PLANCK AND THE BIRTH OF THE QUANTUM HYPOTHESIS

MICHAEL NAUENBERG,
DEPT. OF PHYSICS,
UNIVERSITY OF CALIFORNIA, SANTA CRUZ,

Abstract. Based on the functional dependence of entropy on energy, and on Wien’s distribution for black-body radiation, Planck obtained a formula for this radiation by an interpolation relation that fitted the experimental measurements of thermal radiation at the Physikalisch Technische Reichanstalt (PTR) in Berlin in the late 19th century. Surprisingly, his purely phenomenological result turned out to be not just an approximation, as would have been expected, but an exact relation. To obtain a physical interpretation for his formula, Planck then turned to Boltzmann’s 1877 paper on the statistical interpretation of entropy, which led him to introduce the fundamental concept of energy discreteness into physics. A novel aspect of our account, that has been missed in previous historical studies of Planck’s discovery, is to show that Planck could have found his phenomenological formula partially derived in Boltzmann paper in terms of a variational parameter, but its dependence on temperature was first carried out by Planck.

I. Introduction

One of the most interesting episodes in the history of science was Planck’s introduction of the quantum hypothesis, at the beginning of the 20th century. The emergence of this revolutionary concept in physics is a fascinating story that has been described previously [1, 8], but important aspects of this discovery are generally not found in the description of Planck’s ideas in physics textbooks that discuss quantum mechanics. In particular, most physics textbooks do not give even a clue how the concept of discreteness in energy - the revolutionary concept introduced by Planck to describe the spectrum of black-body radiation-originated in the first place [6]. From Planck’s articles and correspondence on his theory of the spectrum of black-body radiation, it is clear that he took this concept directly from Boltzmann, who in his seminal 1877 paper on statistical mechanics discretized energy as a purely mathematical device in order to be able to count the possible configurations of a molecular gas in thermal equilibrium [9]. But this important connection between Planck’s and Boltzmann’s work has been ignored even in physics textbooks which emphasize a historical approach [10], and in other books by well known physicists. For example, in the description of Planck’s discovery in his biography of Einstein, Abraham Pais concludes that

“His [Planck’s] reasoning was mad, but his madness has that divine quality that only the greatest transitional figures can bring to science” [11].
This comment does not provide any more enlightenment on the origin of the idea of quantization in physics than Richard Feynman’s succinct statement in his well known Lectures on Physics, that

“...by fiddling around [Planck] found a simple derivation [for his formula]” [12].

Most accounts of Planck’s discovery in physics textbooks are historically inaccurate, and Klein’s analysis of Planck’s work [2] debunked some myths contained in these books. For example, according to one of the most common ones in physics textbooks, Planck was responding to the problem in the classical theory of black-body radiation known as the ultraviolet catastrophe. It occurs when the equipartition theorem for a system in thermal equilibrium is applied to the spectral distribution of thermal radiation. But at the time, Planck appears to have been unaware of this problem which was named by Ehrenfest several years after Planck’s discovery. Indeed, the application of the equipartition theorem to black-body radiations, was made by Lord Rayleigh[16] at about the same time that Planck obtained his famous formula for the black-body spectrum, and there isn’t any evidence that Planck was aware of Rayleigh’s result which agreed with new experiments for the long wavelength end of the spectrum observed at that time. Klein concluded, that

“it was probably a very good thing that Planck was not constrained in his thinking by the tight classical web which Rayleigh had woven”

But Planck had introduced a non-classical concept into physics, namely, discrete energies, and therefore he would not have been constrained by Rayleigh purely classical derivation.

Historians of science also have failed to notice that in Boltzmann’s 1877 paper the mean energy of his fictitious molecular ensemble with discrete energies is obtained only in terms of an undetermined variational parameter ( see reference [9], p.181, and the Appendix). But Boltzmann calculated the temperature dependence of this parameter only in the limit relevant to classical mechanics. Moreover, for the past twenty five years, the nature of Plancks’s discovery has become the subject of controversies among historians of science. Recently, some of them have even claimed that contrary to what physicists have always been led to believe, Planck did not introduce the concept of energy discreteness or quantization into physics [4] [5]. This surprising claim originated with the appearance in 1978 of a book by Thomas Kuhn on the history of Planck’s work on black-body radiation, and in his later article on this subject [5], where he disputed the conventional view, held by most physicists and earlier historians of science. In his book, Kuhn concluded that,

“With a single misleading exception, nothing in Planck’s published papers, known manuscripts, or autobiographical fragments suggest that the idea of restricting resonator energies to a discrete set of values had even occur to him as a possibility until others forced it upon him during 1906 and the years following” ( reference [5], p.126 ).

Although some historians of science have criticized Kuhn’s interpretation of Planck’s discovery, others have reached similar conclusions. While Kuhn agreed with many details in the traditional account of Planck’s work, he argued that crucial aspects of it had been
Kuhn’s book was reviewed by Klein [13], Galison [14] and Needell [15], who were all critical of Kuhn’s thesis, but conclusions similar to Kuhn’s were reached by Darrigol, who even quoted Planck for “irrefutable evidence” that he had not meant to introduce energy discontinuities into physics [4]. In 1984, Kuhn responded to his critics in an article defending his interpretation of Planck’s discovery of the quantum hypotheses by stating that

“Part of the appeal of the standard account of Planck’s discovery is, I think, the closeness with which it matches a still cherished view of the nature of science and its developments. Although I appreciate both the charms and the functions of that view, understanding requires that it be recognized as myth” [5].

Unfortunately, many of Planck’s documents and correspondence that could have illuminated the development and early reception of his ideas were destroyed during the bombing of Berlin in 1944. Recently, Darrigol summarized this controversy in an article entitled The Historians’ Disagreement over the Meaning of Planck’s Quantum [4].

In essence, Planck’s approach to the theory of black-body radiation was based on the following steps. Taking advantage of Kirchhoff’s theorem that the black-body distribution is a universal function independent of the nature of the source of radiation, Planck’s first step was to obtain a relation for the energy distribution of this radiation in thermal equilibrium with an ensemble of microscopic Hertzian oscillators with variable frequency $\nu$. By applying Wien’s distribution that fitted the high frequency end of this radiation, Planck obtained an expression for the mean energy of these oscillators. The measurements were made by careful experiments at the Physikalisch Technishe Reichanstali (PTR) which was the center for infrared radiation studies in Berlin at the end of the 19-th century (see Fig. 1). After it was discovered that Wien’s distribution did not fit new data at lower frequencies, Planck obtained a new distribution formula by an interpolation relation based on his application of the relation between entropy and energy for a system in thermal equilibrium. Finally, to obtain a theoretical interpretation for his new formula, Planck turned to the seminal 1877 paper of Boltzmann which formulate the relation between entropy and statistics[9].

In his paper, Boltzmann introduced a relation between the entropy of a molecular gas and the number of microscopic configurations, or complexions as he called them, of the molecules. He defined the state of thermal equilibrium to be the maximum number of these configurations, subject to the constraint of a fixed number of molecules and total energy. At first sight it is surprising that Boltzmann’s ideas that were based on purely classical concepts, and therefore apply to systems having continuous energy could have served as the spring board for Planck’s quantum hypothesis of discrete energy levels. But to implement his statistical ideas, Boltzmann took for his initial example a fictitious model of a gas of molecules having discrete energies in integer multiples of an energy element of magnitude $\epsilon$. For Boltzmann, this discretization of energy was purely a mathematical
Figure 1. Apparatus of Lummer and Kurlbaum to measure the spectrum of black-body radiation. A current heats the filament $E$, located in a tube inside the cylinder $C$, to a fixed temperature $T$, giving rise to black-body radiation inside this cylinder. The spectrum of this radiation is observed by some radiation exiting through the hole at one end along the axis of the cylinder.

artifact that he introduced for the purpose of counting the number of configurations of the molecules. Subsequently, as would be expected, he took the limit of continuous molecular
energy for which $\epsilon$ vanished. But when Planck applied Boltzmann’s discrete model to his ensemble of Hertzian oscillators in thermal equilibrium with radiation, he did not take this continuum limit. Instead, he set Boltzmann’s energy elements to a fixed value $\epsilon = h\nu$, where $\nu$ is the frequency of his oscillators, and $h$ is a new universal constant, now known as Planck’s constant, that relates frequency and energy. It was very fortunate for Planck that Boltzmann initially considered energy as the only degree of freedom of the molecules in his ensemble, because that made possible Planck’s direct extension to and ensemble of linear harmonic oscillators[17].

Planck was aware that with his procedure he was violating the tenets of continuum physics. In his Dec 19, 1900 paper, presented at a meeting of the German Physical Society, he wrote that:

“...all my attempts ...to adapt the theoretical foundations of physics to this knowledge failed completely. It was as if the ground had been pulled from under one, with no firm foundation to be seen anywhere” [21]

In view of this single remark, it is evident that Kuhn’s statement quoted earlier, that the “idea of restricting resonator energies to a discrete set of values had not even occurred to him [Planck]”, is clearly incorrect. But as “irrefutable evidence” that Planck did not have energy discreetness in mind, Darrigol has argued that Planck remarked afterwards that

“...all my attempts ...to adapt the theoretical foundations of physics to this knowledge failed completely. It was as if the ground had been pulled from under one, with no firm foundation to be seen anywhere” [21]
In 1905, however, he resolved the conundrum by assuming that electromagnetic radiation also consisted of discrete energy quanta [20]. But Planck did not show such concern, and instead he attempted for many years to incorporate his new results within the realm of continuum classical physics. Taking again an idea from Boltzmann’s 1877 paper, later Planck considered the energy of the oscillator to be continuous, and \( \epsilon = h\nu \) to be the magnitude of cells of equal probability in the phase space of the oscillators. Otherwise, the derivation of his formula proceeds in precisely the same form as before. In 1906, and again as late as 1909, he presented his derivation in this form in lectures that he gave during his visit at Columbia Univ. But had Planck followed closely Boltzmann’s statistical method, he would have realized earlier that a continuum energy interpretation of his formulae was not feasible. In a card to Ehrenfest in the spring of 1915, Planck wrote: “I hate discontinuity of energy even more than discontinuity of emission”.

Evidently, the disagreement among historians of science and physicists is based on an inadequate understanding of the relation between Boltzmann’s seminal 1877 article establishing the relation between entropy and statistical mechanics, and Planck’s application of this work to justify his black-body formula, which he had obtained previously by an empirical fit to experimental data based on classical thermodynamics and electrodynamics.

The main purpose of this paper is to clarify the relation between Boltzmann’s and Planck’s work by providing a simple but thorough mathematical discussion of it that is absent in the literature on this subject. In the following sections, I discuss Planck’s work as described in some of his publications, his autobiographical recollections [27], his Nobel speech [31], and in some of his correspondence. Section II reviews Planck’s original serendipitous derivation of his well known formula for black-body radiation, which he referred to as his “lucky intuition”. Section III describes his application of Boltzmann’s principles of statistical mechanics, and Section IV describes some of Planck’s recollections on how he discovered his fundamental radiation formula. The relation between Boltzmann’s work and Planck’s application of it is given in an Appendix.

II. Planck’s phenomenological derivation of his black-body formula

An insightful description on how Planck obtained his formula for the spectrum of black-body radiation can be found in his scientific autobiography [27]. While this account was written many years after the occurrence of this event, and may suffer from the usual lapses of memory and the absence of original documents and correspondence\(^1\), it appears to be consistent with Planck’s original publications. Therefore, in this section I will let Planck speak for himself, while, for clarification, I will fill in some of the mathematical details (keeping his original notation) in a form close to his original articles.

Planck wrote:

“While a host of outstanding physicists worked on the problem of spectral distribution, both from the experimental and the theoretical aspects everyone of them directed his efforts solely toward exhibiting the dependence of

\(^1\)Planck’s own papers and correspondence were destroyed when Berlin was bombed in WWII.
the intensity of the radiation on the temperature. On the other hand, I sus-
ppected that the fundamental connection lies in the dependence of entropy
upon energy [my italics]. As the significance of entropy had not yet come
to be fully appreciated, nobody paid any attention to the method adopted
by me, and I could work out my calculations completely at my leisure, with
absolute thoroughness, without fear of interference or competition. Since
for the irreversibility of the exchange of energy between an oscillator and
the radiation activating it, the second differential quotient of its entropy
with respect to its energy is of characteristic significance, I calculated the
value of this function on the assumption that Wien’s law of the Spectral
Energy Distribution is valid - a law which was then in the focus of general
interest; I got the remarkable result that on this assumption the reciprocal
of that value, which I shall call here $R$, is proportional to the energy.”[27]

On theoretical grounds, Wilhelm Wien had proposed [28] that the spectral energy dis-
tribution for black-body radiation with frequency $\nu$ at temperature $T$, had the scaling
form

$$\rho(\nu, T) = \nu^3 f(\nu/T),$$

(1)

where $f$ is a function of a single variable, the ratio of frequency $\nu$ and temperature $T$.
This form satisfies the Stefan-Boltzman relation that the total black-body energy is pro-
portional to the fourth power of the temperature $T$. Originally, this dependence was found
experimentally by Josef Stefan, and later a theoretical derivation was provided in 1884 by
his former student Boltzman [29]. Boltzmann’s method was succinct: applying Maxwell’s
relation between the energy per unit volume $E$, and the pressure $p$ of isotropic radiation,
$p = E/3$, leads to a relation for the entropy $S_R$ per unit volume of this radiation,

$$S_R = \frac{4}{3} \frac{E}{T}.$$  

(2)

Substituting for the temperature $T$ in this relation the thermodynamic condition

$$\frac{1}{T} = \frac{dS_R}{dE},$$

(3)

and integrating the resulting differential equation yields $S_R = c'E^{3/4}$, and eliminating $S_R$
by applying again Eq.2, one obtains the relation $E = \sigma T^4$, known as the Stefan-Boltzmann
law, where $\sigma = (3c'/4)^4$ is a universal constant.

According to Wien’s spectral distribution, Eq.1, integrating the spectrum over all fre-
quencies, and setting $z = \nu/T$ as the variable of integration, one recovers the Stefan-
Boltzmann relation,

$$E = \int_0^\infty \rho(\nu, T)d\nu = \sigma T^4.$$  

(4)

where $\sigma$ is now determined by

$$\sigma = \int_0^\infty z^3 f(z)dz.$$  

(5)
Probably stimulated by earlier phenomenological work by Paschen, Wien assumed for the function \( f(\nu/T) \) the exponential form

\[
(6) \quad f(z) = \frac{8\pi a}{c^3} e^{-bz},
\]

where \( a \) and \( b \) are constants that could be obtained by fitting his theoretical distribution, Eq.1, to the black-body radiation experiments. The constant \( a \) has the dimensions of energy times time and later it will be seen to correspond to Planck’s constant \( h \). For this form of \( f \), according to Eq.5, \( \sigma = 48\pi a/b^4 c^3 \).

Subsequently, in a series of five papers written between 1897 and 1899, Planck discussed the thermal equilibrium between the radiation in a cavity and an ensemble of Hertzian electromagnetic oscillators [32], based on Maxwell theory of electromagnetism. His main result was a relation between the spectral distribution \( \rho(\nu,T) \) and the mean energy \( U(\nu,T) \) of the oscillators,

\[
(7) \quad \rho(\nu,T) = \frac{8\pi}{c^3} \nu^2 U(\nu,T).
\]

Combining this result with Wien’s relation, Eqs.1 and 6, implies that

\[
(8) \quad U(\nu,T) = a\nu e^{-b\nu/T}
\]

In this expression Planck’s constant \( h = a \), which like Wien’s constant \( a \) has the dimension of energy times time. By fitting the data on blackbody radiation obtain in the experiments by Otto Lummer and Ernst Pringsheim at the with radiation emmitted from a small hole in a heated cavity (see Fig.1), Planck obtained \( h = 6.88510^{-27} \text{ erg} \times \text{sec} \), in remarkable correspondence to the modern value \( h = 6.62610^{-27} \text{ erg} \times \text{sec} \), a tribute to the accuracy of the black body radiation experiments at that time. Neither Wien, Planck nor anyone else seemed to notice, however, until it was pointed out by Lord Rayleigh several years later [16], that the Wien exponential law, Eq.6, implied the implausible result that as the temperature \( T \) increases the magnitude of the spectral distribution at a fixed frequency \( \nu \) approaches a constant value \( \rho(\nu,T) = \frac{8\pi h\nu^3}{c^3} \), and \( U = h\nu \), independent of \( T \).

Planck’s next step was to considered the dependence of the entropy \( S(\nu,U) \) of his oscillators on the energy \( U \). Given the relation between the energy \( U \) and the temperature \( T \), Eq. 8, he obtained this dependence from the thermodynamic relation,

\[
(9) \quad \frac{1}{T} = \frac{dS}{dU}
\]

Inverting Eq.8 to obtain \( T \) as a function of \( U \), and substituting the result in Eq.9 gives a first order differential equation for \( S \),

\[
(10) \quad \frac{dS}{dU} = -\frac{1}{b\nu} \ln \frac{U}{a\nu},
\]

and integrating this equation with the boundary condition that \( S \) vanishes when \( U = 0 \),

\[
(11) \quad S = -\frac{U}{b\nu} (\ln \frac{U}{a\nu} - 1).
\]
In the last of a series of five papers by Planck on irreversible radiation processes [32], this expression appears, without any justification, as a **definition** for the entropy of his oscillators. But as it has been shown, it is clear that Planck obtained it in a straightforward fashion from Wien’s relation, Eq.1 and Eq 6. Taking the second derivative of $S$ with respect to $U$, he found that its reciprocal depends linearly on $U$,

$$R = \left( \frac{d^2 S}{dU^2} \right)^{-1} = -b\nu U.$$  

While Planck obtained this simple linear dependence of $R$ on $U$ from Wien’s relation, he attached to it a special significant claiming to have demonstrated that it was **unique**, leading to a **derivation** of the scaling dependence of $U(\nu, T)$ on $\nu$ and $T$, Eq.8. Actually, integrating this equation

$$\frac{dS}{dU} = \frac{1}{b\nu} \ln \frac{U}{\xi(\nu)}$$

where $\xi(\nu)$ is an undetermined function of $\nu$. Hence, the fact that Wien’s relation indicates $\xi(\nu)$ depends linearly on $\nu$ was not justified.

In a paper presented to the Berlin Academy of Sciences on May 18, 1899, Planck stated that:

“I believe it must therefore be concluded that the definition given for the entropy of radiation, and also the Wien distribution law for the energy which goes with it, is a necessary consequence of applying the principle of entropy increase to the electromagnetic theory of radiation, and that the limits of validity of this law, should there be any, therefore coincide with those of the second law of thermodynamics. Further experimental test for this law naturally acquires all the greater fundamental interest for this reason.” [26]

Later on in his autobiography, Planck recalled that

“This relationship is so surprisingly simple that for a while I considered it to possess universal validity, and I endeavored to prove it theoretically. However, this view soon proved to be untenable in the face of later measurements. For although in the case of small energies and correspondingly short waves Wien’s Law continued to be confirmed in a satisfactory manner, in the case of large values of the energy and correspondingly long waves, appreciable divergences were found, first by Lummer and Pringsheim; and finally the measurements of H. Rubens and F. Kurlbaum on infrared rays of fluorspar and rock salt revealed a behaviour which, though totally different, is again a simple one, in so far as the function $R$ is proportional not to the energy but to the square of the energy for large values of the energy and the wave-lengths.” [27]

---

2 Historians of science have followed Klein’s statement that “it seems most likely that Planck was guided by the form Wien’s distribution law”, when if fact there isn’t any other alternative.
By early 1900 the experiments of Otto Lummer and Ernst Pringsheim \(^3\) gave evidence of deviations from Wien’s formula at the longer observed wave lengths of order 10 microns, and at temperature of about 1000 degrees [23]. Further data by Heinrich Rubens and Felix Kurlbaum at a wavelength of 51 microns indicated that the black body radiation dependent linearly on temperature [24]. These experiments were made possible by a new detection technique developed by Heinrich Rubens and his American collaborator Ernst F. N. Nichols which enhanced the low intensity longer wavelengths by resonant scattering from a crystal lattice[25]. Planck was informed of these new results by Rubens himself who visited him with his wife on a Sunday afternoon, Oct. 7, 1900, and he began promptly to reconsider his arguments.

Even before the new data appeared, Lord Rayleigh derived a linear dependence on temperature for the blackbody distribution from the equipartition theorem, applied to classical radiation emitted by charged one dimensional oscillators in a box in thermal equilibrium. To obtain this dependence on temperature Planck found that \(R\) had to dependent quadratically on \(U\) [33]. Supposing that

\[ U(\nu,T) = \eta T, \]

where \(\eta\) is a constant, named \(\alpha\) in Planck’s paper and corresponding to \(k\). Then according to Eq.9

\[ \frac{1}{T} = \frac{dS}{dU} = \eta \frac{U}{U}, \]

and therefore

\[ R = \left( \frac{d^2S}{dU^2} \right)^{-1} = \frac{U^2}{\eta}. \]

In Planck’s own words,

“Thus, direct experiments established two simple limits for the function \(R\): for small energies, \(R\) is proportional to the energy; for larger energy values \(R\) is proportional to the square of the energy. Obviously, just as every principle of spectra energy distribution yields a certain value for \(R\), so also every formula for \(R\) leads to a definite law of the distribution of energy. The problem was to find such a formula for \(R\) which would result in the law of the distribution of energy established by measurement. Therefore, the most obvious step for the general case was to make the values of \(R\) equal to the sum of a term proportional to the first power of the energy and

\(^3\)Alan Franklin wrote that one of the greatest anticlimaxes in all of literature occurs at the end of Shakespeare’s Hamlet. On a stage strewn with noble and heroic corpses - Hamlet, Laertes, Claudius, and Gertrude- the ambassadors from England arrive and announce that “Rosencrantz and Guildenstern are dead”. No one cares. A similar reaction might be produce among a group of physicist, or even among historians and philosophers, werfe someone to announce that “Lummer and Pringstein are dead”[30]
another term proportional to the second power of the energy, so that the first term becomes decisive for small values of the energy and the second term for large values. In this way a new radiation formula was obtained, and I submitted it for examination to the Berlin Physical Society, at the meeting on October 19, 1900."

By such phenomenological considerations, Planck generalized his thermodynamic expression for the dependence of the entropy on the oscillator energy to interpolate between the short wavelength or Wien regime, and the long wavelength or Rayleigh regime. Setting now
\[ R = \left( \frac{d^2S}{dU^2} \right)^{-1} = -\frac{1}{\eta} U(\eta b\nu + U), \]
he obtained his previous linear dependence of $R$ on $U$, Eq. 12, for $U \ll \eta b\nu$ and the quadratic dependence on $U$, Eq. 16, for $U \gg \eta b\nu$. This simple interpolation formula for $R$ turned out, surprisingly, to be valid not only in these two energy regimes, but to be an exact relation for all values of $U$. Integrating this relation by applying the thermodynamic relation between the absolute temperature and the derivative of the entropy with respect to the energy, Eq. 9, and assuming the boundary condition $U \to \infty$ when $T \to \infty$, one obtains
\[ \frac{1}{T} = \frac{1}{b \nu} \ln(1 + \eta b\nu/U), \]
which yields the dependence on temperature $T$ and frequency $\nu$ of the mean oscillator energy,
\[ U(\nu, T) = \frac{\eta b\nu}{(\exp(b\nu/T) - 1)}. \]
Finally, to recover the relation for $U(\nu, T)$ in the Wien limit when $b\nu \gg T$, Planck obtained a relation for the new constant $\eta$
\[ \eta = \frac{h}{b}. \]
Since $b = h/k$, where $k$ is Boltzmann constant, $\eta = k$ in accordance with the equipartition theorem for a one dimensional harmonic oscillator. Substituting this expression in Eq. 7 for his relation between the spectral distribution $\rho(\nu, T)$ and the mean oscillator energy $U(\nu, T)$, Planck then obtained his blackbody formula which he wrote as a function of the wavelength $\lambda$ measured in the experiments [33], where $\lambda = c/\nu$, and $c$ is the velocity of light.
\[ \rho(\lambda, T) = \frac{C \lambda^{-5}}{(\exp(bc/\lambda T) - 1)}, \]
where $C = 8\pi hc$. In the limit $bc/\lambda \gg T$, Planck recovered his earlier result for the Wien spectrum, Eq. 6, while for $bc/\lambda \ll T$ he obtained the linear dependence of $\rho$ on $T$, in accordance with the new experimental results at the PTR. It should be pointed out that it is completely unexpected that by an interpolation procedure to fit experimental
data, Planck obtained a formula for the spectral distribution of black body radiation, that turned out to be exact for all temperatures and wavelengths. His procedure was sensible as a phenomenological data fitting approach, but it is purely accidental that he succeeded in this way to obtain the exact formula for black-body radiation. After all, he did not have any arguments to exclude, for example, cubic or higher powers of $U$ in his expansion, Eq.17, of $R$ in powers of $U$.

After Rubens checked the new radiation formula against his experiments, Planck described his reaction:

“The very next morning I received a visit from my colleague Rubens. He came to tell me that after the conclusion of the meeting, he had that very night checked my formula against the results of his measurements and found a satisfactory concordance at every point... Later measurements too confirmed my radiation formula again and again - the finer the methods of measurement used, the more accurate the formula was found to be.”...In this way a new radiation formula was obtained, and I submitted it for examination to the Berlin Physical Society, at the meeting of October 19, 1900.”[27]

Planck, however, did not refer to Rayleigh’s result, whose work was apparently motivated by the unphysical dependence of Wien’s formula which predicted that the spectral energy $\rho(\lambda, T)$ saturates when $\lambda T >> \frac{hc}{k}$. Rayleigh remarked,

“Nevertheless, the [Wien’s] law seems rather difficult of acceptance, especially the implication that as the temperature is raised, the radiation of given wavelength approaches a limit... The question is one to be settled by experiment; but in the meantime I venture to suggest a modification of the Wien distribution, which appears to me more probable apriori. Speculations upon this subject is hampered by the difficulties which attend the Boltzmann-Maxwell doctrine of the partition of energy [my italics]. According to this doctrine every mode should be alike favoured; and although for some reason not yet explained the doctrine fails in general, it seems plausible that it applies for the graver [longer wavelengths] modes ”[16].

III. Planck’s application of Boltzmann’s relation between entropy and probability in statistical mechanics

This section reviews the critical phase when in order to find the physical significance for his purely phenomenological formula for black-body radiation, Eq.21, Planck turned for enlightenment to Boltzmann’s 1877 seminal paper on the foundations of statistical mechanics. Later, as Planck described it in his autobiography,

“But even if the absolute precise validity of the radiation formula is taken for granted, so long as it had merely the standing of a law disclosed by lucky intuition, it could not be expected to possess more than a formal significance. For this reason, on the very day when I formulated this law, I began to devote myself to the task of investing it with a true physical
meaning. *This quest, as a matter of course, led me to study the interrelation of entropy and probability - in other words to pursue the line of thought inaugurated by Boltzmann.*

In his Nobel speech, Planck added the remark,

> “After a few weeks of the most strenous work of my life, the darkness lifted and an unexpected vista began to appear” [31],

and continued,

> “Since the entropy $S$ is an additive magnitude, but the probability $W$ is a multiplicative one, I simply postulated that $S = k \log W$, where $k$ is a universal constant; and I investigated whether the formula for $W$, which is obtained when $S$ is replaced by its value corresponding to the above radiation law could be interpreted as a measure of probability” [31].

By integrating Eq.18, with the boundary condition that the entropy $S$ vanishes when $U = 0$, Planck obtained

$$S = k[(1 + U/h\nu)\ln(1 + U/h\nu) - (U/h\nu)\ln(U/h\nu)]$$

where $k$ is a constant known as Boltzmann’s constant. This form of the dependence of $S$ on $U$ and $\nu$ satisfies the Wien’s scaling relation, Eq.1, indicating that $S$ depends only on the ratio $U/\nu$.

Having found a phenomenological expression for the entropy $S$ as a function of the mean energy $U$, Planck considered the possibility that the function $W = \exp(S/k)$ could have an interpretation as a measure of the probability for the configuration of his Hertzian oscillators corresponding to that of molecular velocities in Boltzmann’s 1877 formulation of statistical mechanics[9]. In his Nobel speech, Planck remarked,

> “As a result, I found that this was actually possible, and that in this connection $k$ represents the so-called absolute gas constant, referred not to gram molecules or mols, but to the real molecules. Now for the magnitude $W$, I found that in order to interpret it as a probability, it was necessary to introduce a universal constant, which I called $h$. Since it had the dimension of action [energy $\times$ time], I gave it the name *elementary quantum of action*. Thus the nature of entropy as a measure of probability, in the sense indicated by Boltzmann, was established in the domain of radiation, too.” [31].

Further evidence for Planck’s train of thought, can be found in one of the few surviving letters from that period, that he wrote to Otto Lummer on 26 October 1900,

> “If the prospect should exist at all of a theoretical derivation of the radiation law, which I naturally assumed, then in my opinion, this can be the case only if it is possible to derive the expression for the probability of a radiant state, and this, you see, is given by the entropy. Probability presumes disorder, and in the theory I have developed, the disorder occurs in the irregularity with which the phase of the oscillations changes even in the most
homogeneous light. A resonator, which corresponds to a monochromatic radiation, in resonant oscillations will likewise show irregular changes of its phase [and also of its instantaneous energy, which was more important for Planck’s subsequent derivation], and on this the concept and magnitude of its entropy are based. According to my [blackbody radiation] formula communicated on 19 October to the German Academy the entropy of the resonator should become

\[ S = \alpha \ln \left( \beta + U \right) \ln \left( \beta + U \right) / U^U, \]

and this form very much recalls expression occurring in the probability calculus”[36].

The equation that Planck wrote for the entropy \( S \) in this letter, corresponds to Eq.22 for \( \beta = h\nu \), apart from an additive constant \( \alpha \beta \ln \beta \), where \( \alpha \beta = k \).

Planck continued, “After all, in the thermodynamic of gases, too, the entropy \( S \) is the log of a probability magnitude, and Boltzmann has already stressed the close relationship of the function \( \chi \) which enters the theory of combinatorics, with the dynamic entropy. I believe, therefore, that the prospect would certainly exist of arriving at my formula by a theoretical route which would also give us the physical significance of the constants \( C \) and \( c \)”[36].

To illustrate his statistical principles for thermal equilibrium, Boltzmann had considered a gas of molecules, and in order to count the number of configurations or complexions, as he called them, he discretized the energy of the molecules in integral multiples of a unit or energy element \( \epsilon \). Then each configuration is specified by a set of integers giving the number of energy elements of each molecule subject to the constraint that the mean energy \( U \) per molecule satisfies the relation \( U = (P/N)\epsilon \), where \( P \) is an integer, and \( N \) is the total number of molecules. It is fairly plausible, as indicated in his letter to Lummer, that Planck started with his empirically successful relation for the entropy \( S \) of the oscillators, Eq.22, and working backwards proceeded to obtain the quantity \( W = \exp(S/k) \), and associated it with the total number of configurations of his oscillators, in analogy with Boltzmann’s model for a gas of molecules [37]. In this case, replacing the ratio \( U/h\nu \) which appears in Eq.22 for \( S \) by \( P/N \), Planck would have obtained the relation

\[ W = \left( \frac{N + P}{N} \right)^{(N+P)}, \]

which is Stirling’s approximation to the total number of complexions given in Boltzmann’s paper.

In two articles [18] [34] that he presented to the German Physical Society, the first one given on Dec.14, 1900, Planck now derived his entropy formula by starting with Eq.24 for the total number \( W \) of equally probable complexions of his oscillators, and then obtaining the entropy \( S \) from the relation \( S = k \log(W) \) by setting \( P/N = U/h\nu \), which yields Eq.22. In this presentation Planck introduced for the first time the constant \( k = h/\beta \), and pointed out that it is universal and applies to all thermodynamic system including Boltzmann’s
model for a gas of molecules. In this case \( k = R_o/N_o \) where \( R_o \) is the gas constant and \( N_o \) is Avogadro's number. Planck also argued that for each frequency \( \nu \) there would be corresponding values of \( N_\nu \) and \( P_\nu \) such that \( P_\nu/N_\nu = U_\nu/h\nu \) and that the total number of configurations \( W \) which is the product of all \( W_\nu \) should be maximized for a fixed total energy, but he did not carry out such a calculation.

IV. Planck’s 1931 recollections

In the previous section we have shown how Planck obtained a theoretical derivation of his black-body formula based on Boltzmann’s development of statistical mechanics in 1877. The idea of introducing discrete energy elements of magnitude \( \epsilon \) originated with Boltzmann, who applied it as a mathematical device to count configurations for a gas of molecules, but subsequently he reinstated the classical energy continuum by taking the limit \( \epsilon = 0 \). Planck followed Boltzmann’s approach by assigning integral multiples of these discrete energy element to his oscillators of frequency \( \nu \). But then he departed from Boltzmann in a fundamental way by fixing \( \epsilon \) to have a constant value, \( \epsilon = h\nu \), where \( h \) is a new universal constant, instead of letting \( \epsilon \) become vanishingly small as demanded by classical physics. Why Planck did not take this latter step is obvious: the assumption of a discrete energy element of magnitude \( h\nu \) gave him precisely the result he was searching to justify his phenomenological black-body formula. But as would be expected, he did not (and could not) justify such discreteness on a-priori grounds. It is appears that in the short time before Planck submitted his paper to the German Physical Society, he could not have been aware of the broader physical implications of his unprecedented ansatz, because he did not offer any comments indicating that he was departing from the canons of classical physics. Planck turned to Boltzmann’s approach rather late in his research program, and therefore it seems very likely that in 1900 he had not yet fully mastered the foundations of Boltzmann’s statistical mechanics.

In this connection it is interesting to examine Planck’s own account of his discovery some 31 years later. In response to a request from the American physicist Robert Williams Wood, Planck described “the considerations which led him to propose the hypothesis of energy quanta” in a letter he wrote to him on October 7, 1931:

“Briefly summarized, what I did can be described as simply an act of desperation. By nature I am peacefully inclined and reject all doubtful adventures. But by then I had been wrestling unsuccessfully for six years with the problem of the equilibrium between radiation and matter and I knew that the problem was of fundamental importance to physics; I also knew the formula that expresses the energy distribution in the normal spectrum. A theoretical interpretation therefore had to be found at any cost, no matter how high. It was clear to me that classical physics could offer no solution to this problem and would have meant that all energy would eventually transfer from matter into radiation [my italics]. In order to prevent this, a new constant is required to assure that energy does not disintegrate. But the only way to recognize how this can be done is to start from a definite
point of view. This approach was open to me by maintaining the two laws of thermodynamics. The two laws, it seems to me, must be upheld under all circumstances. For the rest, I was ready to sacrifice every one of my previous convictions about physical laws. Boltzmann had explained how thermodynamic equilibrium is established by means of statistical equilibrium, and if such an approach is applied to the equilibrium between matter and radiation, one finds that the continuous loss of energy into radiation can be prevented by assuming that energy is forced at the outset to remain together in certain quanta. This was purely a formal assumption and I really did not give it much thought except that no matter what the cost, I must bring about a positive result.” [38]

The statements in this letter agree with the description that can be gleaned from the papers and earlier correspondence of Planck, but there are also some oddities in this account. Contrary to Planck’s remark that “classical physics could offer no solution to this problem. . . ” he continued to hope for about a decade after his first application of Boltzmann’s discrete statistical approach, that his quantum of action could somehow be justified on purely classical grounds. As he stated in his autobiography,

“ While the significance of the quantum of action for the interrelation between entropy and probability was thus conclusively established, the part played by this new constant in the uniformly regular occurrence of physical processes still remained an open question. I therefore tried immediately to weld the elementary quantum of action $h$ somehow into the framework of classical theory. But in the face of all such attempts, this constant showed itself to be obdurate. . . ” [39]

More likely, originally discreteness in energy for Planck was a “purely formal assumption” that Boltzmann had introduced in 1877 to which Planck “really did not give much thought”, because it brought about “a positive result”, namely, a theoretical derivation of his black-body formula, previously obtained from phenomenological considerations, by an adaptation of Boltzmann’s principles of statistical mechanics to a model of oscillating charges in thermal equilibrium with electromagnetic (black-body) radiation. Planck’s good fortune was that these principles could be applied straightforwardly not only to thermodynamic equilibrium systems which satisfy the laws of classical continuum physics, but also to systems which are described by discrete energy levels such as Hertzian oscillators in equilibrium with thermal radiation. Planck had been correct “to uphold the two laws of thermodynamics under all circumstances” because, as it turned out, only these two laws were not modified by the advent of the new theory of quantum mechanics.

**Summary and Conclusions**

We have shown that:

a) Assuming that the energy distribution for black body radiation is due to an ensemble of Hertzian oscillators in equilibrium with this radiation at a fixed temperature $T$, Planck
calculated the relation of this distribution to the mean energy of these oscillators by applying Maxwell’s theory of electromagnetism, Eq.7.

b) On the basis of Wien’s displacement formula for this distribution, Planck then obtained an expression for the entropy of his Hertzian oscillators, Eq.11.

c) Motivated by new experiments at longer wavelengths that disagreed with Wien’s formula, Planck developed a phenomenological interpolation relation, Eq.22, for his entropy relation that agreed very well with the new and the older experiments, Eq.22.

d) Planck’s then took his “most essential step”, namely, to obtain a theoretical foundation for his new and successful entropy formula. For this purpose, Planck turned to Boltzmann’s 1877 paper on the connection of entropy $S$ with the maximum number of complexions $W_B$ of a system in thermal equilibrium, i.e. $S \propto \ln W_B$. In particular, this paper contained an expression for the total number of complexions $W$ for an ensemble of one dimensional molecules having multiples of a discrete energy element $\epsilon$. Of course, since energy in classical mechanics is continuous, at the end of his calculation Boltzmann took the limit $\epsilon = 0$, corresponding to the result of the equipartition theorem, first published for the distribution of black body radiation by Lord Rayleigh[16]. Planck, however, realized that by setting $\epsilon = h\nu$, he recovered his successful phenomenological relation for the entropy.

Martin Klein [2] and subsequently other historians of science have questioned why Planck departed from Boltzmann derivation, by taking the total number of complexions $W$ instead of its maximum number $W_B$. But the explanation is rather straightforward: Planck had obtained a phenomenological formula for the entropy as the logarithm of a function of the variable $U/h\epsilon$, with $\epsilon = h\nu$, Eq.22, but Boltzmann’s paper did not contain a corresponding expression for the entropy proportional to $\ln W_B$ as a function of this variable, except in the limit $\epsilon = 0$ corresponding to classical mechanics. But his paper contained an expression for the total number of complexion $W$, in terms of the number $N$ of molecules, and an integer $P = E/\epsilon$, where $E$ is the total energy of the system. Substituting $P/N = U/h\nu$ in the Stirling approximation for $W$ leads to an expression for $\ln(W)$ equal to $N$ times Planck’s phenomenological expression for the entropy $S$, apart from the constant of proportionality $k$. In one of most recent articles on this subject [3], Clayton A. Gearhart posed another question also similar to that of other historians of science: “Why should Planck’s complexions, which represent distributions of energy elements, be equally probable?” The answer is that the assumption of equally probable complexions is at the foundation of Boltzmann’s statistical method, and naturally Planck adopted it because it solved his problem. Other historians of science like Kuhn and Darrogol who deny that Planck introduce discreetness of energy into physics, evidently failed to understand how Planck arrived at this insight. A detailed discussion of the relation between Boltzmann’s and Planck’s definition and calculation of entropy is given in the Appendix. In particular, it is shown that $\ln W_P = (1/N)\ln W_B$, an essential relation that Planck could have easily demonstrated.

The occurrence of energy discreteness caused a great deal of concern to Planck’s contemporaries, most notably Einstein, Lorentz and Ehrenfest, and it took several years before the significance of Planck’s quantum of action to physics began to emerge. The acceptance
and further development of Planck’s hypothesis is a very interesting and important subject in its own right, but it will not be pursued here.

**Appendix. Relation between Boltzmann’s and Planck’s treatment of Entropy**

An interesting question that has been raised by several historians of science is to explain why Planck departed from the method introduced by Boltzmann in his 1877 paper on the relation between the second law of thermodynamics and the theory of probability \[?\]. In this paper, Boltzmann considered the state of thermal equilibrium of an ensemble of molecules having discrete energies that are multiples of a energy unit with finite magnitude \(\epsilon\). Therefore, it would appear that by setting \(\epsilon = h\nu\), Planck could just have taken over Boltzmann’s result for the mean energy \(U\) of these molecules, and shown that it corresponded to the result he had obtained previously from his purely phenomenological considerations. It will be shown, however, that the reason why he could not have taken this apparently straight forward step is that Boltzmann had obtained the mean energy \(U\) of his model for molecules with discrete energies in term of a parameter \(x\), but he did not determine the dependence of this parameter on temperature except in the classical limit when \(\epsilon \to 0\), and found that \(U = kT\), corresponding to the result of classical equipartition theorem.

Boltzmann related the entropy \(S\) to the logarithm of the *maximum* number of configurations \(W_B\) of \(n\) molecules of fixed total energy \(\lambda\epsilon\), where \(\lambda\) is an integer, and the energy of each molecule is a multiple of a fixed value \(\epsilon\). In section I of this paper \[?\], he gave an expression for \(W_B\) associated with the distribution of \(\lambda\) discrete energy elements of magnitude \(\epsilon\) among \(N\) molecules, such that a number \(n_j\) of molecules each have energy \(j\epsilon\), where \(j = 0, 1, 2...p\), and \(p\) is an integer. Then

\[
W_B = \frac{n!}{\prod_j n_j!},
\]

and in terms of these variables, Boltzmann’s fundamental principles of statistical mechanics can be stated as follows:

*Equally probable* configurations (Boltzmann called them complexions) are characterized by the set integers \(n_j\) which are subject to the constraints that the total number of molecules \(n\), and the total energy \(E = \lambda\epsilon\) are fixed, where

\[
n = \sum_{j=0}^{p} n_j,
\]

and

\[
\lambda = \sum_{j=0}^{p} jn_j,
\]

where \(p \leq \lambda\), because \(n_j = 0\) or \(1\) for \(j = \lambda\), and \(n_j = 0\) for \(j > \lambda\) \[40\]. In this case the state of thermal equilibrium is obtained by the maximum value of \(lnW_B\), subject to
the constraints that $\delta n = 0$, and $\delta \lambda = 0$. Introducing two undetermined constants, $\alpha$ and $\gamma$ (named $h$ and $k$ by Boltzmann), this condition is satisfied by the requirement that $\delta \ln W_B + \alpha \delta n + \gamma \delta \lambda = 0$. In the Stirling approximation for the factorials that appear in the expression for $W_B$, Eq.25, it becomes

$$\sum_{j=0}^{j=p} \delta n_j (\ln(n_j) + \alpha + \gamma j) = 0,$$

and setting $n_0 = \exp(-\alpha)$, and $x = \exp(-\gamma)$, it is satisfied by

$$n_j = n_0 x^j,$$

Substituting Eq.29 in Eqs. 26 and 27, and taking the limit $p \to \infty$, [41] we have

$$\frac{n_j}{n} = (1 - x)x^j,$$

for the fraction of molecules that have energy $j\epsilon$, and $x$ is determined by the ratio $n/\lambda$

$$x = \frac{1}{1 + n/\lambda}.$$

This result was given explicitly by Boltzmann (reference [?], last line of p.180), and leads the mean energy $U = \lambda \epsilon / n = cx/(1 - x)$ of the molecules in his ensemble as a function of $x$ [42]. Setting $x = \exp(-b\epsilon/T)$, and $\epsilon = h\nu$, this relation corresponds to Planck’s formula, Eq.19, for the mean energy of an oscillators of frequency $\nu$, at the absolute temperature $T$, but in his paper Boltzmann did not calculate the dependence of $x$ on $T$. This dependence is obtained from Boltzmann’s identification of $\ln W_B$ with the entropy $S$ of the system, by applying the second law of thermodynamics relation $1/T = dS/dE$, but he applied it only in the limit of vanishing $\epsilon$ appropriate to classical mechanics.

We now carry out this straight forward calculation for finite $\epsilon$. It would have been easy for Planck too, but there isn’t any evidence that he did it. In the Stirling approximation, the maximum value of $\ln W_B$ is

$$\ln W_B = n\ln(n) - n_0 \sum_{j=0}^{\infty} x^j (j \ln x + \ln(n_0)),$$

and substituting Eqs.26,27,30 and 31 in Eq.32, we obtain $\ln W_B$ as a function of Boltzmann’s parameter $x$,

$$\ln W_B = -n\left[\frac{x}{(1-x)} \ln(x) - \ln(1-x)\right].$$

Substituting for $x$ its dependence on $n/\lambda$,

$$\ln W_B = (n + \lambda)\ln(n + \lambda) - n\ln(n) - \lambda n\lambda.$$
which shows that $W_B$ is the Stirling approximation for the total number of all possible complexions $W$ of Boltzmann’s ensemble either of molecules or Planck’s oscillators, where

$$W = \frac{(n + \lambda - 1)!}{(n - 1)!\lambda!}.$$  

Hence, the only puzzle is why Planck did not carry out this simple calculation missing in Boltzmann’s paper. Setting $U$ equal to the mean energy/molecule, $U = \lambda \epsilon / n$, we have, according to Eq.31,

$$x = 1/(1 + \epsilon/U),$$

and substituting this expression for $x$ in Eq.33, one obtains an explicit dependence of $\ln W_B$ on $U$,

$$\ln W_B = n[(1 + U\epsilon/U) \ln(1 + U\epsilon/U) - U\epsilon \ln U],$$

Apart from the constant of proportionality $k$, this expression with $\epsilon = h\nu$ is equal to $n$ times Planck’s phenomenological expression, Eq.22, for the entropy per oscillator $S = k(\ln W_P)$ of an ensemble of linear oscillators of frequency $\nu$.

Finally, according to the second law of thermodynamics, the absolute temperature $T$ for thermal equilibrium is given by

$$\frac{1}{T} = \frac{dS}{dU} = \frac{k}{\epsilon} \ln(1 + \frac{\epsilon}{U}),$$

which was calculated by Planck for $\epsilon = h\nu$, Eq.18-20. Hence,

$$U = \frac{\epsilon}{\exp(\epsilon/kT) - 1},$$

and according to Eqs.36 and 39, the dependence of Boltzmann’s parameter $x$ on the temperature $T$ is

$$x = \exp(-\epsilon/kT).$$

Substituting this expression for $x$ in Eqs. 29 and 30 yields

$$n_j = n(1 - \exp(-\epsilon/kT)\exp(-j\epsilon/kT)).$$

But in the first section of his paper, Boltzmann did not calculate this relation for $n_j$ except in the limit that $\epsilon << U$, and obtained

$$n_j \approx \frac{n\epsilon}{U} \exp(-j\epsilon/U),$$

without giving the equipartition result of classical thermodynamics, $U = kT$ (see ref. [?] p.186, and ref. [40] p. 49).

Setting $j\epsilon = E$, $\epsilon = dE$, and $n_j/n = dp$, then $p(E)$ is the Maxwell-Boltzmann classical probability distribution for Boltzmann’s molecular ensemble having energy in the interval $E, E + dE$

$$\frac{dp}{dE} = \frac{1}{U} \exp(-E/U),$$
with $U = \int_0^\infty dE p(E) E$, the mean energy/molecule\cite{43}. Boltzmann pointed out, however, that this relation, with $E = (1/2)mv^2$, where $m$ is the mass and $v$ is the velocity of a molecule, is valid only in two spatial dimensions \cite{44}.

ACKNOWLEDGMENTS

I would like to thank Clayton Gearhart for a careful reading of this manuscript and many helpful comments and corrections.

REFERENCES


[17] Boltzmann discrete model was applicable in the limit of vanishing $\epsilon$ only to the motion of molecules in two dimension (ref. [9] p. 190).


[19] In this connection, it is worthwhile to recall Einstein’s observation that “If you want to find out anything from the theoretical physicists about the methods they use, I advise you to stick closely to one principle: don’t listen to their words, fix your attention on their deeds.” Albert Einstein, “Ideas and Opinions” (1954) p. 270.


[37] Leon Rosenfeld, “The first phase of the evolution of the Quantum Theory”, Osiris 2, 148-196 (1936).


[40] But this constraint on $p$ does not apply in the Stirling approximation for the factorials in $W_B$. In his book (ref. [5]), Kuhn sets $p = \lambda$, and states that “standard variational techniques lead directly to the conclusion that for $p \gg n^2$, $n_j$ is given by Boltzmann’s classical expression, Eq. 42. But this claim is incorrect, because to obtain the distribution in the classical limit, Boltzmann set $p = \infty$, and $\lambda \gg n$.

[41] Boltzmann also considered $p$ finite, but the case of interest related to Planck’s formula corresponds to $p \to \infty$.

[42] In the Stirling approximation the resulting values of $n_j$, Eq. 30, are not integers. In this case $\lambda$ and $n$ are also infinite, but the ratio $\lambda/n$ is fixed, and the ratio $n_j/n$ is finite corresponding to the fraction of molecules with energy $j\epsilon$.


[44] Ref. [9], p. 190