

How to teach statistical thermal physics in an introductory physics course

Koo-Chul Lee

*Department of Physics and Center for Theoretical Physics, Seoul National University,
Seoul, 151-742, Korea*

(Received 16 August 1999; accepted 5 May 2000)

We report on several simulation programs (available through <http://phys.snu.ac.kr/howto/> or <http://phys.snu.ac.kr/~kcllee/howto/>) which can be used to teach the statistical foundations of thermal physics in introductory college physics courses. These programs are simple applications of a technique for generating random configurations of many dice with a fixed total value. By merely simulating dice throwing we can demonstrate all the important principles of classical thermodynamics. © 2001 American Association of Physics Teachers.

[DOI: 10.1119/1.1287719]

I. INTRODUCTION

The laws of thermodynamics were formulated long before we became aware of the existence of atoms and molecules.

With regard to intermediate courses in the undergraduate curriculum, for some time now the formal thermodynamics curriculum based on empirical laws has been replaced by courses based on the physical and statistical foundations of the subject. In these courses it is demonstrated how the properties of macroscopic systems are simple consequences of the behavior of their elementary constituents. Unfortunately, however, in introductory physics courses the “traditional” exposition of the subject which follows the historical sequence of events is still prevailing.

In an age when a single atom can be trapped and manipulated the teaching of thermal physics starting from the empirical laws of thermodynamics is a pedagogical scandal. It is now high time to reform the style of teaching thermal physics in introductory college physics courses.

In the majority of introductory college physics textbooks, authors tend to introduce the laws of thermodynamics before embarking on a brief exposition of the statistical basis of the subject, if they mention it at all. They introduce the concept of entropy¹ near the end of the thermal physics section by employing Clausius’s definition

$$dS = dQ/T \quad (1)$$

in conjunction with the Carnot cycle. On the other hand, Boltzmann’s definition of the entropy

$$S = k \ln W \quad (2)$$

usually appears as a passing remark with merely brief comment. Virtually none of the authors makes an attempt to relate the two expressions. The Clausius definition of entropy is one of the most difficult subjects to teach. First of all, T in the above expression (1) is the absolute temperature. If it were not the absolute temperature in expression (1) the Clausius definition would be meaningless. However, the concept of absolute temperature is never fully explained in textbooks that follow the traditional style of instruction. The temperature is usually introduced through an operational definition which cannot explain the significance of the absolute temperature other than as a curious experimental fact. We cannot convey the significance of absolute temperature without a basic understanding of its statistical nature.

Second, dQ is usually introduced in conjunction with the first law of thermodynamics where it is implied to be an arbitrary infinitesimal quantity of heat. However, in the

Clausius definition (1) the left-hand side is an exact or total differential of a state function S . How can we expect students to understand the meaning of the Clausius expression if we do not explain to them clearly and carefully the physical and mathematical bases which allow us to connect an arbitrary infinitesimal quantity to an exact differential of some function, S . This subject is not an easy topic even for intermediate thermal physics courses, where we devote much more time for the discussion of the subject.

In many cases, instructors bypass all these details of the Clausius expression and proceed to apply the formula to thermodynamic processes.

On the other hand, Boltzmann’s definition is simple and clear enough. The only problem is how to relate Boltzmann’s expression to thermodynamic functions in a simple and clear manner. This is the task of statistical mechanics, but the subject is generally considered too difficult to teach in introductory physics courses.

In this paper I propose that teaching statistical mechanics in introductory college physics courses is not at all difficult if we employ a proper method. The method I propose in this paper is a simulation of dice throwing on computers. Dice are familiar objects in our daily lives and every student has experience with throwing dice. Students can learn how thermodynamic principles emerge if they throw a large number of dice, say tens of thousands of dice. Computers came of age to let students throw many dice. With the aid of computer simulations, students can easily understand the statistical foundations of thermal physics.

A Chinese proverb says:² “I hear, I forget; I see, I remember; I do, I understand.”

There is no better way than letting students run the simulations themselves and thereby gain understanding. I wrote all the programs discussed in this paper to run in a JAVA applet and they can be viewed and run from my web site, “<http://phys.snu.ac.kr/howto/>” with a JAVA enabled browser such as NETSCAPE (Version 4.04 with patches or a later version) or MS INTERNET EXPLORER (4.0 or later).

Students should be aware of the fact that the numbers we deal with in statistical mechanics are huge and are not even comparable with astronomical numbers, as we demonstrate in the conclusion of this paper. This fact is intimately connected to a core principle of thermodynamics, namely irreversibility. This fact is also important from a mathematical viewpoint. For example, if we throw N dice simultaneously, the number of possible outcomes is $W = 6^N$. To students who are accustomed to linear quantities, the behavior of an exponential quantity, W , for large N , e.g.,

$$W \cong NW \cong \frac{W}{N}, \quad (3)$$

may seem paradoxical. However, the validity of an approximate relationship (3) within an error of $O(\ln N/N)$ for a quantity $\ln W/N$ is easily tested for a thermodynamic system where a typical value of N is of order 10^{23} . This is because $\ln W$ is proportional to N in general and we are interested only in the specific quantity $s_N = \ln W/N$, which becomes a finite quantity $s_\infty = \lim_{N \rightarrow \infty} s_N$ in the thermodynamic limit. In this limit the approximate relationship (3) merely implies that $s_N \cong s_\infty \pm \ln N/N$, which is certainly valid for large N such as 10^{23} .

However, for small numbers such as $N=3$ or 4 the approximation (3) is certainly nonsense. The validity of statistical mechanical arguments begins at the point where the size of our system becomes large enough to tolerate this approximate relationship (3). Students can learn this fact through computer simulations.

A system of dice is a good analogy to a physical system made of N paramagnetic atoms of spin $J=5/2$ in an external magnetic field.³ Atoms assume $2J+1=6$ equally spaced energy states. We can make the spacing exactly unity by controlling the external magnetic field. Students need not know the details of paramagnetic atoms. It suffices to say that dice represent atoms that assume only six equally spaced energy levels. Therefore, the total value of the dice is the total energy of the system. In all our simulation programs except the first one, we label the value of the dice from 0.

In these simulation programs I emphasize that there are only two principles involved: the conservation of energy which is inherited from mechanics and the ‘‘principle of equal *a priori* probability’’ of accessible microscopic states consistent with the macroscopic specification. We will put in the conservation of energy by hand. This means we are considering the microcanonical ensemble.

In this manner the first law is not something new but a simple application of the conservation of energy to the thermal system. In order to implement the ‘‘principle of equal *a priori* probability,’’ we use the microcanonical Monte Carlo (MC) technique first introduced by M. Creutz⁴ and further elaborated by myself.⁵

The microcanonical MC technique is a simple device which allows us to scan uniformly all accessible microscopic states of fixed total value or total energy.

We generalize the common die to a die of ν faces where ν can be any of the integer numbers, 2,3,...,6,... Although dice with an arbitrary number of faces do not exist in the real world, one can imagine dice in the shape of a tetrahedron ($\nu=4$), octahedron ($\nu=8$), etc. The dice of a polyhedron of order ν may be considered as atoms with ν equally spaced energy levels.

Other parameters of the system that we can control in the simulation programs are the number of dice N and the total value of the dice T , which is the total energy E for the corresponding analogous physical system.

The number N is restricted to a square of an integer L for programming convenience. L ranges from 2 to any number restricted only by computer resources. However, for a value of L larger than 200 ($N > 40\,000$) the dice configurations will not be shown.

Noninteracting atoms cannot relax to a thermal equilibrium state like an ideal gas. We need to introduce a weak

interaction. The microcanonical MC technique used in the simulation program introduces a mechanism which effectively amounts to physical interaction between magnetic atoms. Suppose the spins of the magnetic atom couple to the phonons of the underlying lattice so weakly that only a single phonon interacts with one magnetic atom at a time. In the program, a single atom is randomly selected and exchanges energy with the phonon either by taking in or giving off energy according to a specified rule.

In Sec. II we will explain the microcanonical MC technique by way of simulating the rolling of dice. In order to make students feel at home, we throw only dice in the simulation programs and a phonon is represented by a demon⁴ who carries a bag of energy.

In Secs. III–V we present a dynamical entropy simulation, a statistical temperature simulation, and a Maxwell–Boltzmann distribution simulation. In all these simulations we merely throw dice and observe the consequences. In Sec. VI we present a summary and concluding remarks.

II. PRINCIPLE OF EQUAL *A PRIORI* PROBABILITY AND MICROCANONICAL MONTE CARLO TECHNIQUE

The fundamental postulate of equilibrium statistical mechanics is the principle of equal *a priori* probabilities.

To college freshmen it suffices to say that the internal molecular interaction is chaotic enough that all microscopic states consistent with the macroscopic specification are equally likely to occur. This is analogous to the simple postulate of assigning equal probabilities to heads and tails in the coin tossing experiment. The ultimate validity of the postulate can only be established by checking its predictions against experimental observations.

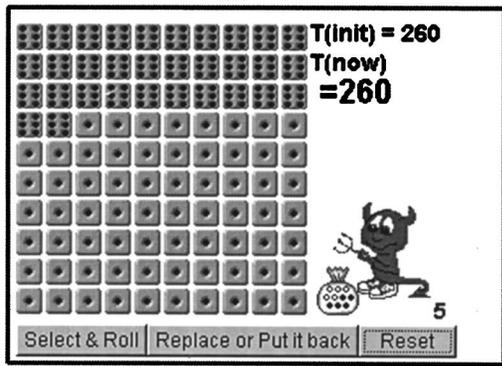
In order to realize the postulate we adopt the microcanonical Monte Carlo technique⁴ which has been successfully used for calculating the thermodynamic functions of a more complex system with precision and efficiency.⁵

We can do this by simulating a throw of N dice on a computer. The total value of the dice is designated by T , which is the sum of the values of the individual dice. We throw dice for a fixed value of T . Since T is the total energy of the analogous physical system, we are generating a microcanonical ensemble.

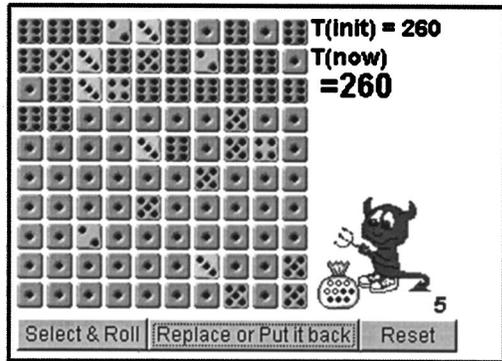
How can we throw N dice whose total value matches some predetermined value T ? In other words, how can we generate a random dice configuration for a given T ? We can think of throwing dice randomly until we get a throw whose total value matches T . This is impractical unless either the number of dice is very small or T is close to $3.5N$. (If we throw N dice randomly we will most likely get T values which are close to $3.5N$.) The microcanonical MC technique is one efficient method of generating random configurations of dice for a given fixed value of T .

For the purpose of illustration let us consider a system of $N=100$ dice and generate a microcanonical ensemble for $T=260$.

Our system consists of 100 dice and a demon who carries a bag with a maximum capacity of 10 units of value or energy. There are initially 5 units of energy in the demon’s bag, as in Fig. 1(a). Therefore the total value of the dice+demon system is 265 and this number remains fixed throughout the simulation. In other words, the system comprising dice and the demon is a closed system.



a



b

Fig. 1. (a) The initial configuration of APPLET 1. If we click the “Select & Roll” button we get a randomly selected die on the margin to the right-hand side and a new face as a result of the roll. The incremental or decremental amount in a value as a result of the roll is indicated. Examining the state of the bag, the decision whether to accept the new configuration or not is also indicated by “Yes” or “No.” By clicking the “Replace or Put it back” button we get a new configuration (if “Yes”) or retain the old one (if “No”). (b) A random dice configuration with $T=260$ which is obtained from the initial configuration of (a) after performing several MC steps.

Thermalization or the MC step consists of the following steps. (1) The demon selects one die out of 100 dice randomly and rolls it. (2) The new configuration generated by the rolled die makes the total value of dice system either increase or decrease. If the increment can be covered by the energy units in the bag or the decrement can be accommodated by enough vacancy in the bag, the new configuration is accepted. Otherwise the move is rejected and the die retains the old value and no new configuration is generated. (3) The demon repeats the same procedure, steps (1) and (2).

This MC procedure amounts to a random walk in the configurational space (analogous to the phase space for the Hamiltonian system) bounded by the energy band $254 < T < 266$, which is analogous to the energy shell for the continuous Hamiltonian system. We define one thermalizing MC step by the whole procedure until the new configuration for the given $T (= 260)$ is generated. This procedure satisfies the condition of detailed balance and can be made ergodic.⁴

It should be noted that the capacity of the bag is the minimum size needed to keep two successive configurations for the given T from being the same. At $T = 260$ the value in the bag is 5 units and there are 5 units of vacancy. Therefore any change in the value of a single die can always be accepted.

Sometimes the rolled dice turn up the old value again but it is still considered as a fresh configuration even though the configuration is unchanged. On the other hand, if the size of the bag is larger than 10, it will take a much longer time to return to the given $T = 260$ state, while for a bag of capacity smaller than 10, there is a chance that two successive configurations might be the same by rejecting the move in step (2) making the statistics deteriorate.⁵

Figure 1(b) is a typical configuration obtained after a few MC steps. In the following and subsequent sections all the simulations use these thermalizing MC steps, the demon will be hidden in the background, and configurations for $T \neq 260$ states will also be suppressed and hidden.

III. THERMAL RELAXATION AND CHANGE OF ENTROPY

Starting from the initial configuration given by the dice arrangement shown in Fig. 1(a) of Sec. II, we thermalize the system by performing the MC steps and watch how the configuration changes in time. The time is the number of MC steps described in Sec. II.

Now we calculate the entropy by Boltzmann’s celebrated formula given by Eq. (2). W in Eq. (2) is the number of configurations and k is Boltzmann’s constant, which defines the unit of thermodynamic entropy. In this paper we set $k = 1$.

From this point onward we relabel the faces of the dice as $\{0, 1, 2, 3, 4, 5\}$ and use different colors to represent these values instead of “dots.” This will make the total value T of Sec. II change to $T - N$.

If we take the view that the total value T is the only macroscopic parameter that specifies our macroscopic state, then $W = W(T)$, where $W(T)$ is the total number of configurations that have the total value T . Then the *equilibrium* entropy of the system is

$$S(T) = \ln W(T). \quad (4)$$

Let us subdivide the microscopic configuration into classes of configurations having a fixed set of values $\{n_0, n_1, n_2, n_3, n_4, n_5\}$, which are the numbers of dice showing 0, 1, 2, 3, 4, and 5, respectively. We will call these numbers “occupation numbers.” These classes may be considered to represent various nonequilibrium states of the system. We designate the numbers of configurations belonging to a fixed set of occupation numbers $\{n_0, n_1, n_2, n_3, n_4, n_5\}$ by $W(\{n_k\})$. Then

$$W(T) = \sum'_{\{n_k\}} W(\{n_k\}), \quad (5)$$

where the prime (') in the above summation implies that the sum is carried out over all possible sets of values $\{n_0, n_1, n_2, n_3, n_4, n_5\}$ consistent with the conditions

$$\sum_{k=0}^5 n_k = N \quad (6)$$

and

$$\sum_{k=0}^5 kn_k = T. \quad (7)$$

In expression (5) the number of summands is at most $(N + 1)^6$ since each n_k runs from 0 to at most N . This implies that at least one of the summands must be $O(\exp(N))$ if $W(T)$ is $O(\exp(N))$ in view of the approximate relations (3), since $6 \ln(N+1)/N$ is still a small number if N is sufficiently large. To be more specific let us consider the following inequalities:

$$W(\{\tilde{n}_k\}) \leq W(T) \leq (N+1)^6 W(\{\tilde{n}_k\}). \quad (8)$$

The inequalities are unchanged if we take the logarithm of the each term and divide by N since they are all positive numbers, i.e.,

$$\begin{aligned} \ln W(\{\tilde{n}_k\})/N &\leq \ln W(T)/N \\ &\leq 6 \ln(N+1)/N + \ln(\{\tilde{n}_k\})/N. \end{aligned} \quad (9)$$

$W(\{\tilde{n}_k\})$ in expression (8) is the largest term among the summands in Eq. (5)

This then allows us to replace the sum (5) by the single largest term, i.e.:

$$W(T) = W(\{\tilde{n}_k\}). \quad (10)$$

The above argument is only valid for equilibrium entropy. Let us now generalize this equilibrium entropy into a generalized entropy which can also represent a *nonequilibrium* state by

$$S^{\text{gen}} = \ln W(\{n_k\}). \quad (11)$$

It should be noted that because of Eq. (10), S^{gen} of Eq. (11) becomes the equilibrium entropy defined by Eq. (4) as our system relaxes to the equilibrium state.

Since the $W(\{n_k\})$ of the generalized entropy (11) is nothing but the multinomial coefficient given by

$$W(\{n_k\}) = \frac{N!}{n_0! n_1! \cdots n_5!}, \quad (12)$$

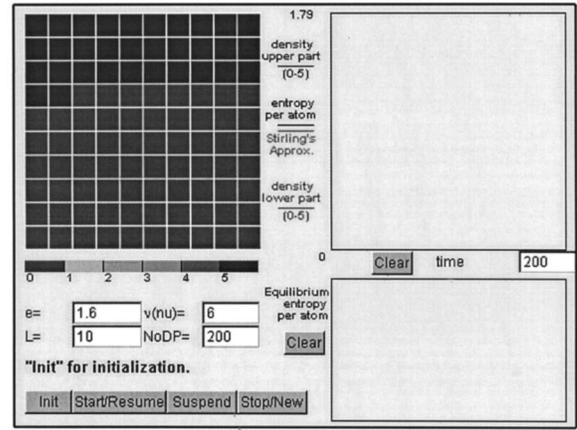
we can write the generalized entropy (11) as

$$S^{\text{gen}} = \ln N! - \sum_{k=0}^5 \ln n_k!. \quad (13)$$

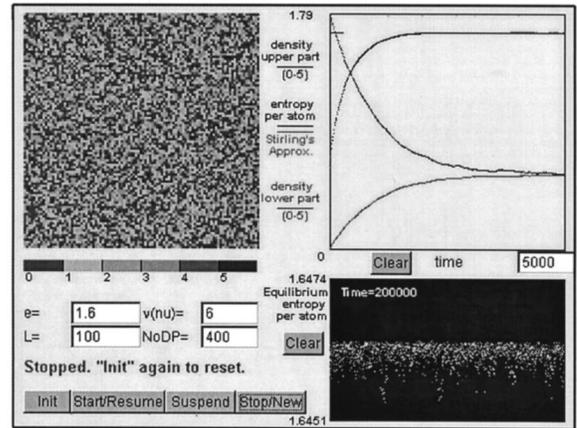
In the simulation programs, the entropy is calculated using formula (13). Figure 2(a) is the default initial configuration of APPLET 2, which is the same dice arrangement as in Fig. 1(a) and Fig. 2(b) is the typical simulation result of APPLET 2 with a different set of system parameters. We see in Fig. 2(b) that the entropy defined by Eq. (13) approaches the equilibrium entropy as time progresses. Students can watch how the configuration changes visually and the computer calculates the entropy at every instant and plots a graph.

As time progresses, we see the ‘‘entropy’’ increasing monotonically, eventually reaching a maximum value and remaining there ‘‘forever’’ apart from small fluctuations which tend to become smaller as the system size grows larger. This is the second law of thermodynamics!

APPLET 2 also calculates the energy density of the two regions, namely the upper region where the dice of value 6 (now 5) were initially occupied and the remaining lower region where dice of value 1 (now 0) were occupied. We will see that the MC steps randomize the configurations in both ‘‘energy’’ and ‘‘configurational’’ space. As the system relaxes to the equilibrium state, the energy density becomes uniform everywhere, manifesting one characteristic of single-phase equilibrium states.



a



b

Fig. 2. (a) The default initial configuration of APPLET 2 which is exactly the same configuration as shown in Fig. 1(a). If we click the ‘‘Start / Resume’’ button we generate a series of configurations such as are shown in Fig. 1(b). (b) Typical simulation result of APPLET 2. The scale of the energy density is (0.0–5.0). Initially the upper region starts out with the energy density 5.0 while the lower region starts out with 0.0.

In the lower rightmost corner in APPLET 2, the equilibrium entropies are plotted. The graph shows the thermal fluctuations in the equilibrium state on an enlarged scale with changing colors. The zoomed-in scale is indicated to the left of the original plotted region. Students can estimate the size of the fluctuations as the size of the system grows.

On the other hand, in the opposite limit where the number of dice is small, say 4, then the average period of return to the initial state is less than $6^4 = 1296$, which students can observe during the experiment. By estimating these probabilities, students can understand the meaning of irreversibility.

In APPLET 2, students can adjust three system variables, namely the energy density, e , equal to T/N , the size of the system using L , which is the square root of N , and the number of energy levels, ν , which is the number of faces of a die (6 for an ordinary die). ‘‘ e ’’ accepts a real number and is converted to the closest integer T by multiplying by N with the possibility of a loss of precision.

There are two more adjustable variables in APPLET 2: These are ‘‘time,’’ the number of thermalizing MC steps,

and the ‘‘NoDP,’’ which is the number of data points to display. Therefore, every ‘‘time/NoDP’’ MC steps, the data generated are displayed or plotted.

For a given T , the initial state is constructed in such a way that T is made of q dice of highest value $\nu-1$ and a single die of value r , where q is the quotient and r is the remainder, satisfying $T=(\nu-1)q+r$. Priority is determined such that dice of a higher value occupy the top and leftmost positions.

Finally the applet also plots the generalized entropy, calculated using the approximate formula

$$S^{\text{Stir}} = -N \sum_{k=0}^{\nu-1} \frac{n_k}{N} \ln \frac{n_k}{N}, \quad (14)$$

in a green color.

S^{Stir} is calculated using Stirling’s approximation for the factorial,

$$\ln n! = n \ln n - n,$$

which is made in the same spirit as the approximation (3). We plot this to show the range of the system size for which the statistical argument based on this approximation becomes valid.

The points to watch in this simulation are as follows.

(A) The visual characteristics of the thermal equilibrium state such as uniformity in various densities.

(B) The relaxation, irreversibility, and second law of thermodynamics. Although the algorithm described above is periodic (any simulation on the computer is periodic for that matter), if N is a moderately large number, say $N=100$, there is absolutely no chance that we will see the configuration return to the original unique arrangement or close to it. This is the irreversibility. It is the sheer enormous length of the recurrence period (Poincaré cycle) which makes reversibility impossible.

(C) The size dependence of fluctuations. Students can watch the fluctuations subside as the system size grows, keeping all other parameters intact. Students can watch the fluctuations in the entropy and local energy densities shrink proportionally to $1/\sqrt{N}$ as N grows.

(D) The entropy versus the energy. The entropy as a function of energy will be shown in Sec. IV. However, in this simulation, students may try various energy densities ‘‘ e ’’ to see how the equilibrium entropy changes as a function of energy. Since the energy levels of this system are finite, the equilibrium entropy does not always increase as the energy density increases. Some students may discover the concept of negative absolute temperature.

(E) The extensibility of the entropy. Students can watch how the entropy density changes or remains constant as the size of the system grows.

IV. THERMAL EQUILIBRIUM AND STATISTICAL TEMPERATURE

In Sec. III we presented a simulation program, APPLLET 2, which demonstrates the process of relaxation of a single system to the equilibrium state and the characteristics of the equilibrium state. In this section we present a simulation program that demonstrates the relaxation behavior of two systems which are thermally interacting with each other. From this point onward we denote the total value of the dice by E instead of T and T will be used to denote the absolute temperature.

Initially we separate the two systems **A** and **B** by a thermally insulating wall so that they cannot exchange energy. We prepare both systems in highly nonequilibrium states as described in Sec. III. When we thermalize the two systems with the insulating wall in place, then the relaxation processes are just the same as in the case of a single system. When we remove the wall, the two systems can exchange energy. The simulation program is designed in such a way that two independent microcanonical MC processes of fixed energies are performed when the two systems are insulated, alternating one MC step each for **A** and **B**. When the wall is removed we use the single microcanonical MC process. In this case we also perform MC steps alternately for **A** and **B**.

In APPLLET 3, we can separately control the sizes, N_A, N_B , the number of energy levels ν_A, ν_B , and the initial energies, E_A, E_B for the two systems **A** and **B**. The duration of the simulation and the number of data points are fixed.

During the simulation process, we can toggle the insulating wall on and off by clicking the check box. This will make the system relax to the final equilibrium state along many different paths.

The equilibrium condition for two systems in thermal contact which allow energy exchange between them is given by the maximum entropy principle. If we maximize total entropy

$$S = S_A(E_A) + S_B(E_B)$$

under the condition that the total energy of the combined system

$$E = E_A + E_B = \text{constant},$$

we obtain

$$\frac{dS_A}{dE_A} = \frac{dS_B}{dE_B}. \quad (15)$$

We emphasize the fact that the equilibrium condition for two systems in thermal contact is that the two slopes of the entropy as a function of energy should be equal. This simulation program explicitly shows that equilibrium is reached when the slopes of the two entropy functions S_A and S_B become equal.

Therefore, the statistical definition of the inverse temperature $1/T$ naturally follows as

$$\frac{1}{T} = \frac{dS}{dE}. \quad (16)$$

Of course, any single-valued function of the entropy slope may be defined as the temperature. However, the concept of absolute temperature was defined before statistical mechanics was discovered. In order to make the previously defined absolute temperature as in the Clausius entropy expression (1) coincide with the statistical definition, the temperature must be defined as in Eq. (16), as we all know.

This temperature can be calculated during the microcanonical MC simulation by examining how often the system visits the energy levels $E-1$ and $E+1$. If in a given period of simulation the frequency of visits are recorded as f_- and f_+ , then $S_+ - S_- = \ln(f_+/f_-)$, and T can be estimated by

$$\frac{1}{T} = \left(\ln \frac{f_+}{f_-} \right) / \Delta E, \quad (17)$$

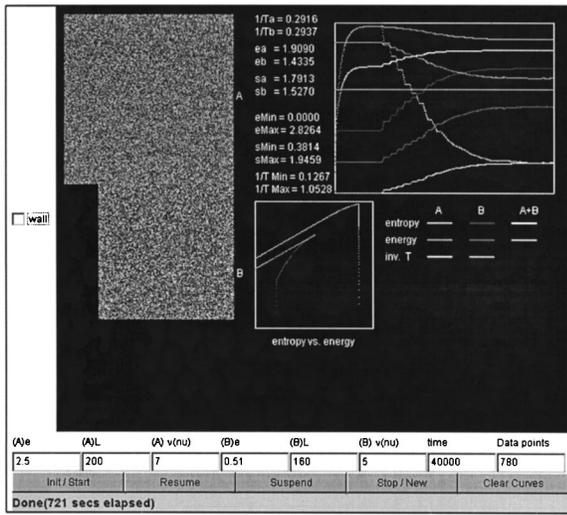


Fig. 3. Typical simulation result of APPLLET 3. The entropy and energy densities as well as the equilibrium canonical inverse temperatures for parts **A** and **B** are plotted in different colors in the same rectangle in the upper right-hand corner. The scales of these quantities are shown in the middle center area. The scale of the plotting area for the entropy vs energy is the same that of the main plotting area. At every instant the values of these data are also displayed during the simulation in the upper middle area. The total entropy and energy densities are simply $(S_A + S_B)/(N_A + N_B)$ and $(E_A + E_B)/(N_A + N_B)$.

where $\Delta E = 2$ in our case. This is the microcanonical temperature.

However, this method of estimating the temperature is time consuming since we have to collect about 1000 data points for f_{\pm} 's to give reasonable statistics at every instant. This makes the simulation almost crawl. Fortunately, we have an analytic expression for the temperature for large systems.³

In APPLLET 3 the inverse temperature $1/T (\equiv \beta)$ is the equilibrium canonical inverse temperature β calculated by solving the following equation for β for a given set of variables $e (= E/N)$ and ν :

$$e = \frac{\nu - 1}{2} - \frac{\nu}{2} \cosh \frac{\nu}{2} \beta + \frac{1}{2} \cosh \frac{\beta}{2}. \quad (18)$$

This is the canonical temperature for large systems. It was tested and verified that the canonical temperature agrees within the statistical error bar with the equilibrium microcanonical temperature calculated using Eq. (17) if the size of the system is large enough so that Stirling's approximation remains valid.

Figure 3 shows a typical simulation result of APPLLET 3. There are many steps for which the thermodynamic functions remain flat. These regions occur when the insulating wall is on so that each system is in a thermal equilibrium state separately.

The entropy as a function of energy is traced and plotted in a small box in the lower center area. At the end of the simulation, a straight line having a slope calculated by solving Eq. (18) is drawn which extends from the end points of the two curves. Students can visually confirm that these slopes are indeed the slopes of the simulated entropy functions at the final equilibrium state.

Points to watch in this simulation program are as follows.

(A) The zeroth law of thermodynamics.

Although we do not have three systems in thermal contact, we can simulate this indirectly by placing two different systems in the position of **B** in thermal contact with one common system, **A**.

(B) Students should watch the small box where the entropy as a function of energy is being traced during the simulation. It shows that the functions are concave downward. If these two systems start at the points where their slopes are different, then they reach thermal equilibrium when their slopes become equal. If they overshoot the equilibrium point, then it will make the total entropy decrease. It is a simple law of probability which keeps this from happening!

(C) Students can also make a rough estimate of the entropy change in the process of heat transfer, by calculating $dS_A = dE_A/T_A$ and $dS_B = dE_B/T_B$ together with the net change $dS = dE/(1/T_B - 1/T_A)$, where $dE = dE_B = -dE_A$.

This can be done by suspending the simulation just before removing the wall and after reinserting the wall. In this way, students can verify that the Boltzmann entropy is indeed the same quantity defined by Clausius (1), $dS = dQ/T$, where $dQ = dE$.

(D) Students can experiment with various paths leading to the thermal equilibrium state by superposing several curves to confirm that the final equilibrium state is indeed a single unique state.

V. MAXWELL-BOLTZMANN DISTRIBUTION

In most textbooks for introductory college physics courses, the authors introduce the Maxwell-Boltzmann (MB) distribution and discuss its consequences. However, none of them derives or even presents a plausible argument for the origin of the distribution.

If we throw and generate a truly *random configuration* of a sufficiently large number of dice, keeping the total value E fixed and examine the distribution of occupation numbers $\{\tilde{n}_k\}$, we get the MB distribution given by

$$\tilde{n}_r = A \exp\left(-\frac{r}{T}\right), \quad r = 0, 1, 2, \dots, 5. \quad (19)$$

Why? It is not so difficult to explain the reason to students in introductory physics courses if they can understand the concepts of probability, entropy, and temperature which we discussed and demonstrated in the previous sections.

Truly random configuration can be realized when the system is in thermal equilibrium. Let us concentrate on one die, say, in the upper leftmost corner. Let us keep performing MC steps after the system of dice reaches thermal equilibrium and then ask how often this particular die will show a value r . This frequency f_r must be proportional to the number of configurations of the rest of the dice having the total value $E - r$, i.e., $W_{N-1}(E - r)$. If N is sufficiently large, $W_{N-1}(E - r)$ can be approximated by $W_N(E - r) = W(E - r)$. Therefore,

$$\begin{aligned} f_r &\propto W(E - r) = \exp(S(E - r)) \\ &\cong \exp\left(S(E) - \frac{dS}{dE} r\right) \\ &= A \exp\left(-\frac{r}{T}\right). \end{aligned} \quad (20)$$

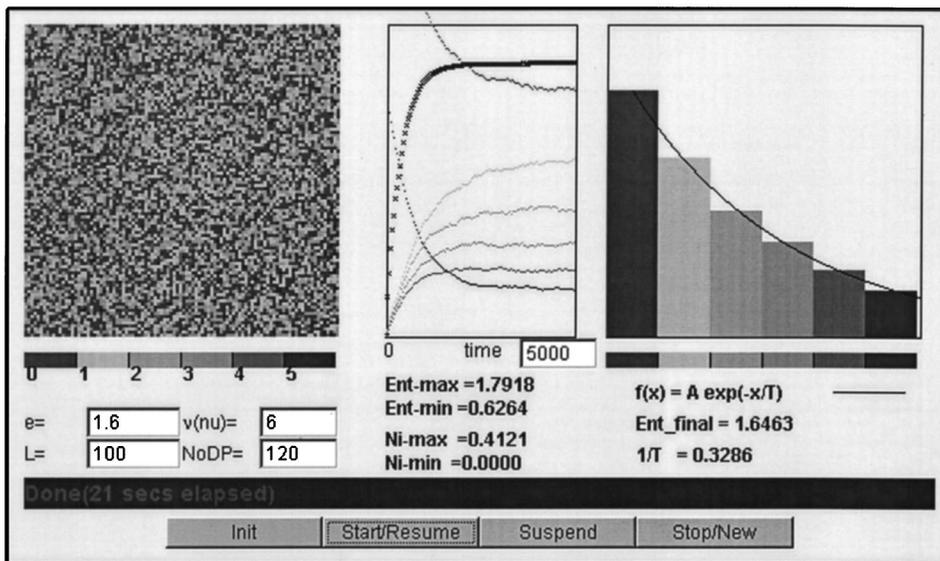


Fig. 4. Typical simulation result of APPLE 4. APPLE 4 is very much the same as APPLE 2. We trace the change of occupation number $\{n_k\}$ in time instead of the energy densities. At the end of the simulation, when equilibrium is reached, we plot the bar diagram of the $\{n_k\}$ using colors representing dice values.

In the above derivation we have used the definition of the temperature (16) and the Taylor series approximation $S(E-r) \cong S(E) - (dS/dE)r$, since $r \ll E$.

This is the standard argument for deriving the canonical distribution⁶ and is not much more difficult than what we have done so far.

In APPLE 4, instead of examining the frequency f_r of a single die, which is time consuming, we examine the distribution of occupation numbers $\{\tilde{n}_r\}$, since every die is equivalent and the outcome of the other dice may be considered as a manifestation of the single die in question in the time series. In other words, this is a demonstration of the ergodic hypothesis that the time average is the ensemble average!

Figure 4 shows a typical simulation result of APPLE 4. In this simulation program, we trace the distribution of occupation numbers from the highly nonequilibrium state to the equilibrium state.

The curve drawn using the cross marks “×” in the middle rectangle is the entropy defined by Eq. (13). It can be used to determine if the system has reached the equilibrium state.

The temperature used for drawing the MB distribution curve given by Eq. (19) is the canonical temperature defined in Eq. (18) and the constant A is determined by the normalization condition (6).

Points to watch in this simulation program are as follows.

(A) Students can perform the experiments for systems with a small number of dice and then increase the size gradually to see from what size the MB distribution is applicable.

(B) Students can watch the fluctuations subside as the size of the system grows.

(C) The entropy of the equilibrium state is represented by a single term of a set of occupation numbers as we saw in Eq. (10). Therefore, $\{\tilde{n}_k\}$ of Eq. (10) are the MB distribution. Students may discover that the MB distribution can also be derived by maximizing S^{gen} (13) under conditions (6) and (7).

VI. CONCLUSION

In this simulation program, I made the point that the statistical foundation of thermal physics is not so difficult to

teach to students in introductory physics courses if we have some simple examples to demonstrate or let students experiment by themselves.

The postulate of equal *a priori* probabilities is an easy concept to teach. Energy conservation is not a new concept. The crux is the realization of the postulate. However, if we use the microcanonical MC technique, we can easily implement the postulate. Without the microcanonical MC technique it would indeed be difficult to generate a truly random configuration for fixed E .

The microcanonical MC is not a difficult technique to follow, as we demonstrated in Sec. II. Once students understand how to generate a truly random configuration for a given fixed total value, E , the simulations that follow are simple applications of these microcanonical MC steps.

Students can easily grasp the laws of thermodynamics, especially the second law together with the concept of irreversibility. The number of configurations even for 100 dice is $6^{100} > 10^{78}$ which far exceeds the age of the universe in seconds which is less than 10^{18} . Therefore it is impossible even for a system of 100 dice in thermal equilibrium to return to the initially prepared highly nonequilibrium state. Even if a computer generates a million configurations a second, the odds of getting the initially prepared nonequilibrium state in a time equal to the age of universe is still 10^{-54} !

Students can also learn and understand that the entropy becomes a maximum at thermal equilibrium and never decreases, apart from fluctuations which will become smaller as the system size grows.

From the maximum entropy principle, it is easy to understand the condition for two systems to come into a thermal equilibrium state. It occurs when the two slopes of the entropy as a function of energy become equal to each other. From this understanding it is natural to introduce the statistical inverse temperature as the slope of the entropy as a function of energy.

The simulation also makes students understand the true meaning of Clausius’s definition of entropy. It is more natural to take the view that Clausius’s entropy formula (1) defines the absolute temperature rather than the entropy.

In most introductory college textbooks, the MB distribution is introduced without derivation. By simply throwing many dice, it is easy to demonstrate and understand why occupation numbers must follow the MB distribution.

In conclusion, it is not so difficult to teach the statistical foundations of thermal physics in introductory college physics courses if we know how to teach statistical concepts. The microcanonical MC technique is one such “how-to.”

ACKNOWLEDGMENTS

This work was supported in part by the Ministry of Education, Republic of Korea through a grant to the Research Institute for Basic Sciences, Seoul National University, and in part by the Korea Science Foundation through a research grant to the Center for Theoretical Physics, Seoul National University.

¹Entropy is discussed in any textbook on thermal and statistical physics. However a web site “Entropy on the World Wide Web” (<http://www.math.washington.edu/~hillman/entropy.html>) contains a variety of useful information on “entropy” used in many different fields as well as links to entropy-related software students may find interesting.

²M. Johnson, D. Johnson, and K. Lee, “A Pedestrian Guide To Reforming The Introductory Physics Course,” <http://www.phy.duke.edu/~mark/reform/>.

³F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, New York, 1965), pp. 257–262.

⁴Michael Creutz, “Microcanonical Monte Carlo Simulation,” *Phys. Rev. Lett.* **50**, 1411–1414 (1983).

⁵The statistical error for N independent samples is proportional to $1/\sqrt{N}$. In taking N samples, if m times are rejected then $N-m$ samples are truly independent. Therefore the statistical error is increased by the factor of $\sqrt{N/(N-m)}$. For further detail see: Koo-Chul Lee, “A new efficient Monte Carlo technique,” *J. Phys. A* **23**, 2087–2106 (1990); “Rejection-free Monte Carlo technique,” **28**, 4835–4842 (1995).

⁶See Ref. 3, pp. 202–206.

DOWNRIGHT PANSIES

Assertiveness and single-mindedness are selected for in every field of science (and business). Only a physicist would be arrogant (or naive) enough to think that physicists are any more aggressive than biologists.

I know this from experience. I did my PhD in experimental low temperature physics at Cornell University, transitioned to biology, and helped launch one of the major Genome Centers for the Human Genome Project. As a result, I have known some of the biggest names in physics and biology. I can assure you that, compared to some of the people that I have dealt with in genomics (or medical genetics), physicists are downright pansies. For that matter, when it comes to aggressiveness, none of these scientists can hold a candle to some of the pharmaceutical executives and venture capitalists that I have dealt with.

Gane Ka-Shu Wang, *APS News* **9** (3), 4 (2000).