

The Boltzmann factor: A simplified derivation

Rainer Müller

Technische Universität Braunschweig, Physikdidaktik, Bienroder Weg 82, D-38106
Braunschweig, Germany

E-mail: rainer.mueller@tu-bs.de

Abstract. We give a simple derivation of the Boltzmann factor from a microcanonical point of view. The derivation is similar to Boltzmann's own approach but avoids the use of Lagrangian multipliers. In the Boltzmann-Einstein model, we consider an assembly of $n + 1$ boxes. One of the boxes represents the system under consideration, the others act as a thermal reservoir. Energy is distributed over the combined system in discrete portions, not necessarily of quantum character. We derive the probability for the system to have a certain number of energy portions by counting the number of compatible states, thus arriving at the Boltzmann factor.

PACS numbers: 01.40.-d, 05.20.-y

1. Introduction

The Boltzmann distribution law for the probability of a system to be in a state with a certain energy E is generally considered as one of the most important laws of statistical physics. Feynman [1] calls it the “summit of statistical mechanics”. Indeed, the applications of the Boltzmann law range from the most theoretical to the most practical: from astroparticle physics over chemistry to solid state technology. Any book on statistical mechanics gives a derivation of the Boltzmann factor. There are a number of different approaches that are common in the literature:

- (i) *Lagrangian multiplier approach*: This method dates back to Boltzmann’s famous 1877 paper [2]. Portions of energy are distributed over the discrete energy levels of an assembly of “atoms” or “oscillators”. The probability of an energy distribution is proportional to the number of ways it can be realized. The most probable distribution is found by a Lagrangian multiplier method which accounts for the constraints of fixed energy and particle number.
- (ii) *Use of symmetry arguments*: There is a method of unparalleled simplicity for finding the Maxwell-Boltzmann velocity distribution. It was already given by Maxwell in his original 1860 paper [3]. Because of symmetry, the probability distribution for the velocity components factors into a product of three terms (one for each spatial direction). Furthermore the distribution must not depend on the vector components themselves but only on the coordinate-independent scalar \vec{v}^2 . With this argument, Maxwell obtains a functional equation for the probability distribution which is only fulfilled for the exponential function.
- (iii) *Quasicontinuum approach*: When passing from the microcanonical to the canonical ensemble, most books in Theoretical Physics consider a small system coupled to a heat bath. The combined system has a quasicontinuous distribution of energy levels. Because the probability factorizes, the probability for the system of being in a state with energy E_j is proportional to the number of microstates available to the reservoir with energy $E_{\text{tot}} - E_j$. A first-order Taylor expansion of the logarithm of the probability leads to the Boltzmann factor.

In addition to the classic approaches described above, there are a number of alternative arguments for the Boltzmann factor: Feynman [1] appealed to the barometric formula; Leff [4] considers a two-level system coupled to a reservoir with the probabilities interpreted as fraction of time the system is in the corresponding state. A version of the quasicontinuum approach is presented Kittel and Kroemer [5], and similarly by Moore [6]. They relate the probability to the entropy via Boltzmann’s formula and then perform a first-order expansion of the entropy. Hannay [7] treats a model with statistical jumps and determines the stationary energy distribution. López-Ruiz, Sañudo and Calbet [8] give a geometrical derivation, and McDowell [9] calculates the Boltzmann factor via the change in entropy when a small portion of energy is added to the system. In addition, there are a number of heuristic arguments (e. g. [10]).

The approach by Boltzmann is probably the one which imparts the deepest insights into the nature of the statistical approach. It sheds light on the stunning fact that the basic operation in statistical mechanics is *counting*. However, it needs the Lagrangian multiplier method which is beyond the mathematical capabilities for most students in introductory courses. In this paper, we give a derivation of the Boltzmann factor which is similar to Boltzmann's own approach but avoids the use of Lagrangian multipliers.

We consider the Boltzmann-Einstein model (also called the Einstein solid) which is mathematically simple but sufficiently abstract to serve as a model for many real systems. Within the model, we describe a small system in thermal equilibrium with a large reservoir. Concentrating on the small system alone, we ask for the probability for a state with energy E_j if the reservoir's temperature is T . In contrast to the third approach mentioned above, we will not assume a quasicontinuum of states because we do not want to abandon the idea of counting discrete states. Within the microcanonical formalism, we arrive at the Boltzmann factor by combinatorial arguments.

The model is similar to the one used, for example, by Callen [11], Dugdale [12], and Moore and Schroeder [13] when they derive a microscopic foundation for thermodynamic variables like entropy and temperature. Essentially the same approach has been taken by Friedman and Grubbs [14]. However, the complicated combinatorial formulas in their paper are likely to overwhelm the students. We have tried to keep the mathematics as simple as possible to make the method more accessible to students.

2. The model

The model we use is often called the Einstein solid [11]. We consider an assembly of n "atoms" or "boxes" which can hold an integral number of energy portions ϵ . Their possible energy values are $0, \epsilon, 2\epsilon, 3\epsilon, \dots$. The model is characterized by two integer numbers: n , the number of atoms, and u , the number of energy portions. The internal energy of the system is $U = u \cdot \epsilon$.

The obvious physical realization of the model is a collection of quantum mechanical harmonic oscillators. Nevertheless, the model is not so quantum mechanical as it may seem. In fact it was introduced already by Boltzmann to model continuous degrees of freedom: He used it as a vehicle to derive the Maxwell-Boltzmann velocity distribution for gas molecules [2]. By choosing ϵ arbitrarily small, we can approximate any continuous energy value by a rational number. For the sake of simplicity, we continue to speak of atoms, although for historical correctness, we will refer to the model as the Boltzmann-Einstein model.

There is no direct interaction between the atoms. In accordance with the basic postulate of statistical mechanics, transitions between states with the same total energy U are thought to be induced by minute interactions with the environment.

Callen [11] gives a clever method to determine the number of ways in which u portions of energy can be distributed among n boxes. Each energy portion is represented by a marble while $n - 1$ match sticks act as "dividers" between two boxes. Thus the

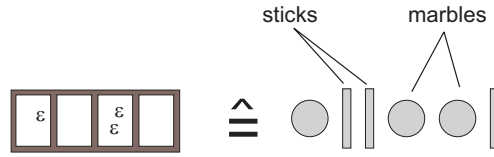


Figure 1. Construction for finding the number of possibilities for distributing u portions of energy among n atoms

problem is equivalent to finding the number of permutations of u marbles and $n - 1$ match sticks arranged in a row (Fig. 1). The number of possible energy distributions is called the multiplicity $\Omega(u, n)$. It is given by:

$$\Omega(u, n) = \frac{(n - 1 + u)!}{u!(n - 1)!}. \quad (1)$$

This quantity is thermodynamically significant because it is related to the entropy via $S = k_B \ln \Omega$, where k_B is Boltzmann's constant. We stress that with this method of counting there is no need to impose additional constraints for keeping the energy and the particle number constant. They are automatically accounted for.

3. An atom plus a reservoir

We now consider an assembly of $n + 1$ atoms. One of them is singled out as a probe, the remaining n atoms constitute the reservoir. In the following the probe atom is simply referred to as “the atom”. Our goal is to determine the probability p_j to find it in a state with energy $E_j = j \cdot \epsilon$ in thermal equilibrium.

There are u energy portions to distribute over the system (which consists of the atom plus the reservoir). We have to ask in how many ways this can be done, given that the atom has a share of j energy portions. Because the atom and the reservoir are only weakly coupled, there are no energy portions that are shared between the atom and the reservoir in the form of binding energy. The multiplicity can therefore be written as a product:

$$\begin{aligned} \Omega(j, u, n + 1) = & \text{number of ways to distribute } j \text{ portions over the atom} \\ & \times \text{number of ways to distribute } u - j \text{ portions over the reservoir.} \end{aligned} \quad (2)$$

The first factor equals 1 because we are dealing with a single state of the atom with a unique distribution of energy. Thus, the multiplicity is just the number of reservoir states with energy $(u - j)\epsilon$. With Eq. (1) we obtain:

$$\Omega(j, u, n + 1) = \frac{(n - 1 + u - j)!}{(u - j)!(n - 1)!}. \quad (3)$$

In a teaching situation, it is helpful to explore the possible energy distributions for small n and u with concrete numbers. Moore and Schroeder [13] point out an insightful









	$\Omega(j, u, n+1)$	ρ_j	
	$j = 0$	1716	0.5
	$j = 1$	924	0.27
	$j = 2$	462	0.13
	$j = 3$	210	0.06
	$j = 4$	84	0.02
	$j = 5$	28	0.008
	$j = 6$	7	0.002
	$j = 7$	1	0.0003

Figure 2. Distributing $u = 7$ energy portions over $n + 1 = 8$ boxes. The total number of states is $\Omega(u, n + 1) = 3432$. The rows represent the number of states with j energy portions in the atom.

way to do this with the help of spreadsheet software. The students are able at this point to qualitatively understand the form of the Boltzmann distribution. If there is more energy concentrated in the atom, there is less energy left to be distributed over the reservoir. Since the multiplicity grows with the number of energy portions, atom states with a higher energy are less probable because there are less available reservoir states. An example with $u = 7$ energy portions distributed over $n + 1 = 8$ boxes is shown in Fig. 2. The multiplicities have been calculated with Eq. (3).

Before making this argument quantitative we have to define a temperature for the system. The entropy of $n + 1$ atoms with internal energy $u \cdot \epsilon$ is:

$$S = k_B \ln \Omega(u, n + 1) = k_B \ln \frac{(n + u)!}{u! n!} \quad (4)$$

(cf. Eq (1)). For large n and u , this expression can be simplified with Stirling's formula $\ln m! \approx m \ln m - m$. We obtain

$$S = k_B \left[n \ln \left(1 + \frac{u}{n} \right) + u \ln \left(1 + \frac{n}{u} \right) \right]. \quad (5)$$

This expression for $S(u, n)$ is the fundamental equation for the model from which all macroscopic thermodynamic information can be derived [11]. The temperature can be defined if we treat u for the moment as a continuous variable:

$$\frac{1}{T} = \frac{\partial S}{\partial(u\epsilon)}. \quad (6)$$

We find[11]:

$$\frac{\epsilon}{k_{\text{B}}T} = \ln\left(1 + \frac{n}{u}\right), \quad (7)$$

which takes a more familiar when we solve for $U = u \cdot \epsilon$:

$$U = \frac{n \cdot \epsilon}{e^{\frac{\epsilon}{k_{\text{B}}T}} - 1}. \quad (8)$$

We now return to our primary question: finding the probability p_j for the atom to be in a state with energy $j \cdot \epsilon$. It is given by ratio of $\Omega(j, u, n + 1)$ to the total number of states available to a system with $n + 1$ atoms and energy $u \cdot \epsilon$:

$$p_j = \frac{\Omega(j, u, n + 1)}{\Omega(u, n + 1)} = \frac{(n - 1 + u - j)! u! n!}{(u - j)! (n - 1)! (n + u)!} = n \cdot \frac{(n - 1 + u - j)! u!}{(u - j)! (n + u)!}. \quad (9)$$

We now apply Stirling's formula to the logarithm of p_j . It is useful to separate beforehand a factor $(n + u)$ from the factorial in the last term of Eq. (9):

$$p_j = \frac{n}{n + u} \cdot \frac{(n - 1 + u - j)! u!}{(u - j)! (n + u - 1)!}, \quad (10)$$

We thereby accomplish that the linear part of $\ln m! \approx m \ln m - m$ does not contribute at all because all the terms cancel. We obtain:

$$\ln\left(p_j \frac{n + u}{n}\right) = (n + u - j - 1) \ln(n + u - j - 1) + u \ln u - (u - j) \ln(u - j) - (n + u - 1) \ln(n + u - 1). \quad (11)$$

Within the slowly varying logarithms, we approximate $u - j \approx u$. This is equivalent to assuming that for a large reservoir, the reservoir does not "feel" the energy taken away by a single atom. A huge cancellation of terms occurs, and we end up with:

$$\ln\left(p_j \frac{n + u}{n}\right) = -j \ln\left(1 + \frac{n}{u}\right). \quad (12)$$

With Eq. (7), the last factor is identified as the inverse temperature of the system so that we can write:

$$\ln\left(p_j \frac{n + u}{n}\right) = -j \frac{\epsilon}{k_{\text{B}}T}. \quad (13)$$

Solving for p_j , we obtain our result,

$$p_j = \frac{n}{n + u} \exp\left(-\frac{j \cdot \epsilon}{k_{\text{B}}T}\right), \quad (14)$$

which is the well-known Boltzmann factor:

$$p_j \propto \exp\left(-\frac{E_j}{k_B T}\right), \quad (15)$$

Expression (14) is already properly normalized if we allow j to vary to infinity:

$$\sum_{j=0}^{\infty} p_j = \frac{n}{n+u} \sum_{j=0}^{\infty} \exp\left(-\frac{j \cdot \epsilon}{k_B T}\right) = \frac{n}{n+u} \sum_{j=0}^{\infty} \left(\frac{u}{u+n}\right)^j = 1. \quad (16)$$

4. Conclusion

Within the Boltzmann-Einstein model, we have derived the Boltzmann factor as the probability for an atom coupled to a reservoir to hold j out of u energy portions. We arrived at our result by the most simple of all mathematical operations—by counting. In a combinatorial analysis we counted the subset of the total available states in which the atom has j energy portions. While the exact formula is given by Eq. (10), the well-known exponential form of Eq. (14) is an excellent approximation if the reservoir is large.

References

- [1] R. P. Feynman, *Statistical mechanics: A set of lectures* (edited by Jacob Shaham), Redwood City: Addison-Wesley (1990).
- [2] L. Boltzmann, Über die Beziehung zwischen dem zweiten Hauptsatz der mechanischen Wärmetheorie und der Wahrscheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht, *Wien. Ber.* **76**, 373-435 (1877).
- [3] J. C. Maxwell, Illustrations of the dynamical theory of gases, *Philosophical Magazine* **19**, 19-32 (1860).
- [4] H. S. Leff, Answer to Question #14, *Am. J. Phys.* **63**, 877-878 (1995).
- [5] C. Kittel, H. Kroemer, *Thermal physics*, New York: Freeman (1980).
- [6] T. Moore, *Six Ideas that shaped physics, Unit T – Some processes are irreversible*, New York: McGraw Hill (2003).
- [7] J. H. Hannay, Answer to Question #14, *Am. J. Phys.* **63**, 877 (1995).
- [8] R. López-Ruiz, J. Sañudo, X. Calbet, A geometrical derivation of the Boltzmann factor, *Am. J. Phys.* **76**, 780 (2008).
- [9] S. A. C. McDowell, A Simple Derivation of the Boltzmann Distribution, *J. Chem. Educ.* **76**, 1393-1394 (1999).
- [10] C. A. Whitney, Answer to Question #14, *Am. J. Phys.* **63**, 876-877 (1995).
- [11] H. B. Callen, *Thermodynamics and an Introduction to Thermostatistics*, 2nd Edition, New York: Wiley & Sons (1985), Chapter 15.
- [12] J. S. Dugdale, *Entropy and its physical meaning*, London: Taylor & Francis (1996), Chapter 7.
- [13] T. A. Moore, Daniel V. Schroeder, A different approach to introducing statistical mechanics *Am. J. Phys.* **65**, 26-36 (1997).
- [14] E. Friedman, W. T. Grubbs, The Boltzmann Distribution and Pascal's Triangle, *Chem. Educator* **8**, 116-121 (2003).