Active learning approach for an intuitive understanding of the Boltzmann distribution by basic computer simulations

Greg Jameson,[1,2] Rafael Brüschweiler[1,2,3*]

[1]Department of Chemistry and Biochemistry, The Ohio State University, Columbus, Ohio 43210, USA

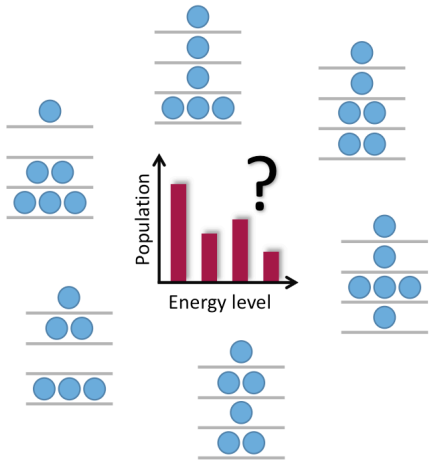
[2]Biophysics Graduate Program, The Ohio State University, Columbus, Ohio 43210, USA

[3]Department of Biological Chemistry and Pharmacology, The Ohio State University, Columbus, Ohio 43210, USA

ABSTRACT

The Boltzmann distribution lies at the heart of essentially all of statistical thermodynamics. In most textbooks, this distribution is introduced either *ad hoc* or it is mathematically derived by constrained entropy optimization using the method of Lagrange multipliers. Unfortunately, when students enroll in a statistical thermodynamics course, many are not familiar with this method from their prerequisite calculus courses. The recent availability of powerful computational devices for essentially all students provides an alternative way to explore the origin of qualitative and quantitative aspects of the Boltzmann distribution. Here we demonstrate a straightforward simulation approach to obtain an intuitive understanding of this distribution. We also make available concise and easily understandable computer programs in MATLAB and Python providing an opportunity for an active learning experience of this fundamental law.

GRAPHICAL ABSTRACT



KEYWORDS

Upper-Division Undergraduate; Graduate Education / Research; Physical Chemistry; Computer-Based Learning; Thermodynamics; Statistical Mechanics; Mathematics / Symbolic Mathematics

The Boltzmann distribution (BD), sometimes also referred to as the Boltzmann distribution law or the Boltzmann factor,

$$p_i = \frac{e^{-E_i/kT}}{\sum_i e^{-E_i/kT}} \quad (1)$$

was introduced by Ludwig Boltzmann 150 years ago. It applies to a system of N identical molecules at absolute temperature T where p_i is the probability that any particular molecule is in state i with energy E_i, where k is the Boltzmann constant. The BD is fundamental and has a remarkably simple form, which makes it one of the most celebrated relationships in all of physical chemistry and physics. Starting out from the Boltzmann distribution many key results of statistical thermodynamics can be derived, such as the Maxwell-Boltzmann velocity distribution, equipartition, free energies, heat capacities, and equilibrium constants of chemical reactions. The central role of the BD is also highlighted in popular Physical Chemistry textbooks, such as "The Boltzmann distribution is the key to understanding a great deal of chemistry. … There is, perhaps, no more important unifying concept in chemistry."[1] or "The Boltzmann factor is one of the most fundamental and useful quantities of physical chemistry".[2]

The BD implies that for a system at thermodynamic equilibrium at positive absolute temperature T, high-energy states have a lower probability than low-energy states and, hence, in the absence of degeneracy the most populated state is always the ground state. Interestingly, when asking physical chemistry students in the classroom what they expect which state has the highest population as a function of temperature, many think that for increasing temperature the *highest populated state* will also have a *higher energy*. This misconception is a consequence of the fact that the BD has little analogy in daily life (perhaps with the exception of the altitude dependence of pressure of the isothermal atmosphere) and, thus, students arrive with little or no intuition even at a qualitative level when they encounter statistical thermodynamics for the first time in college.

In most textbooks, the BD is introduced either *ad hoc* or it is derived using the elegant mathematical Lagrange multiplier method.[3] Unfortunately, when they enroll in a statistical thermodynamics course, an increasing number of students are not familiar with this method from

their prerequisite calculus courses, which makes the derivation of the BD a double challenge that from a didactical point of view should be avoided. Even for students familiar with Lagrange multipliers, the emergence of the BD, which is the energy distribution that maximizes the entropy of a large system when the average energy per molecule (or the total energy of the system E_{tot}) is kept at a predefined value, is perceived as a "*deus ex machina*". As a consequence, students often fail to appreciate the precise mathematical form of the BD and even more importantly they often do not develop much intuition for the qualitative characteristics of this fundamental law.

Over the years, alternative analytical approaches for the derivation or rationalization of the BD in the classroom have been introduced. Some are mathematically similarly abstract as the Lagrange multiplier method,[4] focus on special cases (such as the 2-level system),[5] show that the free energy of the BD is a local minimum,[6] or use analytical combinatorial approaches for finite systems with equally spaced energy levels.[7,8,9]

Fortunately, the recent availability of powerful desktop and portable computational devices available to essentially all students provides an alternative opportunity to explore the origin of the BD. It should be emphasized that the goal here is *not* to re-derive the BD (Eq. (1)). Rather, the goal is to provide students a framework that allows them explore by means of simple "numerical experiments" the most probable energy distributions on their own and compare them with the BD. Once some intuition has been developed for the BD, the students can be acquainted with a formal derivation and other analytical approaches, such as the ones mentioned above.[7,8,9]

The approach we propose here uses computer simulations for the straightforward counting of energy configurations that are consistent with the total energy constraint. The system consists of N distinguishable molecules numbered k = 1, 2, …, N where each of them can have an energy $E_i^{(k)}$ without degeneracy. For simplicity, we assume that the energies available to each molecule are equidistantly spaced, akin to the 1D quantum-mechanical harmonic oscillator, although other energy level diagrams can of course be used instead:

$$\varepsilon_i = i \cdot \varepsilon \qquad \text{where} \quad i = 0,1,2,\ldots,i_{max} \tag{2}$$

where ε is the energy spacing. A microstate is uniquely defined by the energies $\boldsymbol{\varepsilon} = (\varepsilon^{(1)}, \varepsilon^{(2)}, \ldots, \varepsilon^{(N)})$ and has total energy

$$E_{tot} = \sum_{k=1}^{N} \varepsilon^{(k)} \quad (3)$$

We now use a computer program to generate *random* energy vectors $\boldsymbol{\varepsilon}$ by choosing integers i from a uniform random distribution from 0 to i_{max}. Importantly, we only store an energy vector $\boldsymbol{\varepsilon}$ when it fulfills (or nearly fulfills) Eq. (3) and repeat the procedure many times, e.g. $N_{trial} = 10^5$ to 10^8. This key concept our approach is built on uses the "equal a priori probability postulate" of statistical mechanics. This central postulate, which can be motivated but not proven, assumes that each microstate of a microcanonical ensemble has the same probability. It can be viewed as an application of the "principle of indifference" in statistics, which states that in the absence of any further information or constraints, equal probabilities should be assigned to each allowed state. As a result, we obtain a set of unbiased microstates that all have the same total energy E_{tot}, thereby representing a microcanonical ensemble. Next, we count the relative frequency that any particle has energy ε_i according to Eq. (2) and convert this into probabilities p_i. We finally plot p_i vs. ε_i to visualize the statistical energy distribution adopted by the N particles. For sufficiently small N, this procedure can be done manually through exact enumeration (Table 1), but for larger N either some more sophisticated combinatorics methods are needed[7,8,9] or a modern computer as demonstrated here.

Table 1. Allowed energy configurations for $N = 4$ distinguishable molecules (A, B, C, D) with total energy $E_{tot} = 6\varepsilon$. The degeneracy W_j accounts for all possible permutations.

	Energy of each molecule in multiples of ε				
Energy Configuration, j	$E^{(A)}$	$E^{(B)}$	$E^{(C)}$	$E^{(D)}$	*Degeneracy* W_j [a]
1	0	0	0	6	4
2	0	0	1	5	12
3	0	0	2	4	12
4	0	0	3	3	6
5	0	1	1	4	12
6	0	1	2	3	24
7	0	2	2	2	4
8	1	1	1	3	4
9	1	1	2	2	6

[a] Determination of the degeneracies W_j are simple albeit instructive combinatorics problems for the students.

To illustrate the approach, for N = 4 molecules and a total energy E_{tot} = 6ε an explicit enumeration of all energy configurations is possible as shown in Tables 1, 2, and Figure 1A. The degeneracy W_j accounts for all possible permutations of a given energy configuration. Even for such a small molecular ensemble, the populations of the energies are non-uniformly distributed with the lowest energy level having the highest population and populations belonging to increasing molecular energies decaying monotonously, which is consistent with the BD, although the distribution is polynomial rather than exponential because of finite size effects.

Table 2. Total number of occurrences of molecular energy $\varepsilon_i = i\,\varepsilon$ (i= 0, 1, ..., 6) of example of Table 1.

Molecular energy in multiples of ε	*0*	*1*	*2*	*3*	*4*	*5*	*6*
Number of occurrences	112	84	60	40	24	12	4
Probability p_i	0.333	0.250	0.179	0.119	0.071	0.036	0.012

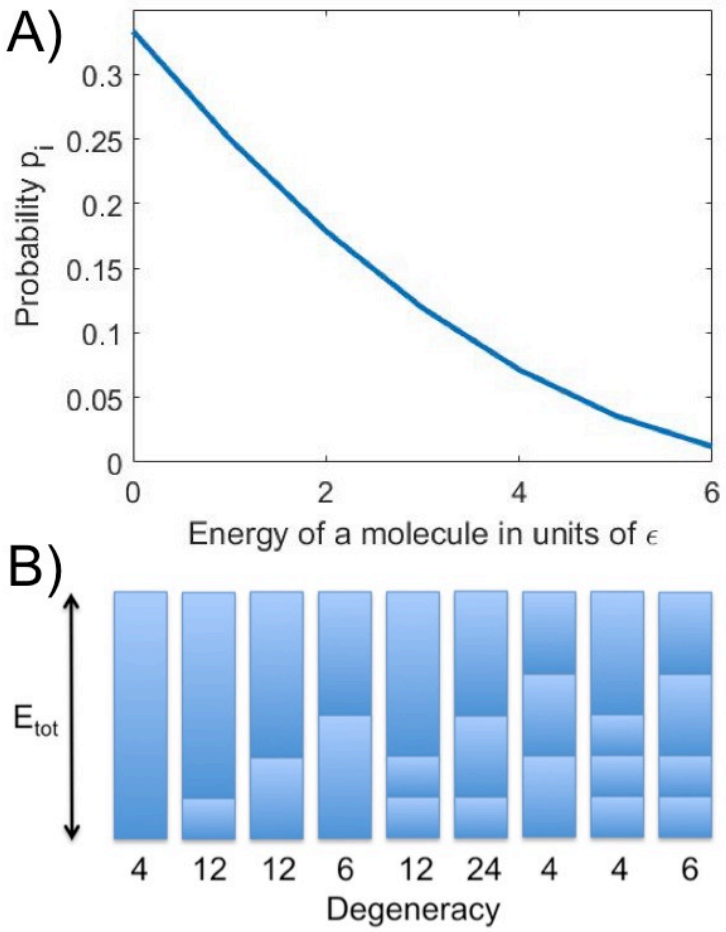
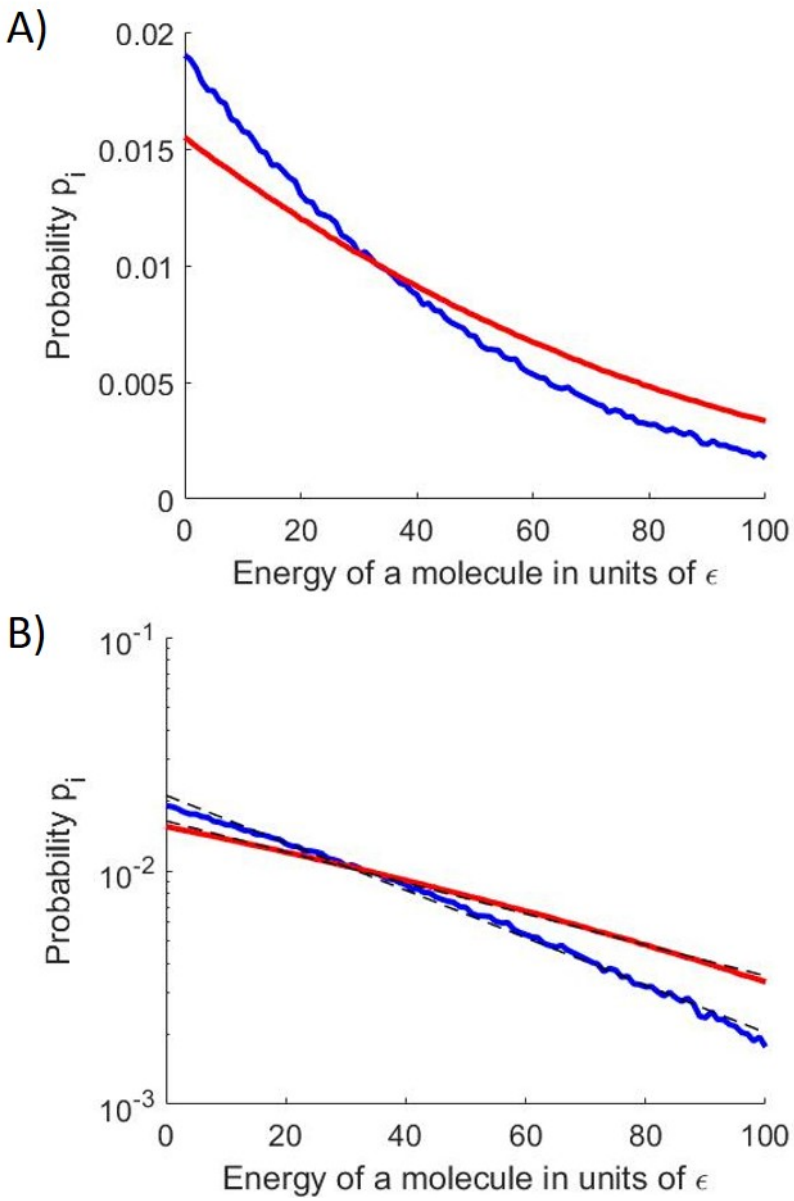


Figure 1. (A) Probability distribution $\{p_i\}$ of Table 2 of molecular energies $\varepsilon_i = i\,\varepsilon$ (i = 0, 1, …, 6) for 4 molecules with total energy $E_{tot} = 6\varepsilon$. (B) Visual representation of the 9 configurations of Table 1 with $E_{tot} = 6\varepsilon$ with their degeneracies W_i indicated at bottom. Note that molecules with zero energy have zero block size and are not depicted in the figure. Even when one does *not* account for degeneracies, the higher propensity for lower energy states is clearly visible, e.g. $\varepsilon^{(k)} = \varepsilon$ occurs 9 times, $\varepsilon^{(k)} = 3\varepsilon$ occurs 4 times, and $\varepsilon^{(k)} = 4\varepsilon$ occurs only 2 times.

For an analogy in daily life, we like to compare the role of the single contribution $\varepsilon^{(k)}$ to the predefined E_{tot} to a mason who needs to construct a wall of predefined height (E_{tot}) by stacking on top of each other a predefined number of rocks (N) of various sizes $\varepsilon^{(k)}$. Each rock size is equally abundant. The mason randomly picks them and builds the wall until it has reached the correct height. If by chance the mason only picks small rocks, he cannot reach the required height and if he only picks large rocks, the wall will be too tall. Therefore, statistically speaking, he must strike a balance between small and large rocks. As the rocks are chosen randomly, a *general preference for smaller rocks will naturally emerge* as they offer more viable combinations to fit them with (on average) fewer larger rocks (Figure 1B).

With larger N, the populations increasingly approach an exponential distribution consistent with the Boltzmann law in the thermodynamic limit as illustrated in Figure 2 with N = 30 for two different total energies E_{tot} using the MATLAB or Python computer programs described below. As usual, the level of exponential behavior can be easiest assessed visually in a log-linear plot. The nearly perfectly straight lines (Figure 2B) demonstrate that a system with N = 30 molecules is already approaching the thermodynamic limit and, hence, temperature T is defined and can be extracted via the slope $-1/kT$. Consistent with intuition, Figure 2B shows how a higher total energy E_{tot} (red lines) leads to a probability distribution $\{p_i\}$ with a shallower slope corresponding to a higher temperature T. The relationship between E_{tot} and T can be explored empirically by computer simulations requiring only minimal modifications of the computer code. The simulations also demonstrate that although the system is isolated (microcanonical ensemble), for large N the energies of individual molecules increasingly adopt a BD as they use the other molecules as a heat bath.



135 Figure 2. Simulation results for probability distribution $\{p_i\}$ of molecular energies $\varepsilon_i = i\ \varepsilon$ (i= 0, 1, …, 100) for N = 30 particles and average energy $E_{tot}/N = 30\varepsilon$ (blue) and $E_{tot}/N = 36\varepsilon$ (red) plotted with (A) linear axes and (B) log-linear axes. The lines in (B) are almost straight (see dashed straight lines for comparison) indicating that the thermodynamic limit has almost been reached.

In order to take full advantage of this empirical computer-assisted counting approach for the 140 rationalization of the BD, students need to know some basic computer programming. However, the availability of the MATLAB and Python programs provided here allows students with even only very little prior programming experience to experiment with input parameters, such as the size N of the ensemble, the total energy E_{tot}, the number of N_{trials} or alternative energy level diagrams. The MATLAB

and Python programs are available via Github (https://github.com/active-learning-boltzmann/boltzmann).

As part of the active learning experience for students, it is useful to discuss the findings also in qualitative and semi-quantitative terms. This includes the discussion of small ensemble effects that give rise to non-exponential behavior of the probability distribution, which are typically not covered in statistical thermodynamics courses, the change of the functional form of the probability distribution as a function of the number of molecules N, the effect of a change of E_{tot} on the probability distribution, the relationship between the slope $-1/kT$ of Figure 2B and E_{tot}, population inversion and negative temperatures when i has a ceiling, and the effect of an energy level diagram other than equidistantly spaced levels (e.g. with energies that scale with a positive or negative power of i or have some other predefined values). Active learning aspects can also encompass basic, but important elements of scientific computing, such as the discussion of uncertainties and convergence for different values of N_{trial}, the meaning of pseudo-random vs. random numbers and the role of the random seed.

ACKNOWLEDGMENT

This work was supported by the U.S. National Science Foundation (grant MCB-1715505). The approach described here was introduced in a 1st year graduate course on Statistical Thermodynamics (CHEM 7550) in Spring 2020 at the Department of Chemistry and Biochemistry at Ohio State as teaching assistant (G.J.) and instructor (R.B.).

AUTHOR INFORMATION

Corresponding Author

*E-mail: bruschweiler.1@osu.edu

REFERENCES

1. Atkins, P.; de Paula, J. *Physical Chemistry,* 10 ed.; Oxford University Press: Oxford, England, 2014.
2. McQuarrie, D. A.; Simon, J. D. *Physical Chemistry, A Molecular Approach,* 1 ed.; University Science Books: Sausalito, CA, 1997.
3. Dill, K. A.; Bromberg, S. *Molecular Driving Forces: Statistical Thermodynamics in Biology, Chemistry, Physics, and Nanoscience,* 2nd ed.; Garland Science: New York, NY, 2011.
4. Hakala, R. W. A New Derivation of the The Boltzmann Distribution Law. *J. Chem. Ed.* **1991** 38 (1), 33.
5. McDowell, S. A. C. A Simple Derivation of the Boltzmann Distribution. *J. Chem. Ed.* **1999** 76 (10), 1393.
6. Russell, D. K. The Boltzmann Distribution. *J. Chem. Ed.* **1996** 73 (4), 299.
7. Friedman, E.; Grubbs, W.T. The Boltzmann Distribution and Pascal's Triangle, W. T. *Chem. Educator* **2003**, 8 (2), 116.
8. Kugel, R. W.; Weiner, P. A. Energy Distributions in Small Populations: Pascal versus Boltzmann. *J. Chem. Ed.* **2010** 87 (11), 1200.
9. Müller, R. The Boltzmann factor: a simplified derivation. *Eur. J. Phys.* **2014** *35* (5), 055002.