The Odd Couple: Boltzmann, Planck and the Application of Statistics to Physics (1900-1913)

Massimiliano Badino*

Abstract

In the last forty years a vast scholarship has been dedicated to the reconstruction of Planck's theory of black-body radiation and to the historical meaning of quantization. Since the introduction of quanta took place for combinatorial reasons, Planck's understanding of statistics must have played an important role. In the first part of this paper, I sum up the main theses concerning the status of the quantum and compare the arguments supporting them. In the second part, I investigate Planck's usage of statistical methods and the relation to Boltzmann's analogous procedure. I will argue that this way of attacking the problem is able to give us some interesting insights both on the theses stated by the historians and on the general meaning of Planck's theory.

1 Introduction: A vexed problem

In his epoch-making paper of December 1900 on blackbody radiation,(Planck 1900b)(Planck 1958, 698-706) for the first time Max Planck (1858-1947) made use of combinatorial arguments. Although it was a difficult step to take, a real "act of desperation" as he would later call it, Planck pondered it deeply and never regretted it. As he wrote to Max von Laue (1879-1960) on 22 March 1934: "My maxim is always this: consider every step carefully in advance, but then, if you believe you can take responsibility for it, let nothing stop you."(Heilbron 1986, 5).

The difficulty involved in this step was the adoption of a way of reasoning that Planck had been opposing for a long time: Ludwig Boltzmann's (1844-1906) statistical approach. However, Planck's application of statistics to the particular problem of finding the spectral distribution of cavity radiation seems to bear only partial resemblance to Boltzmann's original arguments and the opinions of the scholars are split about the interpretation of the relation between Planck's and Boltzmann's procedure.

For discussion's sake, I sort out three basic kinds of outlooks, which I term the discontinuity thesis, the continuity thesis and the weak thesis. According to the discontinuity thesis Planck worked from the very beginning with discrete elements of energy. As early as 1962, Martin Klein, in a series of seminal papers(Klein 1962) (Klein 1963) (Klein 1964) (Klein 1966), argued explicitly that in December 1900 Planck introduced the quantization of energy even

*Max-Planck-Institut für Wissenschaftsgeschichte and Fritz-Haber-Institut der Max Planck Gesellschaft Boltzmannstrasse 22, 14195 Berlin, Germany. E-mail: mbadino@mpiwg-berlin.mpg.de

though he might have not been perfectly aware of the consequences. In recent years, Res Jost has polemically endorsed Klein's classical thesis (Jost 1995) (Koch 1991).

A different outlook emerged in 1978 when Thomas Kuhn(Kuhn 1978) (Kuhn 1984) (Klein, Shimony and Pinch 1979) claimed that in December 1900 Planck was thinking in terms of continuous energy and that the elements were merely a 'shortcut' for continuous distribution over energy (or phase space) cells. For Kuhn, the discontinuity entered physics as late as 1905-1906 through the work of Paul Ehrenfest (1880-1933) and Albert Einstein (1879-1955). I will call this claim the continuity thesis. Both the discontinuity and the continuity thesis argue, in opposite directions, that Planck was definitely committed on the reality of quantum. In a sense also Olivier Darrigol(Darrigol 1988) (Darrigol 1991) can be numbered among the upholders of the continuity thesis, even though his position points less straightforwardly to a clear commitment (Darrigol 2000) (Darrigol 2001).

Recently, the advocates of the weak thesis have claimed that this ascription of commitment is unjustified because Planck restrained himself from taking a clear stand on the status of quantum. The reasons for this absence of decision might be different. Allan Needell, for instance, has convincingly argued that the issue was simply out of the range of Planck's interests. The exact behavior of the resonator belongs to a domain of phenomena, namely the micro-phenomena, which Planck was unwilling to investigate. Hence, the question whether the resonator absorbs and emits continuously or discontinuously was irrelevant in Planck's general approach(Needell 1980).

A similar contention is shared by Peter Galison(Galison 1981) who maintains that, in general, it is not wise to ascribe strong commitments to scientists working in a period of scientific crisis(Galison 1981). More recently, Clayton Gearhart(Gearhart 2002) has suggested another option for the supporters of the weak thesis, holding that, even if interested in the issue of the quantization of energy, for various reasons Planck was unable or unwilling to take a plain position in his printed writings while "he was often more open in discussing their implication in his correspondence"(Gearhart 2002, 192). For Gearhart what Planck lacked was a concluding argument to make up his mind in one way or the other. Similarly, Elizabeth Garber(Garber 1976) stressed some ambiguities in Planck's work when she claimed that his theory escaped the pivotal point: the mechanism of interaction between matter and radiation.

Most significantly, these scholars changed the direction of the debate by claiming that the 'did he or didn't he' question is not the important one. The arguments they brought in order to dismiss the continuity/discontinuity dichotomy eventually suggested new and more fruitful issues.

In this long debate the statistical arguments have played a central role. Literally interpreted, Planck's statements in December 1900 seem to suggest a counting procedure that introduces discontinuous energy elements. In fact, one of Klein's main arguments consists in showing how remarkably Planck's use of combinatorials diverges from Boltzmann's and in hinting that this is the mark of the quantization of energy(Klein 1962).

On the contrary, both Kuhn and Darrigol endeavored to show that the dissimilarities are only superficial or irrelevant. In particular, Kuhn argued that Planck's counting procedure is perfectly consistent with Boltzmann's interpretation based on energy cells and that we need not take Planck's statements too literally and instead must read them in the close historical context of his research program. He found that a discontinuous view of energy was too drastic a step to be justified by a statistical argument only.

This paper aims to contribute to these discussions by analyzing Planck's use and interpretation of statistics. I argue that a deep investigation of this particular issue provides new insights and support to the weak thesis. The argument is articulated in two steps. First I discuss in detail some technical aspects of Planck's statistics and conclude that (1) statistical formalism can be physically interpreted in different ways and (2) Planck followed Boltzmann's original procedure as close as possible. Even though both these conclusions were implicit in Kuhn's analysis, my argument takes a different direction.

For they do not entail the continuity thesis as Kuhn argues. Instead, I suggest that precisely in the neutrality of the formalism Planck saw the major advantage of the new approach. Indeed, he could easily integrate the combinatorial formalism in his overall project of deriving the radiation law without hypotheses on the interaction between matter and radiation just because this formalism does not demand any specific assumption on the micro-process at work during the energy exchange. Thus Planck's uncommitted position resulted from the interplay between the formal features of statistical arguments and his general attitude toward statistics.

In the second part I provide some arguments to endorse this reading. In particular I look at the way in which Planck interpreted the role of statistical considerations in physics and I compare this interpretation to Boltzmann's position on the same issue. The analysis shows that, contrary to Boltzmann, Planck did not consider statistical arguments as effective tools to explore the micro-world.

2 How the story began

Essential for this story is the way in which Boltzmann's normalization factor became Planck's number of complexions. The first step consists therefore in teasing out what exactly this factor meant in Boltzmann's statistics and how Boltzmann handled it. There are two different statistical arguments in Boltzmann's works. The first is presented in his 1868 paper(Boltzmann 1868) (Boltzmann 1909, I, 49-96) devoted to the derivation of Maxwell's distribution law.[1] The structure of this paper is remarkable. In the first part Boltzmann derived Maxwell's law by the analysis of mechanical collisions among molecules. In the second part, he retrieved the same result by means of an original argument based on the calculation of the probability that a molecule is allocated into a certain energy cell. Boltzmann supposed a system of n molecules whose total energy E is divided into p equal elements of size ϵ, so that $E = p\epsilon$. This allows him to define p possible energy cells $[0,\epsilon], [\epsilon, 2\epsilon], \ldots, [(p-1)\epsilon, p\epsilon]$, so that if a molecule is allocated in the i-th cell, then its energy lies between $(i-1)\epsilon$ and $i\epsilon$. The probability that a molecule is allocated in the i-th cell is given by the ratio between the total number of ways of distributing the remaining $n-1$ molecules over the cells defined by the total energy $(p-i)\epsilon$ and the total number of ways of distributing all molecules. In other words, this probability is proportional to the total number of what in 1877 would be called 'complexions' calculated on

[1]An excellent discussion of this section of Boltzmann's paper which, however, does not include an analysis of the combinatorial part of the argument, can be found in (Uffink 2007).

an opportunely defined new subsystem of molecules and total energy.[2] In the general case of n molecules this probability becomes:[3]

$$P_i = (n-1)\frac{q(q+1)\ldots(q+n-3)}{p(p+1)\ldots(p+n-2)}.$$

Letting the numbers of elements and of molecules grow toward infinity, Boltzmann obtained the Maxwell distribution.

Note that Boltzmann defined the energy cells using a lower and upper limit, but in his combinatorial calculation he considered one value of the energy only. This is doubtless due to the arbitrary small size of ϵ, but it also entails an ambiguity in the passage from the physical case to its combinatorial representation and vice-versa. The normalization factor used by Boltzmann in this argument can be obtained by a simple calculation (Costantini, Garibaldi and Penco 1996) (Costantini and Garibaldi 1997):

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

The binomial coefficient in the denominator on the right-hand side is of course the total number of complexions given the energy conservation constraint as Boltzmann's statistical model suggests. More generally, if one considers a statistical model of n distinguishable balls and p distinguishable boxes and assumes that to each box a multiplicity i ($i = 0, \ldots, p-1$) is ascribed, then the total number of possible allocations W consistent with the conditions:

$$\sum n_i = n, \qquad \sum i n_i = p-1,$$

n_i being the number of balls allocated in the i-th box, is given by:

$$W = \binom{p+n-2}{p-1} = \frac{(p+n-2)!}{(p-1)!(n-1)!}. \qquad (2)$$

This formula is exactly the normalization factor written by Boltzmann.

Boltzmann's approach changed in 1877. For technical reasons he moved from cells in the energy space to cells in the phase space and was forced to consider more carefully the vanishing size of the cell. Moreover, to elucidate his statistical view of irreversibility he proposed a brand new combinatorial argument that presumably constituted Planck's main source of statistical insights(Boltzmann 1877) (Boltzmann 1909, II, 164-223). As in 1868, Boltzmann calculated the probability of a state in term of complexions realizing that state, but now he

[2]A complexion is a distribution of distinguishable statistical objects (like balls) over distinguishable statistical predicates (like boxes), namely it is an individual configuration vector describing the exact state of the statistical model.

[3]The dependence on i is contained in q through the relation $p - i = q$.

compared the equilibrium distribution with all possible ones in order to prove that the former is overwhelmingly more probable.

To clarify his ideas as much as possible, Boltzmann presented his procedure in form of a qualitative model. Let us suppose, as before, a system of n molecules whose total energy is divided into elements of size ϵ and imagine an urn with an enormous number of tickets. On each ticket a number between 0 and p is written, so that a possible complexion describing an arbitrary state of the system is a sequence of n drawings where the i-th drawn ticket carries the number of elements to be ascribed to the i-th molecule. Of course, a complexion resulting from such a process could not possibly satisfy the energy conservation, therefore Boltzmann demanded an enormous number of drawings and then eliminated all the complexions that violate that constraint. The number of acceptable complexions obtained after the elimination is still very large. Since a state distribution depends on how many molecules (and not which ones) are to be found in each energy cell, many different complexions are compatible with a single state. By maximizing the state probability, defined as the number of complexions realizing it, Boltzmann succeeded in showing that the equilibrium state has a probability overwhelmingly larger than any other possible distribution.

In this new statistical model Boltzmann had to count again the total number of complexions consistent with the energy conservation constraint. To do so he used the formula (2), but in this particular case the possible allocations of energy are $p+1$ because the cell is defined by a single number of energy elements rather than by lower and upper limits. This means that the normalization factor (2) becomes:

$$\binom{p+n-1}{p} = \frac{(p+n-1)!}{p!(n-1)!}. \tag{3}$$

This is the total number of complexions for a statistical model where distinguishable balls are distributed over distinguishable boxes precisely as in 1868. The only difference results from the new definition of energy cell used in the urn model. Boltzmann actually wrote the formula (3) in passing as an expression of the total number of such complexions(Boltzmann 1909, II, 181). Apparently, he did not consider such a difference as relevant from a combinatorial point of view.

However, the formula (3) can also be interpreted in a completely different way. Instead of describing the state of the system by a specific arrangement of individual balls one can specify how many balls are in a certain box. This kind of description is usually called an 'occupation vector.' Now, formula (3) can equally well express the total number of occupation vectors for a statistical model of p indistinguishable balls and n distinguishable boxes (balls are indistinguishable if their permutation does not give a new configuration). Thus if one wants to calculate the total number of ways of distributing p balls over n boxes counting only how many balls, and not which ones, are put into each box, then the formula (3) gives the total number of such distributions. An ingenious and particularly simple way of proving this statement was proposed by Paul Ehrenfest and Heike Kamerlingh Onnes (1853-1926) in 1915(Ehrenfest and Kamerlingh Onnes 1915). Let us suppose that, instead of n distinguishable boxes, one uses the $n-1$ indistinguishable bars defining the boxes.[4] If both the p balls and the $n-1$ bars were distinguishable, the total number of individual arrangements would

[4]Note that this switch between cells and limits is similar to Boltzmann's with the not negligible difference that Boltzmann's cell limits in 1868 are distinguishable.

be given by $(p+n-1)!$. But the indistinguishability of the balls as well as of the bars forces us to cancel out from this number the $p!$ permutations of the balls and the $(n-1)!$ permutations of the bars. By doing so, one arrives at the total number W_{ov} of occupation vectors:

$$W_{ov} = \frac{(p+n-1)!}{p!(n-1)!}.$$

This ambiguity between formulae and underlying statistical model, as we will see in the next section, is one of the keys to understand Planck's use of statistical arguments.[5]

3 A statistics for all seasons

Planck's application of Boltzmann's statistics to radiation theory has very far-reaching consequences that puzzled his contemporaries and took many years to be completely understood. In this and in the next section, restricting myself to an analysis of Planck's counting procedure and of the general structure of his statistical argument, I will argue that, in combination with other arguments, Planck's use of statistics suggests further support to the weak thesis.

In his December 1900 paper, where he first gave a theoretical justification of the radiation law, Planck was very explicit about his counting procedure(Planck 1900b):[6]

> "We must now give the distribution of the energy over the separate resonators of each [frequency], first of all the distribution of the energy E over the N resonators of frequency ν. If E is considered to be a continuously divisible quantity, this distribution is possible in infinitely many ways. We consider, however — this is the most essential point of the whole calculation — E to be composed of a well-defined number of equal parts and use thereto the constant of nature $h = 6.55 \times 10^{-27}$ erg sec. This constant multiplied by the common frequency ν of the resonators gives us the energy element ϵ in erg, and dividing E by ϵ we get the number P of energy elements which must be divided over the N resonators. If the ratio thus calculated is not an integer, we take for P an integer in the neighborhood. It is clear that the distribution of P energy elements over N resonators can only take place in a finite, well-defined number of ways."

Since Planck is unmistakably speaking of distributing energy elements over resonators, this passage has always been the main weapon of the upholders of the discontinuity thesis.[7] Moreover, in the same paper, Planck wrote the total number of these distributions in the following way:

$$\frac{N(N+1)(N+2)\dots(N+P-1)}{1\cdot 2\cdot 3\dots P} = \frac{(N+P-1)!}{(N-1)!P!}. \tag{4}$$

[5]As said above, this ambiguity is implicit in Kuhn's analysis. It has been stressed also in (Darrigol 1991).

[6]I am following the english translation in (Ter Haar and Brush 1972, 40).

[7]However, the final part of the quotation seems to suggest that he is not considering P to be necessarily an integer. On this point see (Darrigol 2001).

This is the total number of ways of allocating P indistinguishable balls over N distinguishable boxes as Ladislas Natanson (1864-1937) pointed out in 1911(Natanson 1911).

However, Kuhn suggested another reading. As just noted regarding formula (3), formula (4) can also be interpreted as the total number of ways of distributing N distinguishable balls over $P+1$ distinguishable boxes given the energy conservation constraint. This would mean that Planck was using exactly the same statistical model Boltzmann had worked out in his 1877 paper! While according to the 'discontinuous' interpretation the distribution of single indistinguishable energy elements over resonators very naturally suggests that the resonators can absorb and emit energy discontinuously, according to the Kuhn's, the resonators are distributed over distinguishable energy cells with fixed size, but they can be placed anywhere within a given cell thus implying that they can absorb or emit continuously. More importantly, if Planck was faithfully following Boltzmann's original procedure and Boltzmann, as we know, assumed continuity behind this procedure, it is extremely plausible that Planck himself thought about a continuous exchange of energy between resonators and field.[8] Hence, Kuhn's continuity thesis is that in following Boltzmann, Planck ended up with a combinatorial formula that he interpreted as counting the distributions of resonators over continuous energy cells because a discontinuous reading of the same formula would have entailed a contradiction with the first part of his theory, namely the derivation of the electromagnetic relation between energy density in the cavity and average energy of the resonator.

An alternative account, which I suggest here, is that the ambiguity does not speak directly for a commitment of Planck toward continuity or discontinuity. Instead, Planck might have integrated the combinatorial procedure in his general strategy precisely because its formal ambiguity implies that the combinatorial formalism is independent of particular physical assumptions. In other words, since the formal ambiguity leaves completely open the question of what is going on at the micro-level, it allows Planck to switch from one statistical model to another for purely computational reasons. The intrinsic ambiguity in the formalism was therefore instrumental in Planck's approach of leaving aside any detailed assumption on micro-phenomena. This interpretation supports the weak thesis, is in accordance with Planck's general attitude toward micro-processes, and helps explain the coexistence of apparently opposite indications for continuity/discontinuity in Planck's early works.

This point can be strengthened by appealing to the first edition of the *Vorlesungen über Wärmestrahlung* (1906). In the fourth chapter the attempts to remain as close as possible to Boltzmann's original argument and phraseology is patent. First, Planck introduced the concepts of complexion and distribution for a gas and it is clear that, to him, a complexion is an individual allocation of molecules over phase space cells(Planck 1906, 140-143). Then he simply extended this concept to radiation theory, without any, even slight, change of meaning so that it is plausible that this concept kept its general features in the new context as well.

After developing Boltzmann's procedure for a gas, he dealt with the problem of counting complexions in radiation theory(Planck 1906, 151-152):

> "Here we can proceed in a way quite analogous to the case of gas, if only we take into account the following difference: a given state of the system of res-

[8]This ambiguity was exploited in the other direction by Alexander Bach (Bach 1990) who suggested that in 1877 Boltzmann was anticipating a Bose-Einstein statistics. However, I think that Bach's interpretation, though formally correct, cannot be held from a historical point of view.

onators, instead of determining a unique distribution, allows a great number of distributions since the number of resonators that carry a given amount of energy (better: that fall into a given 'energy domain') is not given in advance, it is variable. If we consider now every possible distribution of energy and calculate for each one of these the corresponding number of complexions exactly as in the case of gas molecules, through addition of all the resulting number of complexions we get the desired probability W of a given physical state."

The reference to the "energy domain" is a hint that Planck was thinking of distributing resonators over energy cells and that he considered this method as "a way quite analogous to the case of gas."

But immediately after, he stressed that the same goal can be accomplished by the "faster and easier" ("schneller und bequemer") way of distributing P energy elements over N resonators and then he displayed again the same formula used in December 1900. This passage shows two important points. First, in 1906 Planck was well aware that, from a purely formal viewpoint, both statistical models led to the same result. Initially he stated that the total number of complexions can be calculated in exactly the same way Boltzmann had followed in 1868 to evaluate the normalization factor, namely by adding up the individual configurations corresponding to all possible distributions. Next Planck stressed that this factor can also be obtained by a different statistical model. This shows that he realized that the combinatorial formula underdetermines the physical interpretation.

Second, he considered the distribution of individual energy elements over resonators to be a simpler way of working likely because it does not require any energy conservation constraint, nor, more importantly, any assumption on the distribution of the resonators within the energy cells. Planck subscribed neither of two models and switched from one to the other for purely formal reasons.

Unfortunately, this is the first occasion on which Planck explicitly manifested his knowledge of the subtleties of combinatorics and the objection can be made that this knowledge was the fruit of the intervening years between 1900 and 1906. However, since I do not claim that the combinatorial ambiguity was the only source of Planck's uncommitted position, the exact temporal determination of Planck's awareness of the ambiguity is not decisive. My claim is that the neutrality of the statistical formalism was an additional ingredient in a skeptical attitude toward micro-phenomena in general and statistical tools in particular. At a certain point between 1900 and 1906 Planck realized that his attitude was well confirmed by the variety of models that can be attached to the same statistical formula.

It is plausible that Planck took a prudential position on the meaning of his combinatorial procedure from the very beginning, in 1900-1901, and then found good reasons to maintain this position. In this remark I am following Galison's consideration that it is pointless to attribute strong and definite beliefs to scientists in a period of transition. Moreover, Planck's caution in the matter of micro-phenomena has been pointed out by many scholars, notably Needell and Ed Jurkowitz. In particular, Jurkowitz(Jurkowitz 2007) (Jurkowitz 2008) has argued that Planck pursued a 'framework' approach moving from basic principles to study the micro-world. His picture of Planck's earlier work emphasizes that Planck worked from secure principles toward an explanation of the unknown and as yet obscure micro-world underlying blackbody radiation. In this framework approach the specific features of the micro-processes have a secondary importance.

In section 6 I will show that this general attitude of Planck encompassed the issue of the role of statistics in physics. Before that, however, I extend the study to other technical traits of Planck's statistics. In the next two sections I will show that Planck followed closely Boltzmann in the application of combinatorial tools, but his faithfulness was limited to the formal aspect and did not entail any subscription to Boltzmann's view of the micro-world.

4 To maximize or not to maximize?

In the previous section I have argued that the combinatorial formula used by Planck to compute the total number of complexions underdetermines its physical interpretation because it is compatible with two different statistical models. In this and in the next section I move on and analyze some key points of Planck's statistical procedure in 1900-1901.

The first paper on the quantum was published in December 1900 and the second one in the *Annalen der Physik* in January 1901(Planck 1901). In spite of such a short time interval, there are considerable differences between these two papers, particularly with regard to the structure of the argument.

In the December 1900 paper, Planck considers different classes of resonators characterized by their proper frequency of vibration. Thus we have N_1 resonators at frequency ν_1, N_2 resonators at frequency ν_2, and so on. All resonators of a certain class have the same frequency and do not mutually interact. The energy E_r, namely the fraction of the total energy E ascribed to the resonators only, must be divided over the different frequencies, so that the energetic state of the system is described by the vector:

$$\omega_k = \{E_1.E_2 \dots\}, \tag{5}$$

where E_i is the energy assigned to the frequency ν_i and each vector ω_k must satisfy the condition:

$$E_r = E_1 + E_2 + \dots$$

There are many different ways of distributing the total energy E_r over the possible frequencies in accordance with the condition above, but not all these ways have the same 'probability.' Planck measures the probability that a certain energy E_i is ascribed to the frequency ν_i by the total number of ways of distributing the energy divided into P_i elements of size $\epsilon_i = h\nu_i$ over the N_i resonators. Since the size of the elements depends on the frequency, the division of energy is different in each class of resonators. As we have already seen, this number is:

$$W(E_i) = \frac{(N_i + P_i - 1)!}{(N_i - 1)! P_i!}. \tag{6}$$

Since resonators of different classes do not interact, the probability of the distribution described by the vector (5) is:

$$W(\omega_k) = \prod_i W(E_i). \tag{7}$$

At this point, instead of performing the cumbersome maximization of (7), finding the equilibrium distribution ω_{eq} and the mean energy of a resonator at an arbitrary frequency, Planck mentions that "[a] more general calculation, which is performed very simply, using exactly the above prescriptions shows much more directly" the final result. There is no description at all of what this "general calculation" should look like and Thomas Kuhn has suggested that here Planck has in mind the argument he would present in his *Annalen* paper in January 1901. Kuhn's hypothesis is perfectly reasonable because the maximization procedure does not appear in the 1901 paper. Indeed, in the January article Planck starts by reckoning only a class of resonators[9] and instead of considering a set of possible distributions ω_k among which a particular equilibrium distribution ω_{eq} is to be selected by means of a maximization procedure, he directly presupposes the equilibrium state. After this step, the procedure is similar to the December paper with the calculation of (6) for a single class of resonators and without equation (7). Both arguments present deviations from Boltzmann's original procedure, which have been deeply studied by historians (Klein 1962) (Kuhn 1978) (Darrigol 1988) (Gearhart 2002), but I think that some confusion still remains especially on two specific points: the adaptation of Boltzmann's procedure to the problem of radiation and the issue of maximization. I start with the former point.

Klein has stressed that Planck uses the total number of 'complexions' (ways of distributing the energy over the resonators of a certain class) instead of the number of complexions consistent with a certain distribution like Boltzmann did in 1877. It is doubtless true that, whichever statistics Planck is using, his calculation involves the total number of ways of distribution, but before deeming it a relevant deviation from Boltzmann's procedure we must examine the physical problem Planck is dealing with. To do so effectively, I will depart from Planck's original phraseology and resort to a modern language.

In Boltzmann's theory a macrostate is given by setting down the numbers of molecules that are placed in each phase space interval. In equilibrium the macrostate is described by Maxwell's distribution for a gas. On the contrary, in the case of radiation theory what is physically meaningful is the relation between the energy allocated to a certain frequency and the absolute temperature, and this relation presupposes the calculation of the total energy at each frequency. This distribution of energy over frequencies, condensed in (5), is a macrostate in Planck's theory, while the distribution of energy over a class of resonators (how many resonators are in a certain cell) is still a microstate.[10] The difference between Boltzmann's and Planck's physical situation is outlined by the following scheme:

In Boltzmann's case a microstate is an individual arrangement of molecules over different cells, while a macrostate is given by the number of molecules allocated in each cell. In contrast, in Planck's case, energy is allocated over each frequency and then divided into

[9] This is another difference between Planck's and Boltzmann's procedure: Planck considers only one resonator at a certain frequency, while Boltzmann had considered many molecules. However, this difference is immaterial for the issue in discussion. Resonators do not mutually interact and the exchange of energy takes place with the field only. This means that the simplest unity to be considered is the system 'resonator + field' while in Boltzmann's case the simplest unit is given by two molecules in collision. In other words, Planck's system of resonators bears a close resemblance on a thermodynamic system in a heat bath.

[10] On this important difference see also (Darrigol 2000) and (Darrigol 2001).

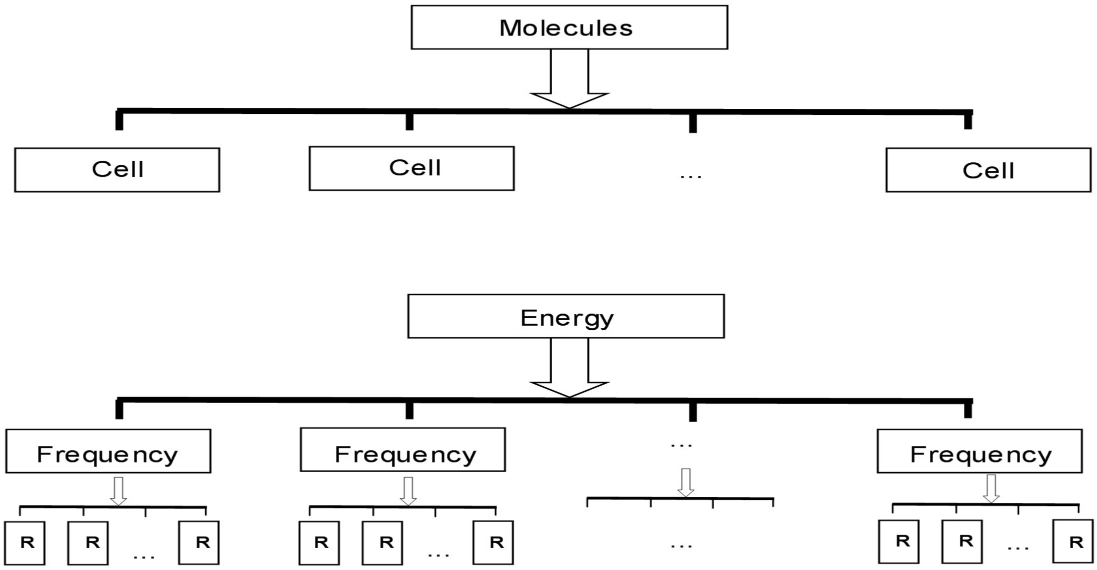

Figure 1: A schematic representation of the difference between Boltzmann's and Planck's distribution model.

elements, which are distributed over the resonators (R) that vibrate at that frequency. A microstate is an individual arrangement of resonators in cells, while a macrostate is the amount of energy allocated to each frequency.

To make this delicate point as clear as possible we can consider where the analogy between Planck and Boltzmann breaks down. To transform Boltzmann's case into Planck's, we can proceed as follows. Let us suppose N molecules, so that a state distribution is a vector:

$$o_k = \{n_0, n_1, \ldots\}.$$

quite analogous to vector (5). This vector tells us that n_i molecules are in the i-th cell. Let us now suppose that no permutations are possible among the different cells, precisely as no exchange of energy is possible among resonators of different frequency. This also means that one should suppose the molecules to be distributed in bunches of dimensionn_i over each cell — as described by the vector o_k — but not singularly, because otherwise one could obtain the same result as a permutation by distributing individual molecules over different cells in an alternative way. This is also due to the fact that the elements of energy only make physical sense when they are associated with a certain frequency; hence we can only ascribe energy as a whole to a frequency and then divide it into P elements, if necessary by taking 'for P an integer in the neighborhood.' Furthermore, since in Planck's model there are many resonators at each frequency, we must suppose that each cell is divided into a number of sub-levels, so that the molecules can be allocated in different ways within each cell. It is clear that in this modification of Boltzmann's model of distributing molecules over cells, a macrostate is defined by the vector o_k, but a microstate is no longer an allocation of individual molecules within the cells (because we cannot speak of individual molecules allocated in the cells), but an allocation of individual molecules over the sub-levels of each cell. In this case, the natural way of defining the probability of a distribution described by the vector ok is to make it proportional to the total number of ways of distributing the molecules over the sub-levels. This is exactly what Planck did. It is actually nothing more than an application of Boltzmann's general rule according to which the probability of a macrostate is proportional to the total number of microstates that leave the macrostate unchanged or that are consistent with that macrostate.

Two comments follow from this account. Firstly, it clearly shows what was really tricky in Planck's use of statistics. If the molecules have to be distributed in bunches, then it does not make sense to talk of the distribution of a single molecule because there is a sort of statistical correlation between molecules located in the same cell. However this crucial point is concealed by the fact that Planck is speaking of energy and his microstate concerns the distribution of resonators.[11] As soon as a corpuscular conception of energy gained footing and the resonators were left aside, the troublesome features of Planck's daring analogy would emerge, because what according to Planck was a macrostate immediately became one of many equiprobable microstates and the bunches of energy became groups of indistinguishable particles. This aspect would become clear when Bose and Einstein introduced the new quantum statistics in 1925. Absolutely crucial in that move was the concept of quantum of light, a concept that Planck refused to accept for many years.

Secondly, as long as one assumes the energy as a continuous quantity and uses resonators as a further level of description, the oddities in Planck's model remain hidden. From this point of view, Planck could have considered his use of statistics as a straightforward application of Boltzmann's doctrine because he ultimately evaluated the probability of a certain macrostate (7) by the total number of microstates consistent with that macrostate. He differed from Boltzmann mainly in his definition of macrostate, but, in turn, this depended more on the particular physical problem, than on the statistical argument used.[12] Planck is very clear on this topic in the first edition of the *Wärmestrahlung*(Planck 1906, 151):[13]

> "In this point lies the essential difference between the [radiation theory] case and that of a gas. Since in the latter the state was defined by the space and velocity distribution among the molecules [...]. Only when the distribution law is given, the state can be considered as known. On the contrary, in the former case, for the definition of the state, the calculation of the total energy E of the N resonators suffices; the specific distribution of the energy over the single resonator is not controllable, it is completely left [anheimgegeben] to chance, to elementary disorder."

More subtle is the issue concerning maximization, because it involves the concept of disorder and Planck's peculiar application of entropy to a single resonator. As early as December 1900, Planck noted that the notion of entropy is closely connected to the chaotic features of the system, but, at the same time, the way in which the disorder enters a system of resonators is radically different from the way in which the same concept is applied in gas theory.

[11] The model also shows that it is statistically irrelevant whether the resonators are distributed over the energy cells or the energy elements are distributed over the resonators. Planck's statistical leap lies elsewhere, and to bring this to light the resonators have to be abandoned.

[12] It should be noted in passing that, even in 1877, Boltzmann marginalized the complexions in the position space. In fact, Maxwell's distribution holds for the velocities only, hence a Boltzmann's macrostate is characterized by the product of the number of the favourable complexions in the velocity space and the total number of complexions in the position space. Of course Boltzmann does not consider this number explicitely because it is an unimportant constant. On this point, see (Hoyer 1980).

[13] A related statement can be found in (Planck 1915, 89).

In the former case, the disorder takes the form of 'natural radiation,' a particular assumption on the incoherent variation of the Fourier components of the waves exciting a resonator. This means that, while in a gas the disorder concerns the mutual interaction of many molecules at a given instant, in radiation theory the disorder is a feature of the interacting field and affects the evolution of a single resonator during a long interval of time. In other words, in a gas the disorder is a characteristic of a set of molecules at a given instant, while in cavity radiation, it is a characteristic of the temporal evolution of individual resonators. It is precisely this shift of meaning that allows Planck to introduce the entropy for a single resonator, a concept that would not make sense in gas theory.

Planck was aware of this fundamental difference from the outset because he mentions it in the introduction of his paper in March 1900(Planck 1900a). However, it is only in the first edition of the *Wärmestrahlung* that one can find an explicit statement on the influence of this difference on the application of the combinatorics to the resonators(Planck 1906, 150):

> "Briefly said: in the thermal oscillations of a resonator the disorder is temporal, while in the molecular motion of a gas it is spatial. However, this difference is not so important for the calculation of the entropy of a resonator as it might appear at first sight; because through a simple consideration can be stressed what is essential for a uniform treatment."

The "simple consideration" comes immediately after:

> "The temporal mean value U of the energy of a single resonator in an irradiating vacuum is evidently equal to the mean value of the energies calculated for a particular instant over a large number N of identical resonators that find themselves in the same radiation field but so far away from each other that their oscillations do not influence one another."

These statements clarify an analogous — but more obscure — passage in the December 1900 paper where Planck suddenly leaps from the temporal disorder of a single resonator to the calculation of the distribution of energy over a set of identical resonators without any justification. By the way, the equivalence between temporal behavior of an individual and instantaneous behavior of a group was rather common in statistical physics. For instance, Rudolf Clausius made use of a very similar assumption in his derivation of the virial theorem as early as 1870(Clausius 1870). Another possible source for Planck's confidence in the 'evidence' of the equivalence may have been Boltzmann's *Gastheorie*. In Section 35 of the second volume, Boltzmann comes up with a qualitative argument to extend the validity of the equipartition theorem proved for a gas to a thermal system in an arbitrary state of aggregation. The presupposition of the argument is a fact of experience: warm bodies reach a stable state of equilibrium. In such a state kinetic energy does not differ appreciably from the mean in the course of time. Moreover, equilibrium is independent of the initial conditions, so that Boltzmann could state(Boltzmann 1898, 310):

> "[W]e can also obtain the same mean values if we imagine that instead of a single warm body an infinite number are present, which are completely independent of each other and, each having the same heat content and the same external

> conditions, have started from all possible initial states. We thus obtain the correct averages values if we consider, instead of a single mechanical system, an infinite number of equivalent systems, which started from arbitrary different initial conditions."

This argument must have pleased Planck very much because it relies solely on the thermal equilibrium as an empirical fact and, as a consequence, on the elementary disorder.

In fact, in the December 1900 paper, we find an important hint in this direction that is missing in the *Wärmestrahlung*. Planck states that a key requirement for adopting his combinatorial derivation is "to extend somewhat the interpretation of the hypothesis of 'natural radiation,' which has been introduced by me into electromagnetic theory'(Ter Haar and Brush 1972, 39). The generalization of the hypothesis of natural radiation Planck is talking about is exactly the broader concept of elementary disorder. One can represent the temporal evolution of a single resonator by means of the combinatorics over a set of many identical resonators precisely because both models are disordered in the same sense and this concept of disorder is shared with gas theory as well. Therefore, one can apply in radiation theory the combinatorial methods that follow naturally from the notion of disorder in gas theory, because there is an analogous notion in radiation theory as well. In other words, the elementary disorder is supposed to bridge the gap between the physical description of a resonator interacting with a field and its combinatorial description as a set of identical copies. One can replace the former with the latter only if the temporal evolution of the system is disordered in the same sense as a distribution over the copies.

However, another ambiguity jumps out of the hat. On the one hand, Planck's analysis of the statistical model fits Boltzmann's procedure of seeking for an equilibrium distribution among the possible ways of allocating energy over the frequencies. On the other hand, if the single resonator is in equilibrium with the field during the long time considered, then the statistical model of N resonators must represent a state of equilibrium as well. If the temporal behavior of a single resonator in equilibrium is equal to the combinatorial behavior of a set of resonators as far as the average values are concerned, then all the configurations calculated in the set must represent equilibrium configurations.[14] From this point of view, a maximization procedure is conceptually unnecessary because all the ways of distributing the energy elements over the resonators are consistent with the equilibrium state. By using the fact that his physical problem (the derivation of the blackbody radiation law) is defined for the equilibrium state and by a daring application of the elementary disorder, Planck could escape the formal necessity of maximization. For these reasons, Planck was free to use or not to use the maximization without affecting the consistency of his reasoning or the analogy with Boltzmann's statistical arguments and he seems to have been aware of this since December 1900.

[14]Again, the temporal evolution of a resonator is analogous to a system in thermal contact with a heat bath. In both cases the exact energy of the system can fluctuate around a mean even though the equilibrium is maintained (Gearhart 2002).

5 Size does matter

In the previous section I have investigated two major differences between Planck's and Boltzmann's statistics, namely the calculation of the state probability by means of Boltzmann's total number of complexion and the absence of maximization. In both cases I have argued that Planck tried to reduce Boltzmann's procedure to basic formal rules in order to extend this procedure from gas theory to radiation theory. Therefore, Planck's subscription to Boltzmann's theoretical method was limited to the formal aspects and did not involve the micro-model.

A further departure of Planck's statistics from Boltzmann's is the fixed size of the energy element. Boltzmann divided the energy space (in 1868) or the phase space (in 1877) into cells, but the size of these cells was arbitrarily small and disappeared from the final result. By contrast, the size of Planck's elementary cells (*Elementargebiete*) was determined by a universal constant that played a crucial role in the final formula. In the *W armestrahlung*, he considers this fact "an essential difference" with the case of gas.(Planck 1906, 153). He also knew that if the size of the elementary cell drops to zero, one retrieves the incorrect Rayleigh-Jeans formula(Planck 1906, 156).

However, even if it was an essential difference in the procedure, Planck tried to assign to the non-vanishing size of the cell a place in his combinatorial approach. In December 1900, the division of the energy into elements of fixed size was performed in the context of the combinatorial calculation; the choice of the constant h was probably due to the law Planck had found in October. However, in the *Wärmestrahlung*, Planck discovers an important meaning of the universal constant. He shows that h can be interpreted as the elementary area, i.e. area of equal probability, in the phase space of the resonator.

This interpretation actually strengthens the link with Boltzmann for two reasons. First, one of the main steps in Boltzmann's argument was the partition of the phase space of a gas in regions of equal volume. Olivier Darrigol (Darrigol 1988) and Ulrich Hoyer (Hoyer 1980) have pointed out that even if these volumes are arbitrarily small, they cannot disappear, otherwise Boltzmann's integrals are doomed to diverge. By shifting the quantization from energy cells to regions of the phase space, Planck was therefore reinforcing the formal analogy between his procedure and Boltzmann's. At the same time, the non-vanishing size of the cell could be understood as an aftermath of the physical problem, an inconvenience of the incomplete analogy, not a flaw in the statistical formalism.

Second, both in December 1900 and in January 1901, Planck had stated that one fundamental assumption of his theory was the equiprobability of the complexions, but he had not clarified the status of this contention, one that, in his opinion, had to be decided empirically. As we will see more clearly in the next section, Boltzmann justified the equiprobability of complexions by appealing to the Liouville theorem and a particular definition of probability. With the special partition of the phase space in 1906, Planck was able to introduce a justification of the equiprobability that relied on general electrodynamics and a universal constant only. Thus, he was convinced that an important gap in his approach had been filled.

In the second edition of the *Wärmestrahlung* (1913), Planck further improved the position of the constant h in his theory by showing that it is closely related to a general thermodynamical result: Nernst's theorem. This theorem, discovered by Walther Nernst in 1906, entails the existence of an absolute definition of the entropy and, from Planck's viewpoint, this implies an absolute definition of probability.[15] But probability hinges upon the partition of the phase space, and that, accordingly, must be also possible in an absolute way. On the necessity of fixing a finite magnitude for the phase cell as a consequence of Nernst's theorem, Planck says(Planck 1959, 125):

> "That such a definite finite quantity really exists is a characteristic feature of the theory we are developing, as contrasted with that due to Boltzmann, and forms the content of the so-called hypothesis of quanta. As readily seen, this is an immediate consequence of the proposition [...] that the entropy S has an absolute, not merely relative, value; for this, according to [$S = k \log W$], necessitates also an absolute value for the magnitude of the thermodynamical probability W, which, in turn [...], is dependent on the number of complexions, and hence also on the number and size of the region elements which are used."

Thus, by exploiting Nernst's thermodynamic result, which he had accepted enthusiastically from the outset, Planck was able to give to the finite magnitude of the cell in the phase space a meaning that was at the same time probabilistic and thermodynamic.

The previous discussion sheds some light upon the most vexed issues of Planck's use of statistics. In particular, it shows that the alleged differences between Planck's and Boltzmann's statistics chiefly stem from the physical problem Planck had to cope with. Thus, the counting procedure is a plain application of Boltzmann's idea of calculating the microstates consistent with a certain macrostate. What is different is the definition of a macrostate. That arises because the spectral distribution required in radiation theory has a different structure from the velocity distribution used in gas theory. Likewise, the maximization procedure was dispensable because only the equilibrium state has an empirical meaning for heat radiation and because of the particular concept of disorder Planck had fostered. Lastly, Planck tried to embody the finite magnitude of the energy elements in his statistical procedure through the interpretation of elementary region of the phase space as an elementary region of probability.

These considerations and the ambiguity of the combinatorial formula discussed above seem to indicate that Planck tried to apply Boltzmann's statistical procedure as closely and correctly as possible but at the same time considered it as a set of formal rules to treat mysterious micro-processes that did not give information on the physics involved in such processes. Thus a careful analysis of Planck's statistics supports neither Klein's opinion that the differences between Planck and Boltzmann suggest a Planck's commitment to discontinuity, nor Kuhn's claim that the similarities speak for a subscription of Planck to continuity. Instead, Planck's treatment of the statistical issues shows that he tried to squeeze the formal essence out of Boltzmann's theory independently of physical assumptions on the micro-phenomena. Moreover, it is significant that the most explicit statements about these issues are to be found in the 1906 *Wärmestrahlung*. This fact suggests that on the problem continuity/discontinuity Planck remained uncommitted during the period 1900-1906 and that the *Wärmestrahlung* should be read as a systematic presentation of the reasons for this lack of commitment. As I will show in the remaining of the paper, this attitude was part of Planck's general views about the relation between statistics and physics.

[15]In fact, the generalization of Nernst's theorem to the entropy was proposed by Planck in 1911. For an overview of the problems connected with Nernst's theorem, see (Kox 2006).

6 Organized disorder

During his scientific career, Boltzmann expounded his philosophical position in various essays and was involved in many scientific disputes especially concerning the necessity of the atomistic hypothesis and the Energeticism, but, unexpectedly, he rarely discusses the role of statistics in his theory. Some hints about his general opinion on this issue can be drawn from the final section of his 1868 paper(Boltzmann 1909, I, 92-96) (Boltzmann 1898, 448-449) where he explains how statistical considerations enter his treatment of mechanical problems.[16] Since his original argument is pretty obscure I will try to reframe it from a more modern perspective. There are three elements:

1. Probability as sojourn time: the probability of a certain physical state represented by a region of the phase space of the system is the ratio between the time the system spends in that region and the total time considered (supposed to be very long).
2. Liouville's theorem: a well-known result of general dynamics stating that if a system evolves according to the Hamiltonian equations of motion, then all the phase regions it passes through have the same volumes.
3. Ergodic hypothesis: a system will pass through all the phase regions consistent with its general constraints (e.g. the conservation of energy) provided that its evolution lasts long enough.

Boltzmann's argument goes as follows. Let us divide the trajectory time of the system into intervals of magnitude Δt, so that a phase trajectory for the system is a sequence of states $\Sigma_t, \Sigma_{t+\Delta t}, \ldots, \Sigma_{t+n\Delta t}$.

From Liouville's theorem, it follows immediately that all these phase volumes are equal. However, since the system spends the same quantity of time Δt in each state, they are also equiprobable by definition (1). Therefore, the probability assigned to a certain state is proportional to the phase space volume of that state. If one now assumes that the ergodic hypothesis holds, then the system will pass through all the phase space regions consistent with its general conditions, and this means that, due to the deterministic evolution, there is only one trajectory filling up all the allowed phase space. Hence, one can describe the long-run behavior of the system by simply dividing the phase space into regions of equal volume (namely of equal probability) and calculating the number of regions corresponding to a certain macrostate. In other words, one can replace the temporal description of the long-run evolution of the system with a combinatorics on the phase space.

Of course, this cannot be considered a formally satisfactory argument because of the problems connected with the ergodic hypothesis, but the general idea is clear enough: Boltzmann presented his use of statistics in mechanics as an account of the behavior of the mechanical system that is as correct as the temporal description that the mechanics itself can provide. It is valid as long as some conditions hold in the system itself, notably the ergodic hypothesis. More importantly, the application of statistics does not rely on our ignorance of

[16]Of course, from a formal point of view, the main goal of the section is a proof of the uniqueness of Maxwell's distribution by using the ergodic hypothesis. But I think that the particular view of the relation between statistics and mechanics implicit in this argument should not be underestimated.

the detailed state of the system, namely on our epistemic status, but on a certain kind of behavior of the mechanical system. The key point is that this justification amounts to an attempt to deeply integrate statistics and mechanics, to seek for the mechanical conditions of an application of statistical arguments. Boltzmann stresses this aspect also in the introduction of his 1872 paper(Boltzmann 1909, I, 317):

> "It would be an error to believe that there is an inherent indetermination in the theory of heat because of the usage of the laws of the calculus of probability. One should not mistake a law only incompletely proved, whose soundness is hence problematic, for a completely demonstrated law of the calculus of probability; the latter represents, like the result of any other calculus, a necessary consequence of given premises, and, if they are true, it is borne out in experience, as soon as many enough cases are observed what is always the case in the theory of heat because of the enormous number of molecules."

Planck's justification goes in exactly the opposite direction. The relation between statistics and electrodynamics is explained at the beginning of the fourth chapter of the first edition of the *Wärmestrahlung* and starts with the following dilemma(Planck 1906, 129):

> "Since with the introduction of probabilistic considerations into the electrodynamic theory of heat radiation, a completely new element, entirely unrelated to the fundamental principles of electrodynamics enters into the range of investigations, the question immediately arises, as to the legitimacy and the necessity of such considerations. At first sight we might be inclined to think that in a purely electrodynamical theory there would be no room at all for probability calculations. Since, as everybody knows, the electrodynamic field equations together with the initial and boundary conditions determine uniquely the temporal evolution of an electrodynamic process, any consideration external to the field equations would be, in principle, unauthorized, and, in any case, dispensable. In fact, either they lead to the same results as the fundamental equations of electrodynamics and then they are superfluous, or they lead to different results and in this case they are wrong."

However, Planck adds that the dilemma arises from an incorrect understanding of the relation between the microlevel and the macrolevel. For the sake of convenience this relation can be summarized by another suitable scheme:

A certain macrostate is combinatorially related with many different microstates (micro 1, micro 2, etc.), which evolve according to dynamical laws. However, Planck claims that we cannot directly apply a dynamical analysis to the system because we do not know which of the many theoretically possible microstates actually hold. Our empirical measurements on the macrostate are able to supply only mean values, which are consistent with many different combinations of exact values and then with many different microstates. Consequently we do not have an unambiguous initial condition to start from. For this reason, the application of dynamical laws to the microstates is ambiguous.

Moreover, even if we know that the result of the dynamical evolution of the microstates is a set of new microstates (micro′ 1, micro′ 2, etc.) whose overwhelming majority is combinatorially related to the equilibrium state, the plain application of combinatorial arguments

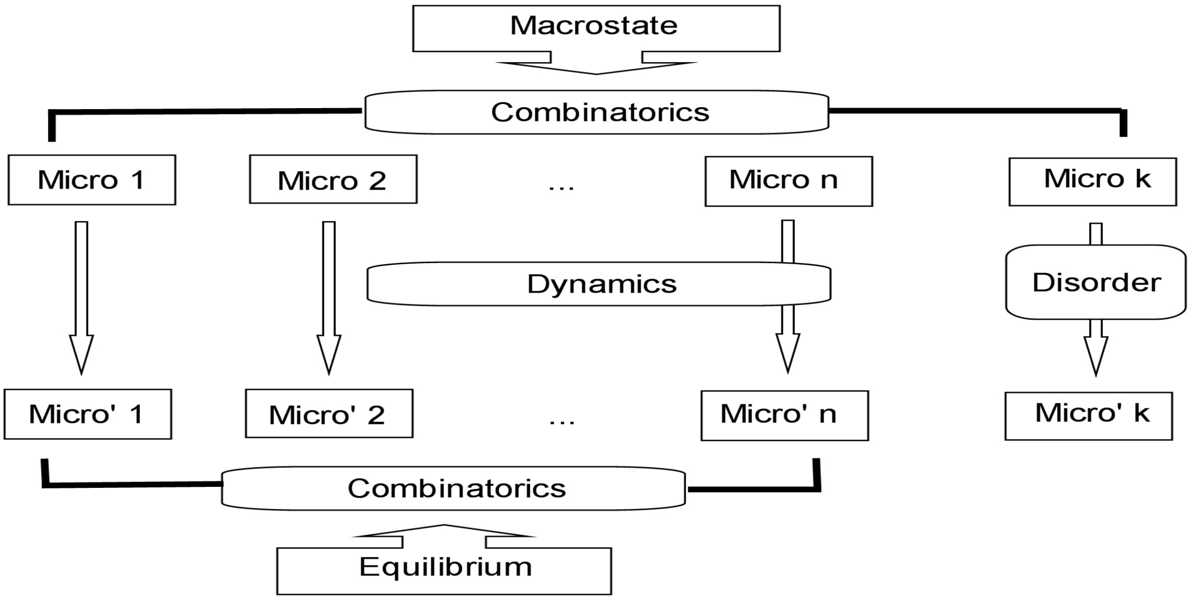

Figure 2: The effect of the molecular disorder in Planck's theory.

is ambiguous as well, because there are some, very few indeed, microstates that might lead to an anti-thermodynamic evolution in which the entropy decreases. Thus, to replace the dynamical arguments with the combinatorial ones, and to retrieve an unambiguous (and deterministic, in Planck's view) picture, one has to reject the anti-thermodynamic microstates. This result is achieved by the hypothesis of the elementary disorder which "states nothing more than that exceptional cases, corresponding to special conditions which exist between the separate quantities determining the state and which cannot be tested directly, do not occur in nature."

So far so good. Nevertheless, scholars agree that Planck's position began to change in 1908 when Lorentz convinced him that there was no chance of including the quantum in the theory of electron. This process eventually led to Planck's overt acceptance of statistical laws in his 1914 speech at the Founder Day of the University of Berlin(Planck 1914). The process was however gradual and piecemeal: Planck's search for acceptable modifications of his original conception resulted in his second theory of radiation in 1911. As Needell has shown, in this theory Planck eventually gives up the individual resonator and uses an approach similar to Gibbs'.[17]

In the course of this maturation Planck's general views about the relation between statistics and physics were the last thing to change. For instance, Planck maintained his view of molecular disorder until the fourth edition of the *Wärmestrahlung* in 1921! In the second edition, written in 1912, Planck expressed the difference between microstates and macrostates in a colorful way(Planck 1959, 121):

"The microscopic state is the state as described by a mechanical or electro-

[17]For many years Planck esteemed Gibbs' approach to statistical mechanics as inferior to Boltzmann's. This was likely due to Planck's peculiar understanding of probability concept in physics and, later, to Ehrenfest's review article. The situation changed in the mid-1910s, when Planck resorted to Gibbs' approach to deal with the quantum theory of monoatomic gas.

dynamical observer; it contains the separate values of all coordinates, velocities, and field strengths. The microscopic processes, according to the laws of mechanics and electrodynamics, take place in a perfectly unambiguous way; for them entropy and the second principle of thermodynamics have no significance. The macroscopic state, however, is the state as observed by a thermodynamic observer; any macroscopic state contains a large number of microscopic ones, which it unites in a mean value. Macroscopic processes take place in an unambiguous way in the sense of the second principle, when, and only when, the hypothesis of the elementary disorder is satisfied."

It is worthwhile noting how deeply Planck's condition of disorder differs from Boltzmann's. Firstly there is a different relation to the statistical formalism: for Boltzmann, disorder is the prerequisite to introducing combinatorial arguments that can completely replace the dynamical ones because disorder permits all theoretically allowed states, even the most improbable ones, to occur. By contrast, for Planck disorder is able to block the improbable states in order to pave the way to the triumph of the mean values. Secondly, while for Boltzmann disorder is a constitutive feature of the system, something that, in a sense, belongs to both the microlevel and the macrolevel, for Planck it belongs exclusively to the microlevel, to a realm populated by hypothetical mechanical and electrodynamical observers who see a completely different world that is utterly concealed from us. As already pointed out by Allan Needell and Olivier Darrigol, Planck's elementary disorder concerns the mysterious interaction of matter and radiation and hence is part of the uncontrollable and inaccessible microworld.

Until 1914 Planck by and large views the statistical arguments not as another, equally correct, way of investigating mechanical problems, but the only way at our disposal to handle the complicated and unreachable realm of the interaction between matter and radiation. This point clearly emerges again in the second edition of the *Wärmestrahlung*. The general architecture of the book is remarkably different from the first edition. In particular, while in the first edition the statistical arguments were a means to overcome the issues left open by the dynamical approach, in the second edition the dynamical part of the theory is constrained to satisfy the general results obtained from the statistical analysis because "the only type of dynamical law admissible is one that will give for the stationary state of the oscillators exactly the distribution densities [...] calculated previously"(Planck 1959, 152).

The new dynamical law is the famous quantum emission hypothesis that Planck introduces with the following words(Planck 1959, 153):

> "[W]e shall assume that the emission does not take place continuously, as does the absorption, but that it occurs only at certain definite times, suddenly, in pulses, and in particular we assume that an oscillator can emit energy only at the moment when its energy of vibration, U, is an integral multiple n of the quantum of energy $\epsilon = h\nu$. Whether it then really emits or whether its energy of vibration increases further by absorption will be regarded as a matter of chance. This will not be regarded as implying that there is no causality for emission; but the processes which cause the emission will be assumed to be of such a concealed nature that for the present their laws cannot be obtained by any but statistical methods."

Even though the statistical part takes the leading role in the second edition of the *Wärmestrahlung*, the judgment on the dynamical assumptions it provides is still suspended.

7 Who cares about the microworld?

The discussion in the previous sections seems to suggest that we should be very cautious in attributing any commitment to Planck on the grounds of his usage of the statistical formalism. Boltzmann had tried to integrate mechanics and statistics by showing that the conditions for applying statistical argument are to be sought in some particular mechanical behavior. However, in Planck's view statistical arguments should be confined to the impenetrable processes taking place at the microlevel: electromagnetism and statistics are completely dis-integrated. This general attitude fits well with Planck's statements on the technical issues I have discussed in the previous sections and reinforces the thesis that somewhere in the period 1900-1906 Planck may have envisioned an underdetemination of the statistical formalism as far as the micro-model is concerned. In this concluding section, I discuss in more detail the concept of virtual observer, the elementary disorder and the ensuing notion of microworld.

Boltzmann's microworld is assembled by means of conceptual elements coming from the macroworld, e.g. molecules as mechanical points or centers of force, elastic collisions and so on, and from the statistical formalism meant as a new tool for investigating physical reality. By integrating a statistical formalism with a mechanics of the macrolevel Boltzmann established a bi-directional relation between two levels of reality. The conceptual apparatus used to describe the microlevel eventually leads to a reinterpretation of the macroscopic laws in terms of the statistical viewpoint. On the contrary, by disintegrating dynamical and statistical formalism and by reducing the latter to a computational device, Planck ends up with a microworld that is shaped by the macroscopic conceptual structure and it is unable to support any reinterpretation of the macrophenomena. Ultimately, statistics is not supposed to give us a description of how the world is, but only how to handle chaotic situations.

We have seen that in the second edition of the *Wärmestrahlung*, Planck uses the concept of 'virtual' (micro- or macro-) observer to figure out a more intuitive definition of the micro- and macrostate. Actually, this concept had made its first appearance in the third lecture of a series of conferences held by Planck at the Columbia University in 1909. Here Planck points out that the contradiction between the reversibility of the micro-phenomena and the irreversibility of the thermodynamical laws stems from different definitions of state. The physical state envisioned by the micro-observer, that is "a physicist [...] whose senses are so sharpened that he is able to recognize each individual atom and to follow it in its motion,"(Planck 1915, 47) is fundamentally different from the state of the usual macro-observer, because the former perceives exact values and the latter only mean values. Clearly the virtual micro-observer is nothing but a projection of a macroscopic one.

The primacy of the macrolevel is mirrored in Planck's notion of elementary disorder. For Boltzmann elementary disorder is a feature concerning the individual configurations of molecules and, more importantly, the ultimate justification of the introduction of statistical arguments. In his *Gastheorie*, Boltzmann distinguishes the concepts of molar and molecular disorder. The former concerns the fact that the mean values of the mechanical quantities, e.g. the molecular velocity, do not vary from a spatial region to another occupied by the gas. The latter is, significantly enough, introduced by means of its opposite, the molecular order(Boltzmann 1898, 40):

> "If the arrangement of the molecules also exhibits no regularities that vary from one finite region to another — if it is thus molar-disordered — then never-

theless groups of two or a small number of molecules can exhibit definite regularities. A distribution that exhibits regularities of this kind can be called molecular-ordered. We have a molecular-ordered distribution if — to select only two example from the infinite manifold of possible cases — each molecule is moving towards its nearest neighbor, or again if each molecule whose velocity lies between certain limits has ten much slower molecules as nearest neighbors."

While Boltzmann's concept of disorder dives straightforwardly into the details of the molecular arrangements, and realizes the possibility of a statistical interpretation of thermodynamics, Planck's analogous notion only deals with the coherence of the Fourier components of radiation (a macroscopic concept) and, even more remarkably, it is supposed to block any anti-thermodynamic evolution. The difference is extremely important. If a system is molecular-ordered in Boltzmann's sense, then only a subset of the theoretically possible states will actually be realized, whereas, if the system is molecular-disordered, there is nothing, in the initial configuration, to prevent all possible states from occurring and this is precisely the condition for applying statistical methods. In other words, Boltzmann's disorder does not block a particular kind of evolution, but simply rules out the occurrence of undesired configurations where only a subset of evolutions is possible.[18]

By contrast, in the crucial years Planck conceives elementary disorder as a limitation on the statistical formalism itself because some of the theoretically possible configurations cannot take place. As a result, one obtains a new definition of microstate(Planck 1915, 50):

> "The micro-observer needs only to assimilate in his theory the physical hypothesis that all those special cases in which special exceptional conditions exist among the neighboring configurations of interacting atoms do not occur in nature, or, in other words, that the micro-states are in elementary disorder. Then the uniqueness of the macroscopic process is assured and with it, also, the fulfillment of the principle of increase of entropy in all direction."

Planck's change of meaning and function of elementary disorder has far-reaching consequences.[19] By means of the virtual micro-observer and of elementary disorder, Planck foists upon the constitution and the formalism of the microworld a series of constraints coming from the macroworld. Since Planck's parasitic microworld is completely shaped by conceptual material and formal requirement coming from the macro-level, and since its characteristic formalism is conceived to be nothing but a set of computational devices, the only conceptual feedback it can warrant are those leading to derivations of the macro-laws, like the

[18]See for example (Boltzmann 1898, 451): "only singular states that *continually* deviate from probable states must be excluded" (italics added).

[19]A further support of the thesis that Planck's notion of elementary disorder differs from Boltzmann's comes again from the third lecture. In a note, he claims that Poincare's recurrence theorem calls for a careful formulation of the hypothesis of the elementary disorder in order to avoid the, even only theoretical, possibility of a low-entropy evolution. In particular, Planck's way out is the statement that "absolutely smooth walls do not exist in nature" (Planck 1915, 51). On the contrary, Boltzmann reckoned the recurrence theorem in its original formulation to be perfectly consistent with his notion of disorder: "[t]he fact that a closed system of a finite number of molecules [...] finally after an inconceivably long time must again return to the ordered state, is therefore not a refutation, but rather indeed a confirmation of our theory" (Boltzmann 1898, 443).

black-body radiation law.[20]

Analogously, Planck's conversion to Boltzmann's point of view initially seems to concern rather the notion of irreversibility as an 'emerging' phenomenon than the statistical interpretation of the macroworld(Planck 1915, 97):

> "[I]rreversibility does not depend upon an elementary property of a physical process, but rather depends upon the ensemble of numerous disordered elementary processes of the same kind, each one of which individually is completely reversible, and upon the introduction of the macroscopic method of treatment."

Of course, the emerging notion of irreversibility is only the premise of Boltzmann's conception, but Planck is not in the position to accept the consequences. Ultimately, his way of justifying the use of statistics in physics is closely related with his prudent use of the statistical formalism as well.

In conclusion, this analysis articulates an argument in support of the weak thesis. We have seen that Planck probably followed Boltzmann's statistical procedure closely, but he also had to adapt it to the new physical problem. This means that Planck was not subscribing any commitment with regards to the reality of the micro-processes involved in the statistical picture. Moreover, his statements on the relation between statistics and physics during the period 1906-1913 buttress this claim. By breaking the Boltzmannian conceptual links between statistics and dynamics, between micro- and macrolevel that would have forced him to take a clear position, by constructing a microworld completely shaped by the macroworld, and by denying an autonomous status to the statistical formalism, he could safely distance himself from dangerous connections between apparently incompatible formalisms.

Even if Planck had adopted the same statistical arguments as Boltzmann, he understood the role of statistics in a completely different way and, more importantly, he was unwilling to integrate the statistical considerations with the physics his theory relied on. Thus, though Planck's and Boltzmann's name are often associated in the history of quantum theory, they seem to be an "odd couple," because, like the characters of the famous movie, their attitudes on the fundamental problems could not have been more different.

This perspective also gives us some clues to understand the relations between Planck and his contemporaries. In fact, Planck's attitude was not a completely idiosyncratic one. On the contrary, he was placing himself within an illustrious thermodynamic tradition including Clausius and Helmholtz. According to this tradition, the hypotheses concerning the uncontrollable microlevel have to be avoided as long as they are not absolutely necessary and, if this is the case, only minimally and cautiously introduced. Famously, Clausius, who was Planck's guiding spirit in thermodynamics, refused to use the distribution function until his late years and when he was forced to give some statistical assumptions on the behavior of molecules, he always limited himself to what was strictly necessary to arrive at his final result. Helmholtz, Planck's predecessor in Berlin, endorsed a pure thermodynamics even when mechanical concepts were used, as in his papers on the monocycle.[21]

[20]Incidentally, the monodirectionality of this relation between micro- and macroworld is part of the reason why Planck, even acknowledging the generality of Boltzmann's approach, did not develop a statistical mechanics. Instead, a decisive move in this direction was performed by Einstein who restored the bidirectional conception of micro- and macroworld and the autonomy of the statistical formalism(Renn 1997) (Uffink 2006).

[21]On this 'Berlin style' in physics, see (Jurkowitz 2002). On Planck's style see (Schirrmacher 2003) and the

On the other side of the river stood Boltzmann, who was not afraid of introducing bold assumptions on the behavior of the molecules and of dealing with them by using the conceptual tools of statistics. Paul Ehrenfest as well as Albert Einstein belonged to the same tradition and, unsurprisingly, they did not understand and therefore sharply critized Planck's use of statistical arguments. This fundamental division affected a large part of the relations between statistical mechanics and the early quantum theory.

Acknowledgement

This paper has been written as part of the History of Quantum Physics Project of the Max Planck Institute for the History of Science and the Fritz Haber Institute of the Max Planck Gesellschaft. A preliminary version was presented at the HQ1 Conference on the History of Quantum Physics (Berlin, 2-6 July 2007). I am grateful to all the participants and collegues at the Max Planck Institute, in particular to Christoph Lehner and Christian Joas, who have given important and interesting comments. I am indebted to Jürgen Renn for such a large body of advice that it would take another paper to mention them all. The paper has benefitted a lot of my long discussions with Ed Jurkowitz. Finally, special thanks go to Clayton Gearhart whose stimulating criticisms have enormously contributed to the final result.

working paper, much longer and with the same title, that can be downloaded from the website of the Munich Center for the History of Science and Technology. I thank Arne Schirrmacher for this information.

References

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

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.: 1877, 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.: 1898, *Vorlesungen über Gastheorie*, Barth, Leipzig.

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

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

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

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

Darrigol, O.: 1988, Statistics and combinatorics in early quantum theory, *Historical Studies in the Physical Science* **19**, 18–80.

Darrigol, O.: 1991, Statistics and combinatorics in early quantum theory, II: Early symptoma of indistinguishability and holism, *Historical Studies in the Physical Science* **21**, 237–298.

Darrigol, O.: 2000, Continuities and discontinuities in Planck's Akt der Verzweiflung, *Annalen der Physik* **9**, 851–860.

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

Ehrenfest, P. and Kamerlingh Onnes, H.: 1915, Vereinfachte Ableitung der kombinatorischen Formel, welche der Planckschen Strahlungstheorie zugrunde liegt, *Annalen der Physik* **46**, 1021–1024.

Galison, P.: 1981, Kuhn and the quantum controversy, *The British Journal for the Philosophy of Science* **32**, 71–85.

Garber, E.: 1976, Some reactions to Planck's law, *Studies in History and Philosophy of Science* **7**, 89–126.

Gearhart, C.: 2002, Planck, the quantum, and the historians, *Physics in Perspective* **4**, 170–215.

Heilbron, J.: 1986, *The Dilemmas of an Upright Man. Max Planck as Spokesman for German Science*, University of California Press, Berkeley.

Hoyer, U.: 1980, Von Boltzmann zu Planck, *Archive for History of Exact Sciences* **23**, 47–86.

Jost, R.: 1995, Planck-Kritik des T. Kuhn, *Das Märchen vom Elfenbeinernen Turm. Reden und Aufsätze*, Springer, Berlin.

Jurkowitz, E.: 2002, Helmholtz and the liberal unification of science, *Historical Studies in the Physical and Biological Sciences* **32**, 291–317.

Jurkowitz, E.: 2007, Resolving the contradictions between principles: Planck's exploration of the micro-world. Unpublished manuscript.

Jurkowitz, E.: 2008, Planck's unification of physics with/in German liberal culture, HQ2 Conference on the history of quantum phyiscs, Utrecht.

Klein, M. J.: 1962, Max Planck and the beginnings of the quantum theory, *Archive for History of Exact Sciences* **1**, 459–479.

Klein, M. J.: 1963, Einstein's first papers on quanta, *Natural Philosopher* **2**, 59–86.

Klein, M. J.: 1964, Einstein and the wave-particle duality, *Natural Philosopher* **3**, 3–49.

Klein, M. J.: 1966, Thermodynamics and quanta in Planck's work, *Physics Today* **19**, 23–32.

Klein, M. J., Shimony, A. and Pinch, T.: 1979, Paradigm lost? A review symposium, *Isis* **70**, 429–440.

Koch, M.: 1991, From Boltzmann to Planck: On continuity in scientific revolutions, *in* W. R. Woodward and R. S. Cohen (eds), *World Views and Scientific Discipline Formation*, Kluwer Academic, Amsterdam, pp. 141–150.

Kox, A.: 2006, Confusion and clarification: Albert Einstein and Walther Nernst's heat theorem, 1911-1916, *Studies in History and Philosophy of Modern Physics* **37**, 101–114.

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

Kuhn, T.: 1984, Revisiting Planck, *Historical Studies in the Physical Science* **14**, 232–252.

Natanson, L.: 1911, Über die statistische Theorie der Strahlung, *Physikalische Zeitschrift* **12**, 659–666.

Needell, A.: 1980, *Irreversibility and the Failure of Classical Dynamics: Max Planck's Work on the Quantum Theory, 1900-1915*, PhD thesis, University of Michigan, Ann Arbor.

Planck, M.: 1900a, Entropie und Temperatur strahlender Wärme, *Annalen der Physik* **4**, 719–737.

Planck, M.: 1900b, Zur Theorie des Gesetzes der Energieverteilung im Normalspektrum, *Verhandlungen der Deutschen Physikalische Gesellschaft* **2**, 237–245.

Planck, M.: 1901, Über das Gesetz der Energieverteilung im Normalspektrum, *Annalen der Physik* **4**, 553–563.

Planck, M.: 1906, *Vorlesungen über die Theorie der Wärmestrahlung*, Barth, Leipzig.

Planck, M.: 1914, *Dynamische und statistische Gesetzmässigkeit*, Barth, Leipzig.

Planck, M.: 1915, *Eight Lectures on Theoretical Physics*, Dover, New York.

Planck, M.: 1958, *Physikalische Abhandlungen und Vorträge*, Vieweg u. Sohn, Braunschweig.

Planck, M.: 1959, *The Theory of Heat Radiation*, Dover, New York.

Renn, J.: 1997, Einstein's controversy with Drude and the origin of statistical mechanics, *Archive for History of Exact Sciences* **51**, 315–354.

Schirrmacher, A.: 2003, Experimenting theory: The proofs of Kirchhoff's radiation law before and after Planck, *Historical Studies in the Physical and Biological Sciences* **33**, 299–335.

Ter Haar, D. and Brush, S. (eds): 1972, *Planck's Original Papers in Quantum Physics*, Taylor and Francis, London.

Uffink, J.: 2006, Insuperable difficulties: Einstein's statistical road to molecular physics, *Studies in History and Philosophy of Modern Physics* **37**, 36–70.

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