

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

MASSIMILIANO BADINO

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 the 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.

A Vexed Problem

In his epoch-making paper of December 1900 on the black-body radiation,¹ for the first time Max Planck made use of combinatorial arguments. Although it was a difficult step to take, a real “act of desperation” as he would call it later, Planck pondered it deeply and never regretted it. As he wrote to Laue 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.”²

The difficulty involved in this step was the adoption of a way of reasoning Planck had been opposing for a long time: Ludwig Boltzmann's statistical approach. But, even after accepting the necessity of introducing statistical considerations into radiation theory, the application of Boltzmann's theory of complexions to the particular problem of finding the spectral distribution of the cavity radiation was not a straightforward one. In fact, the final result seems to bear only a partial resemblance to Boltzmann's original arguments and the opinions of the scholars are split about the correct interpretation of the relation between Planck's and Boltzmann's statistical procedure. The importance of the issue is enhanced by the fact that, in the secondary literature, close relations can be found with the problem of continuity or discontinuity of energy, i.e. whether Planck conceived the quantization of energy as something real or merely as a computational device. With unavoidable simplifications, we can divide the positions about the historical problem of quantization into three main categories: discontinuity thesis, continuity thesis and weak

¹(Planck, 1900b); (Planck, 1958), 698–706.

²Quoted in (Heilbron, 1986, p. 5).

thesis. First, we have the discontinuity thesis according to which Planck worked with discrete elements of energy. As early as 1962, Martin Klein, in a series of seminal papers on this historical period,³ argued along this direction claiming more or less explicitly that in December 1900 Planck introduced the quantization of energy even though he might have not been perfectly aware of the consequences of this step. Furthermore, in recent years, Res Jost has polemically endorsed Klein's classical thesis against the most distinguished upholder of the continuity thesis, namely Thomas Kuhn.⁴

Indeed, in 1978, Thomas Kuhn⁵ claimed that in December 1900, and at least until 1908, Planck was thinking in terms of continuous energy and that his energy elements were merely a shortcut to talk of a continuous distribution over energy cells. The discontinuity entered physics as late as 1905–1906 through the work of Paul Ehrenfest and Albert Einstein. I will call this claim the continuity thesis. Both the discontinuity and the continuity thesis argue for a definite commitment of Planck's on the issue of quantization. In a sense also Olivier Darrigol⁶ can be numbered among the upholders of this thesis, even though his position points less straightforwardly towards a clear commitment and it is halfway between continuity and the third option, the weak thesis.⁷

As a matter of fact, the advocates of the weak thesis claim that we cannot single out any clear-cut position of Planck's on the issue of the reality or the physical significance of quantization. The reasons and the meaning of this absence of decision might be different. Allan Needell, for instance, has convincingly argued that this issue was simply out of the range of Planck's interests.⁸ The exact behaviour of the resonator belongs to a domain of phenomena, namely the micro-phenomena, Planck was unwilling to tackle from the very beginning of his research program. Hence, the question whether the resonator really absorbs and emits quanta of energy was irrelevant in Planck's general approach. Much more important to understand his theory, Needell suggests, is to look at the role played by the absolute interpretation of the second law of thermodynamics.

A similar contention was shared by Peter Galison who maintains that, in general, it is not wise to ascribe strong commitments to scientists working in a period of scientific crisis.⁹ Recently, Clayton Gearhart¹⁰ has suggested another option for the supporters of the weak thesis, holding that even if Planck might have been interested in the issue of the quantization of energy, for various reasons he was unable or unwilling to take a plain position on the status of the energy elements in his printed writings while "he was often more open in discussing their implication in his correspondence."¹¹ His papers in the crucial period 1900–1913 show an incessant shift and change of emphasis between continuity and discontinuity so that both a literal interpretation (the starting point of Klein's thesis) and a re-interpretation (the main tool of Kuhn's approach) of these papers wind up to be misleading. For Gearhart what Planck lacked was not the interest

³(Klein, 1962); (Klein, 1963a); (Klein, 1963b); (Klein, 1964); (Klein, 1966). The original source of the discontinuity thesis is (Rosenfeld, 1936).

⁴(Jost, 1995); see also (Koch, 1991).

⁵(Kuhn, 1978) and (Kuhn, 1984). See also (Klein, Shimony, & Pinch, 1979).

⁶(Darrigol, 1988) and (Darrigol, 1991).

⁷Darrigol's position is wholly presented in his recent papers (Darrigol, 2000) and (Darrigol, 2001).

⁸(Needell, 1980).

⁹(Galison, 1981).

¹⁰(Gearhart, 2002). (Kangro, 1970) holds a weak thesis on Planck's commitment as well.

¹¹(Gearhart, 2002, p. 192).

in the issue of quantization (which is testified by his letters and by his trying different approaches), but rather a concluding argument to make up his mind in a way or in another. Also Elisabeth Garber argued in the direction of a certain ambiguity in Planck's work when she claimed that his theory was perceived as escaping the pivotal point: the mechanism of interaction between matter and radiation.¹²

In this debate, a central role is played by the statistical arguments and, notably, by the comparison between Planck's usage of them and Boltzmann's original doctrine because, in effect, the statistical procedure is the first and the main step in Planck's theory where any discontinuity is demanded. Literally interpreted, Planck's statements in December 1900 and in many of the following papers seem to suggest a counting procedure and a general argument that strongly differ from Boltzmann's, pointing to the direction of real and 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 this makes sensible to think that the physical interpretations behind them must diverge as well.¹³

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

Therefore, both the advocates of the continuity and those of the discontinuity thesis see an intimate connection between Planck's usage of statistical arguments and the issue of the quantization of energy. Nevertheless, Kuhn started from the ambiguity of Planck's combinatorics to claim a commitment on the issue of quantization and this seemed incorrect to the eyes of the advocates of the weak thesis. If it is true that Planck's statistical arguments can be interpreted as not differing from Boltzmann's as much as they seem to do at first sight, then this duality might rather support the thesis of an absence of commitment in the problem of the reality of the quantum. Thus, even after having ascertained whether or not Planck's statistics differ from Boltzmann's, we have to look at the problem from a broader perspective and establish the role that the alleged similarity or dissimilarity might have played in Planck's theory.

As we have seen, Needell suggested that this broader perspective should encompass Planck's view of the laws of thermodynamics, while Gearhart claims that it should take into account the endless shift of emphasis in the printed writings. However, this approach is not decisive as well, because the analysis of the long-term development of Planck's program was among the most original contributions of the advocates of the continuity strong thesis and one of their most effective sources of arguments.¹⁴ In this paper, I will attempt to investigate the way in which Planck tries to justify the introduction of statistical considerations into physics and to compare it with the analogous attempt pursued by Boltzmann. I think that this strategy is able to retrieve a historical perspective on a crucial problem somehow implicit in the previous discussion: what was Planck's attitude towards the relation between statistics and physical knowledge? Put in other

¹²(Garber, 1976).

¹³See in particular (Klein, 1962).

¹⁴For a recent contribution in this direction see (Büttner, Renn, & Schemmel, 2003).

terms: Did Planck think that the statistical approach is another equally correct way of studying macro- and micro-phenomena as the usual dynamic approach? Of course, an answer to this question requires not only an assessment of the similarities or dissimilarities between Planck's and Boltzmann's statistical formalism, but also a clarification of the general status of this formalism in Planck and Boltzmann and the way in which it was related with the rest of the physical knowledge of electromagnetism (in the first case) and mechanics (in the second case).

In the next sections I will attempt to answer to these questions. My main point is that a close investigation of the role of statistics offers a new and original perspective on the debate mentioned above. In particular, I will claim a sort of intermediate position between continuity and weak thesis. My argument to arrive at this conclusion is twofold. I first will analyze Boltzmann's and Planck's combinatorics focussing especially on the counting procedure, on the issue of the maximization and on the problem of the non-vanishing magnitude of the phase space cell. I will argue that, in all these instances, Planck tries to build up his statistical model remaining as close as possible to Boltzmann's original papers, but that the incompleteness of his analogy brings about some formal ambiguities that are (part of) the cause of his wavering and uncommitted position. At any rate, the deviations from Boltzmann's procedure were generally due (or understood as due) to the particularities of the physical problem Planck was dealing with. Next, I will present a sketchy comparison of Planck's and Boltzmann's justification of the introduction of statistical considerations in physics and I will claim that, in this particular aspect, the opinions of the two physicists differ remarkably. This will suggest us an unexpected twist which will bring us to the final conclusion.

How the Story Began

There are two different statistical arguments in Boltzmann's works. The first is presented in the second part of his 1868 paper entitled "Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten" and devoted to the derivation of Maxwell's distribution law.¹⁵ This argument, as we will see soon, is extremely interesting both for its intrinsic features and for the investigation of Planck's combinatorics that follows.¹⁶

In this argument, Boltzmann presupposes that the system is in thermal equilibrium and arrives at an explicit form of Maxwell's equilibrium distribution through the calculation of the marginal probability that a molecule is allocated into a certain energy cell. Let us suppose a system of n molecules whose total energy E is divided in p equal elements ϵ , so that $E = p\epsilon$. This allows us to define p possible energy cells $[0, \epsilon]$, $[\epsilon, 2\epsilon]$, \dots , $[(p-1)\epsilon, p\epsilon]$, so that if a molecule is allocated in the i -th cell, then its energy lies between $(i-1)\epsilon$ and $i\epsilon$. The marginal probability that the energy of 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 the molecules. In other words, the marginal probability is proportional to the total number of what in 1877 would be

¹⁵(Boltzmann, 1868); (Boltzmann, 1909, pp. I, 49–96).

¹⁶An excellent discussion of this section of Boltzmann's paper—that, however, does not include an analysis of the combinatorial part of the argument—can be found in (Uffink, 2007, pp. 955–958).

called “complexions” calculated on an opportunely defined new subsystem of molecules and total energy.¹⁷ This is tantamount to saying that the exact distribution of the remaining molecules is marginalised.

To clarify his procedure, Boltzmann presents some simple cases with a small number of molecules. Let us consider, for instance, the case $n = 3$. Boltzmann estimates the total number of ways of distributing 3 molecules over the p cells by noticing that if a molecule is in the cell with energy $p\epsilon$, there is only one way of distributing the remaining molecules, namely in the cell with energy 0ϵ , if a molecule has energy $(p - 1)\epsilon$ there are two different ways and so on.¹⁸ Therefore, the total number of ways of distributing the molecules is:

$$1 + 2 + \dots + p = \frac{p(p + 1)}{2}.$$

Now, let us focus on a specific molecule. If that molecule has energy $i\epsilon$, then there is $(p - i)\epsilon = q\epsilon$ energy available for the remaining molecules and this means q possible energy cells. There are, of course, q equiprobable ways of distributing 2 molecules over these q cells because once a molecule is allocated, the allocation of the other is immediately fixed by the energy conservation constraint. Hence, the marginal probability that the energy of a molecule lies in the i -th cell is:

$$P_i = 2 \frac{q}{p(p + 1)}.$$

With a completely analogous reasoning one can easily find the result for $n = 4$:

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

Thus, it is clear that, by iterating the argument, in the general case of n molecules this probability becomes:¹⁹

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

Letting the numbers of elements and of molecules grow to infinite, Boltzmann obtains the Maxwell distribution.

Some quick remarks regarding this statistical argument. First, the marginalization procedure leads Boltzmann to express the probability that a certain energy is ascribed to a given molecule in terms of the total number of complexions (for a suitable subsystem) and this procedure is closely related to Boltzmann’s physical task: finding the equilibrium distribution. I will return on this point later on. Second, Boltzmann defines the energy cells using a lower and upper limit, but in his combinatorial calculation he considers one value of the energy only. This is doubtless due to the arbitrary small

¹⁷A complexion is a distribution of distinguishable statistical objects over distinguishable statistical predicates, namely it is an individual configuration vector describing the exact state of the statistical model.

¹⁸Note that, since Boltzmann is working with energy cells, the order of the molecules within the cell is immaterial.

¹⁹The dependence on i is contained in q through the relation $p - i = q$.

magnitude of ϵ , but it also entails an ambiguity in the passage from the physical case to its combinatorial representation and vice-versa. Third, a simple calculation shows that:

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

It is worthwhile noticing that the distribution (1) is a particular case of the so-called Polya bivariate distribution at that time still unknown.²⁰ Moreover, the normalization factor in equation (1), i.e. the binomial coefficient at 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. In fact, let us suppose that there are n distinguishable statistical objects and p distinguishable statistical predicates, and suppose that to each predicate a multiplicity i ($i = 0, \dots, p-1$) is ascribed. An individual description that allocates individual objects over individual predicates is valid if and only if:

$$\sum_i i n_i = p - 1,$$

where n_i is the number of objects associated with the predicate of multiplicity i . Under these conditions, the number W of valid individual descriptions is:

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

namely the normalization factor written by Boltzmann. We will meet this formula again in Boltzmann's 1877 paper and, more importantly, this particular form of the normalization factor will play a major role in explaining some ambiguities of Planck's own statistics.

In 1877 Boltzmann devoted a whole paper to a brand new combinatorial argument which presumably constituted Planck's main source of statistical insights.²¹ There are notable differences between this argument and its 1868 predecessor, starting with the physical problem: in 1877 Boltzmann is facing the issue of irreversibility and his task is not simply to derive the equilibrium distribution, but to show the relation between this distribution and that of an arbitrary state. Thus Boltzmann replaces the marginalization with a maximization procedure to fit his statistical analysis to the particular physical problem he is tackling.

To substantiate his new procedure, Boltzmann puts forward the famous urn model. Let us suppose, as before, a system of n molecules and that the total energy is divided into elements $E = p\epsilon$ and imagine a urn with very many tickets. On each ticket a number

²⁰See (Costantini, Garibaldi, & Penco, 1996) and (Costantini & Garibaldi, 1997).

²¹(Boltzmann, 1877); (Boltzmann, 1909, pp. II, 164–223).

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 demands an enormous number of drawings and then eliminates all the complexions that violate the constraint. The number of acceptable complexions obtained by this procedure still is very large. Since a state distribution depends on how many molecules (and not which ones) are to be found in each cell, many different complexions might be equivalent to a single state and Boltzmann's crucial step is to ascribe a probability to each state in term of the number of complexions corresponding to it. By maximizing the state probability so defined, Boltzmann succeeds in showing that the equilibrium state has a probability overwhelmingly larger than any other possible distribution or that there are far more complexions matching the equilibrium state.

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

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

This is the total number of complexions for a statistical model where distinguishable objects are distributed over distinguishable predicates precisely as in 1868. Boltzmann actually writes the formula (3) in passing as an expression of the total number of such complexions.²² However, it can also be interpreted in a completely different way. It can equally well express the total number of occupation vectors for a statistical model of p indistinguishable objects and n distinguishable predicates. To put it differently, if one wants to calculate the total number of ways of distributing p objects over n predicates counting only how many objects, and not which ones, are ascribed to each predicate, then the formula (3) gives the total number of such distributions. An ingenious and particularly simple way of proving this statement was proposed by Ehrenfest and Kamerlingh Onnes in 1915.²³ Let us suppose that, instead of n distinguishable cells, one has $n - 1$ indistinguishable bars defining the cells.²⁴ If both the p objects and the $n - 1$ bars were distinguishable, the total number of individual descriptions would be given by $(p + n - 1)!$. But the indistinguishability forces us to cancel out from this number the $p!$ permutations of the indistinguishable objects and the $(n - 1)!$ permutations of the indistinguishable 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 purely formal similarity, as we will see in the next section, is one of the keys to understand the ambiguities of Planck's statistical arguments.

²²(Boltzmann, 1877, pp. II, 181).

²³(Ehrenfest & Kamerlingh Onnes, 1915).

²⁴Note 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.

A Statistics for All Seasons

Planck's application of Boltzmann's statistics to radiation theory has very far-reaching consequences which puzzled his contemporaries and which took many years to be completely understood. In this and in the next section I will restrict myself to some problems that have a special bearing on the historical issue of the quantization, namely the analysis of Planck's counting procedure and the general structure of his statistical argument.

In his December 1900 paper, where he first gives a theoretical justification of the radiation law, Planck is apparently very explicit about his counting procedure:

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 neighbourhood. 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.²⁵

In this passage Planck is unmistakably speaking of distributing energy elements over resonators.²⁶ Moreover, in the same paper, Planck writes the total number of ways of this distribution in the following way:

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

This is the total number of ways of distributing P indistinguishable objects over N distinguishable predicates as Ladislav Natanson would point out in 1911.²⁷ But, as I said commenting formula (3), it can also be interpreted as the total number of ways of distributing N distinguishable objects over $P+1$ distinguishable predicates given the energy conservation constraint, namely exactly the same statistical model Boltzmann had worked out in his 1877 paper. While in the first interpretation the distribution of single energy elements over resonators very naturally suggests that the resonators can absorb and emit energy discontinuously, in the second interpretation of the same formula, the resonators are distributed over energy cells that have a fixed magnitude, but they can be placed *everywhere* within a given cell thus implying that they can absorb or emit continuously. More importantly, if Planck was faithfully following Boltzmann in framing his statistical model, and Boltzmann, as we know, assumed a continuous physics behind

²⁵(Planck, 1900b); see (Planck, 1972, p. 40).

²⁶However, 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).

²⁷(Natanson, 1911).

his model, it is extremely plausible that Planck did not feel himself forced to assume any discontinuity as a consequence of his counting procedure.²⁸

Therefore, the formal ambiguity in Planck's way of counting the complexions deprives of cogency one of the fundamental arguments of the advocates of the discontinuity thesis which relies on the original statement in December 1900. This point can be strengthened by appealing to Planck's awareness of this ambiguity. In the fourth chapter of the first edition of the *Vorlesungen über Wärmestrahlung* (1906), his attempt at remaining as close as possible to Boltzmann's original argument and phraseology is patent. First, Planck introduces the concepts of complexion and distribution for a gas and it is clear that, to him, a complexion is an individual allocation of molecules to energy cells.²⁹ Then he simply extends this concept to radiation theory, without any, even slight, change of meaning so that it is plausible that this concept keeps its general features in the new context as well.

After developing Boltzmann's procedure for a gas, he deals with the problem of counting complexions in radiation theory:

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 resonators, 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.³⁰

The reference to the "energy domain" clarifies that Planck is thinking of distributing resonators over energy cells and that he considers this way of doing as "a way quite analogous to the case of gas."

But immediately after, he stresses 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 displays 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. Second, he also knew that the distribution of individual energy elements over resonators was a simpler way of doing because it does not require any energy conservation constraint and, more importantly, no assumption on the distribution of the resonators within the energy cells is needed. Unfortunately, this is the first occasion in which Planck explicitly reveals 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, it seems unlikely that Planck was not aware of such a trivial equivalence already in December 1900 because the decision of adopting combinatorial considerations was probably

²⁸This ambiguity was exploited in the other direction by Alexander Bach who in (Bach, 1990) 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.

²⁹(Planck, 1906, pp. 140–143).

³⁰(Planck, 1906, pp. 151–152).

well pondered by him. Furthermore, the discontinuity thesis loses footing because, as Kuhn pointed out, it is difficult to figure out why Planck, after committing himself to discontinuous emission and absorption in 1900, should have availed himself of a formal equivalence with a combinatorial procedure which presupposes continuous energy cells in 1906.

Further support to the thesis that Planck interpreted the elements of energy in analogy with Boltzmann's procedure is provided by the third lecture delivered by Planck at Columbia University in 1909. The main claim of that lecture is that "irreversibility leads of necessity to atomistics,"³¹ precisely because an atomistic hypothesis is an inescapable prerequisite of the application of probability. But, and this is the point, the only role actively played by the atomistic hypothesis consists of allowing us to distinguish the possible cases that are to be computed. Of course, a combinatorial calculation calls for a separation of the various possibilities and this is manifestly unattainable in case of a continuum, hence:

[I]n order to be able to differentiate completely from one another the complexions realizing [a state], and to associate it with a definite reckonable number, there is obviously no other means than to regard it as made up of numerous discrete homogeneous elements—for in perfectly continuous systems there exist no reckonable elements—and hereby the atomistic view is made a fundamental requirement.³²

Thus Planck conceives the discontinuous elements of energy in the same way as Boltzmann in 1868 and in 1877, namely as formal devices suitable for labelling and then combinatorially manipulating the different cases and, of course, this function can be equally well accomplished by energy cells. Therefore, the formal similarity, Planck's awareness of it and his interpretation of the 'elements' seem to suggest powerfully that he did not perceive any fracture between his theory and classical physics.

On the other hand, the continuity thesis is not so compelling as it can appear at first sight, at least as far as its most ambitious conclusions are concerned. In fact, Planck's 1906 awareness of the formal similarity mentioned above does not straightforwardly support Kuhn's claim that he was committed to a continuous physical model. This inference is possible only after admitting a close interplay between statistics and physics in Planck's theory, otherwise the formal similarity turns out to be nothing but a formal ambiguity and Planck's awareness reduces itself to the possibility of maintaining an uncommitted position, precisely the content of the weak thesis. Such an interplay was typical of Boltzmann's style but, I will argue, absolutely absent in Planck's. Of course, it might be claimed that continuity in the energy exchange between resonators and field would have been much more attractive to Planck than a discontinuity which seemed to entail a contradiction with Maxwell's equations. However, this appeal to Planck's theoretical background is not conclusive because, as Allan Needell has persuasively shown, a fundamental ingredient of this background consisted in avoiding any explicit assumption on the microstructure of the system. Indeed, Planck chose the case of the black-body radiation because there was no need of specifying the detailed internal structure of the

³¹(Planck, 1915, p. 41).

³²(Planck, 1915, p. 45).

resonators. Therefore, Planck's research program might be utilized to support both the continuity thesis and the weak thesis.

A problem with Needell's argument is that Planck's statements about his program mainly belongs to a period before the development of his combinatorics, so that a supporter of the continuity thesis might argue that Planck, after his conversion to Boltzmann's combinatorics, abandoned or weakened the neutrality towards the microstructure that characterized the first phase of his program. This objection would shift the problem to another question: when did Planck's conversion actually take place and how deep was it? There is another way to tackle the issue: as I have said above, the continuity thesis heavily relies on the status ascribed by Planck to the statistical arguments and, fortunately enough, there are clear statements about this topic in the first edition of the *Vorlesungen*. But before analysing them, I will complete the survey of Planck's application of Boltzmann's combinatorics discussing the general structure of his statistical argument.

To Maximize or Not to Maximize?

The first paper on the quantum was published on December 1900 and the second in the *Annalen der Physik* on January 1901.³³ In spite of such a tiny interval of time, there are considerable differences between these two papers, particularly with regard to the structure of the argument.

In the first 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 they do not mutually interact. Next Planck supposes that the total energy of the system consisting of resonators and radiation is E , and E_r is the fraction of the total energy belonging to the resonators only. The energy E_r must be divided over the different frequencies, so that a possible energetic state of the system of resonators is described by the vector:

$$\omega_k = \{E_1, E_2, \dots\}, \quad (5)$$

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 suggests to measure 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 in P_i elements of magnitude $\epsilon_i = h\nu_i$ over the N_i resonators. Since the magnitude 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:

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

³³(Planck, 1901).

Since resonators of different classes do not interact, the distribution described by the vector (5) is a compound event consisting of many independent events, then its probability is:

$$W(\omega_k) = \prod_i W(E_i). \quad (7)$$

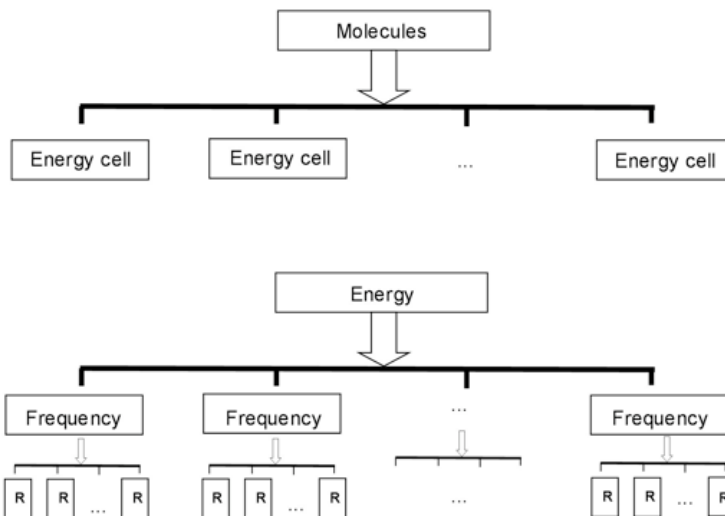
At this point, instead of performing the cumbersome maximization of (7), finding the equilibrium distribution ω_{eq} and the mean energy of a resonator at an arbitrary frequency, Planck mentions that “[a] more general calculation which is performed very simply, using exactly the above prescriptions shows 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 only problem with the calculation above is the maximization procedure which actually does not appear in the 1901 paper. Indeed, in the January article Planck starts by reckoning only a class of resonator and does not consider a set of arbitrary distributions ω_k among which a particular equilibrium distribution ω_{eq} is to be selected by means of a maximization procedure, but 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 historians³⁴ but I think that some confusion still remains as regarding to the justification of such deviations in the context of Planck’s theory.

First, it has been 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. 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 first examine the physical problem Planck is dealing with.

In Boltzmann’s theory a macrostate is given by the number of molecules that are placed in a certain energy interval, a situation that in equilibrium is described by Maxwell’s distribution for a gas. On the contrary, in the case of radiation theory what is physically meaningful is a relation between the energy allocated to a certain frequency and the absolute temperature and this relation presupposes a 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 the energy over a class of resonators (how many resonators lie in a certain energy cell) is still a *microstate*.³⁵ The difference between Boltzmann’s and Planck’s physical situation is outlined by the following scheme:

³⁴See especially (Klein, 1962), (Kuhn, 1978), (Darrigol, 1988), and (Gearhart, 2002).

³⁵On this important difference see also (Darrigol, 2000) and (Darrigol, 2001).



In Boltzmann’s case a microstate is an individual arrangement of molecules over different energy cells, while a macrostate is given by the number of molecules allocated in each cell. Instead, in Planck’s, energy is allocated over each frequency and then divided in elements which are distributed over the resonators (R) that vibrate at that frequency. A microstate is an individual arrangement of resonators in energy cells, while a macrostate is how much energy is allocated to each frequency.

To make this delicate point as clear as possible we can consider where the disanalogy between Planck and Boltzmann stems from. To transform Boltzmann’s case in Planck’s one we can proceed as follows. Let us suppose to have N molecules, so that a state distribution is a vector:

$$o_k = \{n_0, n_1, \dots, \},$$

quite analogous to vector (5). This vector tells us that n_i molecules are in the cell with energy i . Let us now suppose that no permutations are possible among the different energy cells, precisely like no exchange of energy is possible among resonators of different frequency. This also means that one should suppose the molecules as distributed in bunches of dimension n_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 on different cells alternatively. 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 in P elements, if necessary by taking “for P an integer in the neighbourhood.” Furthermore, since in Planck’s model there are many resonators at each frequency, we must suppose that each energy cell is divided in 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 energy 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 o_k is to make it proportional to the total number of ways of distributing the

molecules over the sub-levels. It is exactly what Planck did and 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 model. First, it clearly shows what was really tricky in Planck's usage 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 energy cell. This is exactly the main characteristic of Bose-Einstein statistics. But this crucial point is concealed by the fact that Planck is speaking of energy and his microstate concerns the distribution of resonators.³⁶ As soon as a corpuscular conception of energy gained a sound footing and the resonators were left aside, the troublesome features of Planck's daring analogy would emerge, because what was a macrostate in Planck immediately became one of many equiprobable microstates and the bunches of energy became groups of indistinguishable particles. Not surprisingly, the first step Albert Einstein took in his 1924 paper on quantum statistics was to reduce Planck's macrostate to the rank of a microstate and to define a new sort of macrostate.³⁷

Secondly, however, as long as one keeps assuming the energy as a continuous quantity and using the 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 usage of statistics as a straightforward application of Boltzmann's doctrine because he ultimately evaluated the probability of a certain macrostate (7) through *the total number of microstates consistent with that macrostate*. The difference from Boltzmann rested mainly on the definition of macrostate, but, in turn, it depended more on the particular physical problem, than on the statistical argument used.³⁸ Planck is very clear on this topic in the first edition of the *Vorlesungen*:

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 suffices the calculation of the total energy E of the N resonators; the specific distribution of the energy over the single resonator is not controllable, it is completely left [anheimgegeben] to chance, to elementary disorder.³⁹

More subtle is the issue concerning the maximization because it involves Planck's particular concept of disorder. As early as December 1900, Planck notices that the

³⁶The 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 make it come to light the resonators have to be abandoned.

³⁷(Einstein, 1924, p. 262).

³⁸It should be noticed 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 explicitly because it is an unimportant constant. On this point see (Hoyer, 1980).

³⁹(Planck, 1906, p. 151); an akin statement can be found in (Planck, 1915, p. 89).

notion of entropy is closely connected to the chaotic feature 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. 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 very 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 probably aware of this fundamental difference from the outset because he mentions it in the introduction of his paper in March 1900.⁴⁰ However, it is only in the first edition of the *Vorlesungen* that an explicit statement on the influence of this aspect on the combinatorics of the resonators can be found:

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.⁴¹

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 apparent justification. Obviously Planck was aware of the connection mentioned above already in December 1900. Most probably, Planck’s confidence in the ‘evidence’ of the equivalence—that at first sight is not evident at all!—stems from 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 this state kinetic energy does not differ appreciably from the mean in the course of time. Moreover, this state is independent of the initial conditions, so that Boltzmann can state:

... we 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

⁴⁰(Planck, 1900a); see (Planck, 1958, pp. 668–686).

⁴¹(Planck, 1906, p. 150).

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.⁴²

This argument must have pleased Planck very much because it only relies 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 *Vorlesungen*. 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.”⁴³ The generalization of the hypothesis of natural radiation Planck is talking about is exactly the broader concept of elementary disorder that he considers the foundation of the statistical description of the system. 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 naturally follow from the notion of disorder in gas theory, because, viewed from a general perspective, 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 and to make this leap ‘evident.’ 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 is.

Indeed, to accomplish his final goal, Planck had to find a relation between the energy E_ν allocated on the frequency ν and the absolute temperature T . The entropy S is a concept connecting both quantities by means of the well-known definition of absolute temperature:

$$\frac{\partial S}{\partial E} = \frac{1}{T}.$$

Adopting Boltzmann’s definition of entropy:

$$S = k \log W,$$

one only needs a relation between the probability W and the energy E . In December 1900, Planck writes:

Entropy means disorder, and I thought that one should find this disorder in the irregularity with which even in a completely stationary radiation field the vibrations of the resonator change their amplitude and phase, as long as one considers time intervals long compared to the period of one vibration, but short compared to the duration of a measurement.⁴⁴

⁴²(Boltzmann, 1898, p. 310).

⁴³(Planck, 1900b); see (Planck, 1972, p. 39).

⁴⁴(Planck, 1900b); see (Planck, 1972, p. 38).

Thus, elementary disorder in terms of natural radiation warrants a disordered temporal evolution of the resonator. But Planck still needs, in December 1900, a *measure* of this disorder, a quantitative expression of the disorder involved in the fact that the energy E is allocated on the (resonator of) frequency ν .⁴⁵ This measure, namely the probability W , is provided by the combinatorics on the system of N identical resonators, but the foundation of this bold leap is the general notion of elementary disorder.

In the meantime, another ambiguity jumps out of the hat. On the one hand, Planck's analysis of the statistical model fits Boltzmann's procedure in 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 behaviour of a single resonator in equilibrium is equal to the combinatorial behaviour of a set of resonators as regards the average values, then all the configurations calculated in the set must represent equilibrium configurations.⁴⁶ 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 black-body radiation law) is defined for the equilibrium state only and by a daring application of the elementary disorder, Planck can escape the formal necessity of maximization. In fact, such a procedure is relevant only if one is interested in the problem of irreversibility but, as we will see more clearly below, Planck considers this issue completely solved by the notion of elementary disorder and, from 1900 on, it disappears from his research program. With the riddle of irreversibility put aside, the only problem remaining was the derivation of the radiation law.

For these reasons, Planck was free to use the maximization or not without affecting the consistency of his reasoning or the analogy with Boltzmann's statistical arguments and he seems to be aware of this since December 1900. Furthermore, as we have seen, Boltzmann himself, in 1868, elaborated a statistical argument dealing with the equilibrium state only, which did not make use of the maximization procedure, a technique that would make its appearance only in 1877, in close connection with the problem of irreversibility. Unsurprisingly, given the chance, Planck chose the "faster and easier" way of avoiding maximization.

Size Does Matter

A further departure of Planck's statistics from Boltzmann's is the fixed magnitude of the energy element. Boltzmann divided the energy space into cells, but the magnitude of these cells was completely arbitrary and disappeared from the final result. By contrast, the magnitude of Planck's elementary cells (*Elementargebiete*) was determined by a universal constant that played a crucial role in the final formula. Once again, Planck

⁴⁵As pointed out by Needell, Planck considered probability merely (or mainly) as a measure of disorder and, indirectly, as a way for calculating entropy. This is clear, for instance, in (Planck, 1901) where he says that the combinatorial definition of probability is "the condition [...] which permits the calculation of S ," see (Planck, 1958, p. 719).

⁴⁶Note that the temporal evolution of a resonator is analogous to a system in thermal contact with a heat reservoir. In both cases the exact energy of the system can fluctuate around a mean even though the equilibrium is maintained. On this point see also (Gearhart, 2002).

was conscious of this fact that, in the *Vorlesungen*, he considers “an essential difference” with the case of gas.⁴⁷ He also knew that if the magnitude of the elementary cell goes to zero, one retrieves the incorrect Rayleigh-Jeans formula.⁴⁸

But even if it was an essential difference in the procedure, Planck tried to assign to the unvanishing size of the cell a place in the framework of the combinatorial approach. In December 1900, the division of the energy in elements of fixed magnitude was performed to allow the combinatorial calculation and the choice of the constant h was probably due to the law Planck had found in October.⁴⁹ However, in the first edition of the *Vorlesungen*, 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 and Ulrich Hoyer have pointed out that, even if these volumes can be vanishingly small, they cannot really disappear, otherwise Boltzmann’s integrals are doomed to diverge.⁵⁰ By shifting the quantization from energy cells to regions of the phase space, Planck was therefore reinforcing the analogy between his procedure and Boltzmann’s. At the same time, the non-vanishing magnitude 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 that, in his opinion, had to be decided empirically. As we will see more clearly in the next section, Boltzmann justified the equiprobability of his complexions by appealing to the Liouville theorem and a particular definition of probability. With his special partition of the phase space in 1906, Planck was able to introduce a justification of the equiprobability which relied on general electrodynamics and a universal constant only. He could validly feel that an important gap in his approach had been filled.

In the second edition of the *Vorlesungen* (1913), Planck further improves the position of the constant h in his theory by showing that it is closely related with a general thermodynamical result: Nernst’s theorem. The third law of thermodynamics discovered by Walther Nernst in 1905 entails the existence of an absolute definition of the entropy and, from Planck’s point of view, this implies an absolute definition of probability.⁵¹ But probability hinges upon the partition of the phase space that, of course, must be also possible in an absolute way. Commenting the necessity of fixing a finite magnitude for the phase cell as a consequence of Nernst’s theorem, Planck says:

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

⁴⁷(Planck, 1906, p. 153).

⁴⁸(Planck, 1906, p. 156).

⁴⁹The role played by the universal constants in Planck’s derivation has been stressed in (Badino & Robotti, 2001) and in (Gearhart, 2002).

⁵⁰(Darrigol, 1988), (Hoyer, 1980).

⁵¹For an overview of the problems connected with Nernst’s theorem see (Kox, 2006).

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.⁵²

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

The previous discussion sheds some light on the relation between Planck's and Boltzmann's statistical formalism and on the more general issue of the internal consistency of Planck's theory. One of the main tenets of Kuhn's interpretation is that Planck developed a continuistic understanding of the combinatorial arguments in order to be consistent with the electromagnetic part of his theory that proceeded from the tradition of classical physics. However, Darrigol has pointed out that Planck's departures from Boltzmann's original procedure seem to imply that he was consistent with a part of the tradition he was appealing to and inconsistent with another part.⁵³ This selectivity in consistency looks rather arbitrary and difficult to justify. Martin Klein has proposed to characterize Planck's research program as "uniformly consistent," namely Planck was ready to tolerate some contradictions if only they allowed him to arrive at his final goal, a theoretical justification of the black-body radiation law.

However, I do not think that Planck's usage of the statistical formalism differs from Boltzmann's so remarkably or decisively as it was often assumed by the scholars. Or, at least, I do not think that Planck perceived a real fracture. The previous discussion has shown that Planck tried to apply Boltzmann's formalism through an analogy between a system of resonator and a system of molecules, but, since these two systems are physically different, he was forced to take account of the imperfections of the analogy. I would say that Planck was *analogically consistent* in his usage of Boltzmann's statistical doctrine: he was aware of the differences and that they mainly derived from the particular physical problem he was coping with. For instance, 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, because the spectral distribution Planck has to arrive at, is dissimilar from the velocity distribution of gas theory. In the former case, the distributions over the set of resonators are irrelevant, and they must be marginalized. Likewise, the maximization procedure was dispensable both because only the equilibrium has an empirical meaning for heat radiation and because of the particular concept of elementary disorder Planck had fostered. Lastly, Planck tried to embody the more marked difference—the finite magnitude of the energy element—in his statistical procedure through an interpretation of elementary region of the phase space and, eventually, through Nernst's theorem.

To be sure, the imperfections in the analogy also give rise to a number of ambiguities—especially of formal nature—that Planck exploited to keep an uncommitted position. As we have seen, it was statistically irrelevant, in his theory, if energy elements were distributed over resonators or resonators were distributed over energy cells. Similarly, he

⁵²(Planck, 1913, p. 125).

⁵³(Darrigol, 2001).

had the choice of using a maximization procedure or not. In his writings, Planck keeps shifting from one approach to the other without solving the dilemma: there is hardly more than some modest changes of emphasis. For instance, in the first edition of the *Vorlesungen*, Planck mentions the distribution of resonators over cells, but immediately thereafter he switches to the “faster and easier” original way. By contrast, in the second edition of the same book, published in 1913, Planck follows Boltzmann’s doctrine almost literally, introducing a distribution density function and maximizing it to find the equilibrium case. The original distribution of energy elements has only a marginal role. Once again, this change in emphasis is due to a modification of the physical perspective: in 1911, Planck had brought about his famous ‘second theory’ which relied on the quantum emission hypothesis and on the distribution of resonators over continuous energy cells, and the second edition of the *Vorlesungen* relies heavily on this new method. I will explain this point in the next section.

To sum up, it seems that the weak thesis is correct in suggesting that, as a matter of fact, Planck, for various reasons, adopted an uncommitted attitude towards most of the physical issues emerging from his combinatorial procedure, but the point is that Planck was also justified in doing so by the ambiguities and the differences that his analogical adaptation of Boltzmann’s statistics evoked. However, to complete the picture, we still need an answer to a fundamental question: why was Planck unwilling to draw physical inferences from his combinatorial procedure? This question concerns the main difference between Planck and Boltzmann and demands an analysis of their justifications of the introduction of statistics in physics. This analysis will add a further twist to our discussion.

Organized Disorder

During his scientific life Boltzmann exposed his philosophical position in various essays and was involved in many scientific disputes especially concerning the necessity of the atomistic hypothesis and the so-called Energetic, but, unexpectedly enough, 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.⁵⁴ In the first part of the paper, he undertakes the task of deriving Maxwell’s distribution by a classical analysis of the mechanical collisions and, in the second part, as we have seen above, he accomplishes the same goal employing a statistical argument. In the final section, Boltzmann justifies this two-pronged approach by indirectly explaining how statistical considerations enter his treatment of mechanical problems.⁵⁵

However, his original argument is pretty obscure and I will try to reframe it in a more modern perspective. There are three ingredients:

- (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

⁵⁴(Boltzmann, 1868); see (Boltzmann, 1909, pp. I, 92–96). Another interesting point where Boltzmann displays his opinion on this topic is (Boltzmann, 1898, pp. 448–449).

⁵⁵Of 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 (see for example (Uffink, 2007)), but I think that the particular view of the relation between statistics and mechanics implicit in this argument should not be underestimated.

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:

$$\Sigma_t, \Sigma_{t+\Delta t}, \dots, \Sigma_{t+n\Delta t}.$$

From Liouville's theorem, it follows immediately that all these phase volumes are equal. But, 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 behaviour 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 because all the space is filled up by a system trajectory.

Of course, this cannot be considered a formally satisfactory argument because of the problems connected to the ergodic hypothesis, but the general idea is clear enough: Boltzmann tries to introduce the usage of statistics in mechanics as an account of the behaviour of the mechanical system that is as rightful as the temporal description that the mechanics itself can provide, as long as some conditions hold on the system itself, notably the ergodic hypothesis. More importantly, the application of statistics does not rely on our ignorance of the detailed state of the system, namely on our epistemic status, but on a certain kind of behaviour of the mechanical system. The key point is that this justification amounts to an attempt of deeply *integrating* statistics and mechanics, of seeking for the mechanical conditions of an application of statistical arguments. He stresses this aspect also in the introduction of his 1872 paper:

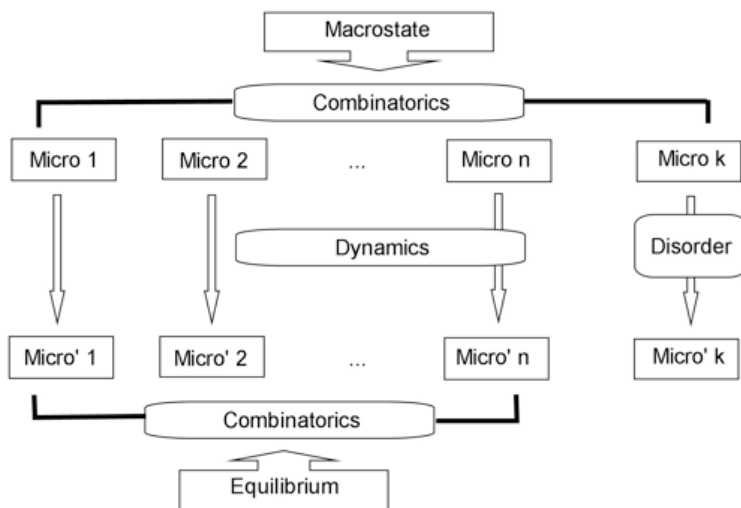
It would be an error to believe that there is an inherent indetermination in the theory of heat because of the usage of the laws of the calculus of probability. One should not mistake a law only incompletely proved, whose soundness is hence problematic, for a completely demonstrated law of the calculus 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.⁵⁶

Planck's justification goes exactly in the opposite direction. The relation between statistics and electrodynamics is explained at the beginning of the fourth chapter of the first edition of the *Vorlesungen* and the starting point is the following dilemma:

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, on the legitimacy and the necessity of such considerations. At first sight we might be inclined to think that in a purely electrodynamic 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.⁵⁷

However, Planck adds, the dilemma crops up 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 which evolve according to dynamical laws, in this case the laws of electrodynamics. However, Planck states, we cannot directly apply a dynamical analysis to the system because we do not know which of the many theoretically possible microstates actually holds. Our

⁵⁶(Boltzmann, 1872); see (Boltzmann, 1909, pp. I, 317).

⁵⁷(Planck, 1906, p. 129).

empirical measurements on the macrostate are able to supply only mean values, which are consistent with very 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 actually know that the result of the dynamical evolution of the microstates is a set of new microstates whose overwhelming majority is combinatorially related to the equilibrium state, the plain application of combinatorial arguments is ambiguous as well, because there are some, very few indeed, microstates that might lead to an anti-thermodynamical 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 hinder the anti-thermodynamical 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."

In the second edition, Planck verbalizes the difference between microstates and macrostates in an even more colorful way:

The microscopic state is the state as described by a mechanical or electro-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.⁵⁸

It is worthwhile noticing how deeply Planck's condition of disorder differ from Boltzmann's. Firstly there is a different relation to the statistical formalism: for Boltzmann, disorder is the prerequisite to introduce combinatorial arguments that can completely replace the dynamical ones because disorder permits all the 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 (and allows the integration of both), for Planck it belongs exclusively to the microlevel, to a realm populated of hypothetical mechanical and electro-dynamical observers which see a completely different world utterly concealed to us. As it was already pointed out by Allan Needell and Olivier Darrigol, Planck's elementary disorder concerns the mysterious interaction of matter and radiation and hence it is part of the uncontrollable and inaccessible microworld.

Therefore, the statistical arguments are not another, equally rightful, viewpoint of looking at mechanical problems, but the only way at our disposal of figuring out what

⁵⁸(Planck, 1913, p. 121).

is going on in the complicated and unreachable realm of the constituents of matter and radiation. Consequently, Planck merely considers the statistical arguments as a conceptual device to represent a chaotic situation and to perform useful calculations, but he does not deem them as a basis for understanding the physical world. This point clearly emerges in the second edition of the *Vorlesungen*. The general architecture of the book is remarkably dissimilar 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.”⁵⁹

The new dynamical law is the famous quantum emission hypothesis that Planck introduces with the following words:

... we 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 pulse, 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.⁶⁰

Even though the statistical part takes the leading role in the second edition of the *Vorlesungen*, the dynamical assumptions it provides are not to be meant as realistic ones. In fact, statistical considerations remain nothing but a convenient mean to paraphrase a hidden reality.

Who Cares for 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. Admittedly, Planck tried to follow Boltzmann’s formalism as faithfully as possible and the deviations are to be ascribed to the imperfections of the analogy, i.e. to the differences in the physical problems, but his attitude towards the statistical arguments was quite the opposite of Boltzmann’s. The Austrian physicist had tried to integrate mechanics and statistics by showing that the conditions for applying statistical argument are to be sought in some particular mechanical behaviour. On the contrary, Planck is quite clear in confining the statistical arguments into the impenetrable processes taking place at the microlevel: in Planck’s view, electromagnetism and statistics are completely dis-integrated. This general attitude and the formal ambiguities of the statistical procedures formed the ground for his contention that there were no conclusions to be drawn from the combinatorial fact that discrete energy elements were distributed

⁵⁹(Planck, 1913, p. 152).

⁶⁰(Planck, 1913, p. 153).

over resonators or, alternatively, that resonators were distributed over continuous energy cells. In this concluding section, I would like to discuss in more detail the concept of virtual observer, the disorder and the resulting relation between the dis-integration of electrodynamics and statistics and the construction of the microworld.

Boltzmann's microworld is assembled by conceptual elements coming from the macro-world, e.g. molecules as mechanical points or centers of force, elastic collisions and so on, and it is also characterized by the emergence of the statistical formalism as a mean of investigating physical reality. By integrating the statistical formalism with the mechanical one, typical of the macrolevel, Boltzmann establishes a new direction for the conceptual flow that eventually leads him to reinterpret the macroscopic laws in terms of the statistical viewpoint. The statistical interpretation of the second law is the most remarkable result of this bi-directional conceptual interaction between micro- and macroworld. The conceptual interaction between micro- and macroworld is bi-directional because Boltzmann integrates the formal ways of describing them and, in particular, integrates statistics with the rest of physical knowledge. On the contrary, by dis-integrating dynamical and statistical formalism and by reducing the latter to a computational device, Planck ends up with a microworld that is completely shaped by the macroscopic conceptual structure and it is unable to support any reinterpretation of the macrophenomena. In fact, what makes up Planck's microworld is only conceived as a way of representing, by means of macroscopic concepts, the mysterious and unobservable business of the interaction between matter and radiation. Furthermore, since statistics is completely separated from the rest of the physical knowledge, the statistical formalism has little to say both on the macroworld and on the microworld: statistics is not supposed to give us a description of how the world is, it is only supposed to give us a way of handling chaotic situations.

The particular relation between macro- and microworld is clearly presented in the Columbia lectures of 1909. As we have seen, in the second edition of the *Vorlesungen*, 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 the series mentioned above as an important element of his general argument for the justification of the usage of statistics.

Again, Planck points out that the contradiction between the reversibility of the microphenomena and the irreversibility of the thermodynamical laws stems from different definitions of state. The physical state envisioned by a 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,"⁶¹ is fundamentally different from the state of a usual macro-observer, because the former observes exact values and the latter only means. However, since the virtual micro-observer is nothing but a projection of a macroscopic one, Planck is shaping the issue of the contradictory relation between micro- and macroworld in terms of the macroworld itself. Thus, it is not surprising that the answer as well embodies the primacy of the macrolevel as can be verified by further comparing Boltzmann's and Planck's notion of disorder.

For Boltzmann the elementary disorder is a feature concerning the individual configurations of molecules and, more importantly, is the ultimate justification of the introduction of statistical arguments and the final warrant of their equivalence with the

⁶¹(Planck, 1915, p. 47).

mechanical ones. 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:

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 nevertheless 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.⁶²

While Boltzmann's concept of disorder directly dives 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-thermodynamical evolution. The difference is extremely important and scarcely stressed in the secondary literature. 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, that prevents 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 occurring of 'conspiratory' configurations where only a subset of evolutions is possible.⁶³

By contrast, Planck conceives the 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:

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.⁶⁴

Planck's change of meaning and function of the elementary disorder has far-reaching consequences.⁶⁵ By means of the virtual micro-observer and of the elementary disorder, Planck foists upon the constitution and the formalism of the microworld a series of

⁶²(Boltzmann, 1898, p. 40).

⁶³Cf. for example (Boltzmann, 1898, p. 451): "only singular states that *continually* deviate from probable states must be excluded" (italics added).

⁶⁴(Planck, 1915, p. 50).

⁶⁵A 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

constraints coming from the macroworld. Since Planck's parasitic microworld is completely shaped by means of conceptual material and formal requirement relying on 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 black-body radiation law.⁶⁶

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

[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.⁶⁷

Of course, the emerging notion of irreversibility is only the premise of Boltzmann's conception, but Planck is not in the position of accepting the consequence. Ultimately, his way of justifying the usage of statistics in physics is a justification of his ambiguous—or prudent—use of the statistical formalism as well.

Thus, we arrive at a similar conclusion to the weak thesis we have discussed above, but on a different ground. Actually, Planck was not interested in committing himself on the issue of continuity (also) because he did not need to. 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 peacefully stay away from dangerous connections between apparently incompatible formalisms.

Our analysis provides a possible explanation why Planck could maintain the reticent position he actually maintained about the quantum in his published writings. But, at the same time, it also provides an unexpected new argument for the continuity view. The monodirectionality of the conceptual exchange by which Planck builds up his microworld suggests that he was unwilling to ascribe to the microphenomena any feature that we do not observe at the macrolevel. Of course, he was not explicit on this point and this cannot be called a clear commitment, because it concerns the relation between micro- and macroworld from a broad methodological viewpoint. Therefore, a slightly different position emerges, a position one can name a 'weak version of the continuity thesis:' Planck did not manifest any commitment on the reality of the quantum and, as

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, p. 51). On the contrary Boltzmann reckoned the recurrence theorem in its original formulation 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, p. 443)).

⁶⁶Incidentally, 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. See for example (Renn, 1997) and (Uffink, 2005).

⁶⁷(Planck, 1915, p. 97).

a matter of fact, exploited all the ambiguities of his statistical formalism to associate his theory as little as possible to any clear statement on this issue, but this strategy demands a separation between statistics and dynamics, micro- and macroworld, and Planck's choice for the macroworld suggests that his preference went to a view of energy as a continuous quantity.

The thrust of this argument is that 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 physical knowledge 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 diverging.

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 bring up some statistical assumptions on the behaviour of the constituents 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, like in his papers on the monocycle.⁶⁸

On the other side of the river stood Boltzmann, who was not afraid of introducing bold assumptions on the behaviour of the molecules and of coping with them using the conceptual tools of statistics. Paul Ehrenfest as well as Albert Einstein belonged to the same tradition and, not unsurprisingly, they did not understand and sharply criticized Planck's usage of statistical arguments. Einstein, for instance, showed his legacy to Boltzmann's train of thought in his light quantum paper where he derives the existence of energy elements of free radiation from their statistical behaviour. That was exactly the kind of inference Planck could not consent. This fundamental fracture affected a large part of the relations between statistical mechanics and the early quantum theory.

References

- Bach, A. (1990). Boltzmann's Probability Distribution of 1877. *Archive for History of Exact Sciences*, 41, 1–40.
- Badino, M., & Robotti, N. (2001). Max Planck and the constants of Nature. *Annals of Science*, 58, 137–162.
- Boltzmann, L. (1868). Studien über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten. *Wiener Berichte*, 58, 517–560.

⁶⁸On this 'Berlin's style' in physics see (Jurkowitz, 2002).

- Boltzmann, L. (1872). Weitere Studien über das Wärmegleichgewicht unter Gas-
molekülen. *Wiener Berichte*, 66, 275–370.
- Boltzmann, L. (1877). Über die Beziehung zwischen dem zweiten Hauptsatze der
mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den
Sätzen über das Wärmegleichgewicht. *Wiener Berichte*, 76, 373–435.
- Boltzmann, L. (1898). *Vorlesungen über Gastheorie* (S. G. Brush, Trans.). Leipzig:
Barth.
- Boltzmann, L. (1909). *Wissenschaftliche Abhandlungen*. Leipzig: Barth.
- Büttner, J., Renn, J., & Schemmel, M. (2003). Exploring the Limits of Classical
Physics: Planck, Einstein, and the Structure of a Scientific Revolution. *Studies in
History and Philosophy of Modern Physics*, 34, 37–59.
- 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., & 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 Quan-
tum. *Centaurus*, 43 (3–4), 219–239.
- Ehrenfest, P., & Kamerlingh Onnes, H. (1915). Vereinfachte Ableitung der kombina-
torischen Formel, welche der Planckschen Strahlungstheorie zugrunde liegt. *An-
nalen der Physik*, 46, 1021–1024.
- Einstein, A. (1924). Quantentheorie des einatomigen idealen Gases I. *Berliner Berichte*,
261–267.
- Galison, P. (1981). Kuhn and the Quantum Controversy. *The British Journal for the
Philosophy of Science*, 32 (1), 71–85.
- Garber, E. (1976). Some reactions to Planck's Law, 1900–1914. *Studies in History and
Philosophy of Science*, 7 (2), 89–126.
- Gearhart, C. A. (2002). Planck, the Quantum, and the Historians. *Physics in Perspec-
tive*, 4, 170–215.

- Heilbron, J. (1986). *The Dilemmas of an Upright Man. Max Planck as Spokesman for German Science*. Berkeley: University of California Press.
- Hoyer, U. (1980). Von Boltzmann zu Planck. *Archive for History of Exact Sciences*, 23, 47–86.
- Jost, R. (1995). Planck-Kritik des T. Kuhn. In *Das Märchen vom Elfenbeinernen Turm. Reden und Aufsätze*. Berlin: Springer.
- Jurkowitz, E. (2002). Helmholtz and the liberal unification of science. *Historical Studies in the Physical and Biological Sciences*, 32 (2), 291–317.
- Kangro, H. (1970). *Vorgeschichte des Planckschen Strahlungsgesetzes*. Wiesbaden: Steiner.
- 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. (1963a). Einstein's First Papers on Quanta. *Natural Philosopher*, 2, 59–86.
- Klein, M. J. (1963b). Planck, Entropy, and Quanta, 1901–1906. *Natural Philosopher*, 1, 83–108.
- 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 (11), 23–32.
- Klein, M. J., Shimony, A., & Pinch, T. J. (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 & R. S. Cohen (Eds.), *World Views and Scientific Discipline Formation* (pp. 141–150). Amsterdam: Kluwer Academic.
- Kox, A. J. (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: Oxford University Press.
- Kuhn, T. (1984). Revisiting Planck. *Historical Studies in the Physical Science*, 14(2), 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*. Unpublished PhD Dissertation, University of Michigan, Ann Arbor.

- Planck, M. (1900a). Entropie und Temperatur strahlender Wärme. *Annalen der Physik*, 4 (1), 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 (4), 553–563.
- Planck, M. (1906). *Vorlesungen über die Theorie der Wärmestrahlung*. Leipzig: Barth.
- Planck, M. (1913). *The Theory of Heat Radiation* (M. Masius, Trans.). New York: Dover.
- Planck, M. (1915). *Eight Lectures on Theoretical Physics* (A. P. Wills, Trans.). New York: Dover.
- Planck, M. (1958). *Physikalische Abhandlungen und Vorträge*. Braunschweig: Vieweg & Sohn.
- Planck, M. (1972). *Planck's Original Papers in Quantum Physics* (D. t. Haar & S. G. Brush, Trans.). London: Taylor&Francis.
- Renn, J. (1997). Einstein's Controversy with Drude and the Origin of Statistical Mechanics. *Archive for History of Exact Sciences*, 51 (4), 315–354.
- Rosenfeld, L. (1936). La premiere phase de l'evolution de la theorie des quanta. *Osiris*, 2, 148–196.
- Uffink, J. (2005). 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 & J. Earman (Eds.), *Philosophy of Physics* (pp. 923–1074). Amsterdam: North Holland.