases of two, three and four particles (n = 2, 3, 4) and computes the total numbers of dispositions which are compatible with the constraint that a certain particle has energy equal to k_1. In a second step, he generalizes to the case of n particles. Boltzmann's problem is to calculate the marginal probability that the energy of the particle is within one of the intervals defined in (5), i.e. to calculate the probability of the event:

$$E_n = \left\{k_1 \in \left[(p-q)\frac{n\chi}{p},(p-q+1)\frac{n\chi}{p}\right]\right\},$$

where $q = 1, \ldots, p$.

Generalizing the argument already developed in the limited cases Boltzmann obtains:

---

[53] In a note Boltzmann mentions another urn model (Boltzmann 1877b, 172): the drawing process takes place like in the first model, but, instead of neglecting the sequences which overcome the energy limit at the end of the process, at each step all the tickets which have become incompatible with the energy condition are eliminated. However, this model is unsatisfactory from a statistical point of view, because it is unable to conserve the exchangeability (cf. Costantini, Garibaldi, Penco 1996, 290-291).

[54] Boltzmann 1868, 84.

$$(6)\qquad \begin{aligned} P(E_n) &= \frac{q(q+1)\ldots(q+n-3)}{(n-2)!}\cdot\frac{(n-1)!}{p(p+1)\ldots(p+n-2)} \\ &= (n-1)\frac{q(q+1)\ldots(q+n-3)}{p(p+1)\ldots(p+n-2)} \end{aligned}.$$

If the number of the energy elements grows to infinite, the marginal probability of the event is:

$$(7)\qquad P(E_n) = \frac{(n-1)(n\chi - k_1)^{n-2}\,dk_1}{(n\chi)^{n-1}} = \frac{n-1}{n\chi}\left(1-\frac{k}{n\chi}\right)^{n-2} dk\,,$$

then, for the limit $n \to \infty$:[55]

$$(8)\qquad f = \frac{1}{\chi}e^{-\frac{k}{\chi}}dk\,,$$

which is the two-dimensional Maxwell distribution.[56]

What is the stochastic model implicit in Boltzmann's argument? Generally speaking, the combinatorial problem is to compute the number of way in which $n-1$ particles, considered indistiguishable, can be distributed on the $p-1$ distinguishable energy intervals defined in (5). A state distribution is a sequence of *occupation numbers*, i.e. how many particles are contained in a certain energy interval. The presupposition of the calculation is that all the *configurations*, i.e. the partitioning of individual particles on individual intervals, are equiprobable.

According to these hypotheses the equation (6) becomes:

$$(9)\qquad P(E_n) = (n-1)\frac{q(q+1)\ldots(q+n-3)}{p(p+1)\ldots(p+n-2)} = \frac{\binom{q+n-3}{q-1}}{\binom{p+n-2}{p-1}}.$$

The equation (9) is a particular case of the multivariate Polya distribution.[57] It can be easily shown that the equation (7) is a differential Beta distribution scaled on $n\chi$ and the equation (8) is a differential Gamma distribution, in accordance with the general theory of statistical distributions.[58]

Some remarks on Boltzmann's stochastic model. First of all it is worth noting that, despite the quantization of the energy in p elements, Boltzmann focusses not on the (indistinguishable) *individual elements*, but on the (distinguishable) *energy intervals* defined by means of the former. Thus, he faces the combinatorial problem of distributing the particles in different "cells" of

[55] Boltzmann requires that p/n remains finite (thermodinamic limit).
[56] The three-dimensional distribution is more problematic because it requires more degrees of freedom. No satisfactory derivation of this distribution can be found in 1868.
[57] Cf. Costantini, Garibaldi, Penco 1996, 286-287.
[58] For details on these analytical relationships see Bach 1990, 10-15 e Costantini, Garibaldi, Penco 1996, 286-288.

energy. In this way, the energy elements merely have the function of "labels" to mark the different cells and to provide a means for a combinatorial elaboration of them. No physical meaning is attached to the energy elements, but only a combinatorial one.[59] Furthermore, this model is completely analogous to the model developed in 1877. The tickets in the urn do not report individual energy elements, but rather "packets" of different dimensions. Thus, in Boltzmann's argument the energy elements play the role of labels useful to distinguish the energy intervals and to weigh them from a combinatorial point of view. An important consequence is that the combinatorial analysis both in equilibrium theory and in non equilibrium theory relies on a common stochastic model: to distribute $n - 1$ particles on $p - 1$ energy intervals.[60]

However, Alexander Bach has suggested that the energy elements may have a more explicit meaning.[61] According to Bach's interpretation, Boltzmann does not assume the equiprobability of the configuration of individual particles on individual intervals, but of the occupation numbers of indistinguishable elements on distinguishable particles. In other words, Bach claims that Boltzmann creates his stochastic model with individual energy elements. Accordingly, Boltzmann's combinatorial problem becomes to calculate the *occupacy numbers*, i.e. how many particles have a certain numbers of elements.[62] To sum up, Bach argues that Boltzmann adopts a Bose-Einstein statistics rather than a Maxwell-Boltzmann statistics.

To be sure, Boltzmann's combinatorial argument is formally unchanged in both cases. Adopting a Maxwell-Boltzmann statistics and focussing on individual intervals or adopting a Bose-Einstein statistics focussing on the individual particles leads to equivalent results. Thus, the question does not concern the formalism, but rather Boltzmann's interpretation of the function of the energy elements. In support of my interpretation I mention two remarks. First, Boltzmann's stochastic model is explicitly developed only in 1877, but the formal similarity suggests that such a model was implicit in 1868 too. There can be little doubt about the model in 1877: the distribution concerns packets of energy and not individual elements. Furthermore, in 1868 Boltzmann defines the elementary intervals as in (5) focussing on them in his combinatorial procedure.

Second, in 1868 Boltzmann states an important requirement: after a drawing process, all the dispositions of energy which break the constraint of the total energy have zero probability.[63] If the energy were drawn in individual element, this requirement would of course be superfluous, because the drawing process stops with the last element. This statement becomes relevant only if the drawing process concerns packets of energy so that the last drawing can eventually breaks the energy constraint. Moreover, a similar requirement can be found also in the urn model of 1877. In conclusion, the possibility to elaborate Boltzmann's argument as an example of Bose-Einstein statistics *ante litteram* is

---

[59] That Boltzmann's quantization cannot be considered a physical quantization is claimed also in von Plato, 1994, 80.
[60] The analogy holds also when Boltzmann introduces the continuum case cf. Boltzmann 1877b, 187.
[61] Bach 1990.
[62] Cf. Bach 1990, 9.
[63] Boltzmann 1868, 83.

not enough to conclude that Boltzmann effectively used such a statistics and leaves open the question about Boltzmann's interpretation of the energy elements.

As we have said, in spite of the similarities between the stochastic models in 1868 and in 1877, the argumentative structure of the two theories are deeply different. In 1868 Boltzmann derives the Maxwell distribution adopting the technique of marginalization, but in 1877, the combinatorial argument relies on minimization. Does it means that Boltzmann changed his mind about probability and combinatorial anaysis? In order to prove that this is not the case, I will briefly review Boltzmann's 1877 argument.

The starting point is the definition of probability of a state distribution as the ratio between the number of complexions corresponding to that state and the total number of possibile complexions. As the latter is a constant, a (non normalized) measure of the state probability is given by the quantity:

$$B = \frac{n!}{w_0!w_1!...w_n!},$$

where w_i are the occupation numbers of each cell in the energy space. To obtain the Maxwell distribution the quantity B must be a maximum. By using a semplified version of the Stirling approximation, the problem is turned into minimizing the denominator:

$$M = w_0!w_1!...w_n!.$$

Or, by using logarithms, to minimize the quantity:[64]

$$\log M = \sum w_i \log w_i - n.$$

To solve this problem of minimization Boltzmann adopts the method – usual in analytical mechanics – of the Lagrangian multipliers showing that the equilibrium distribution has the highest probability.

Why does Boltzmann change his combinatorial argument? The answer is in the close connection between the argumentative structure and the particular physical problems he has to face in 1868 and in 1877. Indeed, the marginalization technique used by Boltzmann in 1868 is not suitable to face the non equilibrium problem. This case requires a comparison between the equilibrium distribution and each possibile non equilibrium distribution, while the marginalization allows to compare the equilibrium with the non equilibrium *as a whole*. Consequently, Boltzmann deals with the non equilibrium problem using the minimization technique. Let us try to better understand the difference.

In 1868 Boltzmann computes a general expression (the equation (6)) expressing the probability that *a certain molecule lies within a certain phase region* provided that all the configurations are equiprobable. Such a constraint is formally represented by the Boltzmann-Liouville theorem, the so-called General Solution to the problem of equilibrium. Thus, the equilibrium distribution obtained by Boltzmann in 1868 is somehow understood as the "implied"

[64] Boltzmann 1877b, 176

distribution given the uniformity constraint or, in more modern terminology, it is the "typical" distribution provided such a constraint. To calculate the typical distribution Boltzmann does not need to compare all the possibile distributions to each other. He only needs the uniformity constraint, i.e. the General Solution. This strategy is perfectly feasible in the equilibrium theory which is focussed on one physical state only.

However, in 1877 Boltzmann faces a different physical problem: to show that the system will reach the Maxwell distribution *independently of the starting distribution*. Boltzmann's aim is to solve this problem showing that the equilibrium distribution is the most probable, but this goal necessarly requires a comparison of the probabilistic weights of all the possible distributions given the uniformity constraint. This aim is clearly expressed in his programmatic statement: «from the ratios of the numbers of the different state distributions their probability could even be calculated».[65] Thus, a new kind of probabilistic argument is required to face a new kind of physical problem.[66]

As a consequence, the different argumentative structure stemming from the common uniformity constraint and from the common stochastic model is not due to a change in Boltzmann's point of view, but simply to the change of the physical context. This fact reinforces the thesis of a deep continuity in Boltzmann's thought and of a close interaction between probabilistic arguments and mechanical problems.

I mention in passing a consequence of this interpretation. Already at first sight can note the formal similarity between the expression for log M and the H-function of 1872, a similarity stressed by Boltzmann himself.[67] In my opinion, this analogy sheds light upon the origin of this function, which was introduced without any justification.[68] Given a certain distribution, the quantity M is the number of the complexions favourable to that distribution. Thus, it derives from the same combinatorial analysis developed by Boltzmann in 1868. Whereas in 1868 Boltzmann worked with the marginalization procedure, in 1877 the same number of "favourable cases" enters in a minimization process. But it is the same strategy. Likely, the combinatorial considerations that in 1877 lead to the introduction of the quantity M find their origin in the article of 1868. In other words, Boltzmann framed the H-function starting from the combinatorial considerations developed in the context of the equilibrium theory. It can be remarked that in 1872 no hints of this origin appear and that H-function is introduced without any argument. But, again, the physical context is relevant. In 1872, as we have seen, the physical aspects are dominant in comparison to the probabilistic ones. Boltzmann's goal was to obtain a non equilibrium theory dealing with the thermodynamical problem related to the entropy and with the trasportation phenomena. It was understood as a generalization of Maxwell's theory of 1867. The explication of the role played by the combinatorial arguments and by the stochastic model will be offered only in 1877, and the dialogue with Loschmidt was very important in this explication,

---

[65] Boltzmann 1877a, 121.

[66] In 1878 Boltzmann will investigate also the probability that an arbitrary distribution differs from the equilibrium one (cf. Boltzmann 1878, 250-264).

[67] Boltzmann 1877b, 179.

[68] In spite of its importance, the problem of the origin of the H-function is in the literature rarely dealt with. Brush claims that Boltzmann discovered it by means of his experience in manipulating entropy definitions and with trial-and-error- methods (cf. Brush 1986, 600).

but they must have had a part in the transition from equilibrium to non equilibrium theory. Likely, Boltzmann understood those arguments as a guide in facing the difficult problem of proving the irreversible behavior of a complex mechanical system of material points.

7. CONCLUDING REMARKS.

In the previous sections it has been pointed out that Boltzmann's thought is characterized by a close interaction between probabilistic and mechanical problems. It has often been claimed that Boltzmann's research programme is something of a generalization of Maxwell's because it progressively tries to remove the starting simplifying assumptions of the early kinetic theories,[69] and that programme proceeds well until Loschmidt's objection proves the statistical nature of the irreversibility. But Boltzmann's research programme is more complex. It cannot be classified as a form of naive redutionism come, at last, to a proved impossibility and to a deserved failure. He was perfectly aware of the statistical nature of thermodynamic laws, and his main aim was not a unidirectional reduction, but rather a sort of "bidirectional analogy", according to which probabilistic concepts can be applied to some particular mechanical systems, and the analysis of these mechanical systems can shed light upon the statistical nature of some natural laws. The lack of a clear ontological and epistemological hierarchy in this programme is mirrored in the conceptual network used by Boltzmann. Thus, the peculiarity of this position is the search of a compatibility between the point of view of the analytical mechanics and the use of probabilistic tools.

I opened this paper stating the Klein thesis; let us see as, according to the ideas above developed, it can be transformed:

(K1*) Boltzmann's view of the role and the meaning of probability does not go through deep alterations during the time. However, there is an evolution and new different nuances come in foreground.

(K2*) The process of evolution of Boltzmann's ideas on probability is mainly an internal one. External influences contributed but not as decisive criticisms.

(K3*) The theory of 1872 shows a deep mixture of mechanical and probabilistic elements. The collision mechanism is an apparent example of a mechanical process analyzed by means of an analogy with a stochastic model.

(K4*) Boltzmann's view of an equilibrium process is statistical from the very beginning. He interprets it as a long-run phenomenon. The *H*-theorem itself is an average result.

(K5*) The combinatorial theory of 1877 was incited by Loschmidt's argument, but it is mainly an explication of elements implicitly present in the theory of 1872 and explicitly expressed in the equilibrium theory of 1868.

[69] See, for example, Clark 1976, 56-57.

Loschmidt'objection was not understood as a decisive criticism, but as a demand of deepening in Boltzmann's view of the mixing process.

To sum up, I am not claiming that Boltzmann's ideas were established once and for all in 1868, but that their variation and evolution are subtle. The concepts of probability, indipendence, diffuse motion undoubtely go through as internal evolution in which, however, changes of perspective or of standpoint are more important than actual revolutions.

In other words, the epistemological dynamics of Boltzmann's thought is not as simple as that described in Klein thesis. (K1)-(K5) suggest a dynamics in which a kind of (exclusively mechanicistic) belief is trasformed in another kind of (statistical) belief by means of an external crucial criticism. On the contrary, I tried to show that the probabilistic notions (equiprobability, diffusion, independence) form a conceptual network which is compatible with the mechanical approach and where no clear epistemological hierarchy can be detected. Accordingly, the epistemological dynamics of Boltzmann's thought cannot be characterized as a transformation from a kind of belief to another, but as an organic development where, in different physical context, every element of this conceptual network can temporaly assume a fundamental role. I think that this interpretation can better valorize the complexity of the great Austrian scientist's thought.

References

Bach, A. 1990, “Boltzmann’s Probability Distribution of 1877”, *Archive for History of Exact Sciences*, 41, 1-40.

Bierhalter, G. 1981, “Boltzmanns mechanische Grundlegung des zweiten Hauptsatzes der Wärmelehre aus dem Jahre 1866”, *Archive for History of Exact Sciences*, 24, 195-205.

Bierhalter, G. 1992, “Von L. Boltzmann bis J. J. Thomson: die Versuche einer mechanischen Grundlegung der Thermodynamik (1866-1890)”, *Archive for History of Exact Sciences*, 44, 25-75.

Boltzmann, L. 1866, “Über die mechaniche Bedeutung des zweiten Hauptsatzes der Wärmetheorie”, *Wiener Berichte*, 53, 195-220, in Boltzmann 1909, I, 9-33.

Boltzmann, L. 1868, “Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten”, *Wiener Berichte*, 58, 517-560, in Boltzmann 1909, I, 49-96.

Boltzmann, L. 1871a, “Über das Wärmegleichgewicht zwischen mehratomigen Gasmolekülen”, *Wiener Berichte*, 63, 397-418, in Boltzmann 1909, I, 237-258.

Boltzmann, L. 1871b, “Einige allgemeine Sätze über Wärmegleichgewicht”, *Wiener Berichte*, 63, 679-711, in Boltzmann 1909, I, 259-287.

Boltzmann, L. 1872, “Weitere Studien über das Wärmegleichgewicht unter Gasmolekülen”, *Wiener Berichte*, 66, 275-370, in Boltzmann 1909, I, 316-402.

Boltzmann, L. 1876, “Über die Austellung und Integration der Gleichungen, welche die Molekularbewegung in Gasen bestimmen”, *Wiener Berichte*, 74, 503-552, in Boltzmann 1909, II, 55-102.

Boltzmann, L. 1877a, “Bemerkungen über einige Probleme der mechanischen Wärmetheorie”, *Wiener Berichte*, 75, 62-100, in Boltzmann 1909, II, 112-148.

Boltzmann, L. 1877b, “Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respective den Sätzen über das Wärmegleichgewicht”, *Wiener Berichte*, 76, 373-435, in Boltzmann 1909, II, 164-223.

Boltzmann, L. 1878, “Weitere Bemerkungen über einige Probleme der mechanischen Wärmetheorie”, *Wiener Berichte*, 78, 1-46, in Boltzmann 1909, II, 250-288.

Boltzmann, L. 1881, “Referat über die Abhandlung von J. C. Maxwell “Über Boltzmanns Theorie betreffend die mittlere Verteilung der lebendingen

Kraft in einem Systeme", *Annalen der Physik*, 5, 403-417, in Boltzmann 1909, II, 582-595.

Boltzmann, L. 1895, "On Certain Questions of the Theory of Gases", *Nature*, 51, 413-415, in Boltzmann 1909, III, 535-544.

Boltzmann, L. 1896, *Vorlesungen über Gastheorie*, Barth, Leipzig, (english translation *Lectures on Gas Theory*, Dover, New York, 1964).

Boltzmann, L. 1909, *Wissenschaftliche Abhandlungen*, 3 voll., Barth, Lipsia.

Brush, S. 1976, "Irreversibility and Indeterminism: Fourier to Heisenberg", *Journal of the History of Ideas*, 37, 603-630.

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

Brush, S. 1986, *The Kind of Motion we call Heat*, 2 voll., North Holland, Amsterdam.

Clausius, R. 1870, "On a Mechanical Theorem applicable to Heat", *Philosophical Magazine*, 4, 40, n. 265, 122-127.

Clark, P. 1976, "Atomismus versus Thermodynamics", in C. Howson (ed.), Method and Appraisal in the *Physical Sciences*, Cambridge University Press, Cambridge, 41-105.

Costantini, D. Garibaldi, U. 1997, "A Probabilistic Foundation of Elementary Particle Statistics. Part I", *Studies in History and Philosophy of Modern Physics*, 28, 4, 483-506.

Costantini, D. Garibaldi, U. 1998, "A Probabilistic Foundation of Elementary Particle Statistics. Part II", *Studies in History and Philosophy of Modern Physics*, 29, 1, 37-59.

Costantini, D. Garibaldi, U. Penco, M. A. 1996, "Ludwig Boltzmann alla nascita della meccanica statistica", *Statistica*, 3, 279-300.

Ehrenfest, P. 1916, "Adiabatische Invarianten und Quantentheorie", *Annalen der Physik*, 51, 327-352.

Ehrenfest, P. e T. 1911, *The Conceptual Foundations of the Statistical Approach in Mechanics*, Dover, New York.

Elkana, Y. 1974, "Boltzmann's Scientific Research Program and its Alternatives", in Y. Elkana (ed.), *The Interactions between Science and Philosophy*, Humanities Press, Atlantic Highlands, 243-279.

Klein, M. J. 1973, "The Development of Boltzmann's Statistical Ideas", *Acta Physica Austriaca*, Supplementum 10, 53-106.

Krüger, L. 1981, “Probability as an Obstacle to Scientific Reduction”, in M. Heildelberger e L. Krüger (eds.), *Probability and Conceptual Change in Scientific Thought*, B Oleine Verlag, Bielefeld, 95-113.

Krüger, L. 1987, “The Probabilistic Revolution in Physics – an Overview”, in L. Krüger, L. J. Dalston, M. Heidelberger (eds.), *The Probabilistic Revolution*, 2 voll., MIT Press, Cambridge, II, 373-378.

Lindey, D. 2001, *Boltzmann’s Atom. The Great Debate that Launched a Revolution in Physics*, The Free Press, New York.

Loschmidt, J. “Über den Zustand des Wärmegleichgewichtes eines Systems von Körpern mit Rücksicht auf die Schwerkraft. I”, *Wiener Berichte*, 73, 128-139.

Maxwell 1867, , “On the Dynamical Theory of Gases”, *Philosophical Transactions of the Royal Society*, 157, 49–88, in Maxwell 1890, I, 26-78.

Maxwell, J. 1879, “On Boltzmann’s Theorem on the Average Distribution of Energy in a System of Material Points”, *Cambridge Philosophical Society Transactions*,12, 547–575, in Maxwell 1890, II, 713-741.

Maxwell, J. 1890, *The Scientific Papers of James Clerk Maxwell*, 2 voll. (ed. W. D. Niven), Cambridge University Press, Cambridge.

Von Plato, J. 1983, “The Method of Arbitrary Functions”, *The British Journal for the Philosophy of Science*, 34, 37-47.

Von Plato, J. 1987, “Probabilistic Physics the Classical Way”, in L. Krüger, L. J. Dalston, M. Heidelberger (eds.), *The Probabilistic Revolution*, 2 voll., MIT Press, Cambridge, II, 379-407.

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 University Press, Cambridge.

Porter, T. 1986, *The Rise of Statistical Thinking – 1820-1900*, Princeton University Press, Princeton.

Watson, H.W. 1876, *Treatise on the Kinetic Theory of Gases*, Clarendon Press, Oxford.