

# Planck's Radiation Law, The Light Quantum, and Indistinguishability in the Teaching of Quantum Statistics

Peter M. Enders<sup>1\*</sup> and Zhumanazar R. Aibekuly<sup>2</sup>

<sup>1</sup>Faculty of Engineering and Natural Sciences, Technical University of Applied Sciences Wildau, Germany, and Higher school of natural science, Pavlodar Pedagogical Margulan University, Kazakhstan; permanent address: D-15712 Königs Wusterhausen OT Senzig, Germany

<sup>2</sup>Institute of Science, Pavlodar State Pedagogical Margulan University, Kazakhstan

## \*Corresponding Author

Peter M. Enders, Faculty of Engineering and Natural Sciences, Technical University of Applied Sciences Wildau, Germany, and Higher school of natural science, Pavlodar Pedagogical Margulan University, Kazakhstan, permanent address: D-15712 Königs Wusterhausen OT Senzig, Germany.  
Email: physics@peter-enders.science

Submitted: 2026, Apr 15; Accepted: 2026, May 25; Published: 2026, Jun 05

**Citation:** Enders, P. M., Aibekuly, Z. R. (2026). Planck's Radiation Law, The Light Quantum, and Indistinguishability in the Teaching of Quantum Statistics. *Int Nat Sci Int Rese*, 1(2), 01-28.

## Abstract

Planck's 1900 introduction of the "energy elements", i.e., electromagnetic radiation quanta of energy  $h\nu$  each, for founding his radiation law marks the origin of quantum theory and is discussed in all courses on this subject. Surprisingly, the question whether Planck really implied quantisation is debated among historians of physics. We present a simplified account of that debate which also sheds light on the issue of indistinguishability and Einstein's light quantum hypothesis. Here, the relevance of the notion '(in)distinguishability' is often overemphasized. It has to be treated together with 'equality', 'identity', and 'interchangeability', where the latter one is the decisive notion in this context. As a result, the difference between Planck's and Einstein's concepts of 'light particles' becomes more easily to grasp. For that and many other reasons, we believe that the teaching of quantum physics benefits from including the material presented here, and that beyond the demand of historical completeness.

Nevertheless, for the sake of our readers' convenience, we almost completely reprint Planck's former students Hettner's remembering of Planck's way to his radiation law, together with its translation into English. It seems that Planck has not really done the calculations described in his December 1900 lecture. For this, following Becker, we show how to overcome the difficulties over there, and we sketch his simplifying approach in his subsequent 1901 article. Strangely enough, in his 1906/07 lectures on the theory of heat radiation, Planck deviates from those calculations and obtains a non-extensive entropy. This, too, is most easily analyzed using Bach's three-level scheme 'configuration. occupation. occupancy' for the counting combinatorics involved. Boltzmann's 1877, Planck's 1900, Einstein's 1905, Bose's 1924, and other treatments are systematized using that scheme.

## 1. Introduction

"Sprache, die für dich dichtet und denkt" (Friedrich Schiller)

En: Language that composes poetry and thinks for you

December 14, 1900 is usually regarded as the "birthday" [88] of quantum theory since, at that day, Max Planck has published the "energy elements" (energy quanta)  $\epsilon = h\nu$  for founding his radiation law during a meeting of the Physical Society at Berlin [76],

$$u(T, \nu) = \frac{8\pi\nu^2}{c^3} \cdot \frac{h\nu}{\exp\left(\frac{h\nu}{kT}\right) - 1} \quad (1)$$

Here,  $u(T, \nu)$  describes the spectral energy density of a ‘black body’ (i.e., an idealized physical body that absorbs all incident radiation) in thermal equilibrium at temperature  $T$ .  $c$  denotes the velocity of light.  $kT$  always means  $k_B T$  ( $k_B$  being the Boltzmann constant).  $h$  is now called Planck’s constant.

There is a broad discussion about how detailed the early development of quantum theory should be included in teaching, e.g. [21,68,92]. Our point, however, is less about the historical details than about the *notions* to be used [62].

While the central notion often considered is that of (in)distinguishability, we will show that the crucial notion is that of *interchangeability*. Thus, the students may be asked the following. Given that the notions are the tools of thinking, discuss the relation between the accuracy of thinking and the accuracy of the notions used.

For this, we propose to discriminate between the notions ‘equal particles’ and ‘identical particles’ on the one hand as well as ‘indistinguishable particles’ and ‘interchangeable particles’ on the other hand, because they often are used as if they were meaning one and the same (for examples, see below).

1. Following Helmholtz [39], *equal* particles agree in *all* their *intrinsic*, state-independent properties such as mass, electrical charge, spin. . .
2. *Identical* particles agree not only in all their intrinsic properties, but also in *all* their *extrinsic* properties such as position, velocity or momentum, being affected by interactions. . . . Identical particles are indistinguishable.
3. *Indistinguishable* (indiscernible) particles cannot be distinguished from another, not at all. Indistinguishable (indiscernible) particles are identical.
4. *Interchangeable* particles can be interchanged among each other without affecting the situation under consideration; they may be distinguishable by some properties.

#### For instance,

- The sodium ions in an ideal NaCl crystal are all equal, at least in the average. They are different in their positions within the crystal; ditto the chlor atoms. Hence, they are not identical, but interchangeable ([36] p. 9, Figure 1.3).
- Classically, the two electrons in the ground state of the He atom are equal, while, quantum-mechanically, they are identical. In both cases, they are interchangeable.
- The red balls in a snooker game are (supposedly) equal in their intrinsic properties of color, size, and mass. They are different in their positions on the snooker table. Within a static position, however, they can be interchanged among each other without affecting the game.

As ions in a crystal are *quantum* particles, while snooker balls are classical bodies, it is obvious that the notions ‘equality’, ‘identity’, ‘indistinguishability’, and ‘interchangeability’ are *not* related to the classical and quantum world, respectively, although many authors, from, say, Natanson 1911 [66] till Saunders 2020 [85] believe that. Saunders rightly comments, “Rarely can terminology in physics have been the source of so much misdirection.” (p. 4) Nevertheless, unfortunately, he does not accurately enough discriminate between those notions, see Subsection 5.1.

Generally speaking, indistinguishable particles are equal classical or quantum particles in a state characterized by a probability measure, a statistical operator respectively, which is invariant under any permutation of the particles under consideration. Accordingly, the particles of classical Maxwell- Boltzmann statistics are indistinguishable [1].

Spalek [89] rightly emphasizes that indistinguishable particles are still identifiable as long as one can count them (p. 428). The same applies to interchangeable particles and cells. Philosophically interested students should be encouraged to study the relationship between identity and indiscernibility.

Often it is claimed that, at the beginning, neither Planck nor anybody else could possibly anticipate the radical implications of the finiteness of the energy elements (see, e.g. [72]). However, while Planck, in his 1900 December talk [76], enthusiastically evokes  $h$  as a novel natural constant, already in his subsequent article submitted very shortly afterwards [77], he sets it back to merely a spectroscopic parameter. For this, we also do not follow Kuhn [54] in “My point is not that Planck doubted the reality of quantization or that he regarded it as a formality to be eliminated during the further development of his theory. Rather, I am claiming that the concept of restricted resonator energy played no role in his thought [...].” (p. 126, quoted after [72]). (see also [4,12,17,19,34,69]). Moreover, in

<sup>1</sup>Students should be aware that expressing the spectral energy density in terms of energy per unit area, steradian and/or wavelength affects shape and peak position of the distribution [62].

---

Section 5, we will prove that Planck's formula (39) does not allow for such a conclusion.

This contribution is not aiming at a detailed account of all the ambiguities in Planck's work which in turn triggered the debate on the possible interpretations. We will bypass many technicalities for making the topic accessible even to undergraduate students. The historical development of statistical mechanics from Boltzmann to Bose merely serves as an illustration to the concepts. Thus, in Section 2, we will briefly review the historical context of the black-body radiation problem and Planck's derivation of his law in which a combinatorial formula prominently figures. For our reader's convenience, we will present an almost complete reprint of Planck's former student Hettner's remembering of Planck's way to his radiation law [40] and its translation into English. It seems that Planck has not really done the calculations described in his December 1900 lecture [76] For this, following Becker, we will present a way to do so, and we will sketch his simplifying calculations in his subsequent article [77].

Section 3 contains the core matter of this article, *viz.*, Boltzmann's [8], Planck's [76,77], and Bose's [10] combinatorics in terms of Bach's [3] three-level description of 'configuration occupation occupancy'. Here, not the notion of (in)distinguishability plays the crucial role, but that of interchangeability.

Section 4 shortly deals with the surprisingly ambiguous reading of Planck's 1900 treatment [76,77] as implying quantization, or not. Surprisingly, because already Planck's 1900 formula (39) uniquely decides that question. On the reading that energy is quantized, the corresponding 'energy quanta' may be compared with the "light quanta" as formulated by Einstein in 1905. This leads us to the prehistory of quantum indistinguishability in Section 5, since already in 1914, Ehrenfest and Kamerlingh Onnes [23] have shown that "distinguishable" light quanta cannot be reconciled with Planck's law. In bypassing, we will point to a footnote there which proves that—contrary to the title "Derivation of . . ."—their scheme is not a proof of formula 39. Within this section, we also will comment on Natanson's 1911/1913 [66,67] and Einstein's 1905 [24] contributions.

Thus, we will largely deal with the distribution of particles over cells, notably, of Boltzmann's 1877 pieces of kinetic energy over molecules [8], Planck's 1900 "energy elements"  $h\nu$  over resonators [76], and Bose's light particles over momentum cells in phase space [10]. The relevant properties of the particles have been described above. The cells can,

- Be interchangeable or not among themselves (Planck's 1900 resonators of equal frequency are interchangeable, while those of different frequencies are not, of of course).
- Have a limited number of particles in them (Fermi-Dirac, parastatistics— not considered here) or not (Maxwell-Boltzmann, Bose-Einstein statistics) [64,41].

Finally, Section 6 will summarize and conclude this contribution.

## 2. The Historical Context and Planck's Derivation of the Radiation Law

"Wenn man das macht, was Alle machen, gelangt man nicht zu einer großen Entdeckung." (Benjamin List, 2021 Nobel Prize in Chemistry [58])

En: If you are doing what all are doing, you will never reach something extraordinary.

The standard picture of the interaction between theory and experiment during the emergence of Planck's radiation law is often a bit too narrow ([41] p. 68). As this article concentrates on Planck's combinatorics and related topics, we will present a short historical account only. However, for the reader's convenience, we will include an almost full account of Hettner's 1922 [40] recapitulations on Planck's last, heuristic steps to his radiation formula in Subsection 2.2 (October 19 talk [75]). Thereafter, we will critically examine its foundation in Planck's December 14 talk [76].

### 2.1 On the Development Till October 1900

In a series of articles in 1859f., Gustav Kirchhoff stated that, for all bodies of any arbitrary material and shape emitting and absorbing thermal electro- magnetic radiation at every wavelength in thermodynamic equilibrium, the ratio of its emissive power to its dimensionless coefficient of absorption is one and the same. Therefore, there should be a *universal*, material-independent function  $u(T, \nu)$  for the spectral energy density of black-body radiation at temperature  $T$  in the frequency interval  $[\nu, \nu + d\nu]$  [49]. That was confirmed by the Stefan-Boltzmann law, according to which the total energy radiated by it per unit surface area per unit time (also known as the radiant exci-tance)  $M$  is directly proportional to the fourth power of the temperature  $T$ ,  $M = \sigma T^4$ . Since the Stefan-Boltzmann law follows from thermodynam- ics and classical electrodynamics, the Stefan-Boltzmann constant  $\sigma$  involves the speed of light  $c$  and the Boltzmann constant  $k_B$ . A dimensional anal- ysis reveals that  $[\sigma] = [k_B]^4 / [c]^2 [\text{action}]^3$  [73]. However, then, there was no indication to assign a universal constant to 'action'. In 1893, on the basis of general thermodynamic considerations, Wilhelm Wien [94] showed that the function  $u(T, \nu)$  should have the most simple form

$$u(T, \nu) = \nu^3 f(\nu/T) \quad (2)$$

(‘Wien’s displacement law’). In 1897, he suggested ‘Wien’s radiation law’,

$$u(T, \nu) = \frac{8\pi\nu^3}{c^3} b \cdot \exp\left\{-\frac{a\nu}{T}\right\}, \quad (3)$$

( $a$  and  $b$  being universal constants) [95].

At that time, Planck begun to work on the radiation problem. Based on electrodynamics, he could show [74] that the spectral energy density  $u(T, \nu)$  is related to the temporal average  $U(T, \nu)$  of the radiation energy of an immersed oscillator as

$$u(T, \nu) = \frac{8\pi\nu^2}{c^3} \cdot U(T, \nu) \quad (4)$$

(for a nowadays derivation see, e.g. [59] pp. 32ff.). Notice, that that relation does not contain the mass, charge or damping factor of the oscillator. Marr and Wilkin [62] recommend that students be taught about  $U(T, \nu)$  instead of Wien’s displacement law (2), and that the decomposition (4) be evoked when the Stefan–Boltzmann law is taught. Indeed, formula (4) is pedagogically valuable because it separates the problem into two distinct parts, *viz.*, (i), the electrodynamics of the radiation field in the factor  $8\pi\nu^2/c^3$  and, (ii), the statistical mechanics of the oscillators in  $U(T, \nu)$ . This separation helps the students to understand that, here, the quantum aspects enter through the statistical behavior of matter in equilibrium with radiation. They can see that the same radiation law could potentially arise from different statistical assumptions about the oscillators—a crucial insight for the comparison of various approaches. In Section 3.5.2, we will consider Bose’s 1924 semi-classical derivation of the factor  $8\pi\nu^2/c^3$ .

## 2.2. Planck’s October 19, 1900 Talk

Now, beginning with the turn of the year 1899/1900, new measurements at low frequencies revealed deviations from Wien’s radiation law (3). In 1922, Planck’s former student Gerhard Hettner [40] recalled, “Als am Sonntag, dem 7. Oktober 1900, *Rubens* mit seiner Frau bei *Planck* einen Besuch machte, . . . erzählte [Rubens], daß bei seinen l’angsten Wellen das ku’rzlich von *Lord Rayleigh* (Phil. Mag. 49, 539, 1900) aufgestellte Gesetz (see fn. 4):

$$E_{\lambda,T} = c_1 \frac{T}{\lambda^4} e^{-c_2/\lambda T} \quad (H-4)$$

gelte . . . . Eine allgemeingu’ltige Strahlungsformel mu’sse jeden- falls fu’r gro’ße  $\lambda T$  in diese Form u’bergehen. Auf dieses Gespra’ch hin stellte nun *Planck* sogleich folgende Rechnung an: Er war gewohnt, nicht mit der Strahlungsformel selbst, sondern mit der Entropie  $S$  eines mit der Strahlung im Gleichgewicht befindlichen Resonators zu rechnen, die als Funktion seiner Energie  $U$  sehr viel einfachere Ausdr’ucke liefert.  $S$  und  $U$  stehen mit der Temperatur  $T$  in der Beziehung:

$$\frac{dS}{dU} = \frac{1}{T} \quad (H-5)$$

Proportionalita’t der Energie mit der Temperatur ergibt also:

$$\frac{dS}{dU} = \frac{\text{const}}{U}$$

oder

$$\frac{d^2S}{dU^2} = -\frac{\text{const}}{U^2} \quad (H-6)$$

Aus dem Wienschen Energieverteilungsgesetz [(3)] folgt dagegen [74,48].

$$\frac{1}{T} = -\text{const} \ln U + \text{const},$$

<sup>3</sup>  $d^2S/dU^2$  ist die ‘Anderung der Entropievermehrung. Letzterer allein schreibt Planck im Ma’rz 1900 eine physikalische Bedeutung zu: sie n’amlich bilde ‘das numerische Maas fu’r die Irreversibilita’t des Processes oder fu’r die incompensierte Verwandlung von Arbeit in W’arme’ ([74] p. 731).’ [48]

---

also

$$\frac{d^2S}{dU^2} = -\frac{\text{const}}{U} \quad (\text{H-7})$$

... *Planck* kam nun auf den überaus glücklichen Gedanken, beide Formeln in der Weise zu verbinden, daß er setzte:

$$\frac{d^2S}{dU^2} = -\frac{\text{const}}{U(U + \text{const})} \quad (\text{H-8})$$

Mit Hilfe der Beziehung (H-5) und des Verschiebungsgesetzes führte ihn diese Formel (H-8) unmittelbar zu dem Energieverteilungsgesetz:

$$E_{\lambda,T} = \frac{c_1}{\lambda^5} \frac{1}{e^{c_2/\lambda T} - 1} \quad (\text{H-9})$$

En.: When *Rubens* visited *Planck* with his wife on Sunday, 7 October 1900, he said that the law recently established by *Lord Rayleigh* (*Phil. Mag.* 49, 539, 1900),

$$E_{\lambda,T} = c_1 \frac{T}{\lambda^4} e^{-c_2/\lambda T} \quad (\text{H-4})$$

holds for long wavelengths . . . . A universal radiation formula must in any case transit to this form for large  $\lambda T$ . After that conversation, *Planck* immediately made the following calculation: He was accustomed to calculating not with the radiation formula itself, but with the entropy  $S$  of a resonator in equilibrium with the radiation, which provides much simpler expressions as a function of its energy  $U$ .  $S$  and  $U$  are related to the temperature  $T$  as

$$\frac{dS}{dU} = \frac{1}{T} \quad (\text{H-5})$$

The proportionality of energy to temperature therefore yields [50,83].

$$\frac{dS}{dU} = \frac{\text{const}}{U},$$

or

$$\frac{d^2S}{dU^2} = -\frac{\text{const}}{U^2} \quad (\text{H-6})$$

In contrast, from Wien's energy distribution law [(3)] it follows that

$$\frac{1}{T} = -\text{const} \ln U + \text{const},$$

Therefore,

$$\frac{d^2S}{dU^2} = -\frac{\text{const}}{U} \quad (\text{H-7})$$

... *Planck* then had the extremely fortunate idea of combining both formulas in such a way that he set

$$\frac{d^2S}{dU^2} = -\frac{\text{const}}{U(U + \text{const})} \quad (\text{H-8})$$

With the help of the relationship (H-5) and the displacement law, this formula (8) led him directly to the energy distribution law

<sup>4</sup>This contradicts Klein [50] according to whom *Planck* did not realize the importance of that note [83].

$$E_{\lambda,T} = \frac{c_1}{\lambda^5} \frac{1}{e^{c_2/\lambda T} - 1}. \quad (\text{H-9})$$

Indeed, formulas (H-6) and (H-7) suggest the linear combination

$$\left(\frac{d^2S}{dU^2}\right)^{-1} = -\alpha \frac{U^2}{\text{const}} - \beta \frac{U}{\text{const}} = -\frac{U(U + \text{const})}{\text{const}}, \quad (5)$$

which is the inverse of formula (H-8). In his first communication of the energy distribution law (H-9), in his October 19, 1900, talk [75], Planck said, “Following this suggestion [finding alternatives to (H-7)] I have finally started to construct completely arbitrary expressions for the entropy which although they are more complicated than Wien’s expression still seem to satisfy just as completely all requirements of the thermodynamic and electromagnetic theory [74,48].

I was especially attracted by one of the expressions thus constructed which is nearly as simple as Wien’s expression. . . .” That expression is (with different  $\alpha$  and  $\beta$  than in (5))

$$\frac{d^2S}{dU^2} = \frac{\alpha}{U(U + \beta)}, \quad (6)$$

cf. (H-8) and (5).

The students may be asked to reconstruct Planck’s calculations. Moreover, this passage from Planck’s October talk provides an excellent opportunity to discuss the nature of theoretical discovery with students. Planck explicitly states that he was “constructing completely arbitrary expressions for the entropy”—a remarkably honest description of the heuristic, trial and error process that often characterizes scientific breakthroughs. Students may also be asked, what guided Planck’s choice among many possible expressions? Why was he “especially attracted” to this particular form? Such questions demystify the creative process in physics and show that even great discoveries involve guesswork guided by physical insight and mathematical elegance. Niels Bohr should have said, “This idea is crazy, but it is crazy enough to be true?”.

### 2.3. Planck’s December 14, 1900 Talk

Thus, the radiation law as presented in October 1900 (H-9) was merely an educated guess and Planck’s immediate ambition was to derive this experimentally confirmed prediction from physical principles. Unfortunately, he follows the strategy sketched above. Instead of searching for a non-classical expression for the mean energy  $U(T, \nu)$  of the resonators of frequency  $\nu$ , he sought for the entropy of the resonators of all frequencies  $\nu$  and of the radiation between them.

#### 2.3.1 The Statistical Setup. “Energy Elements”

Since he had already guessed the correct formula, he could have worked backwards for obtaining the corresponding entropy function and this is suggested by many historians of physics, e.g. [84]. Eventually, he took over some of Boltzmann’s 1877 results [8], in particular, the probabilistic notion of entropy as  $S \propto \ln W$  with  $W$  being the relative probability of the macroscopic state. “. . . a large number of linear, monochromatically vibrating resonators— $N$  of frequency  $\nu$  (per second),  $N'$  of frequency  $\nu'$ ,  $N''$  of frequency  $\nu''$ , . . . , with all  $N$  large numbers—which are properly separated and are enclosed in a diathermic medium with light velocity  $c$  and bounded by reflected walls. Let the system contain a certain amount of energy, the total energy  $E_t$  (erg) which is present partly in the medium as travelling radiation and partly in the resonators as vibrational energy. The question is how in a stationary state this energy is distributed over the vibration of the resonators and over the colors of the radiation present in the medium, and what will be the temperature of the total system.

To answer this question we first of all consider the vibrations of the resonators and try to assign them certain arbitrary energies, for instance, an energy  $E$  to the  $N$  resonators  $\nu$ ,  $E'$  to the  $N'$  resonators  $\nu'$ , The sum

$$E + E' + E'' + \dots = E_0 \quad (7)$$

must, of course, be less than  $E_t$ . The remainder  $E_t - E_0$  pertains then to the radiation present in the medium. We must now give the distribution of the energy over the separate resonators of each group, first of all the distribution of the energy  $E$  over the  $N$  resonators of

<sup>5</sup>”  $d^2S/dU^2$  is the change in entropy increase. Planck attributes physical significance to the latter alone in March 1900: namely, that it forms ‘the numerical measure of the irreversibility of the process or of the uncompensated conversion of work into heat’. ([74] p. 731).” [48]

frequency  $\nu$ . 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  $h = 6 \cdot 55 \times 10^{-27}$  erg sec. This constant multiplied by the common frequency  $\nu$  of the resonators gives us the energy element  $\epsilon$  in erg, and dividing  $E$  by  $\epsilon$  we get the number  $P$  of energy elements which must be divided over the  $N$  resonators. If the ratio thus calculated is not an integer, we take for  $P$  an integer in the neighborhood.” ([80] pp. 39f.) Although the latter prescription was never repeated by Planck ([80] p. 55, no. 32), it has entered the discussion about the interpretation of Planck’s calculations, in particular, whether the energy  $E_0$  (7) is really the sum of discrete energies  $E, E', \dots$  or if  $\epsilon$  is merely the size of energy intervals, see Section 4 below.

### 2.3.2 “Energy Elements” in Matter (Oscillators) or in Radiation (Resonators)?

In his publications before the October 19 talk, Planck uses both “resonator” and “oscillator”. Literally, ‘resonator’ means any vibrating system with a set of eigenfrequencies, or resonant frequencies. It may be, (i), the radiation field in a cavity. Then, the “energy elements” are “light quanta” as taken by Einstein in 1905 [24]. It may also be, (ii), a vibrating electrical charge like the Hertzian oscillator. Then, the “energy elements” are vibration quanta as taken by Einstein in 1907 for calculating the specific heat of solids [26]. According to Planck 1900 [76], in oscillators, the “energy elements” have to account for their mechanical degrees of freedom as well.

However, as we will see below, it is by far simpler to place the resonators in the wall of a cavity with constant temperature, corresponding to real black bodies [41]. Then, the total radiation energy of all resonators  $E_0$  (7) is not prescribed and, most important, the sets of resonators with different frequencies can be dealt with independent of each other, see Subsubsection 2.3.4. This is highly recommended for teaching. It eliminates the need to consider the radiation energy in the medium and allows students to concentrate on the essential combinatorial problem of distributing “energy elements” over resonators, i.e., of particles over cells.

### 2.3.3 The Almost Impossible Task

Planck considers the number  $\mathfrak{R}$  of different ways of sharing  $P$  energy elements  $\epsilon$  among  $N$  oscillators to be the relative probability of their state. For that number, Planck refers to combinatorics (p. 24; combination with repetition, see Table 1),

$$[\mathfrak{R} =] \frac{N(N+1)}{1 \cdot 2} \cdots \frac{(N+P-1)}{P} = \frac{(N+P-1)!}{(N-1)!P!}. \quad (8)$$

Then, Planck uses Stirling’s approximation  $n! \approx (n/e)^n$  to obtain, for  $N, P \gg 1$ ,

$$[\mathfrak{R} =] \frac{(N+P)^{N+P}}{N^N P^P}. \quad (9)$$

The probability for all resonators is (cf. [80] p. 41)

$$\mathfrak{R}_0 = \mathfrak{R} \cdot \mathfrak{R}' \cdot \dots \quad (10)$$

Notice that, usually, multiplying probabilities presupposes the mutual independence of them. By virtue of the constraint (7), however, they are not. This contradiction is not addressed in the literature we are aware of, and we will meet it below, again. In the concrete calculations, it plays no role, see below. Planck seeks the maximum of  $\mathfrak{R}_0$  w.r.t. the energy distributions  $E, E', E'', \dots$  at a given value of  $E_0$  (7). It is equivalent but much simpler to evaluate the maximum of  $\ln \mathfrak{R}_0$  than that of  $\mathfrak{R}_0$  itself, where

$$\begin{aligned} \ln \mathfrak{R} &= (N+P) \ln(N+P) - N \ln N - P \ln P \\ &= N \left\{ \left(1 + \frac{P}{N}\right) \ln \left(1 + \frac{P}{N}\right) - \frac{P}{N} \ln \frac{P}{N} \right\} \quad \text{etc.} \quad (11) \end{aligned}$$

Since  $P/N$  is independent of  $N$  (see formulas (16)), the corresponding entropy is extensive, as it should be.

We set  $E = E^{(1)}, E' = E^{(2)}, \dots, P = P^{(1)}, P' = P^{(2)}, \dots$ , etc. and recall  $E^{(i)} = P^{(i)} \epsilon^{(i)}, \epsilon^{(i)} := h\nu^{(i)} (i = 1, 2, \dots, \infty)$ . The constraint (7) becomes

$$\sum_{i=1}^{\infty} E^{(i)} = \sum_{i=1}^{\infty} P^{(i)} \epsilon^{(i)} = \sum_{i=1}^{\infty} P^{(i)} h\nu^{(i)} = E_0. \quad (12)$$

The students may be asked, whether a physically reasonable solution to this equation exists, if  $P$  is integer-valued, since the set  $\{\nu^{(i)} | i = 1, 2, \dots, \infty\}$  is not limited from above. Possibly, Planck has seen this contradiction and for this wrote, “. . . we take for  $P$  an integer in the neighborhood.” ([80] pp. 39f.) as quoted above.

For the sake of generality, the total number of “energy elements” may be given, too.

$$\sum_{i=1}^{\infty} P^{(i)} = P_0. \quad (13)$$

Then, the maximum of  $\ln \mathfrak{R}_0$  under the constraints (12), (13) means

$$\frac{d}{dP^{(i)}} \left[ \sum_{j=1}^{\infty} \left( (N^{(j)} + P^{(j)}) \ln(N^{(j)} + P^{(j)}) - N^{(j)} \ln N^{(j)} - P^{(j)} \ln P^{(j)} \right) - \beta \left( \sum_{j=1}^{\infty} P^{(j)} \epsilon^{(j)} - E_0 \right) - \alpha \left( \sum_{j=1}^{\infty} P^{(j)} - P_0 \right) \right] \stackrel{!}{=} 0; \quad i = 1, 2, \dots \quad (14)$$

( $\alpha, \beta$  being Lagrangian multipliers). Hence,

$$\begin{aligned} \ln(N^{(i)} + P^{(i)}) + 1 - \ln P^{(i)} - 1 - \beta \epsilon^{(i)} - \alpha &\stackrel{!}{=} 0, \\ \ln \left( \frac{N^{(i)}}{P^{(i)}} + 1 \right) - \beta \epsilon^{(i)} - \alpha &= 0, \quad \frac{N^{(i)}}{P^{(i)}} + 1 = e^{\beta \epsilon^{(i)} + \alpha}. \end{aligned} \quad (15)$$

Therefore, the *most probable* values are

$$P^{(i)} = \frac{N^{(i)}}{e^{\beta \epsilon^{(i)} + \alpha} - 1}; \quad E^{(i)} = P^{(i)} \epsilon^{(i)} = \frac{N^{(i)} \epsilon^{(i)}}{e^{\beta \epsilon^{(i)} + \alpha} - 1}. \quad (16)$$

The denominators exhibit the ‘Planckian form’ (with  $\alpha = 0$  and provided that  $\beta > 0$ ).

Here, the usual analytical calculation of the unknown parameters by means of the constraints (12), (13) and the thermodynamic relation  $1/T = dS/dE$  is not possible. (Students may be asked to try that to get the experience that well-established solution methods not always work and that these cases are gateways to novel approaches.) Indeed, as Planck [76] remarks, “Es würde nun freilich sehr umständlich sein, die angegebenen Rechnungen wirklich auszuführen . . .” ([81] p. 60, [80] p. 10)

En: “It would, to be sure, be very complicated to perform explicitly the above-mentioned calculations. . .” ([80] p. 42) For this, largely following Becker ([5] (49.13) ff.), we take from eq. (14) the function

$$\begin{aligned} R_{\max} &:= \ln \mathfrak{R}_0 - \beta \sum_{j=1}^{\infty} P^{(j)} \epsilon^{(j)} - \alpha \sum_{j=1}^{\infty} P^{(j)} \\ &= \sum_{j=1}^{\infty} \left( (N^{(j)} + P^{(j)}) \ln(N^{(j)} + P^{(j)}) - N^{(j)} \ln N^{(j)} - P^{(j)} \ln P^{(j)} \right. \\ &\quad \left. - \beta P^{(j)} \epsilon^{(j)} - \alpha P^{(j)} \right). \end{aligned} \quad (17)$$

For in the conventional maximization of  $\mathfrak{R}_0$  using eqs. (14), the terms  $\beta E_0$  and  $\alpha P_0$  fall out. We have

$$\frac{dR_{\max}}{d\beta} = -\sum_{j=1}^{\infty} P^{(j)} \epsilon^{(j)} = -E_0; \quad \frac{dR_{\max}}{d\alpha} = -\sum_{j=1}^{\infty} P^{(j)} = -P_0. \quad (18)$$

Further, assuming  $\alpha$  and  $\beta$  to be functions of  $E_0$  and  $P_0$ , we have

$$\frac{dR_{\max}}{dE_0} = \frac{\partial R_{\max}}{\partial \alpha} \frac{\partial \alpha}{\partial E_0} + \frac{\partial R_{\max}}{\partial \beta} \frac{\partial \beta}{\partial E_0} = \frac{\partial \ln \mathfrak{R}_0}{\partial E_0} - \beta - E_0 \frac{\partial \beta}{\partial E_0} - P_0 \frac{\partial \alpha}{\partial E_0}. \quad (19)$$

Therefore,

$$\frac{\partial \ln \mathfrak{R}_0}{\partial E_0} = \beta \quad \frac{\partial R_{\max}}{\partial \alpha} = -P_0; \quad \frac{\partial R_{\max}}{\partial \beta} = -E_0. \quad (20)$$

Analogously,

$$\frac{dR_{\max}}{dP_0} = \frac{\partial R_{\max}}{\partial \alpha} \frac{\partial \alpha}{\partial P_0} + \frac{\partial R_{\max}}{\partial \beta} \frac{\partial \beta}{\partial P_0} = \frac{\partial \ln \mathfrak{R}_0}{\partial P_0} - E_0 \frac{\partial \beta}{\partial P_0} - \alpha - P_0 \frac{\partial \alpha}{\partial P_0} \quad (21)$$

and

$$\frac{\partial \ln \mathfrak{R}_0}{\partial P_0} = \alpha \quad \frac{\partial R_{\max}}{\partial \alpha} = -P_0; \quad \frac{\partial R_{\max}}{\partial \beta} = -E_0. \quad (22)$$

The latter two equations repeat those in eqs. (20).

Now, thermodynamics tells us for the entropy  $S_0 = k \ln \mathfrak{R}_0$

$$\frac{\partial S_0}{\partial E_0} = \frac{1}{T}; \quad \frac{\partial S_0}{\partial P_0} = -\frac{\mu}{T} \quad (23)$$

( $\mu$  being the chemical potential). Hence,

$$\beta = \frac{1}{kT}; \quad \alpha = -\frac{\mu}{kT}. \quad (24)$$

Inserting that into formulas (16) yields

$$P^{(i)} = \frac{N^{(i)}}{e^{h\nu^{(i)}/kT - \mu/kT} - 1}; \quad E^{(i)} = P^{(i)} h\nu^{(i)} = \frac{N^{(i)} h\nu^{(i)}}{e^{h\nu^{(i)}/kT - \mu/kT} - 1}. \quad (25)$$

### 2.3.4 Treating the Sets of $N^{(i)}$ Resonators Separately

“. . . danach zeigen Resonatoren von großer Schwingungszahl eine besondere Habgier nach Energie (wobei es ihnen dann beim Austausch der Energieelemente geschehen kann, dass sie besonderes wenige davon bekommen).” (Hendrik Antoon Lorentz 1910 [60] 5th lecture, p. 238)

En.: . . . according to this, resonators with a large number of vibrations show particular greed for energy (whereby it can then happen to them during the exchange of the energy elements that they get particularly few of them).

The calculations become straightforward if one considers the common radiation field in all resonators of frequency  $\nu^{(i)}$  as a subsystem in thermal equilibrium with all other sets of resonators. Then, one can require—as implicitly done by Planck in 1901 ([77] 5) and explicitly, e.g., by Larmor in 1910 [57]—that

$$\frac{1}{T} = k \frac{d \ln \mathfrak{R}_1}{dE_1} = k \frac{d \ln \mathfrak{R}_2}{dE_2} = \dots \quad (26)$$

According to formula (11), the reduced entropy  $S/k$  per resonator of frequency  $\nu$  equals (cf. [77] (6))

$$\frac{S}{k} = \frac{1}{N} \ln \mathfrak{R} = \left\{ \left( 1 + \frac{U}{\epsilon} \right) \ln \left( 1 + \frac{U}{\epsilon} \right) - \frac{U}{\epsilon} \ln \frac{U}{\epsilon} \right\}. \quad (27)$$

From Wien's displacement law (2) Planck derives the formula ([77] (10))

$$S = f \left( \frac{U}{\nu} \right). \quad (28)$$

Comparing that with formula (27) yields  $\epsilon = h\nu$  with some constant  $h$ . Furthermore (cf. [77] (9)),

$$\frac{1}{T} = \frac{dS}{dU} = \frac{k}{h\nu} \ln \left( 1 + \frac{h\nu}{U} \right). \quad (29)$$

Consequently ([77] (11)),

$$U = \frac{h\nu}{e^{h\nu/kT} - 1}. \quad (30)$$

Inserting that expression into formula (4) gives Planck's radiation law (1). Notice that the entropy (27) is *extensive*, since  $U/\epsilon$  is independent of  $N$ .

This is *not* because the "energy elements" (photons) are indistinguishable (what they are not if they are in different resonators), but because the energy elements of one and the same frequency are interchangeable. This point about extensiveness is crucial for teaching. Students often assume that quantum statistics is required to explain why entropy is extensive. Here we see that interchangeability—a concept that applies to classical particles as well—already yields extensivity. This prepares the ground for a more nuanced discussion of (in)distinguishability in Section 5.

For Bose's 1924 resolution of those difficulties, see Subsubsection 3.5.2.

### 3. Boltzmann's, Planck's, and Bose's Combinatorics in Terms of Bach's 3-level Description

Planck's physical foundation of his radiation law [76] represents a good occasion, (i), to discuss the relationship between physics and mathematics in general and, (ii), to exemplify it by means of combinatorics in particular. At once, it proves the rule that the study of the masters of the past can be replaced only with more studies of the masters of the past.

In view of no other possibilities, Planck [76] eventually turns to combinatorics<sup>9</sup>. However, while Boltzmann's exploitation of combinatorics remained within classical physics, Planck's use of it went beyond that. Why, how?

For the sake of systematization of the various attempts of explanation ([65] pp. 557ff.), we propose to describe Planck's and Boltzmann's (and Bose's) combinatorics in terms of Bach's 3-level description [3]. It provides an exceptionally clear framework for teaching combinatorial statistics for the purpose of statistical mechanics for three-dimensional systems, i.e., Maxwell-Boltzmann, Bose-Einstein, and Fermi-Dirac statistics; parastatistics and anyons being beyond the scope of this contribution. By distinguishing between configuration, occupation, and occupancy, students can see how different physical assumptions lead to different counting methods and, consequently, different statistical distributions. We recommend to introduce this scheme early in statistical mechanics courses and returning to it whenever new distributions are encountered. We will see that Planck in 1900 uses the same "complexions" as Boltzmann in 1877 but differently calculates the entropy from them. For the same sake, in Subsubsection 3.4.2, we will sketch Planck's different calculation of state probability and entropy in his 1906 lectures on the theory of heat radiation [78]. Surprisingly enough, he unnoticed switches from occupations to configurations and ends up with a wrong formula for the entropy. Finally, in Subsubsection 3.5.2, we will comment on Bose's 1924 treatment [10], in which the difficulties of Planck's 1900 approach [76][77] described in Subsection 2.3 have been overcome.

#### 3.1 Few Counting Combinatorics. The Six-fold Way

We begin, however, with the six simplest cases within counting combinatorics, where three of them are directly related to Bach's

<sup>9</sup>For crash courses on combinatorics, see, e.g. [11][56]. Feller's classic [32] is comprehensive, but the sequence of examples and theorems is unconventional.

scheme. As they are a subset of the ‘twelve-fold way’ [97], we propose to term them the ‘six-fold way’.

Selection of $k < n$ particles				Selection of all $n$ particles	
combination		variation		permutation	
w/o rep.	with rep.	w/o rep.	with rep.	w/o rep.	with repetition
$\binom{n}{k}$	$\binom{n+k-1}{k}$	$\frac{n!}{(n-k)!}$	$n^k$	$n!$	$\frac{n!}{k!}$

**Table 1: The Six-fold Way, cf. [90]**

Those six cases are the following ones, see Table 1.

### 1. Combination *without* Repetition

Draw of  $k$  particles from a set of  $n > k$  particles. The drawn  $k$  particles

are *not* returned to the set. There are

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \quad (31)$$

different cases/possibilities. The  $k$  drawn particles and the  $(n-k)$  not- drawn particles are both and *interchangeable* in their order. Despite the numerator  $n!$  this interchangeability determines the number of cases.

Example: Lotto 6 out of 49:  $\binom{49}{6} = 13, 983, 816$ . Here, the individual

Example: Fermi-Dirac distribution ([5] (49.6)). Here, the individual numbers does not play any role.

### 2. Combination *with* repetition

Draw of  $k$  particles from a set of  $n > k$  particles. Here, after having been drawn, the particles *are* put back into the cell and hence can be drawn, *again*. There are

$$\binom{n+k-1}{k} := \frac{(n+k-1)!}{k!(n-1)!} = \frac{n}{1} \frac{n+1}{2} \cdots \frac{n+k-1}{k} \quad (32)$$

different cases/possibilities.

Example: The number of ways to distribute  $k$  interchangeable “energy elements” among  $n$  non-interchangeable resonators (Planck 1900 [76][77]).<sup>10</sup>

Example: The number of different occupation number vectors (Level 2 in Bach’s scheme, see Subsection 3.4), e.g., Boltzmann’s 1877 [8] and Planck’s 1900 [76][77] complexions.

### 3. Variation *without* Repetition

Draw  $k$  out of  $n > k$  not interchangeable particles. The particles drawn are not returned into the cell (urn). The number of possibilities equals

$$\frac{n!}{(n-k)!} = \binom{n}{k} k! \quad (33)$$

The order of the not drawn particles plays no role.

Example: 4 particles are selected from 6 particles that are not inter- changeably. There are  $6!/(6-4)! = 360$  ways to select and arrange those 4 particles.

### 4. Variation *with* Repetition

The number of ways of arranging  $k$  particles selected from a set of  $n$  non-interchangeable particles. where selected particles are put back, equals

<sup>10</sup>Planck did not put the precondition  $n > k$ , i.e.,  $N > P$ , see Subsection 2.3. According to formulas (25),  $0 < P/N < \infty$ . The students may be asked to tackle the special case of Bose-Einstein condensation using combinatorics.

$$n^k \tag{34}$$

For each draw, there are  $n$  possibilities to select one of the  $n$  particles in the cell/urn. Since the selected particles are put back into the cell/urn, the draws are independent of each other. For this, there are  $k$ -times  $n$  possibilities. The selected particles are not interchangeable, hence, their order is important.

Example: The number of configurations, Level 1 in Bach’s scheme (see Subsection 3.3).

### 5. Permutation without Repetition

The number of arrangements of  $n$  particles that cannot be interchanged equals

$$n! \tag{35}$$

Example:

1					
1	2				
2	1				
1	1	2	2	3	3
2	3	1	3	1	2
3	2	3	1	2	1

**Table 2: Permutations Without Repetition;  $n = 1, 2, 3$ ; the Recursion Formula  $(n + 1)! = (n + 1) n!$  is Easily Seen**

### 6. Permutation with Repetition

Here,  $k_1, k_2, \dots$  particles are *interchangeable* among themselves. Hence, their mutual interchanges do not create different sequences of the  $n$  particles. The number of arrangements of those  $n$  particles thus equals

$$\frac{n!}{k_1!k_2! \dots} \tag{36}$$

Example:  $\bullet_1 \otimes \odot \otimes \otimes \bullet_2 = \bullet_2 \otimes \odot \otimes \otimes \bullet_1$

Example: The number of different occupancies (Level 3 in Bach’s scheme, see Subsection 3.5), e.g. Boltzmann 1877 [8], Bose 1924 [10].

### 3.2 Bach’s Three-level Scheme ‘Configuration Occupation Occupancy’

According to Bach ([2], [3] sects. 3.2.1, 5.1), Boltzmann [6] has invented the scheme  $(P, N)$  for the problem of distributing  $P$  equal particles (Boltzmann: pieces of kinetic energy  $\varepsilon = \frac{m}{2}v^2$ , Planck: “energy elements”  $\epsilon = h\nu$ ) onto  $N$  equal cells (Boltzmann: molecules, Planck: resonators). There are three possibilities and corresponding levels of description (cf. [3]; in the literature, the names ‘configuration’ *etc.* are not uniquely used):

**1. Configuration:** Both the particles and the cells are not interchangeable. The question is, in *which* cell (in cell no. 1, 2, . . . ,  $N$ ) is *which* particle (particle no. 1, 2, . . . ,  $P$ )?

- *Which* and the number set  $\{1, 2, \dots\}$  belong to ‘not interchangeable’.

- Example: At Christmas,  $N$  persons shall obtain  $P = N$  gifts of equal price, one gift each. *Which* person shall receive *which* gift?

**2. Occupation:** The particles are interchangeable, while the cells are still not interchangeable. The question is, *how many* particles (0, 1, 2, . . . ,  $P$ ) are in cell no. 1, 2, . . . ,  $N$ ?

- *How many* and the number set  $\{0, 1, 2, \dots\}$  belong to ‘interchangeable’.

- Example: *How many* equal chocolates are in a small, medium, or large box? Which chocolate is in which box plays no role.

**3. Occupancy:** Both the particles and the cells are interchangeable. The question is, how many cells (0, 1, 2, . . . ,  $N$  cells) are there with how many particles (with 0, 1, 2, . . . ,  $P$  particles)?

- Example: How many bags of 0 kg (bags damaged during transport), 1 kg, 2 kg, . . . potatoes were delivered?

### 3.3 Configuration

A configuration describes, in which cell is particle 1 till in which cell is particle  $P$ . It represents a complete description of the distribution of the  $P$  individual, distinguished, not interchangeable particles onto the  $N$  individual, distinguished, not interchangeable cells. This

information is recorded in the configuration number vector  $\vec{j} = (j_1, j_2, \dots, j_P)$ , where  $j_s$  ( $s = 1, 2, \dots, P$ ) means that particle  $s$  is in cell  $j$  ( $j = 1, 2, \dots, N$ ).

In contrast to the occupation and occupancy number vectors below, there are no constraints imposed on  $\vec{j}$ . For this, the  $P$  particles are distributed independent of each other, so that each of the  $P$  particles has  $N$  different possibilities of being in one of the  $N$  cells. The number of different configurations thus equals

$$Q_1(P, N) := N^P, \quad (37)$$

see ‘variation with repetition’ in Subsection 3.1, Table 1.

In terms of micro- and macro-states, a configuration describes a micro-state, since, within this 3-level scheme, there is no more detailed description.

We will meet an example in Subsubsection 3.4.2 and in Section 5.

### 3.4 Occupation. Boltzmann. Complexion. Planck

Often, the particles are interchangeable. For instance, the red balls in a snooker game are distinguishable by their positions on the table. Nevertheless, an interchange of any two of them does not influence the outcome of the game. For this, they are interchangeable in the sense of the rules of the game.

In this case, it is not relevant, *which* particles are in a cell under consideration, but only *how many* particles are in it. This information is recorded in the occupation number vector  $\vec{p} = (p_1, p_2, \dots, p_N)$ , where  $p_r$  ( $p_r = 0, 1, \dots, P$ ) is the number of particles in cell  $r$  ( $r = 1, 2, \dots, N$ ). There is the obvious constraint

$$p_1 + p_2 + \dots + p_N = P. \quad (38)$$

Notice, that this constraint makes the  $P$  particles to be distributed *not in-dependent* of each other. This difference to the configurations is crucial for the difference between Planck’s 1900 “energy elements” and Einstein’s 1905 “light quanta”, see Subsection 5.3 below.

#### 3.4.1 Boltzmann 1877 and Planck 1900

“Any such distribution, in which the first molecule may have a kinetic energy of e.g.,  $2\varepsilon$  [ $p_1 = 2$ ], the second may have  $6\varepsilon$  [ $p_2 = 6$ ], and so on, up to the last molecule, we call a complexion. . .” (Boltzmann [8] p. 1977)<sup>11</sup>

1	2	3	4	5	6	7	8	9	10
7	38	11	0	9	2	20	4	4	5

**Table 3: Example of a Complexion According to Planck 1900 [76][77]; 100 Energy Elements in 10 Resonators. Upper Row: Resonators 1, 2, . . . , 10; Lower Row: Number of Energy Elements Hosted,  $p_1, p_2, \dots, p_{10}$ . The resonators are Still not Interchangeable so that the Sequence of the Figures in the Second Row Matters**

The number of such complexions equals the number of occupation number vectors  $\vec{p}$  which obey the constraint (38). It is the *multiset number* (cf. [3] p. 59, [32] (5.2))

$$\left( \binom{N}{P} \right) := \binom{N-1+P}{P} = \frac{(N-1+P)!}{P!(N-1)!} =: Q_2(P, N), \quad (39)$$

see ‘combination with repetition’ in Subsection 3.1, Table 1. This is Boltzmann’s 1877  $J$  ([8] p. 1983) and Planck’s expression (8). In terms of particles in states,  $N$  is the degeneracy of the state under consideration [89].

Referring to Kries [52], Planck 1901 ([77] 4) remarks that all complexions are supposed to be of equal likelihood.

Later, Ehrenfest and Kamerlingh Onnes [23] proposed the following “intuitive interpretation” ([89] p. 427). The distribution of  $P$  energy elements  $\varepsilon$  over  $N$  resonators is symbolized as a string of  $\varepsilon$ ’s and symbols ‘|’ for the boundaries of the resonators, e.g.

<sup>11</sup>Boltzmann’s tables are rather large and have to be rewritten for fitting into Bach’s scheme, see [30][31].

<sup>12</sup>For an explanation of Kries’ arguing in modern terms, see Reiche’s Comment 4 to Planck 1901 [77] in Ostwalds Klassiker 206.

<sup>13</sup>Contrary to some authors (e.g. [65] p. 559, [81] Remark V [5] p. 87) and, notably, the word “derivation” in the title of Ref. [23], it is not really a deduction. For Ehrenfest and Kamerlingh Onnes have introduced  $N - 1$  partitions between the  $N$  resonators just to reflect the  $N - 1$  in (8) and (39), see p. 298, fn. \*.

$$\epsilon \mid \epsilon \epsilon \mid \mid \epsilon . \quad (40)$$

In this example,  $P = 4$  energy elements  $\epsilon$  are distributed among  $N = 4$  resonators such that the first resonator contains one element, the second one two, the third one zero, and the fourth one one. Such a string contains  $P$  times  $\epsilon$  and  $N - 1$  times '|'. Hence,  $(P + N - 1)!$  possible permutations exist.

However, the  $P!$  permutations among the  $\epsilon$  and the  $(N - 1)!$  permutations among the '|' correspond to *one and the same* distribution.

The Ehrenfest-Kamerlingh Onnes visualization is a powerful teaching tool that makes abstract combinatorics tangible. We recommend having students construct such diagrams for small values of  $P$  and  $N$  (e.g.,  $P = 3$ ,  $N = 2$ ) and enumerate all possible distributions. This hands-on activity helps them to understand why formula (7) takes the form it does. Why the cell boundaries in (40) are interchangeable, while the cells themselves are not? Talented students may be asked to develop a real proof of formula (39). They may, (i), begin with the examples just mentioned and use mathematical induction, (ii), apply generating functions [96], (iii), use mathematical and physical intuition.

Therefore, Planck's 1900 "complexion" agrees with Boltzmann's "complexion" at the beginning of his 1877 article as quoted above. Boltzmann, however, a little bit later, moves from occupation to occupancy numbers, see his formulas (1) . . . (3) with the two constraints which exactly correspond to the constraints (49) *etc.* in Subsection 3.5.

The reason for Boltzmann's move is this: The (relative) probability of a state as given by  $Q_2(P, N)$  (39) is static. Hence, the corresponding entropy is that of an *equilibrium* state. In contrast, the occupancy numbers in Subsection 3.5 allow for a motion from non-equilibrium to equilibrium.

The number of complexions for the distribution of  $P$  particles (energy pieces/elements) over  $N$  cells (molecules, resonators) equals the number of complexions for the distribution of  $N - 1$  cells (molecules, resonators) over  $P + 1$  particles (energy elements). Obviously, Planck has in mind the distribution of the energy elements  $h\nu$  over the resonators, not the reverse one. We will return to this issue in Section 4.

### 3.4.2 Planck 1906

For substantiating the discussion below, let us add some arguing from Planck's 1906 lectures on the theory of heat radiation [78]. In contrast to his December 14, 1900 talk [76] and subsequent 1901 article [77] (see Subsection 2.3 and Subsubsection 3.4.1), he includes the configurations for calculating the probability of a "complexion". This makes the complexion used here quite different from the complexion used there. As a consequence, the resulting entropy is not extensive, but displays Gibbs's paradox of mixing gases of equal molecules.

Thus, Planck considers a very large number  $P$  (Planck's  $N$ , though we keep the above notation) of equal molecules distributed over  $N$  small, but finite *space elements* of a given volume, each containing a specified number of molecules. That specification defines the macroscopic distribution in space (no. 122). The manner in which the molecules are distributed within the every separate space element is immaterial.

The space elements are not interchangeable and enumerated by numbers,  $1, 2, 3, \dots, N$ . The numbers of molecules in them be  $p_1, p_2, p_3, \dots, p_N$  (Planck's  $N_1, N_2, N_3, \dots$ ). Planck's constraint

$$N_1 + N_2 + N_3 + \dots = N \quad (\text{Planck-165})$$

is our constraint (38). Hence,  $\vec{p} = (p_1, p_2, p_3, \dots, p_N)$  is the corresponding occupation number vector as above.

He gives the following example (cf. formulas (168) f.;  $P = 10$ ,  $N = 7$ ):

$$p_1 = 1, p_2 = 2, p_3 = 0, p_4 = 0, p_5 = 1, p_6 = 4, p_7 = 2. \quad (41)$$

<sup>14</sup>This method is also known as 'stars and bars' ([32] p. 38) and other names [96].

1	2	3	4	5	6	7	8	9	10
6	1	7	5	6	2	2	6	6	7

**Table 4: A Complexion According to Planck 1906 [78]**

Upper row: molecules no. 1 . . . 10; lower row: volume elements  $r$ , in which the molecules lie,  $r = 1 \dots 7$ . Notice, that this table displays not an occupation, but a configuration as there is the configuration number vector  $\vec{j} = (6, 1, 7, 5, 6, 2, 2, 6, 6, 7)!$

Here, Planck implicitly changes from level 2, occupation (how many particles are in cell  $r$ ), to level 1, configuration (which particle is in which cell, see Subsection 3.3). “The number of all possible complexions required is now easily found by inspecting the lower of the two lines of figures in (169) [Table 4].”

For each given occupation number vector  $\vec{p}$ , there are

$$Q_3(\vec{p}, P) := P_P(\vec{p}) := \binom{P}{p_1 \dots p_N} := \frac{P!}{p_1! p_2! \dots p_N!} \quad (42)$$

configurations ([3] p. 59). This is Planck’s formula (170) for the probability  $W$  of a “complexion”. Obviously, this “complexion” is no longer the complexion of Boltzmann 1877 and Planck 1900 in the foregoing Subsubsection 3.4.1!

For large numbers  $P$  and  $p_1 \dots p_N$ , Stirling’s formula allows for simplifying formula (42) for  $W$  as

$$W = Q_3(\vec{p}, P) \approx \left(\frac{P}{p_1}\right)^{p_1} \left(\frac{P}{p_2}\right)^{p_2} \dots \left(\frac{P}{p_N}\right)^{p_N}. \quad (43)$$

This approximation immediately yields the corresponding entropy as

$$S = k_B \ln W = k_B \sum_{r=1}^N p_r \ln \frac{P}{p_r} = -k_B P \sum_{r=1}^N \bar{p}_r \ln \bar{p}_r; \quad \bar{p}_r := \frac{p_r}{P}. \quad (44)$$

It looks extensive, but it is *not* as we will show next.

In contrast to the static probability  $Q_2(P, N)$  (39), the probability  $Q_3(\vec{p}, P)$  (42) allows for a development from a non-equilibrium distribution  $\vec{p}$  to the equilibrium distribution  $\vec{p}|_{S=\max.}$ . (Planck himself writes nothing about the maximization of the entropy (44) w.r.t.  $\vec{p}$ .) Accounting for the constraint (38), this means

$$\vec{p}|_{S=\max.} = \max_{\{p_r\}} \left\{ S + \lambda \left[ \sum_{r=1}^N p_r - P \right] \right\} \quad (45)$$

( $\lambda$  being a Lagrangian multiplier). The result is  $p_1|_{S=\max.} = \dots = p_N|_{S=\max.} = P/N$ . The molecules are evenly distributed over the volume elements (cf. [60] p. 250). This corresponds to maximum disorder as there is not any internal structure.

Thus, the maximum entropy (44) becomes

$$S = k_B P \ln N. \quad (46)$$

Obviously, it is *not extensive*. Dividing the volume into two equal parts with  $N/2$  cells and  $P/2$  particles each, the sum of the entropies of both halves equals

$$S_1 + S_2 = 2 \left( k_B \frac{P}{2} \ln \frac{N}{2} \right) = S - k_B P \ln 2. \quad (47)$$

This is called Gibbs’s paradox of mixing equal gases, see Subsection 5.1.

That artifact can be overcome through partitioning not the position space, but the momentum space. For otherwise free equal classical

<sup>15</sup>The microscopic state is the state as described by the coordinates and velocities of the molecules and the electromagnetic field strengths (no. 121).

particles in a box, which have got equal momentum (velocity) vectors are interchangeable in the sense of Newton's (Euler's) notion of state [28][29]. Halving the spatial volume does not affect the momentum distribution. Hence, if  $N$  denotes the number of cells in momentum rather than position space, we have

$$S_1 + S_2 = 2 \left( k_B \frac{P}{2} \ln N \right) = S \quad (48)$$

instead of (47). Of course, that contradicts the use of configurations in Table 4 and  $Q_3(\vec{p}, P)$  (42). For this, one has to turn to occupancy numbers, cf. Bose's treatment in Subsubsection 3.5.2 below. We will return to this issue in Section 5.

### 3.5 Occupancy. Boltzmann. Bose

Finally, not only the particles, but also the cells may be interchangeable. In this case, the cells solely differ by the number of particles hosted, so that the question consists in how many cells host zero, one, . . . ,  $P$  particles. This question is answered by the *occupancy* number vector  $\vec{n} = (n_0, n_1, \dots, n_P)$ , where  $n_s$  is the number of resonators with  $s$  particles ( $s = 0, 1, \dots, P$ ). The occupancy numbers are subject to the following *two*, obvious constraints,

$$n_0 + n_1 + n_2 + \dots + n_P = N \quad (49a)$$

$$0 \cdot n_0 + 1 \cdot n_1 + 2 \cdot n_2 + \dots + P \cdot n_P = P \quad (49b)$$

Notice, that those constraints make the  $P$  particles to be distributed *not* independent of each other. For given  $N$  and  $P$ , the number of different occupancy number vectors  $\vec{n}$  equals

$$Q_3(\vec{n}, N) = P_N(\vec{n}) = \binom{N}{n_0 \dots n_P} = \frac{N!}{n_0! n_1! \dots n_P!}, \quad (50)$$

cf. 'permutation with repetition' in Subsection 3.1, Table 1. For there are  $N!$  permutations of the  $N$  cells. All cells with the same number of particles hosted are interchangeable, so that their permutation does not yield a different occupancy number vector  $\vec{n}$ .

According to Spalek [89], Natanson, in 1911 [66], was the first to relate formula (50) to the issue of indistinguishability. For more details, see Section 5.

#### 3.5.1 Boltzmann 1877

Eventually, formula (50) has been used by Boltzmann (cf. [8] p. 1979, eqs. (1) f.). "The most likely state distribution will be for those  $w_0, w_1, \dots$  values [our  $n_0, n_1, \dots$ ] for which  $P$  [our  $Q_3(\vec{n}, N)$ ] is a maximum. . ." (p. 1977).

Here, of course, the two constraints (49) have to be accounted for.

The solution is ([8] (7))

$$n_s = n_0 x^s; \quad s = 0, 1, \dots, P. \quad (51)$$

The two constants  $n_0$  and  $x$  are determined by the two constraints (49) as

$$N = n_0 \sum_{s=0}^P x^s = n_0 \frac{1 - x^{P+1}}{1 - x},$$

$$P = n_0 \sum_{s=0}^P s \cdot x^s = n_0 x \frac{1 - (1 + P - Px)x^P}{(1 - x)^2}. \quad (52)$$

Now, by virtue of the constraints (49), one has  $0 < x < 1$ . For this, for large values of  $P$ , one can set  $x_p = 0$ . Then, one obtains

$$n_0 = \frac{N}{1 + P/N}, \quad x = \frac{P/N}{1 + P/N}. \quad (53)$$

<sup>16</sup>Actually, Boltzmann's molecules have a maximum kinetic energy of  $pe$ , where  $p \leq \lambda$ , i.e.,  $p \leq P$ . This is an unnecessary complication which we discard.

Here,  $x^p \rightarrow e^{-P} \rightarrow 0$  as  $P \rightarrow \infty$ ,  $P > N$ . This justifies  $x^p = 0$ . Of course, the correct limit transition is  $P \rightarrow \infty$ ,  $N \rightarrow \infty$ , where  $P/N = \text{const}$ .

Eventually, Boltzmann makes the analogous simplification, see his formulas (15). Unfortunately, he did not proceed to the corresponding entropy (54). Admittedly, there was no reason for him to keep  $\varepsilon$  finite.

The entropy divided by  $k_B$  and  $N$  becomes, using Stirling's approximation and setting  $x^p = 0$ , again,

$$\frac{S}{k_B N} = \frac{1}{N} \ln Q_3(\vec{n}, N) = \left(1 + \frac{P}{N}\right) \ln \left(1 + \frac{P}{N}\right) - \frac{P}{N} \ln \frac{P}{N}. \quad (54)$$

It is compatible with Planck's formula (27).

Nauenberg [68] asks, ". . . why Planck did not carry out this simple calculation that is missing in Boltzmann's paper" (Appendix A, p. 717). Perhaps, unfortunately, starting from his new relationship (4) between the mean electromagnetic energy of an oscillator and its surrounding electromagnetic field, Planck considered not only the field in the resonators but also the field in the medium around them. The question, why Planck—contrary to all experimental black bodies [41]—has imagined that setup in 1900, is beyond the scope of this article. Anyway, it is much easier—and much closer to the experiments [41]—to prescribe the temperature rather than field energies.

### 3.5.2 Bose 1924

Bose [10] overcomes the difficulties of Planck's 1900 probabilistic treatment (see Subsubsection 2.3.3) as follows.

1. The cells are not resonators, but the parts of a phase space of size 'spatial volume  $V$  times spherical shell  $4\pi h^3 \nu^2 d\nu/c^3$  in momentum space', see formula (56). The probabilities are connected with the radiation itself. Notice, that working with frequency intervals  $d\nu$  avoids the issue of line width.
2. The factor  $8\pi\nu^2/c^3$  in Planck's radiation law (1) (Bose's  $A/d\nu$ ) is obtained as the number of light quanta (Planck's "energy elements") in phase space cells of size  $h^3$ .
3. All frequencies  $\nu, \nu', \dots$  are dealt with at once, from the beginning till the end, while there is no additional radiation in space (in Planck's "diathermic medium").
4. Bose uses occupancy (Subsection 3.5) rather than occupation numbers (Subsection 3.4).
5. Eventually, the number of particles (light quanta) is set to infinity (though in a not rigorous manner).

We will outline Bose's treatment, (i), to illustrate Bach's Level 3 and, (ii), to point to some necessary corrections, a major methodical and a minor typo ones.

Bose's setup At the beginning, Bose stresses that, so far, the factor  $8\pi\nu^2/c^3$  has not yet been derived without classical theory. He himself, however, assumes the light quanta to move along *classical trajectories*.

"Der Momentanzustand des Quantums wird charakterisiert durch seine Koordinaten  $x, y, z$  und die zugehörigen Momente  $p_x, p_y, p_z \dots$ " (p. 179)

En: The current state of the quantum is characterized by its coordinates  $x, y, z$  and the corresponding momenta  $p_x, p_y, p_z \dots$

Thus, the factor  $8\pi\nu^2/c^3$  in Planck's radiation law (1) is semi-classically calculated as follows. A radiation quantum of frequency  $\nu$  is supposed to move in the restricted phase space  $\{x, y, z, p_x, p_y, p_z\}$  of spatial volume  $V$  and momentum sphere

$$p^2 \equiv p_x^2 + p_y^2 + p_z^2 = \frac{h^2 \nu^2}{c^2}. \quad (55)$$

Accordingly, the phase space volume belonging to the frequency interval  $d\nu$  equals

$$\begin{aligned} \iiint_V dx dy dz \iiint_{dp=(h/c)d\nu} dp_x dp_y dp_z &= V \int_{dp=(h/c)d\nu} 4\pi p^2 dp \\ &= V \cdot 4\pi \frac{h^2 \nu^2}{c^2} \frac{h d\nu}{c} = 4\pi \frac{h^3 \nu^2}{c^3} V d\nu. \end{aligned} \quad (56)$$

<sup>17</sup>Planck 1906 [78] had considered cells in space  $\{x, y, z\}$  instead of resonators, see Subsubsection 3.4.2. The resulting entropy is not extensive, exhibits Gibbs' paradox of mixing equal gases. Bose's partition of the momentum space  $\{p_x, p_y, p_z\}$  into cells avoids Gibbs' paradox as exemplified at the end of Subsubsection 3.4.2.

Now, this phase space volume is divided into cells of size  $h^3$  each. Moreover, the two polarization directions of electromagnetic radiation in free space (being a classical notion, too!) are accounted for. Then, the number of cells belonging to the  $i$ -th frequency interval  $d\nu^{(i)}$  ( $i = 1, 2, \dots$ ) becomes

$$N^{(i)} = 8\pi V \frac{(\nu^{(i)})^2}{c^3} d\nu^{(i)} \quad (57)$$

(Bose's  $A^s$ ).

**The probabilities of distributions** For each frequency interval  $d\nu^{(i)}$ , the number of possible distributions and thus the relative probability is postulated to equal

$$W^{(i)} = \binom{N^{(i)}}{n_0^{(i)} \dots n_P^{(i)}} := \frac{N^{(i)!}}{n_0^{(i)}! n_1^{(i)}! \dots n_P^{(i)}!} \quad (58)$$

This corresponds to formula (50), where our occupancy numbers  $n_0^{(i)}, n_1^{(i)}, \dots$  are Bose's  $p_0^s, p_1^s, \dots$ .

Altogether, the relative probability of a state as given by all  $n_{s^{(i)}}^{(i)}$  is postulated

$$W = \prod_i \frac{N^{(i)!}}{n_0^{(i)}! n_1^{(i)}! \dots n_{P^{(i)}}^{(i)}!} = \prod_i W^{(i)}. \quad (59)$$

Usually, the product formula  $W = \prod_i W^{(i)}$  assumes the probabilities  $W^{(i)}$  (58) to be *independent* of each other. This independence would contradict their interdependence according to the constraint (61c). The possible contradiction eventually disappears, see below.

$W$  (59), or, applying Stirling's approximation as usual,

$$\ln W = \sum_i N^{(i)} \ln N^{(i)} - \sum_i \sum_{s^{(i)}=0}^{P^{(i)}} n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)} \quad (60)$$

should be maximum at the auxiliary conditions

$$n_0^{(i)} + n_1^{(i)} + n_2^{(i)} + \dots + n_{P^{(i)}}^{(i)} = N^{(i)}, \quad (61a)$$

$$0 \cdot n_0^{(i)} + 1 \cdot n_1^{(i)} + 2 \cdot n_2^{(i)} + \dots + P^{(i)} \cdot n_{P^{(i)}}^{(i)} = P^{(i)}, \quad (61b)$$

$$\sum_i P^{(i)} h\nu^{(i)} = E, \quad (61c)$$

where  $E$  is the given total energy of the radiation. When compared with the constraints (49), eq. (61c) represents an additional constraint which inter- connects the distributions  $i = 1, 2, \dots$  (the frequency). It leads to the special formula (72) for the entropy, while we will obtain Boltzmann's expression (54), too.

As mentioned before, for establishing the general formulas for the distribution and eventually the radiation law, it is by far preferable not to prescribe energies. In all experiments we are aware of [41], the temperature of the black body rather than the energy of the radiation is given.

**Maximizing the probability** Now, the variations  $\delta n_s^{(i)}$  of all  $n_s^{(i)}$  in formula (60) and eqs. (61) yield (all sums over  $s^{(i)}$  run from 0 to  $P^{(i)}$ )

<sup>18</sup>That point has been overlooked in Master's [63] and Nolte's [70] appraisals, too.

$$-\sum_i \sum_{s^{(i)}} \delta n_{s^{(i)}}^{(i)} \left( \ln n_{s^{(i)}}^{(i)} + 1 \right) = \delta \ln W, \quad (62a)$$

$$\sum_{s^{(i)}} \delta n_{s^{(i)}}^{(i)} = 0, \quad (62b)$$

$$\sum_{s^{(i)}} s^{(i)} \delta n_{s^{(i)}}^{(i)} = \delta P^{(i)}, \quad (62c)$$

$$\sum_i \delta P^{(i)} h\nu^{(i)} = 0, \quad (62d)$$

Generally speaking, the variation  $\delta P^{(i)}$  (62c) is a remarkable difference to Planck 1900. Finally, however, as indicated above, it will play no role.

Further, using condition (62b) in eq. (62a) and eq. (62c) in condition (62d) leads to

$$\sum_i \sum_{s^{(i)}} \delta n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)} = \delta \ln W, \quad (63a)$$

$$\sum_{s^{(i)}} \delta n_{s^{(i)}}^{(i)} = 0, \quad (63b)$$

$$\sum_{s^{(i)}} s^{(i)} \delta n_{s^{(i)}}^{(i)} = \delta P^{(i)}, \quad (63c)$$

$$\sum_i \sum_{s^{(i)}} s^{(i)} \delta n_{s^{(i)}}^{(i)} h\nu^{(i)} = 0, \quad (63d)$$

Therefore, the maximization condition can be written as (we write  $\beta (= 1/kT)$  for Bose's  $1/\beta (= kT)$ )

$$\sum_i \sum_{s^{(i)}} \delta n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)} + \sum_i \lambda^{(i)} \sum_{s^{(i)}} \delta n_{s^{(i)}}^{(i)} + \beta \sum_i h\nu^{(i)} \sum_{s^{(i)}} s^{(i)} \delta n_{s^{(i)}}^{(i)} = 0. \quad (64)$$

Since all variations  $\delta n_{s^{(i)}}^{(i)}$  are independent of each other, one has

$$\ln n_{s^{(i)}}^{(i)} + \lambda^{(i)} + \beta h\nu^{(i)} s^{(i)} = 0. \quad (65)$$

With  $\exp(-\lambda^{(i)}) =: B^{(i)}$ , this makes Bose's solution

$$n_{s^{(i)}}^{(i)} = B^{(i)} e^{-\beta h\nu^{(i)} s^{(i)}} \quad (66)$$

obvious. It corresponds to Boltzmann's solution (51) with  $B = n_0$ ,  $e^{-\beta h\nu} = x$  (Boltzmann did not consider a set of ensembles  $\{i \mid i = 1, 2, \dots\}$ ).

As always, the Lagrangian multipliers are calculated by means of the constraints, as done by Boltzmann 1877, see Subsubsection 3.5.1. Bose writes, (i), using constraint (61a),

$$N^{(i)} = \sum_{s^{(i)}} B^{(i)} e^{-\beta h\nu^{(i)} s^{(i)}} = \frac{B^{(i)}}{1 - e^{-\beta h\nu^{(i)}}}, \quad (67)$$

<sup>19</sup>For the sake of accuracy, we write  $s(i)$  rather than  $s$  (Bose's  $r$ ).

In turn,

$$B^{(i)} = N^{(i)} \left( 1 - e^{-\beta h \nu^{(i)}} \right). \quad (68)$$

(ii), inserting formula (66) into the constraint (61b),

$$P^{(i)} = \sum_{s^{(i)}} s^{(i)} B^{(i)} e^{-\beta h \nu^{(i)} s^{(i)}} = B^{(i)} \frac{e^{-\beta h \nu^{(i)}}}{(1 - e^{-\beta h \nu^{(i)}})^2} = \frac{N^{(i)}}{e^{\beta h \nu^{(i)}} - 1}, \quad (69)$$

or,

$$\frac{P^{(i)}}{N^{(i)}} = \frac{1}{e^{\beta h \nu^{(i)}} - 1}. \quad (70)$$

This ratio  $P^{(i)} / N^{(i)}$  corresponds exactly to Planck's  $P^{(i)} / N^{(i)}$  (??), where  $\beta = 1/kT$  will be obtained in what follows.

In both cases, it is unclear whether Bose—like Boltzmann—abandons terms proportional to  $\exp\{-\beta h \nu^{(i)} P^{(i)}\}$ , or lets  $P^{(i)} \rightarrow \infty$ . The limit transition  $P^{(i)} \rightarrow \infty$  contradicts the constraint (61c), according to which  $P^{(i)} \sim E / h \nu^{(i)}$  for  $\nu^{(i)} \rightarrow \infty$ .

**Radiation energy and entropy** Inserting formula (69) into eq. (61c) and accounting for formula (57), one obtains the total radiation energy as (cf. p. 181)

$$E = \sum_i E^{(i)} = V \sum_i \frac{8\pi h (\nu^{(i)})^3}{c^3} \frac{d\nu^{(i)}}{e^{\beta h \nu^{(i)}} - 1}. \quad (71)$$

“Using the results obtained so far, one further finds” (p. 181)

$$S = k_B \left[ \beta E - \sum_i N^{(i)} \ln \left( 1 - e^{-\beta h \nu^{(i)}} \right) \right] \quad (72)$$

(correct is  $e^{\beta h \nu^{(i)}}$  see below). Applying  $\frac{\partial S}{\partial E} = \frac{1}{T}$ , it follows that  $\beta = \frac{1}{k_B T}$ .

Inserting that into formula (71), one obtains

$$E = \sum_i E^{(i)} = V \sum_i \frac{8\pi h (\nu^{(i)})^3}{c^3} \frac{d\nu^{(i)}}{e^{\beta h \nu^{(i)}} - 1}. \quad (73)$$

This “is equivalent to Planck's formula” (end of Bose's article, p. 181). In- deed, the energy density per frequency interval  $d\nu$  becomes (eventually omit- ting the superscript  $(i)$ )

$$u(T, \nu) = \frac{E^{(i)}}{V d\nu^{(i)}} = \frac{8\pi h \nu^3}{c^3} \frac{1}{e^{h\nu/kT} - 1}. \quad (74)$$

This is Planck's radiation law (1).

**Bose's entropy more detailed** As an addendum, let us explore in more detail Bose's expression (72) for the entropy. According to formula (60), we have

$$S/k_B = \ln W = \sum_i N^{(i)} \ln N^{(i)} - \sum_i \sum_{s^{(i)}=0}^{P^{(i)}} n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)}. \quad (75)$$

<sup>20</sup>In eq. (62a), Bose has 0 instead of  $\delta \ln W$ . That is corrected in eq. (64) below.

Here, using (66) and (68), and omitting terms containing  $\exp\{-\beta h\nu^{(i)}P^{(i)}\}$  (see above),

$$\begin{aligned}
 -\sum_{s^{(i)}} n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)} &= -\sum_{s^{(i)}} B^{(i)} e^{-\beta h\nu^{(i)}s^{(i)}} \ln \left( B^{(i)} e^{-\beta h\nu^{(i)}s^{(i)}} \right) \\
 &= -B^{(i)} \ln B^{(i)} \sum_{s^{(i)}} e^{-\beta h\nu^{(i)}s^{(i)}} + B^{(i)} \beta h\nu^{(i)} \sum_{s^{(i)}} s^{(i)} e^{-\beta h\nu^{(i)}s^{(i)}} \\
 &= -\frac{B^{(i)} \ln B^{(i)}}{1 - e^{-\beta h\nu^{(i)}}} + B^{(i)} \beta h\nu^{(i)} \frac{e^{-\beta h\nu^{(i)}}}{(1 - e^{-\beta h\nu^{(i)}})^2} \\
 &= -N^{(i)} \ln N^{(i)} (1 - e^{-\beta h\nu^{(i)}}) + N^{(i)} \beta h\nu^{(i)} \frac{e^{-\beta h\nu^{(i)}}}{1 - e^{-\beta h\nu^{(i)}}}. \tag{76}
 \end{aligned}$$

In the last term, we replace  $N^{(i)}$  with the expression (57) and obtain

$$\begin{aligned}
 -\sum_{s^{(i)}} n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)} &= -N^{(i)} \ln N^{(i)} - N^{(i)} \ln(1 - e^{-\beta h\nu^{(i)}}) \\
 &\quad + 8\pi V \frac{(\nu^{(i)})^2}{c^3} d\nu^{(i)} \frac{\beta h\nu^{(i)}}{e^{\beta h\nu^{(i)}} - 1}. \tag{77}
 \end{aligned}$$

Inserting that into formula (75) and accounting for formula (73) yields Bose's entropy (72) with the correction indicated there.

Alternatively, setting, in formula (76),  $\exp\{\beta h\nu^{(i)}\} = 1 + N^{(i)}/P^{(i)}$  (see formula (69)), we obtain

$$\begin{aligned}
 -\sum_{s^{(i)}} n_{s^{(i)}}^{(i)} \ln n_{s^{(i)}}^{(i)} &= -N^{(i)} \ln N^{(i)} \\
 &\quad + (N^{(i)} + P^{(i)}) \ln \left( 1 + \frac{P^{(i)}}{N^{(i)}} \right) - P^{(i)} \ln \left( \frac{P^{(i)}}{N^{(i)}} \right). \tag{78}
 \end{aligned}$$

Inserting that into formula (75) yields

$$\begin{aligned}
 S &= \sum_i S^{(i)}, \\
 S^{(i)} &= k_B N^{(i)} \left[ \left( 1 + \frac{P^{(i)}}{N^{(i)}} \right) \ln \left( 1 + \frac{P^{(i)}}{N^{(i)}} \right) - \frac{P^{(i)}}{N^{(i)}} \ln \left( \frac{P^{(i)}}{N^{(i)}} \right) \right]. \tag{79}
 \end{aligned}$$

This agrees with Planck's 1901 formula (27) for the entropy.

As sketched at the end of the foregoing subsection, the same result can be obtained much simpler when treating all frequency intervals  $d\nu^{(i)}$  individually.

#### 4. On the Interpretation of Planck's Combinatorics

"Most textbooks take it for granted that  $\epsilon = h\nu$  marks the origin of energy quantisation, but this might be an overhasty interpretation of the mathematical procedure," Passon and Grebe-Ellis warn [72]. According to them, Kuhn "provides strong evidence" that Planck proposed a "physically structured phase space rather than discontinuous energy levels" ([55] p. 187). This view is claimed to be supported by Planck's remark:

"If the ratio [of the total energy  $E$  to the energy element  $\epsilon$ ] thus calculated is not an integer, we take for  $P$  an integer in the neighborhood." ([80] pp. 39f.)

This is understood as if  $\epsilon$  can also be viewed as describing an *interval* of the still continuous energy. On the contrary, the upholders of Planck having intended a *physical* quantization usually quote his remark, that the energy elements were the “most essential point of the whole calculation” [76]. As always, a single quote taken out of context cannot prove anything ([72] p. 6, fn. 3). As stressed at the end of Subsubsection 2.3.1, that ambiguity is removed when placing the resonators in the wall and taking the temperature as given.

That debate originates from the following ambiguity (after [72] p. 6). Planck’s combinatorics is claimed to can be interpreted in two distinct ways (cf. [16] pp. 243ff., [18]):

**1. Discontinuity reading:** Eq. (8) gives the number of ways how  $P$  energy elements  $\epsilon$  can be distributed over  $N$  resonators. This view suggests that the absorption and emission is discontinuous, cf. [13][47][50] [51].

**2. Continuity reading:** Eq. (8) describes the ways to distribute resonators over ‘energy cells’ (taking care that energy conservation is not violated<sup>21</sup>). According to this view, the resonators are placed in energy cells of finite size, where they can be put anywhere inside this cell (e.g., [15] p. 55, [23], [54] p. 118). Hence, a continuous emission and absorption is compatible with this view while only the specific size of Planck’s energy cells is mysterious.

We agree with Badino ([4] pp. 54f.) that Planck’s overt statement in the December 1900 talk [76] supports the discontinuity reading. This interpretation is seconded by the fact that immediately after that talk, Planck tried to return to continuous absorption and/or emission, cf. [33]. Already in his subsequent article [77] submitted January 9, 1901, i.e., only three weeks after his December 14, 1900 lecture,  $h$  was set back to merely a spectroscopic parameter at the end of the text. Still in 1913, in his recommendation for Einstein to be elected into the Prussian Academy of Science [79], Planck acknowledges Einstein’s 1907 quantum theory of specific heat, but rejects Einstein’s light quantum hypothesis.

“Daß er [Einstein] in seinen Spekulationen gelegentlich auch ein-mal uber das Ziel hinausgeschossen haben mag, wie z. B. in seiner Hypothese der Lichtquanten . . .” [79].

En: Although he [Einstein] may occasionally have “overshot the mark” in his speculations, as, for example, in his hypothesis of light quanta. . .

To corroborate Badino’s conclusion ([4] pp. 54f.) referred to above, let us add the following. For the distribution of  $P$  particles over  $N$  cells, the number of occupations equals Planck’s  $\binom{N-1+P}{P}$  (39). It can be rewritten as

$$\begin{aligned} \left( \binom{N}{P} \right) &:= \binom{N-1+P}{P} = \frac{(N-1+P)!}{P!(N-1)!} = \frac{(P+1)-1+(N-1)!}{(N-1)!P!} \\ &= \binom{(P+1)-1+(N-1)}{N-1} = \binom{P+1}{N-1}. \end{aligned} \quad (80)$$

Hence, Planck’s formula (39) also describes the distribution of  $N-1$  resonators over  $P+1$  energy elements. For the distribution of  $N$  resonators over  $P$  energy elements, however, the number of complexions equals  $\binom{P}{N} \neq \binom{N}{P} = \binom{P+1}{N-1}$ . Obviously, that rules out the ‘continuity reading’.

“Man merkt doch immer gleich, ob ein Historiker oder Philosoph der Physik selbst aktiv physikalisch geforscht hat.” (Hans-Ju“rgen Treder [91])

En: You can always tell right away whether a historian or philosopher of physics has actually conducted physics research them- selves.

Altogether, we second Passon and Grebe-Ellis [72] in that, for physics education, it is more important whether black-body radiation as such pro- vides a clear indication for discontinuity, i.e., whether today’s students should accept the claim that this phenomenon needs the introduction of a discontinuous energy (our answer is clearly, yes). This helps also to understand why Planck’s work was soon acknowledged as an experimentally confirmed radiation law while a debate on quantization did not follow immediately ([44] pp. 23f.). While we strongly support the study of the masters of the past— being common in arts and musics—the history of the physics of black- body radiation seems not to be the best place for learning physical thinking.

## 5. On the Issue of (in) Distinguishability

### 5.1. General Considerations

“For 80 years it has seemed natural that, to find what Gibbs had to say about this [the extensive property of entropy], one should turn to his Statistical Mechanics. For 60 years, text- books and teachers (including, regrettably, the present writer) have impressed upon students

<sup>21</sup>This condition is not necessary if the temperature rather than the energy is prescribed, see above.

<sup>22</sup>This view had been adopted by Boltzmann in his 1868 pioneering paper [6], in which he—at the age of 24—invented statistical mechanics. However, it has led him to contradictions [31].

<sup>23</sup>Einstein did not use  $h$  in his 1905 [24] and 1907 [26] articles, although he has explicitly applied Planck’s “energy elements”.

<sup>24</sup>The same applies to all other authors who see no essential difference between the dispersion of energy quanta over resonators or, vice versa, that of resonators over energy levels, e.g. Reiche in his Comment 4 to Planck 1901 [77] in Ostwalds Klassiker 206. Correct is that the numerical difference between the resulting entropies becomes the smaller, the larger the values of  $N$  and  $P$  are.

---

how remarkable it was that Gibbs, already in 1902, had been able to hit upon this paradox which foretold—and had its resolution only in—quantum theory with its lore about indistinguishable particles, Bose and Fermi statistics, etc. . . . In short, quantum theory did not resolve any paradox, because there was no paradox.” (Edwin Thomson Jaynes 1996 [45] p. 2)

In other words, there is no “indistinguishability principle for quantum particles” as ascribed to Natanson [66] and recently claimed, again, by Spa-lek ([89] p. 427). It still prevails in German-language pages of Wikipedia. Spalek [89] is right, however, in that interchangeable particles may still be identifiable and countable (p. 428). In the English literature, ‘equality’ and ‘identity’ often are not discriminated in the necessary manner. Saunders [85] even writes, “Indistinguishable particles have the same state-independent properties, but may differ wildly in state-dependent properties.” Obviously, that is a *contradictio in adjecto*.

As a matter of fact, Planck’s (27) and Bose’s (72) entropies are extensive. That is due to their assumption that the particles [10][76] and the cells [10] are interchangeable and the cells can be occupied by arbitrary many particles. Actually, that holds not only for bosons, but is possible for classical particles, too. For instance, their (stationary) states of motion can be described by their momenta (Newton) and velocities (Euler), respectively, while their positions are not accounted for. The set of momenta/velocities corresponds to the set of cells. Then, all equal particles with the same momentum resp. velocity are interchangeable, and their number is not limited. Essentially, that is what Boltzmann 1877 [8] used. Agreeing with Einstein [26], the crucial difference consists in that, for classical particles, the set of stationary states (momenta/velocities) is continuous, while that for quantum particles in bound states is discrete.

Using Newton’s or Euler’s notion of state rather than Lagrange’s, Laplace’s, or Hamilton’s one, Gibbs’ paradox in the mixing entropy of equal gases is avoided [29], classical statistical mechanics being self-consistent [27]. Another example is the interchangeability of the red balls in a snooker game mentioned at the beginning of Subsubsection 3.4.1.

We thus emphasize to teach the necessity of the ‘labor of notion’. For the notions are the tools of thinking. Inaccurate notions lead to inaccurate thinking. Perhaps, the best known historical examples are force and energy. Still Helmholtz’s 1847 pioneering talk establishing the conservation of energy is titled ‘On the Conservation of Force’ [38].

## 5.2. Natanson 1911

In 1911, Natanson [66] scrutinized the statistical assumptions underlying Planck’s law and introduced the concept of indistinguishability ([64] p. 151, fn. 211, [89] Subsection 2.2, p. 429). He discriminated between three cases:

1. Both the units of energy (particles, say, Planck’s energy elements) and the “receptacles of energy” (cells, say, Planck’s resonators), are not interchangeable;
2. Only the cells are not interchangeable;
3. Only the particles are not interchangeable.

Case 1 corresponds to Bach’s level 1 ‘configuration’ (Subsection 3.3), Case 2 to level 2 ‘occupation’ (Subsection 3.4), while Case 3 is to be replaced with level 3 ‘occupancy’ (Subsection 3.5)<sup>27</sup>. According to Spa-lek ([89] p. 429), he considered “indistinguishable wave packets” in “distinguishable” “receptacles” (wave modes, resonators), i.e., level 2 ‘occupation’. According to Mehra and Rechenberg ([65] p. 559), he used the occupancy numbers of level 3 (ditto in his 1913 book [67]). He correctly stated that, in either case, a different combinatorics needs to be applied and that Planck’s expression (8) assumes Case 2.

## 5.3. Einstein 1905 versus Planck 1900

Einstein 1905 [24] was the first to take Planck’s quantization<sup>28</sup> of light seriously. His reference to Planck<sup>29</sup> suggests that his “light quanta” differ from Planck’s “energy elements” solely in that they exist not only in resonators, but also in free space. However, this is not the case [22,53,98].

Let us demonstrate that the concepts of configuration (Subsection 3.3) and occupation (Subsection 3.4) allow for an immediate clarification of that issue.

From Wien’s radiation law (3), Einstein derives the formula

$$W_E = (V/V_0)^{E/h\nu} = 1/N^P \quad (81)$$

for the probability of a fluctuation of all ‘light particles’ into the sub-volume  $V$  of the total volume  $V_0$ . Here,  $V_0/V = N$  corresponds to the number of cells (assuming  $V$  to be sufficiently small) and  $E/h\nu = P$  to

<sup>25</sup>For pedagogical purposes let us add that we do not fully agree with Jaynes’ question, “Why did Gibbs fail to give this explanation in his *Statistical Mechanics*?” Admittedly, the description of how to deal with indistinguishable (more exactly, interchangeable) particles in its Chapter XV [35] is rather obscure, cf. [85].

<sup>26</sup>In 1914f., Wolfke [98] and Krutkow [53] published on similar ideas, see [65] p. 559.

---

the number of particles distributed over them.  $N^P$  is the number of all configurations of such a distribution, see formula (37). If all configurations are of equal probability (what is usually assumed, since the  $P$  particles are independently of each other distributed over the  $N$  cells, see Subsection 3.3), then WE (81) is the probability of each single configuration.

As just mentioned, within a configuration, the particles are distributed independent of each other. In contrast, within an occupation (Planck 1900 [76]), the particles are *not* distributed independent of each other, see Sub-section 3.4. For Ehrenfest and Kamerlingh Onnes [22,23], that is a crucial difference (see also [16,82]).

Unfortunately, Ehrenfest and Kamerlingh Onnes [23] do not properly distinguish the distribution of  $P$  particles over  $N$  cells and the distribution of  $N$  cells over  $P$  particles. The number of configurations equals  $N^P$  and  $P^N$ , respectively. The number of occupations equals  $\binom{N-1+P}{(N-1)P}$  and  $\binom{P-1+N}{(P-1)N}$ , respectively, as stressed in Section 4.

Anyway, as we have demonstrated in Subsubsection 3.4.2, the partition of space into cells leads to a non-extensive, i.e., wrong formula for the entropy. For a more detailed account of Einstein's relation to the theory of heat radiation, see [13].

## 6. Summary and Concluding Remarks

Historians of science usually agree about that 'discoveries' are rarely attributable to a particular moment in time and sometimes not even to single individuals [72]. They are rather extended processes which involve the interaction of several if not many researchers as in large accelerator experiments. Nevertheless, it is often possible to single out individuals who have finally pushed this development to a point from which there could be no retreat.<sup>30</sup> According to Kuhn ([54], p. 369), Planck is an example for that. After his derivation of the black-body radiation law, the recognition of discontinuity eventually was inevitable, although he himself was reluctant to draw this conclusion [79]. The debate about its interpretation is instructive as it is typical for a discovery of first rank. However, to make this point needs to introduce Planck's original approach, i.e., not his 1901 article [77], but his December 1900 talk [76] (see Subsection 2.3). For the concrete calculations, his 1901 article [77] (see Subsubsection 2.3.4) and Bose's treatment [10] are preferable, the latter one together with the corrections and simplifications proposed in Subsubsection 3.5.2.

Bose's [10] succeed because his partition of phase space concerned solely the momentum space, not the configuration space, see Subsubsection 3.5.2. In contrast, a partition of the configuration space—as done by Einstein in his 1905 famous elaboration of Planck's idea of elementary quantities of light [24]—leads to a non-extensive formula for the entropy. Related to that is Ehrenfest's [22] simple objection against Einstein's 1905 light quanta as "in-dependent", distinguishable particles. This shows that the early rejections of the light quantum were more rational than commonly presented; Planck and Einstein have not obtained the Nobel award before World War I. However, without wishing to diminish Ehrenfest's numerous extraordinary contributions to physics, renaming 'Bose-Einstein statistics' to "Ehrenfest-Natanson- Bose-Einstein statistics" [72] seems not yet to be justified (for Natanson, see Subsection 5.2).

Bach's [3] three-level scheme 'configuration—occupation—occupancy' (see Section 3) provides the simplest and at once most powerful tool we are aware of for systematizing and teaching probability topics in heat radiation. This suggests to explore its applicability in other areas of physics. In particular, we expect it to apply to statistics, in which the maximum occupation of a state is larger than one (Fermi-Dirac), but smaller than infinity (Bose-Einstein). Since Planck's radiation law is such a fundamental result, it can be and has been derived in various different ways. Many of today's textbooks follow a different route, e.g., using Einstein's A and B coefficients [37]. However, the latter one needs much more presuppositions than Planck's [76,77] and Bose's [10] treatments.

Irons [43] criticizes that many treatments lack a mechanism for the thermalization of the radiation, in particular, the setups with perfectly reflecting walls. And he sees only questionable attempts to solve that problem. However, black-body radiation is radiation in thermal equilibrium, so that its description is independent of the process of reaching thermal equilibrium. The latter one is connected with a maximization of the entropy. That is involved in Boltzmann's [8] and Bose's [10] calculations, see Subsection 3.5. Given that black-body radiation is a state of equilibrium between the radiation and the emitting and absorbing bodies, it opened up two different research lines at the same time. In a confusing and perplexing sequence of events, the notion of quantized matter and quantized radiation developed almost simultaneously. According to Passon & Grebe-Ellis [72], it would be a very unfavorable teaching strategy to follow those twists and turns. For this, it is the common choice to introduce non-relativistic quantum mechanics first and to deal with relativistic quantum theory and the quantum theory of radiation only later. However, many textbooks (and popular representations) on quantum theory cannot resist the temptation to introduce light quanta already in connection with Einstein's explanation of the photo-electric effect [24]. As demonstrated in Subsection 5.3 in an elementary manner, Einstein's light quanta should not be confused with Planck's ones, i.e., with the current photon concept. Moreover, contrary to Einstein's "fuzzy ball" concept [86], the photon of nowadays QED is neither distinguishable nor localizable. Cum grano salis, the same applies to the Compton effect [86].

<sup>27</sup>Generally speaking, if the cells are interchangeable, the particles have to be interchangeable, too.

<sup>28</sup>Students may be asked to discuss the difference between discretization and quantization. Discretization is caused by boundary or periodicity conditions and concerns classical waves and diffusion modes. There, the wavelength and, consequently, the frequency assume discrete values, while the energy of a wave remains to be continuous.

<sup>29</sup>Einstein—then a yet little known patent engineer—was clever enough to hide himself behind whose authority.

<sup>30</sup>On the other hand, the personality of a researcher should not be reduced to its distinguished scientific results.

Unfortunately, not seldom in the history of physics, concepts of notions were changed without changing the name of the notion. One of the most serious cases is the change of the notion of state from Newton and Euler to Lagrange and Laplace [93]. It had a large impact on the (mis)understanding of the (in)distinguishability of classical and quantum particles, see Sections 3 and 5. For this, against the usual teaching tradition, the notions of (in)distinguishability and the like should be introduced not only within advanced many-particle theory and quantum statistics. Bach's [3] three-level description adopted in Section 3 provides a straightforward tool for that (see also the Introduction). The crucial notion is not indistinguishability, but interchangeability.

Boltzmann's 1877 statistics [8] is quantum statistics as long as the possible energies of the molecules are held finite (for which Boltzmann had no reason, of course). As stressed in [28], that is confirmed in Einstein's pioneering article on the specific heat of solids [26]. For this, we strongly disagree with Spa-lek [89] in that "... only the explicit inclusion of the indistinguishability principle [by Natanson [66]] enlightens the difference between the original approach due to Boltzmann, defining classical statistics, and its quantum correspondent" (p. 430). On the contrary, again, the crucial notion is interchangeability, not indistinguishability [28]. Altogether, "Indistinguishability is not an intrinsic property of particles but a property of their state." (Bach 1997[3] p. 8)

Finally, suppose the students to be familiar with the density operator  $\hat{\rho}$  for a system with Hamiltonian  $\hat{H}$ , which is in thermal equilibrium at temperature  $T$ ,

$$\hat{\rho} = \frac{\exp(-\hat{H}/kT)}{\text{Tr}\{\exp(-\hat{H}/kT)\}}. \quad (82)$$

For a quantized electromagnetic field of wave vector  $\vec{k}$  and polarization  $s$  in a cavity, the Hamiltonian is

$$\hat{H} = \sum_{\vec{k}, s} \left( \hat{n}_{\vec{k}, s} + \frac{1}{2} \right), \quad (83)$$

and the mean photon number of mode  $(\vec{k}, s)$  is Boltzmann distributed as [61]

$$\langle \hat{n}_{\vec{k}, s} \rangle = e^{-h\nu/kT}. \quad (84)$$

### Acknowledgments

This article has partially been written during his second stay of one of us (PE) at the Pavlodar State Pedagogical Margulan University, Kazakhstan. The great hospitality of the ruling bodies, friends, and colleagues over there, in particular, Yerzhan Amirbekuly, Zhanar Baiseitova, Nurbala Ubaidulayeva, and Assemgul Kissabekova, is truly acknowledged. Moreover, he gratefully acknowledges the email exchange with Rolf Muczinski as well as the discussion with Hassan Bolouri, Jan Helm, Axel Kilian, Michael Brieger, Herbert Capellmann, and Rudolf Germer during a presentation of the main content to them. He feels indebted to Harry Paul for pointing his attention to Ref. [73] and more.

Last but not least, this work would not have been possible without the numerous people in the internet generously sharing their knowledge, e.g., the German L<sup>A</sup>T<sub>E</sub>X society Dante e.V., and making original texts accessible for free.

The authors have no conflicts to disclose.

### References

1. Bach, A. (1988). The concept of indistinguishable particles in classical and quantum physics. *Foundations of Physics*, 18(6), 639-649.
2. Bach, A. (1990). Boltzmann's probability distribution of 1877. *Archive for History of Exact Sciences*, 41(1), 1-40.
3. Bach, A. (1997). *Indistinguishable Classical Particles*; Springer: Berlin/Heidelberg, Germany, (Lect. Notes Phys. m44).
4. Badino, M. (2015). *The Bumpy Road: Max Planck from Radiation Theory to the Quantum (1896-1906)*, Springer.
5. Becker, R. (1966). *Theory of Warmth* (Berlin Heidelberg New York 1966).
6. Boltzmann, L. (1866). *Über die mechanische Bedeutung des zweiten Hauptsatzes der Wärmetheorie*, Wiener Ber. LIII, 195-220.
7. Boltzmann, L. (1868). *Studien Über das Gleichgewicht der lebendigen Kraft zwischen bewegten materiellen Punkten*, Wiener Ber. 58, 517-560, Vol 9(5). I, pp. 49-96.

8. Boltzmann, L. (1877). Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht Wien. Ber. II LXXVI, 373–435, Vol (9)II, pp. 164–223.
9. Boltzmann, L. (1909). Wissenschaftliche Abhandlungen (F. Hasenöhr, ed.), I. Bd. (1865-1874), II. Bd. (1875—1881), reprint: 2012 Cambridge Univ. Press, III. Bd. (1882-1905), (Leipzig: Barth).
10. Bose. (1924). Plancks gesetz und lichtquantenhypothese. Zeitschrift für Physik, 26(1), 178-181.
11. Brinken, A. (2007). Einführung in die Kombinatorik, revised by Plotke, H 2011.
12. Brush, S. G., Segal, A. (2015). Making 20th Century science: how theories became knowledge. Oxford University Press.
13. Büttner, J., Renn, J., Schemmel, M. (2000). Exploring the Limits of Classical Physics—Planck, Einstein, and the Structure of a Scientific Revolution, in: [14] pp. 22–46.
14. Büttner, J., Darrigol, O., Hoffmann, D., Renn, J., Schemmel, M. (2000). Revisiting the quantum discontinuity. (Max Planck Institute for the History of Science, Preprint 150).
15. Darrigol, O. (1988). Statistics and combinatorics in early quantum theory. Historical studies in the physical and biological sciences, 19(1), 17-80.
16. Darrigol, O. (1991). Statistics and combinatorics in early quantum theory, II: Early symptoms of indistinguishability and holism. Historical Studies in the Physical and Biological Sciences, 21(2), 237-298.
17. Darrigol, O. (1992). From c-numbers to q-numbers: The classical analogy in the history of quantum theory (Vol. 8). Univ of California Press.
18. Darrigol, O. (2000). The Historians' Disagreements over the Meaning of Planck's Quantum, in: [14] pp. 3–21.
19. Darrigol, O. (2001). The historians' disagreements over the meaning of Planck's quantum. Centaurus, 43(3-4), 219-239.
20. Dorling, J. (1971). Einstein's Introduction of Photons: Argument by Analogy or Deduction from the Phenomena?. The British Journal for the Philosophy of Science, 22(1), 1-8.
21. Dougal, R. C. (1976). The presentation of the Planck radiation formula (tutorial). Physics Education, 11(6), 438-443.
22. Ehrenfest, P. (1911). Welche Züge der Lichtquantenhypothese spielen in der Theorie der Wärmestrahlung eine wesentliche Rolle?. Annalen der Physik, 341(11), 91-118.
23. Onnes, H. K., Ehrenfest, P. (1914). Simplified deduction of the formula from the theory of combinations which Planck uses as the basis of his radiation-theory. In KNAW, Proceedings (Vol. 17, pp. 870-873). Phil. Mag., Ser. 6, 29 (170), 297–301.
24. Einstein, A. (1905). Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt Ann. Phys. 17, 132–148.
25. Einstein, A. (1906). Zur theorie der lichterzeugung und lichtabsorption. Annalen der Physik, 325(6), 199-206.
26. Einstein, A. (1907). Die plancksche theorie der strahlung und die theorie der spezifischen wärme. Annalen der Physik, 327(1), 180-190.
27. Enders, P. Is Classical Statistical Mechanics Self-Consistent?(A paper in honor of CF von Weizsäcker, 1912-2007). (Progr. Phys.) 3, 85-87.
28. Enders, P. (2006). Equality and Identity and (In) distinguishability in Classical and Quantum Mechanics from the Point of View of Newton's Notion of State. ICFAI Univ. J. Phys. 2008, I, 71–78.
29. P. Enders, Gibbs' Paradox in the Light of Newton's Notion of State, Entropy 11 (2009) 454-456.
30. Enders, P. (2016). Historical perspective: Boltzmann's versus Planck's state counting—Why Boltzmann did not arrive at Planck's distribution law. Journal of Thermodynamics, 2016(1), 9137926.
31. Enders, P. (2021). Statistical Mechanics and Thermodynamics: Boltzmann's versus Planck's State Definitions and Counting. Entropy, 23(7), 875–909.
32. Feller, W. (1971). An introduction to probability theory and its applications. New York: Wiley. Vol. 1.
33. Galison, P. (1981). Kuhn and the quantum controversy. British Journal for the Philosophy of Science. 32(1), 71-85.
34. Gearhart, C. A. (2002). Planck, the Quantum, and the Historians. Physics in perspective, 4(2), 170-215.
35. Gibbs, J. W. (1902). Elementary principles in statistical mechanics: developed with especial reference to the rational foundations of thermodynamics. C. Scribner's sons.
36. Glazer, M., Wark, J., Schmittmann, B. (2002). Statistical mechanics: a survival guide.
37. Haken, H., Wolf, H. C. (2004). Atom-und Quantenphysik: eine Einführung in die experimentellen und theoretischen Grundlagen (Vol. 8). Berlin: Springer.
38. Helmholtz, H. (1982). Über die Erhaltung der Kraft (Vol. 1). Walter de Gruyter GmbH & Co KG.
39. Helmholtz, H. v. (1903). Introduction to the Lectures on Theoretical Physics, Leipzig: Barth, 10.
40. Hettner, G. (1922). Die Bedeutung von Rubens Arbeiten für die Plancksche Strahlungsformel. Naturwissenschaften, 10(48), 1033-1038.
41. Hoffmann, D. (2001). On the Experimental Context of Planck's Foundation of Quantum Theory I. Centaurus, 43(3-4), 240-259.
42. Irons, F. E. (2004). Reappraising Einstein's 1905 application of thermodynamics and statistics to radiation. European journal of physics, 25(2), 269-277.

- 
43. Irons, F. E. (2005). Why the cavity-mode method for deriving Planck's law is flawed. *Canadian journal of physics*, 83(6), 617-628.
  44. Jammer, M., Merzbacher, E. (1967). The conceptual development of quantum mechanics.
  45. Jaynes, E. T. (1996). The Gibbs Paradox, pp. 1–22; gibbs 7/8, 1–19.
  46. Jeans, J. H. (1905). XI. On the partition of energy between matter and Æther. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 10(55), 91-98.
  47. Jost, R. (1995). Planck-Kritik des T. Kuhn. *Das Märchen vom Elfenbeinernen Turm: Reden und Aufsätze*, 67-78.
  48. Kango, H. (1972). Anmerkung Nr. 13 in [80], p. 19; En: Note no.13 in [80], p. 50.
  49. Kirchhoff, G. (1860). Ueber das Verhältniss zwischen dem Emissionsvermögen und dem Absorptionsvermögen der Körper für Wärme und Licht. *Ann. Phys.*, 185, 275–301.
  50. Klein, M. J. (1961). Max Planck and the beginnings of the quantum theory. *Archive for History of Exact Sciences*, 1(5), 459-479.
  51. Klein, M. J., Shimony, A., Pinch, T. J. (1979). Paradigm lost? A review symposium. (*Isis*) 70, 429–440.
  52. Kries, J.V. (1927). *The Principles of Probability Theory: A Logical Investigation*. p. 36.
  53. Krutkow, G. (1914). Aus der Annahme unabhängiger Licht-quanten folgt die Wiensche Strahlungsformel. *Physikalische Zeitschrift*, 15, 133-136; Krutkow, G. (1914). Bemerkung zu Herrn Wolfkes Note: Welche Strahlungsformel folgt aus der Annahme der Lichtatome. *Physikalische Zeitschrift*, 15, 363-364.
  54. Kuhn, T. S. (1987). *Black-body theory and the quantum discontinuity, 1894-1912*. University of Chicago Press.
  55. Kuhn, T. S. (1980). Einstein's critique of Planck. *Some Strangeness in the Proportion*, 186.
  56. Kuryshova, Y. V. *Principles of combinatorics*.
  57. Larmor, J. (1910). On the statistical theory of radiation *London, Edinburgh, and Dublin Phil. Mag. and J. of Sci. Ser. 6*, 20:116, 350-353.
  58. List B Der Konkurrenzkampf motiviert zu Höchstleistungen [Competition motivates people to perform at their best], talk with Paulus Muller, radio Deutschlandfunk, 2026, 13:30 h.
  59. Longair, M. (2013). *Quantum concepts in physics: An alternative approach to the understanding of quantum mechanics*. Cambridge University Press.
  60. Lorentz, H. A. (1910). *Alte und neue Fragen der Physik*. 11, 1234.
  61. Mandel, L., Wolf, E., Shapiro, J. H. (1996). Optical coherence and quantum optics.
  62. Marr, J. M., Wilkin, F. P. (2012). A better presentation of Planck's radiation law. *American Journal of Physics*, 80(5), 399-405.
  63. Masters, B. R. (2013). Satyendra Nath Bose and Bose-Einstein Statistics. *Optics and Photonics News*, 24(4), 40-47.
  64. Mehra, J., Rechenberg, H. (2000). *The historical development of quantum theory (Vol. 1)*. Springer Science & Business Media. Pt(1).
  65. Rechenberg, H. (1982). *The historical development of quantum theory (Vol. 1)*. Springer Science & Business Media. Pt(2).
  66. Natanson, L. (1911). On the statistical theory of radiation *Bull. Acad. Sci. Cracovie A* 134–148; *Über die statistische Theorie der Strahlung Phys. Z.* 12, 659–666.
  67. Natanson, W. (1913). *Zasady Teorii promieniowania. Prace Matematyczno-Fizyczne*, 24(1), 1-88.
  68. Nauenberg, M. (2016). Max Planck and the birth of the quantum hypothesis. *American Journal of Physics*, 84(9), 709-720.
  69. Needell, A. A. (1980). *Irreversibility and the failure of classical dynamics: Max Planck's work on the quantum theory 1900-1915*. Yale University.
  70. Nolte, D. D. (2024). 100 Years of Quantum Physics: The Statistics of Satyendra Nath Bose (1924), blog Galileo Unbound.
  71. Norton, J. D. (2006). Atoms, entropy, quanta: Einstein's miraculous argument of 1905. *Studies in History and Philosophy of Science Part B: Studies in History and Philosophy of Modern Physics*, 37(1), 71-100.
  72. Passon, O., Grebe-Ellis, J. (2017). Planck's radiation law, the light quantum, and the prehistory of indistinguishability in the teaching of quantum mechanics. *European Journal of Physics*, 38(3), 035404.
  73. Paul, H., Greenberger, D. M., Stenholm, S. T., Schleich, W. P. (2015). The Stefan–Boltzmann law: two classical laws give a quantum one *Phys. Scr.* 165, 014027.
  74. Planck, M. (1900). Entropie und temperatur strahlender wärme. *Annalen der Physik*, 306(4), 719-737.
  75. Planck, M. (1900). Über eine Verbesserung der Wien'schen Strahlungsgleichung. *Verhandlungen der Deutschen Physikalischen Gesellschaft*. 202–4.
  76. Planck, M. K. E. L. (1900). Distribution of energy in the normal spectrum. *Verhandlungen der Deutschen Physikalischen Gesellschaft*, 2, 237-245.
  77. Planck, M. (1901). On the law of distribution of energy in the normal spectrum. *Annalen der physik*, 4(553), 1.
  78. Planck, M. (1914). *The theory of heat radiation*. Blakiston.
  79. Planck, M., Nernst, W., Rubens, H., Warburg, E. (1913). Wahlvorschlag von Max Planck für Albert Einstein zum ordentlichen Mitglied der physikalisch-mathematischen Klasse der Preußischen Akademie der Wissenschaften. *Archiv der Berlin-Brandenburgischen Akademie der Wissenschaften*. II–III–36, Bl. 36–37.
  80. Planck, M., Kangro, H., Haar, D. T., Brush, S. G. (1972). *Planck's original papers in quantum physics: German and English edition*.
  81. Planck, M. (1997). *Die Ableitung der Strahlungsgesetze (1895–1900)*.

- 
82. Provost, J. P., Bracco, C. (2008). Einstein's quanta and the 'true' volume dependence of the black-body entropy. *European journal of physics*, 29(5), 1085-1090.
  83. Rayleigh, L. (1900). LIII. Remarks upon the law of complete radiation. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, 49(301), 539-540.
  84. Cohen, R. S., Stachel, J. J. (1979). The First Phase in the Evolution of the Quantum Theory [1936b]. In *Selected Papers of Léon Rosenfeld* (pp. 193-234). Dordrecht: Springer Netherlands.
  85. Saunders, S. (2020). The Concept 'Indistinguishable', *Stud. Hist. Phil. Mod. Phys.* 71 Aug. 37–59.
  86. Scully, M. O., Sargent, M. (1972). The concept of the photon. *Physics Today*, 25(3), 38-47.
  87. Sharp, K., Matschinsky, F. (2015). Translation of Ludwig Boltzmann's Paper "On the Relationship between the Second Fundamental Theorem of the Mechanical Theory of Heat and Probability Calculations Regarding the Conditions for Thermal Equilibrium. *Entropy* 2015, 17, 4, 1971–2009.
  88. Sommerfeld, A. (1922). *Atombau und Spektrallinien*, Braunschweig.
  89. Spałek, J. (2020). The Bose-Einstein statistics: Remarks on Debye, Natanson, and Ehrenfest contributions and the emergence of indistinguishability principle for quantum particles. *Studia Hist. Sci.* 19, 423–441.
  90. Studienkreis, *Permutation—Variation—Combination*.
  91. Treder, H. J. (1977). *Vorlesungen zur Geschichte der Physik*, Humboldt University at Berlin, unpubl.
  92. Weinberg, S. (2015). *Lectures on quantum mechanics*. Cambridge University Press.
  93. C. F. V. (2002). *Weizsacker, Aufbau der Physik*, Munchen: dtv 4, p. 224.
  94. Wien, W. (1893). Eine neue Beziehung der Strahlung schwarzer Körper zum zweiten Hauptsatz der Wärmetheorie. In *Von Kirchhoff bis Planck: Theorie der Wärmestrahlung in historisch-kritischer Darstellung* (pp. 156-165). Wiesbaden: Vieweg+ Teubner Verlag.
  95. Wien, W. (1896). Ueber die Energievertheilung im Emissionsspectrum eines schwarzen Körpers. *Annalen der Physik*, 294(8), 662-669.
  96. Wikipedia, Stars and bars (combinatorics).
  97. Wikipedia, Twelfefold way.
  98. Wolfke, M. (1914). Welche Strahlungsformel folgt aus der Annahme der Lichtatome?. *Physikalische Zeitschrift*, 15, 308-310; Antwort auf die Bemerkung Herrn Krutkows zu meiner Note: Welche Strahlungsformel folgt aus der Annahme der Lichtatome?. *Physikalische Zeitschrift*, 15, 463-464.

*Copyright:* ©2026 Peter M. Enders, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.