

Article

Introducing Quantum and Statistical Physics in the Footsteps of Einstein: A Proposal [†]

Marco Di Mauro ^{1,*} , Salvatore Esposito ^{2,‡}  and Adele Naddeo ^{2,‡} 

¹ Dipartimento di Matematica, Università di Salerno, Via Giovanni Paolo II, 84084 Fisciano, Italy

² INFN Sezione di Napoli, Via Cinthia, 80126 Naples, Italy; sesposit@na.infn.it (S.E.); anaddeo@na.infn.it (A.N.)

* Correspondence: madimauro@unisa.it

[†] This paper is an extended version from the proceeding paper: Di Mauro, M.; Esposito, S.; Naddeo, A.

Introducing Quantum and Statistical Physics in the Footsteps of Einstein: A Proposal. In Proceedings of the 1st Electronic Conference on Universe, online, 22–28 February 2021.

[‡] These authors contributed equally to this work.

Abstract: Introducing some fundamental concepts of quantum physics to high school students, and to their teachers, is a timely challenge. In this paper we describe ongoing research, in which a teaching–learning sequence for teaching quantum physics, whose inspiration comes from some of the fundamental papers about the quantum theory of radiation by Albert Einstein, is being developed. The reason for this choice goes back essentially to the fact that the roots of many subtle physical concepts, namely quanta, wave–particle duality and probability, were introduced for the first time in one of these papers, hence their study may represent a useful intermediate step towards tackling the final incarnation of these concepts in the full theory of quantum mechanics. An extended discussion of some elementary tools of statistical physics, mainly Boltzmann’s formula for entropy and statistical distributions, which are necessary but may be unfamiliar to the students, is included. This discussion can also be used independently to introduce some rudiments of statistical physics. In this case, part of the inspiration came from some of Einstein’s papers. We present preliminary, qualitative results obtained with both teachers and selected pupils from various high schools in southern Italy, in the course of several outreach activities. Although the proposal was only tested in this limited context for now, the preliminary results are very promising and they indicate that this proposal can be fruitfully employed for the task.

Keywords: physics teaching; history of physics; quantum physics; statistical physics



Citation: Di Mauro, M.; Esposito, S.; Naddeo, A. Introducing Quantum and Statistical Physics in the Footsteps of Einstein: A Proposal. *Universe* **2021**, *7*, 184. <https://doi.org/10.3390/universe7060184>

Academic Editor: Gerald B. Cleaver

Received: 30 April 2021

Accepted: 31 May 2021

Published: 4 June 2021

Publisher’s Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The world around us is becoming increasingly more complex and technological, making basic scientific literacy essential for citizens. It is also likely that, in the near future, quantum technologies will be at everyone’s disposal. Thus, it is becoming more and more important that average educated people have at least a basic understanding of quantum physics, which today is still a prerogative of physics graduates. This is also desirable in view of the great cultural significance of quantum physics. It is not surprising then that, in many countries, quantum physics is now part of high school curricula, along with elements of other important parts of twentieth century physics, such as the theories of relativity, nuclear and particle physics, astrophysics and cosmology, and chaos theory. This state of things is currently challenging teachers, and also researchers in physics education, to develop educational tools aimed at introducing high school students, who have studied the basics of classical physics, to the main concepts of modern physics and of quantum theory in particular [1]. This is a highly non trivial task (not only at the high school level,

I have thought a hundred times more about quantum problems than I have about general relativity. (A. Einstein to O. Stern)

but also at university), considering the intrinsic difficulty of the subject, which moreover has the reputation of being awkward and counterintuitive. To make things more difficult, there is the problem that high school teachers themselves may lack a proper education in modern physics. For example, in the case of Italy, some of the teachers have a major in mathematics, but until recent years, the curriculum of mathematics majors, in particular of those who intended to pursue a career in school teaching, did not include any quantum theory. While things are changing, the gap remains, and in the attempt of filling it, many universities organize dedicated courses aimed at teachers or directly at their students. It is therefore important to design appropriate teaching–learning sequences which can give them a solid grounding in quantum physics.

It is well established that the history of physics can be helpful to teaching in multiple ways (see e.g., [2]). A first advantage of using history surely consists in the possibility of putting physics topics in context, by embedding them in the time in which they were developed and linking them with other disciplines, especially in the humanities, whose teaching is intrinsically historical. Then there is the possibility of taking inspiration from the reading of original texts by the founders of the subject, which can be useful especially for modern physics, which involves subtle conceptual steps. In fact, due to the particularly counterintuitive nature of the subject, and to the large number of new concepts involved, which required a very long gestation, the teaching of quantum physics is very often developed by following a historical path. In fact, a typical teaching–learning sequence for quantum physics, both at the high school and the undergraduate level (see, for example, the textbooks [3–6]), will start from Planck’s formula for black body radiation, and gradually introduce the main physical concepts that in the end will be coherently subsumed in the theory of quantum mechanics, roughly in the same order in which they were discovered for the first time (of course complemented with a discussion of all the relevant experiments, in which the phenomena which motivated the introduction of these concepts were discovered). We do not have any objection towards this way of developing the subject but we are also conscious that history is far more complex and richer than any didactic presentation. While it is of course inconceivable to present the full complexities of history in a didactic course, it may still be the case that some of the parts of history, that are generally excluded in modern streamlined presentations, may provide additional precious insight. In the present paper we suggest that such insight can be found in some works by Einstein in quantum physics¹. As historians know very well [8,9], Einstein’s contributions to quantum theory were many and multifaceted, and several of them were actually groundbreaking. In fact, many of the fundamental concepts that nowadays are at the core of quantum mechanics have actually been introduced in some seminal paper by Einstein. Despite this, much of this work is not included in usual curricula. As an example, we may think of the way light quanta are introduced. After discussing Planck’s radiation formula and its explanation in terms of the energy quantization formula $\varepsilon = h\nu$, it is customary to go on immediately to the explanation of the photoelectric effect in terms of light quanta, as if there was no conceptual leap between Planck’s statistical hypothesis of energy quantization and the interaction of individual light quanta with electrons in a metal. Students in this way may get the idea that light quanta were introduced by Planck, and Einstein’s contribution was limited to an application of this idea, while actually the concept was born in Einstein’s work. The explanation of the photoelectric effect is of course a revolutionary contribution² (after all, it earned Einstein a Nobel prize) but it is, more often than not, the only contribution of Einstein that is mentioned. The deep difference between Planck’s hypothesis of energy quantization and Einstein’s hypothesis of light quanta is not emphasized enough. Nor is there any mention of Einstein’s compelling statistical reasoning (in fact, much simpler than Planck’s one) in order to put forward the hypothesis that radiation itself is quantized [11]. The already mentioned great conceptual leap involved in going from both energy quantization and the light quanta hypotheses, which are statistical in nature, to the application of the latter to the photoelectric effect, in which individual quanta are involved, is often not acknowledged. The photoelectric effect

was not the only phenomenon which was explained by Einstein with his hypothesis. The application of light quanta to fluorescence and Stokes' rule, which is very insightful despite being so simple, is often not considered. Yet, it provides a link between quantum physics and everyday phenomena, which is highly desirable in teaching.

Apart from the mentioned issues, we recall that, historically, between the light quanta hypothesis of 1905 and Bohr's atomic theory of 1913 (which is typically the next topic in courses), in 1909 there was the pivotal introduction by Einstein of the idea of wave–particle duality for light, which was again the result of a statistical reasoning applied to the black body radiation formula [12] (see also [13])³, and it shows that in fact both the wave and the particle nature of light (i.e., not just the latter) are necessary for accounting for the observed spectrum. This brilliant paper is usually left out of curricula.

Going some years forward, we arrive in 1916 at Einstein's derivation of the Planck formula in terms of atomic transitions [15] (also discussed by him in [16,17]). This result actually sometimes does find its place in courses. However, the emphasis is nearly always on the process of stimulated emission, which is of course a very important aspect to be considered, in view of the fact that it was introduced for the first time in this paper, and that it is the theoretical basis of modern lasers. On the conceptual level, the truly revolutionary aspect of that paper was the recognition of the intrinsically probabilistic and causality violating nature of the process of spontaneous emission. This was in fact the first appearance of what can surely be considered one of the most puzzling, misconceived and characteristic aspects of quantum physics, that is, the fact that quantum processes are intrinsically probabilistic⁴.

In the literature, there are actually several references which do discuss one or more of the aspects we listed at the undergraduate or graduate level, for example [18–24], even if the practical necessity of going straight to full quantum mechanics and its applications does not leave much time for exploring them. Typical high school books, on the other hand, completely neglect these developments. It is our opinion instead that also high school students should benefit from their discussion, if of course they are presented in a suitably simple way. The same is true for teachers, with the advantage of their wider mathematical and physical background, which allows them to appreciate deeper and more complete treatments.

Most of Einstein's papers are uniformly models of deep and clear physical thinking, and they are also full of marvellous insight and philosophical discussions. Unlike many papers by contemporaries, these works turn out to be very readable also for modern readers. This is famously true for Einstein's first paper about special relativity, dating back to 1905, which is still quite useful to modern students. The same can be said about the above cited pivotal papers about light quanta and quantum theory in general. Einstein's papers can therefore be a wonderfully stimulating reading for both teachers and, in a suitable selection, students, who can then be enriched also from the general cultural point of view. Luckily, they are also easily accessible: all of Einstein's writings (up to May 1927 at the moment) are freely available in English translation on the web [25], and translations in many other languages are often republished. Among the enormous scholarly literature on Einstein's work, several excellent expositions of Einstein's ideas on quantum physics are available. For example, concerning the aspects we discuss, we mention [26,27].

Inspired by the above considerations, we have extracted a didactic path to quantum physics from the papers by Einstein [11,12,15], which all deal with the quantum theory of radiation. This topic is therefore the Ariadne thread of the path. This path could fruitfully supplement and complement a standard history-based introduction to the subject. The path is divided in three parts, each of which focusing on one of the papers by Einstein cited above. In the present paper we describe the path in detail, by considerably extending the outline we gave in [28].

In each part, we have tried to stay as close as possible to Einstein's original reasoning, while at the same time simplifying the mathematics whenever possible, and paying attention to not losing physical insight. Rigorous derivations that would be too demanding for

the high school level have been substituted by heuristic, intuitive arguments. Of course such a task requires that emphasis be put on clarity rather than on mathematical rigor. Moreover, we have used modern notation. These choices should make the presented topics and derivations affordable by students in the second half of their last high school year (which is when they are typically exposed to quantum physics anyway), which have a background in classical physics, including elementary kinetic theory and electromagnetic waves⁵, and have already been exposed to elementary differential and integral calculus. This way of presenting the material should also be useful for teachers (although in their case one could well settle at a higher level, and in some case we give suggestions for more advanced topics which can be introduced), since it puts it in a form that can then be proposed by them in class.

Since basically all of Einstein's fundamental ideas on light quanta came from statistical considerations, we also included an introduction to the necessary tools of statistical physics, again trying to explain them in as simple a way as possible, starting from the elementary notions of kinetic theory that students can be assumed to possess. In particular, we introduce Boltzmann's entropy and its computation for the ideal gas case, which is necessary for understanding part 1, and after that we discuss the Maxwell-Boltzmann distribution, which is needed in part 3, and the Gibbs distribution together with its application to the computation of statistical fluctuations, which are used in part 2. Einstein was a master of statistical physics, and he gave fundamental contributions to it besides using it to uncover quantum mysteries. Both in his papers on statistical and quantum physics it is customary to find very clear explanations of statistical tools. Therefore we naturally took some of these as inspirations for this part, when possible.

The statistical physics part could as well stand on its own to constitute the core of a proposal to teach the elements of this subject to high school students. In fact, the teaching of basic statistical physics is itself a very active area of research, with many different proposals (see e.g., [29–33] for a sample; of these, the last one is especially suitable for high school students).

Parts of this program are being currently tested during various outreach initiatives aimed at introducing students and teachers to 20th century physics. We still do not have quantitative results concerning the effectiveness of this approach (this will be the object of a forthcoming publication). However, very encouraging preliminary results were obtained, which prompted us to publish the general idea of our approach.

The paper is organized as follows: in Section 2 we give a summary of Planck's law, limiting the choice of topics to what is needed in the following. In Section 3 we describe the statistical reasoning which led Einstein to the light quanta hypothesis, and discuss the application to the Stokes rule of fluorescence. In Section 4, we discuss the statistical computation which led Einstein to conceive wave–particle duality for light. In Section 5, after a lightning resume of the postulates of Bohr's theory, Einstein's discussion of the Planck's distribution in terms of atomic processes is described. In Section 6 we develop the needed tools of statistical physics, starting from Boltzmann's formula for entropy and its application to the ideal gas, and then going on to statistical distributions and energy fluctuations. In Section 7 we describe how we are implementing our program, and the preliminary results we got thus far. In Section 8 we give some suggestions about some ways in which our program can be used and complemented.

2. Setting the Stage. Cavity Radiation and Planck's Law

As most treatments on quantum physics, ours begins with a discussion of radiation inside a cavity at thermal equilibrium, which as well known is an effective model of a black body, that is, an object which can absorb electromagnetic radiation of any wavelength with perfect efficiency. A black body is a remarkably good model to describe the electromagnetic radiation emitted by a body in thermal equilibrium with its surroundings. This topic can be treated in a standard way, and in this section we limit ourselves to highlighting the main points which must be touched in order to understand the following discussions.

Black body radiation has a well known experimental spectrum, which is of course described by Planck's law

$$u(\nu, T) = \frac{8\pi h \nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1}, \quad (1)$$

where c is the light velocity, k_B is the Boltzmann constant and h is the Planck constant. The quantity $u(\nu, T)d\nu$ is the energy density of radiation⁷ of frequency ν (or, more rigorously, with frequency included in the very narrow interval between ν and $\nu + d\nu$) inside the cavity, which is in equilibrium at temperature T . This law can be introduced as a phenomenological function which fits experimental data on thermal radiation⁸.

Some discussion about the attempt to explain the above law by using classical physics should be included. In particular, a heuristic discussion about how principle of equipartition of energy (which students should know from elementary kinetic theory), when applied to electromagnetic waves in a cavity, does not reproduce Planck's law, but rather the Rayleigh-Jeans law⁹

$$u(\nu, T)d\nu = \frac{8\pi \nu^2}{c^3} k_B T d\nu, \quad (2)$$

should be given. The factor $Z(\nu) = \frac{8\pi \nu^2}{c^3} d\nu$ can be intuitively introduced as the "number of electromagnetic waves" of frequency ν per unit volume, with $\bar{\varepsilon} = k_B T$ being the average thermal energy associated with each of them, hence equipartition tells us that the average energy density per unit volume is given by (2). Intuitively, the "ultraviolet catastrophe" can be understood by observing that the numbers of higher and higher frequency waves grows unboundedly, so equipartition assigns them exceedingly large amounts of energy. This will allow a heuristic explanation of why energy quantization, that is, the hypothesis that energy can be exchanged between radiation and the walls of the cavity only in multiples of the quantity $\varepsilon = h\nu$, affects the principle of equipartition of energy and allows us to obtain Planck's formula. A very nice discussion can be found in [34].

For what follows, it is vital to point out that the Planck distribution is essentially a function of the ratio $\frac{h\nu}{k_B T}$, that is of the energy quantum over the average thermal energy per degree of freedom. Energy quanta associated with frequency low enough to ensure that $\frac{h\nu}{k_B T} \ll 1$ (the temperature, which sets the scale of what is meant by "low enough", is considered fixed), therefore, are much smaller than the average thermal energy, hence the effect of energy quantization should be negligible. Indeed, we can use the fact that¹⁰, for small x , $e^x \approx 1 + x$, to show the distribution (1) reduces to the Rayleigh-Jeans law (2), which was obtained by a reasoning which ignored energy quanta. Notably, Planck's constant h is erased in the process. For this reason, this situation may be dubbed the "classical limit".

On the other hand, energy quanta associated with a frequency high enough that the condition $\frac{h\nu}{k_B T} \gg 1$ is achieved, are much bigger than average thermal energy. In this case the effects of energy quantization are expected to be maximally evident, so that this limit can be christened the "extreme quantum limit". In this limit the exponential in the denominator becomes much bigger than the -1 , which can be neglected, and Planck's law reduces to the so-called Wien's law¹¹:

$$u(\nu, T) = \frac{8\pi h}{c^3} \nu^3 e^{-\frac{h\nu}{k_B T}}. \quad (3)$$

In fact, as we shall see in the next section, it was by studying Wien's distribution law that Einstein was led to the hypothesis of light quanta. After this discussion, which shows that this law is valid precisely when quantum effects are most evident, this should feel a bit less surprising.

3. Part 1. Einstein 1905: Light Quanta

In his 1905 paper [11], Einstein computed in an ingenious way the thermodynamic entropy associated with thermal radiation described by Wien's distribution (3), and combined the result with the Boltzmann principle in order to understand its statistical origin. The computation involves ordinary thermodynamics and the integration of a logarithmic function, so it can be followed by students.

3.1. Entropy of Thermal Radiation

Let us start from the first principle of thermodynamics

$$TdS = dE + p dV, \quad (4)$$

which implies that, when the volume is constant, that is, $dV = 0$,

$$\frac{dS}{dE} = \frac{1}{T}. \quad (5)$$

As announced, we are going to apply this formula to thermal radiation in a cavity, in the extreme quantum limit, which is described by Wien's law. Since the radiation components of each frequency are independent by the superposition principle, we restrict to a very narrow interval of frequencies $[\nu, \nu + d\nu]$, that is, we consider monochromatic radiation. What we shall say will hold for any value of the frequency ν (such of course that the condition $\frac{h\nu}{k_B T} \gg 1$ remains valid). Since we are at equilibrium, the energy of the radiation in the cavity is uniformly distributed over the volume, so we can express the total energy contained in the radiation in the considered interval of frequency as:

$$E = V u d\nu, \quad (6)$$

where u is given by Wien's law (3). We express the entropy in an analogous way:

$$S = V \varphi d\nu, \quad (7)$$

where $\varphi d\nu$ is the entropy density of radiation in the chosen frequency interval. Substituting (6) and (7) into (5) we get

$$\frac{d\varphi}{du} = \frac{1}{T}. \quad (8)$$

From this equation, we can compute φ . Inverting Wien's law with respect to $1/T$, we have

$$\frac{1}{T} = -\frac{k_B}{h\nu} \ln \frac{u c^3}{8\pi h\nu^3} = \frac{d\varphi}{du}, \quad (9)$$

so that

$$\varphi(u, \nu) = -\frac{k_B}{h\nu} \int \ln \left(\frac{c^3 u}{8\pi h\nu^3} \right) du. \quad (10)$$

This integral can be easily done by putting $\frac{c^3}{8\pi h\nu^3} u = x$ and integrating by parts:

$$\varphi(u, \nu) = -\frac{k_B}{h\nu} u \left(\ln \frac{c^3 u}{8\pi h\nu^3} - 1 \right) + C, \quad (11)$$

where C is an integration constant. This expression, recalling Equation (7), allows us to write down the entropy of the radiation in the cavity as a function of the volume:

$$S(V) = V \varphi dv = -\frac{k_B}{h\nu} V u dv \ln \left(\frac{c^3 E}{8\pi h\nu^3 V dv} - 1 \right) + C', \quad (12)$$

where $C' = VCdv$. Our aim is to compute the entropy variation under an isothermal and adiabatic transformation. Of course the calculation can be repeated for a cavity with different volume V_0 , with the same E and T . The variation of entropy between these two configurations is in fact given by a much simpler expression (in particular, the integration constant C' is the same in both cases, so it is erased):

$$S(V) - S(V_0) = -\frac{k_B E}{h\nu} \ln \frac{V_0}{V} = \frac{k_B E}{h\nu} \ln \frac{V}{V_0} = k_B \ln \left(\frac{V}{V_0} \right)^{\frac{E}{h\nu}}. \quad (13)$$

Strikingly, this has exactly the same form of the Boltzmann entropy variation for an analogous transformation of an ideal gas, that is, the Joule free expansion, which is given by: $S(V) - S(V_0) = k_B \ln \left(\frac{V}{V_0} \right)^N$ (this equation is derived in Section 6.1 from the Boltzmann principle), provided we identify

$$\frac{E}{h\nu} = N \quad \text{or} \quad E = Nh\nu. \quad (14)$$

This identification has been boldly interpreted by Einstein by saying that the radiation behaves statistically as if it were made up of N independent “quanta”, each of which has energy given by Planck’s expression $\epsilon = h\nu$, so that the total energy is the sum of the energies of the individual quanta (which are independent just like the molecules of an ideal gas).

As stated in the introduction, is important to emphasize that this statement is much more revolutionary than Planck’s hypothesis, according to which only the exchange of energy between the radiation and the walls of the cavity are discrete, but the radiation propagates as waves as usual between emission and absorption. Here instead it is argued that radiation itself is quantized, at least in the limit where quantum effects dominate and Wien’s distribution law is valid. This hypothesis meant a complete divorce with Maxwell’s very successful theory of electromagnetism and light, and a reconciliation would have begun only later with the concept of wave–particle duality (see next section), which showed that far from the Wien limit radiation is actually more complex than an ideal gas of independent quanta.

It may be also appreciated that in this work Einstein used the Boltzmann entropy in a peculiar and illuminating way [8], namely exploiting the expression of entropy coming from thermodynamics to infer the statistical behavior of radiation, rather than the opposite as was customary at the time. The statistical behavior that came out is just the same as that of a bunch of molecules.

3.2. Some Applications of Light Quanta

The argument of the preceding section shows that cavity radiation behaves from a thermodynamical point of view as an ideal gas of light quanta. There are however more direct ways of confirming this conclusion, based on experimental evidence. These applications actually go one step further because they assume that not only the statistical behavior of the radiation is given by its corpuscular nature, but that the physical effects of individual quanta can be measured. Hence they show that quantization of radiation is a fundamental feature of Nature. It is to be stressed that these applications could not have been imagined thinking about Planck’s theory, since, once again, in that theory radiation is not quantized. In his paper [11], Einstein considered three applications, one of which is the famous one to the photoelectric effect, which as we already remarked must be included in

the track; here we omit it, as it can be treated in a standard way (e.g., as in [3]). However, Einstein gave two more applications in his paper [11], which are never mentioned. These are ionization by UV quanta and the explanation of Stokes' rule of fluorescence. We have found that the second one in particular can be useful and instructive. Such rule, in fact, states simply that, when a fluorescent material absorbs light, it re-emits it with a lower frequency. A well-known effect is the violet glow of white clothes which are exposed to ultraviolet light, which is a phenomenon students have likely experienced (if they like going to the disco for example). This rule is readily understandable in terms of light quanta. Namely, a fluorescent material can absorb a light quantum, whose energy is $\varepsilon = h\nu$, and emit it back. Since part of the energy is absorbed, the light quantum is emitted back with a lower energy $\varepsilon' < \varepsilon$. However, $\varepsilon' = h\nu'$, hence $\nu' < \nu$, which is precisely Stokes' rule. Thus, the latter comes as an immediate consequence of the proportionality of energy and frequency of light quanta. The last application, ionization by UV light quanta, can be considered as well. In fact, the understanding of light as being constituted by light quanta can give student an intuitive grasp of why radiation of short wavelength and high frequency, such as UV light (and even more, X-rays and γ -rays), can be dangerous for the health. Indeed, one speaks of ionizing radiation. The ionizing power of such radiation is unexplained by the wave picture, while it is promptly understood by realizing that the energy of high frequency radiation is distributed in a small number of "bullets", which therefore have enough energy to tear electrons away from atoms.

4. Part 2. Einstein 1909: Wave-Particle Duality for Light

At this point an important observation can be done. We have seen that the Planck distribution reduces, in the classical limit, to the Rayleigh-Jeans law, which can be obtained by applying the equipartition theorem to classical electromagnetic waves, while in the "most quantum limit" it reduces to Wien's law, which as showed above leads to considering radiation as an ideal gas of non-interacting light quanta. The full Planck distribution contains both these limits, and this means that, while for extreme values of the frequencies either a purely wave or a purely corpuscular description of radiation can be used, for intermediate values both aspects are relevant. In 1909 Einstein, in a brilliant paper [12], gave a quantitative foundation to this observation. In fact, he again applied his mastery of statistical physics to thermal radiation, this time to the full Planck distribution (1). This time, instead of entropy, he computed quadratic fluctuations of the energy of thermal radiation. This computation can be reproduced by using elementary statistical physics methods, namely the Gibbs distribution, which implies that quadratic energy fluctuations in a system at thermodynamic equilibrium are given by the equation:

$$\overline{(\Delta E)^2} = k_B T^2 \frac{d\bar{E}}{dT}, \quad (15)$$

where \bar{E} is the average internal energy of the system. We introduce the Gibbs distribution and prove the above equation in Section 6.3. In this case, the average energy of radiation of frequency ν is given by

$$\bar{E} = V u(\nu, T) d\nu = V \frac{8\pi h \nu^3}{c^3} \frac{1}{e^{\frac{h\nu}{k_B T}} - 1} d\nu, \quad (16)$$

where $u(\nu, T)$ is Planck's distribution. Substituting this in (15) gives, after a straightforward computation:

$$\overline{(\Delta E)^2} = \left[h\nu u + \frac{c^3}{8\pi} \frac{u^2}{\nu^2} \right] V d\nu, \quad (17)$$

which was Einstein's main result in [12]. We see that (17) is the sum of two terms. It is apparent that the second term is the dominating one for low frequencies. This is the only

one we would have got if the Planck distribution we had used the Rayleigh-Jeans one. In fact, the form of this term is consistent with thermal radiation being made up of waves: if we have a bunch of randomly superimposed waves in a cavity, at a certain point a wave of a given frequency can interfere with a wave of slightly different frequency, creating beats which in turn cause energy fluctuations. Thus, in such a picture fluctuations are given by interference effects. This leads to their being quadratic in the energy density. Intuitively, the latter fact can be inferred by the fact that the energy of a wave is proportional to the square of its amplitude, so a fluctuation which doubles the amplitude makes the energy four times bigger. This was originally argued by Einstein in [12] using dimensional analysis, while explicit computations can be found in [20,22,23].

Concerning the first term, it is the leading one on high frequencies, where as we saw the quantum aspect dominates. In fact, had we used the Wien distribution in place of the Planck one, we would have only had this term. Consistently, it describes the fluctuations of an ideal gas of corpuscles, all with the same energy $\epsilon = h\nu$. This can be seen by observing that in such a gas there will be on average $\bar{N} = \frac{\bar{E}}{h\nu}$ corpuscles, so the quadratic energy fluctuations can also be computed by $(\Delta E)^2 = \epsilon^2 \bar{N} = (h\nu)^2 \bar{N} = h\nu \bar{E} = h\nu V d\nu$, which is just the first term of (17)¹².

Equation (17) thus shows that the energy fluctuations for the full Planck distribution involve both a particle contribution and a wave contribution, and that both are equally important, apart from special limits in which one or the other dominates.

In this way, Einstein found that the statistical properties of thermal radiation are a mixture of those expected from a wave theory and those expected from a particle theory, with both contributions equally important at not too low or not too high frequencies. He was led to suggest that this was leading to

a theory of light which can be interpreted as a kind of fusion of the wave and the emission [that is, the corpuscular] theory

and he stated that

the wave structure and the quantum structure [...] are not to be considered mutually incompatible.

This was the first time wave–particle duality was hypothesized, even if only for light. This idea is particularly striking, since in classical physics the concepts of wave and of particle seem to be mutually exclusive. In 1925 Einstein extended this argument to matter particles [14], thus putting light and matter on the same footing.

Further elaboration allows us to introduce at this stage a fundamental principle of quantum mechanics, namely the principle of correspondence, which is usually introduced in connection with the Bohr atom (introducing it in that context would in fact be more adherent to history, but we allowed ourselves to take licenses from history for didactic purposes). In fact, light quanta are in a sense “smaller” if they are associated with radiation of lower frequency, which means that radiation of very low frequency contains a lot of them. Thus, in the expression $E = N h \nu$ for the energy, the number N , which may be considered a first instance of a quantum number, assumes a very large value in this limit. This is thus a situation in which a classical picture (the wave one) emerges when a quantum number becomes large, which is precisely the content of the correspondence principle.

It is to be stressed that, while the argument presented in this section is statistical in nature, the concept of wave–particle duality actually applies to the individual elementary constituents of radiation. In fact, radiation can be regarded as being made of light quanta for all frequencies, and these light quanta which have a double nature, of both waves and particles. In fact, a frequency is associated with light quanta in all regimes. In particular, in the Wien limit, such quanta have high energy, hence they are not many, and they are very far away from each other; furthermore, they are associated with waves whose frequency is very high, and hence of very short wavelength. Hence the wave nature is not evident, consistently with the fact that radiation behaves as an ideal gas of light quanta. On the other hand, as stressed in the preceding paragraph, in the classical limit the number of

individual light quanta is much bigger, so they are on average much closer to each other. Moreover, their associated waves have long wavelength, so they overlap and interfere with each other, and classical electromagnetic waves emerge. The same double picture holds for material particles such as electrons, in which case one speaks of matter waves. This picture is still quite heuristic though, since in the full theory of quantum mechanics, the famous Born rule interprets these waves as probability waves, which describe the probability for the associated quantum to be in a certain position, while clearly wavelike behavior manifests itself only when many quanta are involved, as it happens in classical electromagnetic waves, or in electron beam diffraction experiments. Moreover, direct measurements can never highlight wave behavior and particle behavior at the same time (this is an instance of Bohr's principle of complementarity).

5. Part 3. Einstein 1916: Probability

In another breakthrough paper [15], written in 1916, Einstein considered black body radiation again, this time using concepts introduced by Bohr in his theory of atomic structure, which had been introduced in 1913. In this paper, Einstein did not start from the spectral distribution of thermal radiation, rather he investigated how a bunch of atoms interacting with radiation, could be in thermal equilibrium with the radiation itself. In this way he showed that equilibrium is achieved precisely when radiation obeys Planck's law. He thus had found, to use his words "an amazingly simple derivation of Planck's formula, I should say *the* derivation (Letter to M. Besso, 11 August 1916, emphasis in original)".

In what follows, we shall not strictly follow Einstein's treatment, rather we shall adapt it to our purposes. In fact, our aim is not to derive Planck's formula (although Einstein's derivation is in fact likely the simplest one in the literature, and could be affordable by high school students), but rather to use it in order to gain information on the nature of the elementary processes by which atoms interact with radiation.

Of course, Bohr's theory is included in any course of quantum theory, and it can actually be treated in a standard way (as e.g., in [3]). For what follows all we have to recall is that, according to Bohr, electrons in atoms and molecules can only occupy one of a set of stationary states with discrete energies ε_1 , ε_2 , and so forth, and transitions between these states can occur (these are the famous quantum leaps). According to the Bohr theory postulates, in the transition between a state with energy ε_m and another state with energy $\varepsilon_n < \varepsilon_m$, the electron emits radiation with frequency given by the equation

$$\varepsilon_m - \varepsilon_n = h\nu, \quad (18)$$

which represents the conservation of energy in the process. The inverse transition can be induced if the electron absorbs radiation with the same frequency. This allowed Bohr to explain qualitatively (quantitatively in the case of one electron atoms) line spectra of atoms and molecules.

As said, Einstein considered an ensemble of atoms or molecules, which continuously absorb or emit light quanta. He investigated in much greater detail than Bohr the elementary processes of atom-light interactions, singling out three of them, namely absorption of a light quantum, stimulated emission (in which a light quantum hits an atom, stimulating its decay to a lower energy state with emission of another quantum), and spontaneous emission (in which an atom spontaneously decays to a state with lower energy emitting a quantum). The process of stimulated emission was introduced for the first time by Einstein in this work, and it is at the basis of modern lasers. As we shall see, it is crucial to include it in order to account for the whole Planck distribution. Since the system is assumed to be at thermal equilibrium, the various states of the atoms will be distributed according to the Maxwell-Boltzmann distribution, and moreover this distribution should not change with time, meaning that on average each transition from a state to another must happen with the same rate as the inverse transition. This translates in a detailed balance equation, from which the Planck distribution for the energy density of radiation follows. An interesting historical fact underlines the revolutionary nature of these developments. When Bohr

developed his atomic theory, he did not believe in light quanta; for him, in a transition, an atom emitted a bunch of classical monochromatic electromagnetic waves. Einstein, on the other hand, assumed that in any atomic transition, radiation is emitted or absorbed as a single light quantum, whose energy matches the energy difference between the two states involved in the transition, and he actually gave substantial evidence for that in the second part of his paper (which however is too advanced to be considered here)¹³.

Let consider the three processes in more detail, starting with the process of absorption. When a light quantum of frequency ν hits the molecule, if it has the right energy, it can trigger a transition from a state Z_n to a higher energy state Z_m . Then the light quantum yields all its energy to the atom, disappearing; in other words, it is absorbed by the atom. We assume that the transition happens with a certain probability, which is proportional to the spectral energy density of radiation $u(\nu, T)$ (such a process can be expected to occur more often when there are more light quanta around). Specifically, the probability that the transition happens in the time δt is given by:

$$\delta W_{n \rightarrow m}^a = B_{mn} u \delta t. \quad (19)$$

Stimulated emission occurs when a light quantum hitting a molecule in the state Z_m , triggers a transition to a lower energy state Z_n , with the emission of a second light quantum. Since also this process relies on light quanta hitting the atom, the probability for it to occur in the time δt is again proportional to the energy density:

$$\delta W_{m \rightarrow n}^{\text{st}} = B_{nm} u \delta t, \quad (20)$$

where again B_{mn} is a constant.

Finally, spontaneous emission occurs when a molecule is in an excited state Z_m . Then it will decay to a lower energy state Z_n spontaneously, that is, without an external stimulation, with the emission of a light quantum. Since this process does not depend on the presence of external light quanta, the probability for it to occur in the time δt is not proportional to the energy density:

$$\delta W_{m \rightarrow n}^{\text{sp}} = A_{nm} \delta t. \quad (21)$$

In Equations (19)–(21), the quantities B_{mn} , B_{nm} and A_{nm} are constants, depending on the two states involved, which include the detailed information about the molecule.

As announced, let us now consider a bunch of atoms or molecules in thermal equilibrium with radiation, continuously emitting and absorbing radiation by means of the above described elementary processes. In order to avoid unnecessary complications, we shall assume that only one species of atom or molecule is present.

The hypothesis of thermal equilibrium means two things. First, the states of the molecules will be distributed according to an exponential distribution, that is, the probability for the atom to be in a stationary state with energy ε_n will be given by (this is justified in Sections 6.2 and 6.3):

$$p(\varepsilon_n) = A p_n e^{-\varepsilon_n / k_B T}, \quad (22)$$

where A is an inessential normalization constant and p_n is the number of stationary states which have energy ε_n , which is called the degeneracy of the energy level. In the following, for simplicity, we shall assume that energy levels are not degenerate, that is, that $p_n = 1$ for any n . The inclusion of the degeneracies can be achieved in a straightforward manner, but it would unnecessarily complicate the algebra. Second, being an equilibrium distribution, (22) should not change with time, that is, it should not be altered despite the atoms change state all the time. The latter condition means that any transition must

occur with the same rate as the inverse transition, which in turns requires the validity of a detailed balance equation. This condition is written as:

$$e^{-\varepsilon_n/k_B T} \delta W_{n \rightarrow m}^a = e^{-\varepsilon_m/k_B T} (\delta W_{m \rightarrow n}^{\text{st}} + \delta W_{m \rightarrow n}^{\text{sp}}), \quad (23)$$

or, more explicitly,

$$e^{-\varepsilon_n/k_B T} B_{mn} u(\nu, T) = e^{-\varepsilon_m/k_B T} (B_{nm} u(\nu, T) + A_{nm}). \quad (24)$$

This relation must hold at any temperature T (as long as this is not so high as to ionize the molecules, which anyway happens at a very high temperature). Solving Equation (24) with respect to $u(\nu, T)$ we obtain

$$u(\nu, T) = \frac{A_{nm}}{B_{mn} e^{(\varepsilon_n - \varepsilon_m)/k_B T} - B_{nm}} \quad (25)$$

From Bohr's postulate we have $\varepsilon_n - \varepsilon_m = h\nu$. We thus see that the above expression looks like the Planck distribution (1). The latter is thus reproduced only if the A and B coefficients satisfy the following relations:

$$B_{mn} = B_{nm}, \quad (26)$$

and

$$\frac{A_{nm}}{B_{nm}} = \frac{8\pi h \nu^3}{c^3} \quad \text{or} \quad A_{nm} = \frac{8\pi h \nu^3}{c^3} B_{nm}. \quad (27)$$

Thus, by studying atoms in thermal equilibrium with radiation in terms of elementary atomic processes, and by requiring that the result is in agreement with experiments (which happens if the energy density is described by Planck's law), we get much information about the atomic processes themselves, through the relations (26) and (27)¹⁴. The first tells us that the probabilities per unit time of the processes of absorption and stimulated emission are equal. These processes are thus to be considered the inverse of each other, and the equality of the probabilities signals their reversibility. The second relation expresses the spontaneous emission coefficients as the product of the stimulated emission coefficient associated with the same transition, multiplied by the factor $\frac{8\pi h \nu^3}{c^3}$ which, as we recall, describes the "number of waves" of frequency ν in the cavity (per unit volume). Intuitively, this factor may be regarded as counting how many possibilities there are for emitting a light quantum, or in other words how many possible states there are for the emitted light quantum. The larger this number, the more probable is spontaneous emission, since there are more possibilities for the light quantum to be emitted. The factor $\frac{8\pi h \nu^3}{c^3}$ evidently grows with the energy $h\nu$ of the emitted photon¹⁵, meaning that spontaneous emission dominates over stimulated emission in the high frequency limit. Since the temperature is fixed, this is precisely what we dubbed the "extreme quantum limit". A confirmation of this comes from the observation that, if we assume that there is no stimulated emission, Equation (24) reduces to:

$$e^{-\varepsilon_n/k_B T} B_{mn} u(\nu, T) = e^{-\varepsilon_m/k_B T} A_{nm}, \quad (28)$$

which, upon solving with respect to $u(\nu, T)$, and using the Bohr postulate has the same form of Wien's distribution law (3), which is fully reproduced if $A_{nm} = \frac{8\pi h \nu^3}{c^3} B_{mn}$. This relation replaces (26) and (27) in this regime.

On the other hand, since for any frequency $u(\nu, T)$ grows with temperature, for T high enough the spontaneous emission term will be negligible compared with the others (in other words, there are so many light quanta around that stimulated emission occurs much more frequently than spontaneous emission). This is precisely the regime where the

Rayleigh-Jeans law (2) holds. In fact, the limit of large T with fixed frequency corresponds to the regime in which the classical limit is valid.

The last two observations tell us that the extreme quantum regime and the classical regime are dominated respectively by spontaneous emission and stimulated emission, which in turn means that the quantum features are captured by spontaneous emission. In fact, the process of spontaneous emission has a fundamental difference with respect to the other two, since it is not triggered by the interaction with a light quantum. Since there is no trigger, there is no way of predicting when spontaneous emission will take place. The process seems to have no apparent cause, and it is necessary to assume that it occurs, in a given interval of time, with a given probability, which is all we can know or predict about it¹⁶. Since the radiation is emitted in a single quantum, this quantum has to be emitted in a given direction. The direction in which the emerging light quantum is emitted is unpredictable as well, that is, it ruled by a probabilistic law (which is simply a uniform distribution since all directions are equiprobable). Thus, spontaneous emission is an intrinsically probabilistic process, which seems to violate the principle of causality. This behavior is actually analogous to that of radioactive decay of a nucleus, and in fact (21) was introduced in analogy with the probability law for a radioactive decay. The upshot of all this is that the full Planck distribution can only be reproduced if intrinsically probabilistic and causality violating processes occur, and moreover these processes capture the “most quantum” part of it, while they are suppressed in the classical limit. This means that the quantum behavior, which, as we know very well by now, is encoded in the Planck distribution, is inextricably linked with probability. This fact, which emerged in Einstein’s work for the first time¹⁷, then became a pillar of the complete theory of quantum mechanics, in fact constituting one of its most characteristic and mind-blowing features. It is to be emphasized that this probabilistic aspect is very different from the one which characterizes statistical physics (see next section). The latter, in fact, is due to the ignorance of the observer of the underlying microscopic dynamics, and hence it is not intrinsic to the latter. On the other hand, the probabilistic aspect of spontaneous emission is considered to be fundamental, that is, intrinsic to the microscopic dynamics. It is an unavoidable characteristic of the elementary atomic processes. Since a very common misconception is that probability in quantum mechanics is due to instrumental limitations, just like probability in statistical physics, this aspect has to be properly emphasized in the course.

6. Tools from Statistical Physics

Since basically all of Einstein’s fundamental ideas on light quanta came from statistical physics, it is desirable to enhance our path by introducing the necessary tools of this subject, starting from the elementary notions of kinetic theory that are part of the standard curriculum. It is therefore the aim of this section to review, in a pedagogical way, some tools of elementary statistical physics which are needed for discussing Einstein’s papers. As stated in the introduction, also for these topics we take inspirations from Einstein’s papers themselves. This part can be considered on its own as part of an introduction to elementary statistical physics to high school students, who are already familiar with elementary kinetic theory of gases as it is usually taught at school, and in fact the first half of it was used in this way by us. The level of this section is comparable to that of the rest of the paper.

We start by introducing Boltzmann’s postulate for the entropy (which is also part of the standard high school curriculum) and by using it for the computation of the entropy variation of an ideal gas undergoing free expansion. This result is of course needed in Part 1 of the track for comparison with the entropy of radiation. For this topic to be followed, some basics of probability theory, which should be part of the toolbox of any last year pupil, are needed. For Parts 2 and 3 a slightly more sophisticated tool is needed, namely the probability distribution of the energies of states of a system in thermal equilibrium, and its application to computing averages and fluctuations. This is introduced starting from the Maxwell distribution of velocities of the molecules in an ideal gas at equilibrium, which again is part of standard curricula. This tool is used in Part 2 to compute the quadratic

fluctuations of energy and in Part 3 to describe the equilibrium distribution of atomic states. Some familiarity with the use of probability distributions, which again should be part of the standard mathematics curriculum, is desirable.

6.1. Boltzmann's Formula and the Entropy of an Ideal Gas

In kinetic theory and statistical physics, an ideal gas is modelled as a set of N non-interacting molecules¹⁸, which for a monatomic gas are considered Newtonian point particles. It is of course not our aim to dwell on the various conceptual subtleties involved in this topic. Instead of computing the entropy of the ideal gas in a given state (the so-called Sackur-Tetrode entropy), which requires quite sophisticated tools, we shall actually limit ourselves to computing the entropy variation under an adiabatic and isothermal expansion, since this is all we need for dealing with light quanta. We shall give an intuitive, yet quantitative treatment, which is an elaboration of that presented by Einstein in 1905 (cf. [11], Section 5).

One of the basic postulates of statistical physics (called Boltzmann's principle) says that the entropy of a given thermodynamic configuration M of a thermodynamic system, which for us is our ideal gas, is given by

$$S(M) = k_B \ln W(M) + C, \quad (29)$$

where C is a constant which depends on the system, and the quantity $W(M)$ (which may be called multiplicity) is the number of possible configurations of the gas molecules (which are called microstates) which build up the thermodynamic state M (which is called macrostate) in question. Another fundamental principle of statistical physics is that all microstates are equiprobable.

The quantity $W(M)$ can be intuitively considered a measure of the disorder of the state, in the sense that a disordered state can be made up in more ways, therefore more information is hidden to the macroscopic observer. The precise definition of $W(M)$ is somewhat subtle, but this is not actually a problem for us, for the following reason. In view of the principle of equiprobability of microstates, we can define the probability of the macrostate M by using the naive frequentist definition

$$P(M) = \frac{W(M)}{\text{total number of microstates}}. \quad (30)$$

The "total number of microstates" is actually a subtle quantity whose definition is highly nontrivial. However, for what will concern us here, it will cancel out without giving any problem (the same considerations apply to the constant C in (29), by the way).

As is known from thermodynamics, an equilibrium state of the ideal gas is specified by the variables p , V and T . However, just two of them are independent as a consequence of the equation of state $pV = nRT$, where R is the universal constant of ideal gases. If we choose to use V and T , then Boltzmann's principle allows us to write down the entropy associated with the thermodynamic state (V, T) as:

$$S(V, T) = k_B \ln W(V, T) + C. \quad (31)$$

Accordingly, the quantity $W(V, T)$ represents the number of configurations of the N molecules compatible with the fact that a macroscopic observer measures a volume V and a temperature T for the gas.

Let us now consider the following situation: an ideal gas is initially in equilibrium at temperature T in a volume V_i . We denote this macrostate as $M_i = (V_i, T)$. Then, while being isolated from the outside world, the gas is allowed to freely expand to a volume $V_f > V_i$. This is known as Joule expansion, and the thermodynamic of the ideal gas tells us that this is an adiabatic and isothermal process. Therefore, when the gas reaches

equilibrium again, it will have the same temperature T . This means that the final macrostate will be $M_f = (V_f, T)$. The variation of entropy in the process is thus given by:

$$\Delta S = S(V_f, T) - S(V_i, T) = k_B \ln \frac{W(V_f, T)}{W(V_i, T)} = k_B \ln \frac{P(V_f, T)}{P(V_i, T)}. \quad (32)$$

We see that the troublesome “total number of microstates” cancels out when computing ratios of the W 's, which are therefore the same as ratios of the probabilities. The constant C cancels out as well. This is the main advantage in computing an entropy variation rather than an entropy. Now ratios of probabilities are computed straightforwardly by an elementary symmetry argument, based on the postulate that all microstates are equiprobable¹⁹. Suppose we only have one molecule. Since the molecule a priori has the same probability of being in any point of space, the ratio of the probabilities that it be inside two volumes is the same as the ratio of the volumes themselves

$$\frac{P_{1 \text{ molecule}}(V_f, T)}{P_{1 \text{ molecule}}(V_i, T)} = \frac{V_f}{V_i}. \quad (33)$$

This is true for any molecule of the gas. Since the gas is ideal, so by hypothesis there is no interaction between the molecules, all the molecules are independent. Therefore we have, for N molecules

$$\frac{P(V_f, T)}{P(V_i, T)} = \left(\frac{V_f}{V_i} \right)^N, \quad (34)$$

by the law of multiplication of probabilities of independent events. Thus

$$\Delta S = k_B \ln \frac{P(V_f, T)}{P(V_i, T)} = k_B \ln \left(\frac{V_f}{V_i} \right)^N = N k_B \ln \frac{V_f}{V_i}. \quad (35)$$

If we recall from kinetic theory that $N = nN_A$ and $N_A k_B = R$, where N_A is the Avogadro number, we see that we have recovered the well known result from thermodynamics that $\Delta S = nR \ln \frac{V_f}{V_i}$ for the Joule expansion. The intuitive meaning of this formula is that if the volume occupied by the gas increases, the disorder grows, since there are more ways of putting N molecules in a bigger volume than in a smaller one, therefore the entropy variation is positive. Viceversa, if the volume decreases (we can think of a configuration in which all the molecules occupy only a small fraction of the available volume), the entropy decreases. To use an ever more intuitive picture (which will be familiar to untidy students), we can think of a room in which many objects are scattered all around, as opposed to the situation in which all the objects are put in order in drawers, cupboards, and so forth.

6.2. The Maxwell-Boltzmann Distribution

As already noticed, the computations performed in Einstein's papers of 1909 and 1916, involve the statistical distribution of energies at thermal equilibrium. This tool can be introduced and discussed at various levels. A standard derivation from Boltzmann's postulate involves Stirling's approximation and Lagrange multipliers, hence it is not suitable to high school students²⁰. However, for our purposes it is sufficient to have an intuitive insight into it. In this subsection and in the next, we seek to provide such an insight, with no claim of mathematical rigor, for the case of the ideal gas. In particular, we shall not attempt to prove that the results we get also hold for non-ideal systems like thermal radiation far from the extremely quantum limit, limiting ourselves to assume that.

The energy distribution can be thought as a generalization of the Maxwell distribution for the velocities of the molecules in an ideal gas at equilibrium. The probability that the

velocity of a gas molecule is comprised in the very narrow interval between $\mathbf{v} = (v_x, v_y, v_z)$ and $\mathbf{v} + d\mathbf{v} = (v_x + dv_x, v_y + dv_y, v_z + dv_z)$ is given by²¹ $dP(\mathbf{v}) = f(\mathbf{v})dv_x dv_y dv_z$, where

$$f(\mathbf{v}) = Ae^{-\frac{mv^2}{2k_B T}}, \quad (36)$$

where A is an inessential normalization factor, whose role is to ensure that the probabilities for all possible events sum up to one, and m is the mass of the molecules. For simplicity, we consider a gas made up of a single species of molecules, which therefore are all identical. Since the distribution depends only on the modulus of the velocity, it can be cast in a way that gives the probability of v being comprised between v and $v + dv$. In this way, the Maxwell distribution tells us that

$$dP(v) = Ae^{-\frac{mv^2}{2k_B T}} 4\pi v^2 dv. \quad (37)$$

This is the form of the distribution which is usually shown to students²². One way to understand intuitively the factor $4\pi v^2$ is the following. The fact that there are many possible velocity vectors that share the same modulus v has to be taken into account. These vectors in fact point in all possible directions and have the same length, hence they are like the radii of a sphere of whose radius length is v . Such a sphere has a surface $4\pi v^2$; this quantity gives a measure of the number of velocity vectors with modulus v . It is what is commonly called the density of states in more advanced treatments.

We recognize, in the exponent of the distribution, the kinetic energy of a molecule with speed v , that is, $\varepsilon = \frac{1}{2}mv^2$. Instead of the distribution of velocities then, we may consider the distribution of energies, and give the probability that the a molecule has kinetic energy in the narrow interval between ε and $\varepsilon + d\varepsilon$ is:

$$dP(\varepsilon) = Ae^{-\frac{\varepsilon}{k_B T}} \omega(\varepsilon) d\varepsilon. \quad (38)$$

where by a slight abuse of notation we again denote by A the normalization constant, and $\omega(\varepsilon)$ is essentially the number of possibilities for a particle to have kinetic energy ε (it is the analog of the $4\pi v^2$ in (37)²³). This form of the distribution is sometimes referred to as the Maxwell-Boltzmann distribution. The probability of being in an interval that is not narrow is of course $P(\varepsilon_1 < \varepsilon < \varepsilon_2) = A \int_{\varepsilon_1}^{\varepsilon_2} e^{-\frac{\varepsilon}{k_B T}} \omega(\varepsilon) d\varepsilon$.

While we have considered the ideal gas case, the Maxwell-Boltzmann distribution in fact describes the equilibrium distribution of the energies ε of the elementary constituents of more general thermodynamic systems. It is possible for example that the elementary constituents can occupy only a discrete set of levels. In that case the statistical distribution is expressed as

$$P(\varepsilon_n) = Ap_n e^{-\frac{\varepsilon_n}{k_B T}}, \quad n = 1, 2, \dots, N. \quad (39)$$

where p_n is the number of levels which share the same energy (it is the analog of the function $\omega(\varepsilon)$), and N is the number of elementary constituents. In the 1916 paper [15], Einstein considered in fact atoms or molecules with discrete energy levels according to Bohr's theory, and thus used the Maxwell-Boltzmann distribution in this form.

Once we have the probability distribution of the energies, we can use it to compute the average value of the energy of an elementary constituent (i.e., of a molecule in the case of the ideal gas). As with any probability distribution, this is computed by:

$$\bar{\varepsilon} = \sum_n \varepsilon_n P(\varepsilon_n) = A \sum_n \varepsilon_n p_n e^{-\varepsilon_n/k_B T} \quad (40)$$

for the case of (39), and by

$$\bar{\varepsilon} = \int \varepsilon P(\varepsilon) \omega(\varepsilon) d\varepsilon = A \int \varepsilon e^{-\varepsilon/k_B T} \omega(\varepsilon) d\varepsilon \quad (41)$$

for the case of (38). Here the integral is extended to all possible values of the energies. In the case of the ideal gas, it is possible to show that this computation reproduces the result from elementary kinetic theory that $\bar{\varepsilon} = \frac{3}{2}k_B T$ (if the gas is monatomic).

Let us now keep in mind the case of the ideal gas for simplicity. While the average energy per molecule is $\bar{\varepsilon}$, any particular molecule will have an energy which is different from that value: either higher, or lower. It is said that the energies of the molecules fluctuate around the average value $\bar{\varepsilon}$. To quantify these fluctuations we could consider all the possible differences $\Delta\varepsilon = \varepsilon - \bar{\varepsilon}$ and average them with the Boltzmann distribution. Actually, since by definition $\overline{\Delta\varepsilon} = \overline{\varepsilon - \bar{\varepsilon}} = 0$ (the contribution of molecules with lower than average energies compensates that of molecules with higher than average energy) this will work if we take the squares $(\Delta\varepsilon)^2 = (\varepsilon - \bar{\varepsilon})^2$. The average of this quantity, $\overline{(\Delta\varepsilon)^2}$ is the so-called quadratic fluctuation, or equivalently its square root $\sqrt{\overline{(\Delta\varepsilon)^2}}$ if we want to keep the same dimensions, measures how much the energies of the molecules tend to be different from the average, that is, it measures how wide the probability distribution is. In 1909 Einstein actually computed the energy fluctuations of the whole radiation contained in the cavity. In the ideal gas case, this corresponds to computing the thermal fluctuations of the energy of the whole gas. In the next subsection we shall see what that means, and how it can be done.

6.3. The Gibbs Distribution and Energy Fluctuations

In order to understand Einstein's 1909 computation, one more tool is needed, namely, the Gibbs (ensemble) distribution, which describes the distribution of energies of the whole system, rather than those of its elementary constituents, like the Maxwell-Boltzmann distribution. A rigorous discussion of the Gibbs distribution would require introducing the ensemble picture of statistical physics, in which the focus is not on the many elementary constituents of a single copy of the system, but rather on many different copies of the system. This topic has subtle foundations, and introducing it properly would require quite a large conceptual leap. Therefore, in this subsection, we introduce this distribution in a non-rigorous, yet intuitive way, for the case of the ideal gas.

The starting point is the observation that even if the gas is enclosed into a container, it will likely not be isolated from the external environment. Rather, the gas will be able to exchange energy with the external environment through the walls of the container. In fact, it is exactly this energy exchange which allows the gas to be in thermal equilibrium with the environment. When this is the case, the energy that the gas absorbs from the environment has to be on average equal to that that it gives back to the environment, otherwise it will heat up or cool down, in contradiction with the hypothesis of equilibrium. However, since the exchange of energy between the gas and the environment is ultimately due to atoms hitting the wall, and such atoms are not infinitely small, the energy of the gas will not be exactly equal to its average, but it will fluctuate around it, hence it will be described by a probability distribution, just like the energies of the individual molecules. This is the Gibbs distribution. In fact, for the ideal gas case, the Gibbs distribution can be obtained by a quite straightforward extension of the Maxwell-Boltzmann one, and the two distributions share the same mathematical form. According to the Gibbs distribution, the total energy of the system is comprised in the narrow interval between E and $E + dE$ with the probability

$$dP(E) = \mathcal{A} e^{-\frac{E}{k_B T}} \omega(E) dE \quad (42)$$

where as usual \mathcal{A} is a normalization constant and $\omega(E)$ denotes the number of possible ways for the system of having energy E . Again, the actual computation of this factor involves several subtleties, but fortunately its exact form is not needed for our discussion. The standard way of interpreting (and proving) this distribution consists in imagining a set (in French, an ensemble) of identical copies of the container with the ideal gas, all at thermal equilibrium with the environment, and each one with a value of the total energy²⁴. Then

Equation (42) describes the distribution of the values of the energy among all the copies of the system. For this reason this is also referred to as the Gibbs ensemble distribution.

A possible way of justifying this expression is the following. Due to the absence of interactions and correlations, an ideal gas can be thought as an ensemble of copies of a single molecule. The energies of the molecules, hence of the copies, as we saw in the previous subsection, are distributed according to the Maxwell-Boltzmann distribution (38). By drawing a parallel between the molecules in a single copy and the ensemble copies of the ideal gas, one may intuitively grasp why the two distributions have the same mathematical form.

The Gibbs distribution allows us to compute the average value of the energy of the system and its quadratic fluctuations. Since it has exactly the same mathematical form of (38), the average energy will be given by a formally identical expression:

$$\bar{E} = \int E P(E) \omega(E) dE = \mathcal{A} \int E e^{-E/k_B T} \omega(E) dE \quad (43)$$

The energy fluctuations can be computed by an effective formula, which was derived by Einstein himself in 1904 [41]²⁵. As we noticed above, we have $\Delta \varepsilon = \varepsilon - \bar{\varepsilon} = 0$ for the single molecule energy. Analogously, we have $\Delta E = E - \bar{E} = 0$ for the energy of the whole system. This equality can be written explicitly as

$$0 = \mathcal{A} \int (E - \bar{E}) e^{-E/k_B T} \omega(E) dE. \quad (44)$$

Now, we can differentiate both sides of this equation with respect to the temperature T :

$$\begin{aligned} 0 &= \mathcal{A} \int \left(-\frac{d\bar{E}}{dT} e^{-E/k_B T} + (E - \bar{E}) e^{-E/k_B T} \frac{E}{k_B T^2} \right) \omega(E) dE \\ &= \mathcal{A} \int \left(\frac{E^2}{k_B T^2} - \frac{E\bar{E}}{k_B T^2} - \frac{d\bar{E}}{dT} \right) e^{-E/k_B T} \omega(E) dE = \frac{\bar{E}^2}{k_B T^2} - \frac{(\bar{E})^2}{k_B T^2} - \frac{d\bar{E}}{dT}, \end{aligned} \quad (45)$$

where in the last equality we used Equation (43). This can be rearranged as:

$$k_B T^2 \frac{d\bar{E}}{dT} = \bar{E}^2 - (\bar{E})^2. \quad (46)$$

The second term is nothing but the quantity we are looking for, since it can be expressed as:

$$\overline{(\Delta E)^2} = \overline{E^2 - 2E\bar{E} + (\bar{E})^2} = \bar{E}^2 - 2(\bar{E})^2 + (\bar{E})^2 = \bar{E}^2 - (\bar{E})^2, \quad (47)$$

so we obtain the final formula²⁶

$$\overline{(\Delta E)^2} = k_B T^2 \frac{d\bar{E}}{dT}, \quad (48)$$

This result has the interesting meaning (greatly emphasized by Einstein himself) that the quadratic energy fluctuations are proportional to Boltzmann's constant k_B . This means that, in typical situations, they are to be expected to be small. In other words, the size of fluctuations is related to that of atoms. Moreover, we notice that the derivative in the right hand side of Equation (48) describes the variation of the internal energy of the system with the temperature. In other words, it is the specific heat of the system, which describes the ability of the system to absorb energy without raising its temperature too much. From the macroscopic point of view, as well-known, when a system absorbs energy, much of which is then turned into heat (a typical example being friction), this energy is said to have been dissipated. Hence, Equation (48) links the fluctuations of the energy to the ability of the system to absorb energy and dissipate into heat. It is thus an example of a so-called

fluctuation-dissipation relation. Such relations are ubiquitous in statistical physics, and indeed the earliest examples of them are found in Einstein's works [8].

Although our heuristic discussion holds only for the case of the ideal gas, the Gibbs distribution actually is valid for much more general systems in thermal equilibrium with their surroundings. In particular, it holds in situations where there are interactions and/or correlations between the elementary constituents of the system. In fact, in more advanced courses (a standard reference is [43]) one typically derives this distribution first, and then applies it to the ideal gas, where it is shown that the energies or the velocities of the individual molecules obey the Maxwell-Boltzmann distribution by exploiting the fact that (up to a combinatorial subtlety) the Gibbs distribution factorizes. The distribution of discrete energy levels of elementary constituents (39) can be derived from the Gibbs distribution in the hypothesis that there are no correlations among the elementary constituents. It is possible to prove that this is exactly the case for a bunch of atoms in interaction with radiation, since (apart from special situations such as lasers) light thermalizes much more quickly than atoms. This justifies the use of the distribution (39) in Section 5.

The justification of the use of the Gibbs distribution and of formula (48) in Section 4 comes then from the fact that thermal radiation is a system at equilibrium as well. It is worth noticing that thermal radiation is an ideal gas only in the Wien limit, while in general it is not describable as a system of independent light quanta²⁷. In [12] Einstein computed the thermal fluctuations of thermal radiation for all regimes using the fluctuation Formula (48). We discuss this computation in Section 4. In the next subsection, instead, following [41], we apply the fluctuation formula to the ideal gas case.

6.4. Thermal Stability of the Ideal Gas

To use Formula (48), we need to know the average internal energy for the ideal gas. Let us consider for simplicity the monatomic case. In that case, we know from kinetic theory that $\bar{\epsilon} = \frac{3}{2}k_B T$. Then the average internal energy is given by $\bar{E} = N\bar{\epsilon} = \frac{3}{2}Nk_B T$, so we get

$$\overline{(\Delta E)^2} = \frac{3}{2}N(k_B T)^2. \quad (49)$$

The quantity that measures whether fluctuations are small (in Einstein's words, the thermal stability of the system) is actually the ratio of the fluctuations with the energy, that is, the relative fluctuations. This may be defined as

$$\zeta = \frac{\sqrt{\overline{(\Delta E)^2}}}{\bar{E}}. \quad (50)$$

According to Einstein, if this quantity is large, fluctuations of the systems are very large and can be observed. For the ideal gas we have

$$\zeta = \sqrt{\frac{2}{3N}} \sim \frac{1}{\sqrt{N}}. \quad (51)$$

This shows that the relative strength of the fluctuations decreases as the number of molecules N grows, in agreement with the expectation that for very large number of molecules, fluctuations with respect to the average values of thermodynamic quantities should be unobservable. Historically [8], it was likely the realization that fluctuations would be very difficult to observe in typical systems such as gases that led Einstein to study both thermal radiation and the Brownian motion²⁸. In fact, ref. [41] was the very first instance in which Einstein computed energy fluctuations of thermal radiation, although in that case he got a wrong result [8].

7. Implementation and Preliminary Results

The track was tested with both teachers and selected students attending their last year. Students were initially exposed only to part 1, and to the Boltzmann entropy part. More recently, we began introducing them also to the Maxwell-Boltzmann distribution (without the fluctuation formula) and part 3. Teachers instead were also exposed to part 2, and in some cases to all three parts, including a review of the statistical tools. This was done in the course of outreach activities performed in various high schools in southern Italy. The preliminary results are very encouraging. Informal tests were administered to teachers and students before and after the lectures, in order to evaluate their progress. Before the course, a part of the teachers admitted having some difficulties with quantum physics. After the course, most teachers claimed having obtained a more intuitive grasp of the relevant quantum concepts. Two facts seem to have concurred to this result. First, the concepts were not introduced in an abstract way, but rather with constant reference to a concrete physical system, namely thermal radiation; the continuous confrontation with the still more visualizable model of the ideal gas also helped. Second, the historical development was really seen as accompanying an increasing clearer and more detailed picture of quantum phenomena, specifically those concerning light and light-matter interaction.

The responses of the teachers emphasized how they managed to see quantum physics from a new point of view, which helped them to overcome some criticalities in their understanding. All of the tested teachers, including those whose major was in physics and therefore had had a more complete education in quantum physics, appreciated the fact that the presented material was new for them, since it was not presented in the textbooks they studied, nor in the ones they use in teaching. While most teachers were in fact familiar with the idea of light quanta, they considered the statistical picture of wave-particle duality as a useful complement to the usual one that is taught, namely the De Broglie one. In fact, before being exposed to Part 2 of the track, many teachers did not actually realize that Planck's law does not imply in general that light is made up of an ideal gas of quanta: this picture is actually valid only in the Wien limit, while in general it is more correct to think at thermal radiation as a manifestation of both the wave and the particle natures of light (or by a quantum ideal Bose gas, although this goes beyond our track). Usual treatments do not clearly state this, hence generating this diffuse misconception.

The introduction in Part 3 of probability in analogy with radioactive decay was judged useful as well. In fact, as we already remarked, a diffuse misconception is the belief that the probability of spontaneous emission is conceptually the same as the probability in classical statistical physics, that is, due to the limited amount of information available on the inner workings of atoms, rather than an intrinsic feature of Nature. After all, this view was advocated by Einstein himself. We observe that analogous misconceptions are often observed concerning Heisenberg's uncertainty principle, which often is thought to be due to instrumental limits (see e.g., [1,44]) rather than fundamental. Therefore, proper emphasis has to be devoted to this issue while developing Part 3. If this is not done, indeed, there is the risk that the above misconception is propagated. Stressing this aspect, instead, allows attendees to reach an increased consciousness of the intrinsic nature of probability in quantum mechanics, and should equip them to appreciate also the related fundamental nature of the uncertainty principle, when they tackle it in their later studies.

The students, for their part, all claimed to be intrigued by quantum physics and by the possibility of understanding it better, and also by the fact of being exposed to less well-known parts of Einstein's work. In fact, while most of them did know that Einstein contributed to quantum physics, they did not perceive how great his contribution was. Students admitted that they found the study of the presented material quite demanding, but also rewarding. Students, like their teachers, claimed that the deep study of a definite physical system had been helpful. In fact, some of the students had already been exposed to the concept of thermal radiation and of energy quantization, and had found them obscure and detached from subsequent topics. Einstein's treatment had the result of clarifying

that. As a final outcome, we also observed that students found the explicit application of Boltzmann's entropy to the ideal gas and to radiation useful and illuminating.

8. Discussion and Conclusions

In this paper, we have described a teaching–learning sequence aimed at introducing high school teachers to some relevant concepts in quantum and statistical physics. Our inspiration was found in the history of physics, which has often been considered as useful in teaching. Preliminary investigations with both teachers and selected students, performed in the course of outreach activities, gave encouraging results, clearly showing that the track helped the attendees clarify some conceptual issues and to get some intuitive grasp of the involved concepts. This has stimulated both a larger scale research, also with the inclusion of more students, which should begin in the near future, and also the extension of the track presented to students by the incorporation of more statistical mechanics. We plan to present quantitative results and a more complete study in a forthcoming publication. It is clear to us that Part 3 and especially Part 2, with the related discussion of the Boltzmann distribution and of fluctuations, are a bit more demanding for students than Part 1. A full course including all the material can easily exceed time constraints, especially if proposed in the regular school hours (but it can still be considered in outreach activities). However, the reduced path consisting of only Part 1 and the discussion of Boltzmann's entropy has already proven very useful. In fact, this was the original seed from which the whole path grew up. We also believe that a very good compromise for high school students is the combination of Part 1 plus Part 3, with a discussion of the material in Section 6.1 and of the first half of Section 6.2, omitting the discussion on the Gibbs distribution and of fluctuations. However, we aim at testing the response of students to the remaining parts to students in a future publication. Preliminary results instead seem to indicate that a full discussion of Parts 1, 2 and 3, with the related statistical tools and various complements can instead safely be proposed to teachers undergoing in-service or pre-service formation. We also think a slightly modified form, our track can also form the basis of an undergraduate course emphasizing the more historical aspects of quantum physics. In this context, one can of course also exploit the excellent dedicated literature [20–23].

Of course, as it is developed in this paper, our track does not constitute a complete course in quantum physics but rather, as explicitly stated in the introduction, it is meant to supplement such a course. Obviously, the parts that we streamlined in this presentation, because we thought that they could be treated in a standard way, must be fully developed in a complete course.

The track is centered on (or better limited to) the most important contributions by Einstein to thermal radiation theory, but of course these were not the only ones. Our choices were dictated by didactic, not historical, purposes even though in the course of developing the path a good number of interesting historical facts can be told. While we aimed at making some parts of Einstein's physics more well-known among students and teachers, our intention was not to do justice to all of his contributions to quantum theory. Such an exposition would require a long and far too advanced course. However, many more of Einstein's contributions could be of didactic relevance, also at the high school level. For example, Einstein's explanation of the low temperature behavior of the specific heat of solids in terms of quantum theory [45] is at least as ground-breaking as the work on radiation, proving in particular that quantum effects were pervasive and not specific to thermal radiation. We are currently considering the possibility of modifying the path by including this and other results.

More importantly, the path entirely unfolds in the realm of old quantum theory. After it is completed, an introduction to the basics of the full theory of quantum mechanics must be given. For this, a very rich literature has developed (see e.g., [1] and references therein), hence there are various ways in which our proposal can be complemented. Currently, the merging with other paths which are more oriented towards the full theory of quantum mechanics and more modern and trendy topics, such as entanglement, teleportation

and quantum computation (and of course the EPR argument), is under study. We have developed our own proposal for such a path in [46], again taking some inspiration from the heroic history of quantum mechanics.

Author Contributions: Conceptualization, M.D.M.; Data curation, S.E. and A.N.; Formal analysis, A.N.; Investigation, M.D.M. and A.N.; Methodology, S.E. and A.N.; Project administration, S.E.; Supervision, S.E.; Writing—original draft, M.D.M.; Writing—review & editing, M.D.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank A. Drago for suggesting in the first place that they investigate the pedagogical value of Einstein's 1905 paper, and the anonymous Referees for giving us useful suggestions for improving the presentation, especially of Section 5.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

- ¹ Interestingly, Einstein himself adopted a point of view concerning the conceptual use of history in teaching similar to that expressed by the present authors in the wonderful book [7].
- ² It has even been proposed that it should be the starting point for teaching quantum physics [10] (this is also the teaching strategy of [3], but only in newer editions), since it avoids the subtleties of statistical physics. We may agree with this claim if the teaching–learning sequence is severely time constrained, but we think that students can benefit from a suitable treatment of statistical physics, and it is our point here that such a treatment can indeed help in grasping some quantum concepts. Moreover, thermal radiation is a very important topic in its own right, with many applications.
- ³ This result was later extended by him to the case of matter along the same lines [14]. Thus Einstein introduced wave–particle duality independently of De Broglie, and his arguments rest on a completely different basis. The two approaches can be regarded as complementary, since Einstein's one is statistical in nature, while De Broglie's one applies to individual (light or matter) quanta.
- ⁴ It is fair to recall that Einstein himself, unlike other founders of quantum mechanics, did not regard this introduction of probability as a fundamental one, but rather due to our ignorance of the true internal working of atoms. Hence, for him, this probabilistic aspect was not fundamentally different from those emerging in statistical physics. He was very disturbed by the lack of causality ("I, at any rate, am convinced that *He* is not playing at dice." Letter from A. Einstein to M. Born, 4 December 1926, emphasis in original), and famously kept this viewpoint for all his life.
- ⁵ We point out in passing that for a better grasp of the topics of the path, and of elementary quantum theory in general, at least the qualitative idea that accelerated electric charge radiate electromagnetic waves has to be introduced. Because of the complexity of a full treatment, this topic is usually omitted from high school courses. However, this idea underlies thermal radiation and it was the main motivation for the inadequacy of Rutherford's atom and its substitution with Bohr's model. Hence, when presenting the material to students, we typically include this topic as a preliminary, with a brief discussion of Larmor's formula, without derivation of course.
- ⁶ A word about notation: in this paper we deal with thermodynamic systems made up of a large number of elementary constituents, hence different notions of energy come into play. We denote by E the total energy of the system, by u the spectral (i.e., per unit frequency) energy density, and by ϵ the energy of an elementary constituent, such as a gas molecule or a light quantum.
- ⁷ Students should have the notion of the volumic energy density of an electromagnetic field.
- ⁸ In fact, before explaining it with energy quantization, Planck wrote it down by fitting the experimental curve.
- ⁹ We point out that a very clear discussion can be found in the first section of [11], to the point that some authors, such as [8], refer to (2) as the Rayleigh–Einstein–Jeans law.
- ¹⁰ To avoid discussing Taylor series this can be easily explained by saying that the exponential function is approximated by its tangent in the point $x = 0$ in a neighborhood of that point.

- 11 In fact this law (not to be confused with Wien's displacement law) was obtained before Planck's one by a semi empirical reasoning based on the Maxwell distribution, which indeed has a very similar mathematical form (cf. Section 6.2).
- 12 A more rigorous explanation of why this result has the same form as that expected for a bunch of independent molecules involves the Poisson distribution, which describes "counting" particles, and can be found for example in [35] or in [36].
- 13 The crucial idea is that, unlike the emission of a spherical electromagnetic wave, the emission of a light quantum is intrinsically directional, hence the atom recoils after the emission.
- 14 At this level, there is no way of computing the A and B coefficients. The computation of these coefficients by first principles was achieved in 1927 by Dirac, when quantum mechanics was fully developed and applied to radiation.
- 15 Another thing to notice is that, since there are many possible states in which the emitted quantum can go, the process of spontaneous emission is irreversible in a statistical sense. Actually, by engineering appropriate cavities, it is actually possible to change the number of states, which through Equation (27) allows us to actually manipulate the spontaneous emission probability, which can be suppressed or enhanced. For example, by considering a cavity which is so small that, so to speak, there is no room for the waves to be into it (i.e., if the dimensions of the cavity are smaller than the wavelength), the number of possible states can be greatly reduced, and this in fact can suppress spontaneous emission. In extreme cases, the number of possible states can be reduced to one, thus making the process of spontaneous emission reversible. Radiation stays in the cavity long enough that it can be reabsorbed by the atom before being dissipated, and this generates oscillations between lower and higher atomic states. Such phenomena are the object of a very active research area nowadays (see e.g., [37] for a very clear review).
- 16 The other two processes depend on a well defined cause, that is the molecule being hit by a light quantum, so in that case the probability is related to the occurrence of that event. Hence the probabilistic nature of spontaneous emission is on a very different footing with respect to that of absorption and stimulated emission.
- 17 True, the same law appeared even earlier in radioactive decay, but for many years nobody could tell whether the nucleus obeyed the same quantum laws which hold at the atomic scales.
- 18 This model of course is good for describing the equilibrium state, not the process of relaxation to it, for which it is necessary to consider collisions between molecules; the equilibrium state is then characterized as non-changing under collisions. Here we only consider equilibrium states, so we do not need to concern about this complication.
- 19 This is a simple generalization of the argument by which we say a priori that in tossing a coin we get heads with probability $1/2$.
- 20 This derivation, originating from [38], is excellently presented in [39], and it can be used in place of the arguments in this subsection with a more advanced audience.
- 21 Usually the Maxwell distribution is expressed in terms of the number of molecules with velocity comprised in that interval, which is $dN(\mathbf{v}) = NdP(\mathbf{v})$, where N is the total number of molecules.
- 22 Typically high school students are exposed to the Maxwell distribution without proof. In that case, instructors may consider devoting some time to a simple proof of it. A very nice and instructive one which can be considered is that given by Maxwell himself in 1860 [40]. This proof uses the fact that an ideal gas is isotropic, and the statistical independence of the probabilities associated with each component of the velocity. The latter observation implies that we may write the function f as a product of functions, each one expressing the probability associated with that component; by isotropy, the distribution of the velocities along the three directions must be the same, hence these three functions must be equal. Hence we may write: $f(\mathbf{v}) = g(v_x)g(v_y)g(v_z)$, for some function g . However, isotropy also means that $f(\mathbf{v})$ can depend on \mathbf{v} only through its modulus v . These two conditions are satisfied by the function (36), since $Ae^{-\frac{mv^2}{2k_B T}} = Ae^{-\frac{m(v_x^2 + v_y^2 + v_z^2)}{2k_B T}} = Ae^{-\frac{mv_x^2}{2k_B T}}e^{-\frac{mv_y^2}{2k_B T}}e^{-\frac{mv_z^2}{2k_B T}}$, hence $g(x) = \sqrt[3]{Ae^{-\frac{mx^2}{2k_B T}}}$.
- 23 It can be computed by changing variables from v to ε in the differential, since it is defined by $4\pi v^2 dv = \omega(\varepsilon)d\varepsilon$; however its precise form is of no interest to us. In fact $\omega(\varepsilon) \sim \sqrt{\varepsilon}$.
- 24 This is the canonical Gibbs distribution, which is the appropriate one to use if the system is exchanging only energy with the environment. This is the only case we consider here.
- 25 Einstein gave a derivation of this formula also in his 1909 paper [12], however that derivation involves Taylor series, and therefore it is not suited for high school students. It can nevertheless be employed when proposing this material to a more advanced audience. The 1909 derivation is actually quite interesting also because it again involves a reversal of Boltzmann's principle, analogous to that he used in 1905, which we saw in Section 3 [42].
- 26 We notice that, again because of the fact that the Maxwell-Boltzmann and the Gibbs distribution have the same form, a formally identical equation can be derived to describe the fluctuations of the energies ε of the molecules of the gas around the average value $\bar{\varepsilon}$.

- 27 It is a so-called quantum ideal (Bose) gas, which like the usual (classical) ideal gas is made up by many identical particles. However, in this case the latter are not independent far from the Wien limit, because of quantum correlations (cf. the discussion at the end of Section 4).
- 28 It is in fact safe to assert that fluctuations constitute the Leitmotiv of most of Einstein's work in statistical physics [8].

References

1. Krijtenburg-Lewerissa, K.; Pol, H.J.; Brinkman, A.; van Joolingen, W.R. Insights into teaching quantum mechanics in secondary and lower undergraduate education. *Phys. Rev. Phys. Educ. Res.* **2017**, *13*, 010109. [\[CrossRef\]](#)
2. Buongiorno, D.; Michelini, M. The conceptual contribution of the history in learning physics: The case of optical spectroscopy. In Proceedings of the 37th Annual Conference of the Italian Society of Historians of Physics and Astronomy, Bari, Italy, 26–29 September 2017; Campanile, B., De Frenza, L., Garuccio, A., Eds.; Pavia University Press: Pavia, Italy, 2019; pp. 377–408.
3. Walker, J. *Fundamentals of Physics*, 10th ed.; Wiley: New York, NY, USA, 2014.
4. Eisberg, R.; Resnick, R. *Quantum Physics of Atoms, Molecules, Solids, Nuclei and Particles*; Wiley: New York, NY, USA, 1985.
5. Bohm, D. *Quantum Theory*; Prentice-Hall: New York, NY, USA, 1951.
6. Messiah, A. *Quantum Mechanics*; Elsevier: Amsterdam, The Netherlands, 1961.
7. Einstein, A.; Infeld, L. *The Evolution of Physics. The Growth of Ideas from Early Concepts to Relativity and Quanta*; Simon & Schuster: New York, NY, USA, 1938.
8. Pais, A. 'Subtle is the Lord...' *The Science and the Life of Albert Einstein*; Oxford University Press: Oxford, UK, 1982.
9. Pais, A. Einstein and the quantum theory. *Rev. Mod. Phys.* **1979**, *51*, 863–914. [\[CrossRef\]](#)
10. Bethe, H.A. My experience in teaching physics, Hans Bethe's acceptance speech for the 1993 Oersted Medal presented by the American Association of Physics Teachers, 13 April 1993. *Am. J. Phys.* **1993**, *61*, 972–973. [\[CrossRef\]](#)
11. Einstein, A. Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt (On a Heuristic Point of View Concerning the Production and Transformation of Light). *Ann. Phys.* **1905**, *17*, 132–148. [\[CrossRef\]](#)
12. Einstein, A. Zum gegenwärtigen Stande des Strahlungsproblems (On the Present Status of the Radiation Problem). *Phys. Zeit.* **1909**, *10*, 185–193.
13. Einstein, A. Entwicklung unserer Anschauungen über das Wesen und die Konstitution der Strahlung (On the Development of Our Views Concerning the Nature and Constitution of Radiation). *Phys. Z.* **1909**, *10*, 817–825.
14. Einstein, A. Quantentheorie des einatomigen idealen Gases. 2. Abhandlung (Quantum Theory of the Monatomic Ideal Gas, Part II). *Sitzung. Preuss. Akad. Wiss. Phys. Math. Klasse* **1925**, *1925*, 3–14.
15. Einstein, A. Quantentheorie der Strahlung (On the Quantum Theory of Radiation). *Mitteilungen Phys. Ges. Zürich* **1916**, *16*, 47–62.
16. Einstein, A. Strahlungs-emission und-absorption nach der Quantentheorie (Emission and Absorption of Radiation in Quantum Theory). *Verhandlungen Deutsch. Phys. Ges.* **1916**, *18*, 318–323.
17. Einstein, A. Zur Quantentheorie der Strahlung (On the Quantum Theory of Radiation). *Phys. Z.* **1917**, *18*, 121–128.
18. Duncan, A. *The Conceptual Framework of Quantum Field Theory*; Oxford University Press: Oxford, UK, 2012.
19. Paffuti, G. *Note sulla Nascita della Meccanica Quantistica (Notes on the birth of Quantum Mechanics)*; Pisa University Press: Pisa, Italy, 2013.
20. Cheng, T.-P. *Einstein's Physics. Atoms, Quanta and Relativity Derived, Explained and Appraised*; Oxford University Press: Oxford, UK, 2013.
21. Kennedy, R.E. *A Student's Guide to Einstein's Major Papers*; Oxford University Press: Oxford, UK, 2012.
22. Longair, M.S. *Theoretical Concepts in Physics. An Alternative View of Theoretical Reasoning in Physics*, 3rd ed.; Cambridge University Press: Cambridge, UK, 2020.
23. Longair, M.S. *Quantum Concepts in Physics. An Alternative Approach to the Understanding of Quantum Mechanics*; Cambridge University Press: Cambridge, UK, 2013.
24. Milonni, P.W. *The Quantum Vacuum. An Introduction to Quantum Electrodynamics*; Academic Press: San Diego, CA, USA, 1994.
25. The Collected Papers of Albert Einstein. Available online: <https://einsteinpapers.press.princeton.edu> (accessed on 29 April 2021).
26. Pauli, W. Einstein's contribution to quantum theory. In *Albert Einstein. Philosopher-Scientist*; Schilpp, P.A., Ed.; Open Court: La Salle, IL, USA, 1970; pp. 147–160.
27. Born, M. Einstein's statistical theories. In *Albert Einstein. Philosopher-Scientist*; Schilpp, P.A., Ed.; Open Court: La Salle, IL, USA, 1970; pp. 161–177.
28. Di Mauro, M.; Esposito, S.; Naddeo, A. A proposal for teaching introductory quantum physics in the footsteps of Einstein. In Proceedings of the 1st Electronic Conference on Universe, Online, 22–28 February 2021; doi:10.3390/ECU2021-09320. [\[CrossRef\]](#)
29. McDowell, S.A. A Simple Derivation of the Boltzmann Distribution. *J. Chem. Ed.* **1999**, *76*, 1393–1394. [\[CrossRef\]](#)
30. Müller, R. The Boltzmann factor: A simplified derivation. *Eur. J. Phys.* **2014**, *35*, 055002. [\[CrossRef\]](#)
31. Engelhard, L.; del Puerto, M.L.; Chonacky, N. Simple and synergistic ways to understand the Boltzmann distribution function. *Am. J. Phys.* **2015**, *83*, 787–793. [\[CrossRef\]](#)
32. Zhang, B. Coconuts and Islanders: A Statistics-First Guide to the Boltzmann Distribution. *arXiv* **2019**, arXiv:1904.04669.
33. Borsini, I. Una esperienza analogica sulla meccanica statistica dell'equilibrio termico (An analog experience on the statistical mechanics of thermal equilibrium). *Giornale Fisica* **2019**, *60*, 31–42.
34. Gamow, G. *Thirty Years That Shook Physics. The Story of Quantum Theory*; Doubleday & Co. Inc.: New York, NY, USA, 1966.

-
35. Varró, S. Einstein's fluctuation formula. A historical overview. *Fluct. Noise Lett.* **2006**, *6*, R11–R46. [[CrossRef](#)]
 36. Boffi, S.; D'Anna, M. *Le Radici del Dualismo Onda-Corpuscolo (The Roots of the Wave-Particle Duality)*; Bibliopolis: Napoli, Italy, 1999.
 37. Haroche, S.; Kleppner, D. Cavity Quantum Electrodynamics. *Phys. Today* **1989**, *42*, 24. [[CrossRef](#)]
 38. Boltzmann, L. Über die Beziehung zwischen dem zweiten Hauptsatze der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung resp. dem Satzen über das Wärmegleichgewicht (On the relationship between the second fundamental theorem of the mechanical theory of heat and probability calculations regarding the conditions for thermal equilibrium). *Wiener Berichte* **1877**, *76*, 62–100. English translation available: *Entropy* **2015**, *17*, 1971–2009.
 39. Fermi, E. *Molecules, Crystals and Quantum Statistics*; W. A. Benjamin: New York, NY, USA, 1966.
 40. Maxwell, J.C. Illustrations of the dynamical theory of gases. Part I. On the motion and collisions of perfectly elastic spheres. *Phil. Mag.* **1860**, *19*, 19–32. [[CrossRef](#)]
 41. Einstein, A. Allgemeine molekulare Theorie der Wärme (On the General Molecular Theory of Heat). *Ann. Phys.* **1904**, *11*, 170–187.
 42. Klein, M. Einstein and the wave–particle duality. *Nat. Philos.* **1964**, *3*, 1–49.
 43. Pathria, R.K.; Beale, P.D. *Statistical Mechanics*, 3rd ed.; Elsevier: Oxford, UK, 2011.
 44. Ayene, M.; Kriek, J.; Damtie, B. Wave-particle duality and uncertainty principle: Phenomenographic categories of description of tertiary physics students' depictions. *Phys. Rev. Spec. Top. Phys. Educ. Res.* **2011**, *7*, 020113. [[CrossRef](#)]
 45. Einstein, A. Plancksche Theorie der Strahlung und die Theorie der Spezifischen Wärme (Planck's Theory of Radiation and the Theory of Specific Heat). *Ann. Phys.* **1907**, *22*, 180–190. [[CrossRef](#)]
 46. Di Mauro, M.; Naddeo, A. Introducing quantum mechanics in high schools: A proposal based on Heisenberg's Umdeutung. *Universe* **2021**, under review.