

Understanding The Second Law.

Addison Ault, Department of Chemistry, Cornell College, Mount Vernon, IA

The business of science is the discovery of the “rules of nature”. Two of the most important of these rules are the first and second laws of thermodynamics. The first law states that only those changes *can* occur in which energy is conserved. The second law states that only those changes *will* occur in which entropy increases.

The most famous statements of these two rules were made in 1865 by Rudolf Clausius (1822-1888). Clausius reached these conclusions through his consideration of Carnot’s analysis of heat engines:

“Die Energie der Welt ist Konstant. Die Entropie der Welt strebt einem Maximum zu.”

“The energy of the world is constant. The entropy of the world tends toward a maximum.”

The first sentence recognizes the conservation of energy. The second sentence recognizes the directionality of change.

The first law

We can easily accept conservation laws. We easily understand the conservation of mass, for example, probably because we see many instances of the conservation of mass in everyday life. It is not hard to extend that idea to atoms. The conservation of energy is not so obvious, but if we look closely and think carefully we can find many examples of the conservation of energy in everyday life. Again, it is not hard to extend that idea to the behavior of atoms and molecules.

The second law

The second law is more subtle. Things happen, but the connection to an

increase in entropy is not obvious.

The efficiency of a heat engine

An alternative statement of the second law is that it is *impossible in a cyclic process to completely* convert thermal energy to mechanical energy.

Sadi Carnot (1796-1832) showed that the maximum efficiency of a steam engine (or any heat engine) is *entirely* determined by the *difference in temperature* between the source of the heat that makes the steam, and the sink for the heat released by the condensation of the steam.

Carnot realized, for example, that if 6,000 units of thermal energy is *taken in* from a source at 600 K only half as much thermal energy, 3,000 units, need be *transferred out* to a heat sink at 300 K in order that the two thermal transfers involve the *same amount of entropy*. Thus the original entropy “debt” of 6,000 energy units / 600 K = 10 entropy units due to the import of thermal energy can be “paid back” with only *half* the original amount of thermal energy, 3,000 energy units / 300 K = 10 entropy units. The balance of the original input of thermal energy in a maximally efficient heat engine, the other 3,000 energy units in this example, is available for conversion to mechanical energy, or “work”. The maximum possible efficiency of a heat engine can now be represented in this way:

$$\text{Maximum possible efficiency} = \frac{\text{work}_{\text{out}}}{\text{heat}_{\text{in}}} = \frac{T_{\text{high}} - T_{\text{low}}}{T_{\text{high}}} = 1/2 = 50\%$$

$$\text{Maximum possible efficiency} = \frac{3000}{6000} = \frac{600 - 300}{600} = 1/2 = 50\%$$

Carnot, as was typical of the time, believed that heat was a substance, and that its “fall” from a higher temperature to a lower temperature was analogous to the fall of water from a higher level to a lower level. It was actually Clausius

who realized that during a “heat fall” to a lower temperature some of the thermal energy is removed and converted to mechanical energy.

Other consequences of the second law

Imagine that you make some iced tea; let tea bags stand in water for a while. Do you worry that if you let the mixture stand for too long the color will go back into the bag?

Or imagine that you and your friend each has a glass of iced tea. You put a cube of ice in your glass; your friend does not. Imagine that both glasses are *well but not perfectly insulated*. What will be the most likely state of things at the end of the day? Will there be an ice cube in your glass only? Will there be an ice cube in your friend's glass only? Will there be an ice cube in each glass? Will there be an ice cube in neither glass? How will you decide?

How do we know which is the way to equilibrium?

I believe we all know the answers to those questions. We might say “That’s what happens.” Or we might use the word entropy in our answer, and hope that no one asks “What is entropy?” Or we might say that the final state, the equilibrium state, is the state of *maximum probability*. We would certainly be asked for an explanation.

Understanding the second law

Hoping to be able to understand the connection between the equilibrium state, which can be determined by experiment, and the state of maximum probability, which, in principle, can be determined by calculation from the properties of molecules, I decided to read some books.

What did I learn from the books?

Microstates and macrostates

First, I learned about microstates and macrostates. It was Feynman who said “We measure [entropy] by the number of ways that the insides can be

arranged, so that from the outside it looks the same.” The “insides” are called *microstates*, and they are the atoms and molecules of the system as distributed over their accessible energy levels. It is the *most probable distribution* that is responsible for the macroscopic properties of the system at equilibrium, its “state” or *macrostate*, or how it looks from the outside.

P, V, T, S, and E versus Ω , W, and Q

While the important quantities for classical thermodynamics are Pressure, Volume, Temperature, Entropy, and Energy, or P, V, T, S, and E, the most important numbers for statistical thermodynamics are Ω , W, and Q where Ω is the *total number of microstates* for a macrostate, W is the *number of microstates* for any *one distribution* of a macrostate, and Q, the partition function, is the *total of states effectively accessible* to the atoms or molecules of a system at equilibrium (Dill and Bromberg, p. 178).

The partition function, Q

We will see that the partition function, Q, is a measure of the manner and degree to which the particles of a system are spread out, or *distributed*, over possible energy levels. We will also see that the partition function, Q, depends upon the temperature, T, increasing with increasing T, and that equilibria favor states for which Q is large.

Ω , W, and Q are dimensionless

Each of these three numbers is dimensionless, just as dozen is dimensionless, and just as Avogadro’s number, the “chemist’s dozen” is dimensionless. The fact that these numbers are dimensionless makes it possible to reexpress each of them as the corresponding log.

The tomb of Ludwig Boltzmann (1844-1906)

Boltzmann’s tomb bears the following inscription.

$$S = k \log W$$

It was Boltzmann's genius that led him to believe that the thermodynamic entropy, S , should be *proportional* to the log of the largest number of microstates for any macrostate. Since W is dimensionless, as is $\log W$, we see that it is k , the *constant of proportionality*, known as Boltzmann's constant, that converts the dimensionless number $\log W$ to S , to the thermodynamic entropy, which has the dimensions of energy per degree.

Why should it be log W?

The thermodynamic entropy, S , is an *extensive* property. That is, a larger sample will have a larger entropy than a smaller sample. In contrast, the temperature, T , is an *intensive* property; a larger sample will have the same temperature as a smaller sample. Thus since thermodynamic entropy is extensive, its measure must be additive. It is the *log* of W that is additive (Bent, page 138; Dill and Bromberg, page 84). Although combinations of possible outcomes, or microstates, are multiplicative, and W is a measure of possible outcomes, the logs of the combinations, are additive.

Another point of view

Sometimes it can be helpful to separate the purely statistical part of statistical thermodynamics from its application to chemical systems. Thus the relationship between thermodynamic entropy and the number of microstates can be rewritten as

$$\frac{S}{k} = \log W$$

This equation separates a quite general statistical formulation, $\log W$, from the specific application to thermodynamics, thermodynamic entropy. Other specific applications, we are told, include reconstruction of satellite photos or

elimination noise from electronic signals. In all of these applications the goal is to make maximally useful (unbiased) inferences from insufficient information. We shall see that in chemistry this application is made by way of the Boltzmann distribution and the partition function, Q .

A deeper interpretation

Dill and Bromberg go on to say (Dill and Bromberg, pages 99-101) that “For more than a century, entropy has been a controversial concept. One issue is whether entropy is a tool for making predictions based on the knowledge of an observer, or whether it [has an existence that] is independent of an observer.” These two possibilities can be represented by these *alternative forms of the same relationship*.

$$\frac{S}{k} = \ln W \qquad \frac{S}{k} = -\sum_i p_i \ln p_i$$

The first expression emphasizes entropy as a counting problem. According to Schrödinger, page 1, “There is, essentially, only one problem in statistical thermodynamics: the *distribution* of a given amount of energy E over N identical systems. Or perhaps better: to determine the *distribution* of a assembly of N identical systems over the possible states in which this assembly can find itself, given that the energy of the assembly is a constant E .”

The second expression embodies the view that the *distribution* of particles over energies is determined by and reflects constraints on a system, constraints such as constant temperature, or constant energy, or constant volume, or constant number of particles. The system, through its probabilities, or p_i s, will adjust itself to maximize S , and, as this happens, S can be seen as a force of nature. In fact, for thermodynamics, this *is* the force of nature. This is *why* things happen: the probabilities, or multiplicities, will

maximize.

Dill and Bromberg show on page 82 that these relationships are two ways of saying the same thing with a different emphasis.

A set of microstates is a distribution

A simple example of a *distribution*, or set of microstates, would be the set of ways in which the atoms of a monoatomic ideal gas might be distributed over the accessible translational energy levels.

While there are many *conceivable* ways in which atoms might be distributed over accessible energy levels, only those ways that meet the following two conditions are considered to be *acceptable* ways.

The two conditions are 1) the sum of the *number of atoms* of each energy must equal the total number of atoms, and 2) the sum of the *energies of each atom* must equal the total energy of all the atoms. Not every conceivable way will meet the standards, but those that do meet the standards are acceptable and are therefore *possible* microstates. The distribution with the largest number, the largest set, of acceptable microstates is the most probable distribution.

Counting the microstates

The primary goal of statistical mechanics to identify, classify, and count the microstates. We will start with the number of acceptable microstates in *all* distributions, Ω , and we will then consider the number of acceptable microstates in any *one* distribution, W .

The microstates in all the distributions, Ω

If the energy levels are evenly spaced, the number of microstates in *all the distributions* is given by Equation 1 where N is the number of particles in the system and n is the number of units of energy, ϵ , possessed by the particles of the system (Bent, page 151; Dill and Bromberg, page 13).

$$\Omega = \frac{(n + N - 1)!}{(N - 1)! n!} \quad \text{Eq. 1}$$

Since the total energy of the system is E , the number of units of energy, n , equals E divided by the size of each unit of energy, which is ϵ . Thus $n = E/\epsilon$. Replacement of n by E/ϵ in Equation 1 gives Equation 1a (Bent, page 151).

$$\Omega = \frac{\left(\frac{E}{\epsilon} + N - 1\right)!}{(N - 1)! \frac{E}{\epsilon}!} \quad \text{Eq. 1a}$$

Equation 1a indicates that the total of all microstates is *entirely* determined by the *number* of particles, N , and the *ratio* between the total energy of all the particles, E , and the spacing, ϵ , between the energy levels of the particles.

The microstates in a particular distribution, W

The number of microstates in a *particular distribution* is given by Equation 2 where N is the number of particles and the terms in the denominator indicate the number of particles at each level of energy in a particular distribution (Bent, page 163).

$$W = \frac{N!}{a_1! a_2! a_3! \dots a_i!} \quad \text{Eq. 2}$$

A very simple example

If 6 quanta of energy are to be distributed over 6 particles we will see that there are 462 acceptable ways to do this (462 microstates) over 11 possible distributions. That is, for this system $\Omega = 462$.

We will also see that the 11 possible distributions, or sets of microstates, have the following individual weights, or multiplicities, or numbers of microstates, or values of W : 6, 30, 30, 60, 120, 90, 60, 30, 20, 15, and 1. The

sum of these weights equals 462.

Finally, we will see that the distribution with the most microstates, the most probable distribution, has 120 microstates. This is shown in **Table 1**.

6	1											
5	0	1										
4	0	0	1	1								
3	0	0	0	0	1		1				2	
2	0	0	1	0	1	2	0	1	3	0		
1	0	1	0	2	1	2	3	4	0	0	6	
0	5	4	4	3	3	2	2	1	3	4	0	
W	6	30	30	60	120	90	60	30	20	15	1	

Table 1

The column on the left indicates the possible quantum states for a system that has 6 particles and 6 quanta of energy.

The other 11 columns indicate the 11 possible sets of *occupation numbers* that meet the two conditions. The conditions are 1) the *sum of the numbers* of particles of each energy, the sum of the occupation numbers, must equal 6, and 2) that the *sum of the energies* of the 6 particles must equal the total energy of the system, which is 6 quanta.

The bottom line indicates W, the multiplicity, or weight, or number of acceptable microstates for each of the 11 sets. In this example the sum of the 11 values for W equals 462. $\Omega = 462$.

The most *likely distribution* is 3 particles in the ground state, one particle with 1 quantum, 1 particle with 2 quanta, and 1 particle with three quanta.

This is because there are 120 ways that this can happen. There are 120 microstates for this distribution out of a total of 462 microstates for all the distributions. Thus $(120/462) \times 100 = 26\%$ of the time this system of 6 atoms will have its 6 quanta of energy distributed as just described.

Other, less probable, distributions

The first distribution, where 1 particle has all the energy, could be realized in 6 ways; any one of the 6 particles could be the one with all the energy. Equation 2 for this state would be

$$W = \frac{6!}{5! 1!} = 6$$

The last distribution, where each particle has 1 quantum of energy can be realized in only 1 way. Equation 2 for this state would be

$$W = \frac{6!}{6!} = 1$$

The most probable distribution, described above, has a weight or multiplicity of 120. Equation 2 for this state is

$$W = \frac{6!}{3! 1! 1! 1!} = 120$$

The general pattern is the more “spread out” the quanta the greater the multiplicity, or weight, or number of microstates, or probability, or value for W . More specifically, W will be larger when more states are occupied. The reason for this is that the value of the numerator of Equation 2 will be the same for each distribution (it will be $N!$), but the value of the denominator will depend upon the particular distribution. The factorials of the denominator will be more numerous but smaller individually and collectively when more states are occupied. A smaller denominator will lead to a larger

value for W .

What happens when N is sufficiently enormously large

This very simple example accurately represents the general approach to counting for statistical mechanics, but it does not indicate the outcome when extremely large numbers of particles are involved.

Six quanta distributed over 12 particles

When we consider more particles than quanta we see new possibilities. First, if we consider that 6 quanta are distributed over 12 particles, or $n=6$ and $N=12$, Equation 1 becomes

$$\Omega = \frac{(6 + 12 - 1)!}{(12 - 1)! 6!} = \frac{(17)!}{(11)! 6!} = \frac{8\,910\,720}{720} = 12,376$$

The number of acceptable microstates increases from 462 for $N=6$ to 12,376 for $N=12$.

Second, if we consider that 6 quanta are distributed over 12 particles, or $n=6$ and $N=12$, Equation 2 for the distribution in which 6 particles each have one quantum and 6 particles are in the ground state becomes

$$W = \frac{N!}{a_1! a_2! a_3! \dots a_i!} = \frac{12!}{6! 6!} = \frac{665\,280}{720} = 924$$

The number of microstates for this most spread out distribution has gone from 1, when the number of particles is equal to the number of quanta to 924, when the number of particles is twice the number of quanta. If the number of particles goes to three times the number of quanta, the number of microstates for this most spread out distribution goes to 18,564.

The most amazing consequence of large N : W can equal Ω

When the numbers of atoms and quanta are *sufficiently enormously large*,

of the order of magnitude of, say 10^{23} or Avogadro's number, the number of microstates, \mathbf{W} , for the *most probable* macrostate is essentially *equal* to $\mathbf{\Omega}$, the *total of all* microstates for that macrostate. It is as if the "margin of victory" is equal to the number of votes cast. At equilibrium the system will be in its most probable distribution for 100.000% of the time. This, again, is due to the enormous largeness of the numbers of particles and quanta that are involved, numbers on the order of 10,000,000,000,000,000,000,000.

Back to the example of 6 quanta distributed over 12 particles

We can actually see this starting to happen in the example of 6 quanta, $n=6$, distributed over 12 particles, $N=12$.

If N were to become extremely enormously large, the numerator in Equation 1 after division by $(N-1)!$ would become $(N+5)(N+4)(N+3)(N+2)(N+1)N$ and can be approximated by N^6 , and $\mathbf{\Omega}$ would become $N^6 / 720$.

$$\mathbf{\Omega} = \frac{(n + N - 1)!}{(N - 1)! n!} \quad \text{Eq. 1}$$

Similarly if N were to become extremely enormously large, the numerator in Equation 2 after division by $(N-6)!$ would become $N(N-1)(N-2)(N-3)(N-4)(N-5)$, which can *also* be approximated by N^6 , and \mathbf{W} would *also* become $N^6 / 720$.

$$\mathbf{W} = \frac{N!}{a_1! a_2! a_3! \dots a_i!} = \frac{N!}{(N-6)! 6!} \quad \text{Eq. 2}$$

Thus if N were on the order of 10^{20} , $\mathbf{\Omega}$ and \mathbf{W} would be equal and on the order of $10^{120} / 10^3 = 10^{117}$, an extremely large number.

A similar example is worked out in detail in Bent, pages 336 and 337, and in Davies, pages 35-39. This trend is also discussed in Nash, pages 7-11.

Relevance of the examples

Although the distributions in these examples are not Boltzmann

distributions and although they have involved only a small numbers of particles, N , or small numbers of quanta, n , and evenly spaced energy levels the examples accurately represent the quantum mechanical approach to counting microstates and thinking about distributions.

We turn next to the Boltzmann distribution.

The Boltzmann distribution

Boltzmann's insight was that the *most probable state*, the *equilibrium state* of any atomic / molecular system, corresponds to a *particular distribution of a unique nature*, which has come to be known as the Maxwell-Boltzmann, or Boltzmann distribution.

The Boltzmann distribution is the most probable distribution; it is the distribution of particles over energy states that will predominate at equilibrium. The Boltzmann distribution is also the distribution toward which any system will evolve if the system is not at equilibrium.

The nature of the Boltzmann distribution

The nature of the distribution is that the *populations of the states of the particles decrease* as the *energies of the states increase*. For a certain number of particles in the ground state, there will be a fraction of that number in the first excited state, a fraction of that fraction in the second excited state, and so on. More specifically, the trend of the decrease population is *exponential*.

Deriving the Boltzmann distribution

A derivation of the Boltzmann distribution is given in most books about statistical mechanics. Since the Boltzmann distribution will be the distribution with the largest multiplicity, W , a common approach is to look for the maximum value for W as a function of $a_1, a_2, a_3, \dots a_i$.

$$W = \frac{N!}{a_1! a_2! a_3! \dots a_i!} \quad \text{Eq. 2}$$

This is often done equivalently and more easily by way of the logarithmic form of Equation 2 in which the continuing products of the numerator and denominator become continuing sums of log terms that are represented by Stirling approximations.

We've seen for the example of 6 quanta distributed over 12 particles that the most probable distribution is the one in which no particle has more than one quantum of energy. The typical derivation of the Boltzmann distribution assumes that the number of particles, N , is finite, that the total energy, E , is finite, and that the energy quanta, ϵ , are of such a size that there are more quanta, n , than particles, N (that $E/\epsilon = n \gg N$). The constraints are, then, that the occupation numbers, a_i , must add up to N , and the individual energies ϵ_i , of the particles must add up to the total energy of the system, E .

Mathematically the problem is presented as: maximize $\log W$ in this expression.

$$\log W - \alpha \sum a_i - \beta \sum a_i \epsilon_i$$

That is: find an α and a β such that when the α and β terms are subtracted from $\log W$ the net value for the expression will be maximal. Thus one looks for the largest net value of W that remains after taking the α and β terms, the constraints, into account.

While at this point α and β are mathematical factors, the “Lagrange multipliers”, α and β each has a physical significance. The first, α , corresponds to a_0 , often represented as n_0 , the *number of particles in the ground state*. Alpha corresponds to the population constraint.

The second, β , represents a property that systems in equilibrium with one another must have in common. That common property is *temperature*, and β

is a function of temperature: $\beta = 1/kT$, where k is Boltzmann's constant and T is the thermodynamic temperature in Kelvins, K . The units of β are therefore $1/\text{energy}$. Beta corresponds to the energy constraint.

Limitations of the Boltzmann distribution

The Boltzmann distribution itself is actually quite limited in its application. We have specified already that the system must have a constant number of identical particles, each with its own "private" energy, the system must have a constant energy, and the system must be at equilibrium. We have also assumed that the energy levels do not change, which means that the system must have a constant volume since translational energy levels depend upon the volume of the sample. Thus the exact application of the Boltzmann distribution is limited to a low pressure, constant volume sample of an monatomic ideal gas near room temperature. Helium gas at room temperature and pressure could serve as an example.

We shall see, however, that when a system is at equilibrium, or close enough to equilibrium, the rules of molecular behavior exemplified by the Boltzmann distribution apply to everything.

Representing a Boltzmann distribution

We start by letting N represent the total number of particles, we let n_0 represent the number of particles in the ground state, we let n_1 represent the number in the first excited state, we let n_2 represent the number in the second excited state, and so on ...

$$N = n_0 + n_1 + n_2 + n_3 + n_4 \dots \quad \text{Eq. 3}$$

Next, we represent the numbers of particles in the excited states as *fractions* of the number in the ground state.

For the first excited state the fraction is $(1/e^{\epsilon_1/kT})$, where ϵ_1 is the energy of the first excited state. Thus the number of particles in the first excited

state is n_0 times this fraction: $n_0(1/e^{\epsilon_1/kT})$.

For the second excited state the fraction is $1/e^{\epsilon_2/kT}$. This second fraction is smaller than the first fraction because ϵ_2 is greater than ϵ_1 . Thus the number of particles in the second excited state is n_0 times this second fraction: $n_0(1/e^{\epsilon_2/kT})$.

Continuing in this way we convert Equation 3 into Equation 4.

$$N = n_0 + n_0(1/e^{\epsilon_1/kT}) + n_0(1/e^{\epsilon_2/kT}) + n_0(1/e^{\epsilon_3/kT}) + n_0(1/e^{\epsilon_4/kT}) \dots \quad \text{Eq. 4}$$

The terms on the right are the Boltzmann distribution.

What the Boltzmann distribution tells you (Dill and Bromberg, page 175)

“The Boltzmann distribution says that more particles will have low energies and fewer particles will have high energies. Why? Particles don’t have an intrinsic preference for lower energy levels. Fundamentally, all energy levels are equivalent. Rather, *there are more arrangements* [microstates] *of the system that way*. It is extremely unlikely that one particle would have such a high energy that it would leave all the others no energy. There are far more arrangements [microstates] in which most particles have energies that are relatively low, but non-zero. If each particle takes only a small fraction of the total, it leaves a great many more ways for the other particles to distribute the remaining energy.” (Dill and Bromberg, page 175)

The significance of the Boltzmann distribution

Since most real atomic / molecular systems will consist of a suitably enormously large number of particles, the weight, or multiplicity, or number of microstates, W , for this most probable distribution, the Boltzmann distribution, will equal Ω , and therefore the microstates of the Boltzmann distribution will include “all” of the microstates of the system: $W = \Omega$.

Also, as the *most probable distribution*, the Boltzmann distribution will represent the *equilibrium state* of the system, and thus the thermodynamic

entropy, S , will be known for this (equilibrium) state:

$$\text{Thermodynamic entropy} = S = k \log W$$

But that's not all

Starting with the Boltzmann distribution it is possible, in principle, to calculate the values of *all* of the thermodynamic properties of a chemical system when the system is at equilibrium.

These calculations are made by way of the partition function, Q .

The partition function, Q

The partition function is a *way of counting* the particles of a system when the particles are distributed over the available energy levels in accord with the Boltzmann distribution.

Counting for the partition function, Q

Start with Equation 4.

$$N = n_0 + n_0(1/e^{\epsilon_1/kT}) + n_0(1/e^{\epsilon_2/kT}) + n_0(1/e^{\epsilon_3/kT}) + n_0(1/e^{\epsilon_4/kT}) \dots \quad \text{Eq. 4}$$

The terms on the right are the Boltzmann distribution

First, factor out n_0 to give Equation 5.

$$N = n_0 (1 + 1/e^{\epsilon_1/kT} + 1/e^{\epsilon_2/kT} + 1/e^{\epsilon_3/kT} + 1/e^{\epsilon_4/kT} \dots) \quad \text{Eq. 5}$$

Then divide both sides of Equation 5 by n_0 to convert it to Equation 6.

$$\frac{N}{n_0} = 1 + 1/e^{\epsilon_1/kT} + 1/e^{\epsilon_2/kT} + 1/e^{\epsilon_3/kT} + 1/e^{\epsilon_4/kT} \dots \quad \text{Eq. 6}$$

The *sum of terms on the right* of Equation 6 is called the *partition function*.

When you let Q stand for the partition function you can rewrite Equation 6 in a very concise form, as Equation 7. As you can see, Q is dimensionless.

$$\frac{N}{n_0} = Q \quad \text{Eq. 7}$$

You can see also that the partition function indicates the degree to which all

states are “effectively accessible” (could be occupied) by particles of the system relative to the occupation of the ground state (Dill and Bromberg, p. 178).

The partition function is the *number of times larger* the total number of particles could be than the number of particles in the ground state.

While individual Boltzmann factors indicate the ratio of particles in *one* excited state relative to the ground state, the partition function does the same thing for *all* the states.

Another way to say it is that the partition function is a total number of particles expressed as the sum of the numbers of particles of each energy, with the population of the ground state being the unit of measure of population.

The partition function is a measure of the degree to which the particles can be spread out over, *or partitioned among*, the energy levels.

$$Q = 1 + 1/e^{\epsilon_1/kT} + 1/e^{\epsilon_2/kT} + 1/e^{\epsilon_3/kT} + 1/e^{\epsilon_4/kT} \dots \quad \text{Eq. 8}$$

The population of the ground state (the most highly populated state) is taken to be 1, and the populations of other energy states are expressed as fractions of the population of the ground state. Considered in this way the partition function can be seen as a sum of the relative occupancies of states. It is for this reason that the partition function is also called the sum-over-states, or *Zustandsumme*.

The size of the partition function

The *size* of the partition function depends upon the sizes of the fractional populations, and the size of each fraction depends upon both the *energy* of that excited state and the *temperature* of the system.

Energy dependence of Q

The size of the *fractional population* of each energy level depends upon the

energy of the level, and the fractional population of the level *decreases* as the energy, ϵ , of the level increases. This is because the energy, ϵ , is in the *numerator* of the exponent of e . A larger ϵ makes a larger exponent and a smaller fraction.

We can say this another way: the size of the partition function depends upon the *spacings* of the energy levels. The larger the spacings of the energy levels the smaller the partition function.

Temperature dependence of Q

The size of the fractional population of each energy level also depends upon the temperature of the system, and the *fractional population* of the level *increases* as the temperature, T , of the system increases. This is because the temperature, T , is in the *denominator* of the exponent of e . A larger T makes a smaller exponent and a larger fraction. The fraction, however, can never become as large as 1.¹

The range of the partition function

The lowest possible value for the partition function is 1, the value it would have if the temperature of the system were 0 K, the temperature at which *all* particles in the system would be in the ground state and the fractions in the excited states would be zero.

The value for the partition function at a high temperature could be very large, much larger, even, than the number of particles. At a high temperature the number of states available for occupancy, Q , could greatly exceed the number of occupants, N . At an infinite temperature, *assuming* an unlimited number of states, the probability for occupancy of *an* excited state infinitely exceeds the probability of occupancy of the ground state; as T approaches infinity the partition function, Q , also approaches infinity (Bent, page 352, answer to Problem 4).

If a system has only t states, however, Q will have a limiting value of t at infinite T or vanishing ϵ_i because all of the finite number of states, t , will be equally occupied (Dill and Bromberg, page 178).

Possibilities versus probabilities

The reason the partition function can have a value greater than the number of particles is that the partition function is fundamentally a sum of *possibilities*, as implied by the phrase “effectively accessible” or the phrase “could be occupied”. However, as soon as a number of particles is named, these possibilities are converted to probabilities by dividing by the number of particles. (Compare Dill and Bromberg, page 100).

An illustration

The following table, **Table 2**, illustrates the relationships between energy levels, the absolute temperature, the Boltzmann distribution, and the partition function for a sample of 10,000 particles.

In this example the energy levels are uniformly spaced, and the difference in energy between successive levels is kT . The first column lists the levels in order of increasing energy, from the bottom to the top. The second column gives the ratio of the energy of each level, ϵ , to kT . The third column gives the fractional population of each state relative to the population of the ground state. And the fourth column gives the Boltzmann distribution of 10,000 particles over these energy levels. The sum of the numbers in the third column, 1.58198, is Q ; the sum of the numbers in the fourth column, 10,000, is N ; and the number at the bottom of the fourth column, 6321, is n_0 .

You can see that the ratio of the total number of particles, N , to the number in the ground state, n_0 , does equal the partition function, Q .

$$\frac{N}{n_0} = Q$$

$$\frac{10,000}{6321} = 1.58198$$

Alternatively, you can see that the total number of particles, N , does equal the product of the number of particles in the ground state, n_0 , and the partition function, Q .

$$N = n_0Q$$

$$10,000 = (6321)(1.58198)$$

If we were to consider 10^{23} particles, rather than just 10^5 particles, each number in the two columns on the right would be 10^{18} times larger. You would then see that the 56th level would be the lowest level that did not contain an average of at least one particle (Davies, page 58).

If the spacing of the energy levels were ten times greater, only the first 6 levels would be populated, and if the energy levels were 100 times greater only the ground state would be populated (Davies, page 58),

The importance of the partition function

Quoting Andrews: “The sum Q has a central role in statistical thermodynamics because once it is known as a function of the variables on which it depends, all thermodynamic quantities may be calculated from it directly” (Andrews, page 35). Q is a function of the energies, ϵ , of the quantum states and of the absolute temperature, T . “The set of states in turn is found in principle by using quantum mechanics and is a function of only the number and nature of the particles comprising the system and the constraints” (Andrews, page 35).

The thermal energy, E , of N particles, for example, can be calculated with the help of the partition function.

The thermal energy, E , of N particles

The total thermal energy, E , of N particles is the sum of the thermal energies of each of the N particles. The approach is to determine the total energy by *multiplying* the *fraction* of particles of each energy state by that energy, and adding up the contribution of each set of particles.

The fraction of particles of a particular energy

For each energy state the *fraction* of particles in that energy state will be the number in the ground state, n_0 , times the fraction in each energy state relative to the number in the ground state, $e^{-\epsilon_i/kT}$.

$$n_0 (1/e^{\epsilon_i/kT}) = n_0 e^{-\epsilon_i/kT}$$

fraction of particles in state i relative to n_0

The fraction of particles in each state can be reexpressed in terms of N and Q since $Q = N / n_0$, and so $n_0 = N / Q$

$$n_0 e^{-\epsilon_i/kT} = \frac{N}{Q} e^{-\epsilon_i/kT}$$

fraction of particles in state i relative to n_0

The energy of the fraction of particles in each state

The *energy of the fraction* of particles of energy ϵ_i in state i is then ϵ_i times the fraction of particles in that state.

$$\frac{N}{Q} \epsilon_i e^{-\epsilon_i/kT}$$

the contribution of the fraction of particles in state i to the total thermal energy

The total thermal energy, E

The *total thermal energy* is the sum of the contributions of the fractions of particles of each energy.

$$E = \frac{N}{Q} \left(\varepsilon_0 e^{-\varepsilon_0/kT} + \varepsilon_1 e^{-\varepsilon_1/kT} + \varepsilon_2 e^{-\varepsilon_2/kT} + \varepsilon_3 e^{-\varepsilon_3/kT} + \varepsilon_4 e^{-\varepsilon_4/kT} \dots \right) \text{ Eq. 9}$$

We now have in Equation 9 an expression for the *total thermal energy*, E , of N particles that are distributed over a set of energy levels, ε_i according to the Boltzmann distribution. The thermal energy is thus a weighted average of the contributions of particles of each energy. The *contribution* is ε_i , and the *weight* is n_0 times the exponential factor. The common factor of n_0 has been factored out and expressed as N/Q . The terms of the sum depend only upon the *energies* of the particles and the *temperature* of the system. Either a closer spacing of the levels or a higher temperature will increase the thermal energy of the system.

The implications of Equation 9

The first thing to notice about Equation 9 is that the first term, the contribution to the *thermal energy* of the particles in the ground state, is 0. This is consistent with the concept of a *ground state*: a state of no *thermal energy*, but a state that can have zero-point energy.

The second thing to notice about Equation 9 is that the weighting factor, the exponential factor, falls off rapidly with increasing spacing of the energy levels. As the equal spacing of the levels in Table 1 indicates, while the energies of the excited states increase in a linear way the populations of the excited states decrease exponentially. The exponential always wins, and eventually the contributions of the higher terms become insignificant.

What if the spacings are not equal?

All the examples to this point have invoked a small number of generic and equally spaced energy levels. The reason for this is that with a small number of equally spaced level it is much easier to do the math, and to tell the story. The crucial thing is not that the levels are equally spaced or of a certain

number or of a certain type but that their weightings are *exponential*; that their weighting factor is of the form $e^{-\epsilon_i/kT}$; that their distribution over the possible energies is a *Boltzmann* distribution. It is then possible, when both the energies of the states, ϵ_i , and their weightings, $e^{-\epsilon_i/kT}$, are known, to determine the total thermal energy of the system, E , by adding together the weighted contributions of particles of each energy .

The contributions to thermal energy

Many particles can have more than one type of thermal energy, and therefore many particles can contribute more than one type of energy level to the system. We will consider gases at normal temperatures.

Translational energy and translational energy levels

All gas particles will have translational energy. The number of translational energy levels available at room temperature can be very large, perhaps on the order of 10^{20} . These translational energy levels will be extremely closely spaced, and their spacing, or energies, ϵ_i , will *depend upon the volume* available to the particles. It is for this reason that all the examples to this point have specified that the volume does not change. If the volume were to change, the *spacings* of the translational energy levels would change.

Modes of energy flow: heat and work

We need to mention at this point that there are exactly two ways by which energy can be added to or removed from a system. The first is by changing the *populations* of the energy levels of the system, and the second is by changing the spacing or the *energies* of the energy levels of the system.

Up until now we have considered only *changes in the populations* of the energy levels. That is, up until now we have considered only transfer of energy as *heat*.

The other of the two possible modes of energy transfer is *work*. The typical example of work is the action of a change in pressure to produce a change the volume, or vice-versa. For this type of work, the change at the atomic / molecular level is a change in the *spacings* of the translational energy levels.

Thus the internal energy of a system, E , can be changed by either of two possible adjustments to energy levels. The first adjustment is through a change in the *populations* of the energy levels, through a change in thermal energy alone, or q . The second adjustment is through a change in the *spacings* of the energy levels, through a change in work alone, or w .

Most changes in energy are a combination of these two modes of energy transfer, and this combination of effects can be represented in this way

$$E = q + w \quad \text{Eq. 10}$$

or in this way

$$dE = dq + dw \quad \text{Eq. 10a}$$

A change in E , the energy, is, in general, a result of changes in both the *populations* of energy levels (“ q effects”) and in the *spacing* of energy levels, (“ w effects”). A given change in energy can be the result of *many combinations* of q and w . In classical thermodynamics E , or the sum of q and w , are functions of state, while w and q individually are not state functions.

Because a change in internal energy, E , can be the result of a combination of changes in q and w , changes in E are often measured under conditions in which the observed change will be due entirely to a transfer of heat, q , or due entirely to a transfer of work, w . Under constant volume, or “no work” conditions, Equation 10 becomes $E = q$, and under adiabatic, or “no heat transfer” conditions, Equation 10 becomes $E = w$.

Constant volume has been specified or assumed in all our examples. This

is why the only atomic / molecular effects that had to be considered were changes in the *populations* of the energy levels. When $w = 0$ a change in the internal energy, E , is due entirely to q , a change in thermal energy.

Vibrational energy and vibrational energy levels

We now continue our consideration of molecular energy levels. In addition to translational energy and translational energy levels, molecules, or particles with more than one atom, can have vibrational energy and vibrational energy levels. These energy levels will be more closely spaced for molecules with heavy atoms or flexible bonds, or not so closely spaced for molecules with light atoms and stiff bonds. Lower vibrational energy levels are evenly spaced, and, since they are due to relationship *within* the molecule, the spacing does *not* change with a change in the volume of the sample. Figure 3-8 on page 67 of Davies presents five examples of vibrational spacings. These examples are based on experimental data.

Rotational energy and rotational energy levels

Molecules, or particles with more than one atom, can also have rotational energy and rotational energy levels. Rotational levels are much more closely spaced than are vibrational levels, and they are not evenly spaced, the spacing of the higher levels being slightly greater than that of lower levels. Figure 3-5 on page 62 of Davies presents five examples of rotational spacings. These examples are also based on experimental data. Again, since the spacings are determined by relationships *within* the molecule rotational spacings do not depend upon the volume of the sample.

Electronic energy and electronic energy levels

Electronic energy levels are typically high enough that no particles will be present in an electronically excited state at room temperature.

Summary of contributors to thermal energy at room temperature

We've just seen that particles can make more than one kind of contribution to the thermal energy of a gas at ordinary temperatures. All particles will contribute translational energy, polyatomic particles will also contribute both vibrational and rotational energy, but there will normally be no contribution of electronic energy. Correspondingly, accessible energy levels will include an enormous number of translational levels, a very much smaller number of rotational levels, a still smaller number of vibrational energy levels, and no electronic energy levels.

Quantum mechanics and statistical mechanics

It is the job of quantum mechanics to determine from the structure of the particles of the sample the types of energy levels that will exist and their spacings. It is then the job of statistical mechanics to determine how the particles of the sample will be distributed over the energies available to the molecules. When the system is at equilibrium, or close enough to equilibrium, the distribution will be a Boltzmann distribution.

Another way to express the internal energy

While expressing the thermal energy of a system as a sum of contributions from sets of particles provides an insight into the nature of thermal energy, there is a transform of Equation 9 that will provide insight as to why things happen.

Transforming Equation 9

It is the exponential nature of the Boltzmann distribution makes it possible to transform Equation 9

After rewriting Equation 9 as Equation 11 where the first term in the numerator, the term that represents the “thermal energy of the ground state” is zero, and the denominator, Q , is shown explicitly as its sum of terms you can see that the sum of terms in the numerator is the *derivative* with respect

to $1/kT$ of the sum of terms in the denominator.

$$E = N \frac{0 + \epsilon_1 e^{-\epsilon_1/kT} + \epsilon_2 e^{-\epsilon_2/kT} + \epsilon_3 e^{-\epsilon_3/kT} + \epsilon_4 e^{-\epsilon_4/kT} \dots}{1 + e^{-\epsilon_1/kT} + e^{-\epsilon_2/kT} + e^{-\epsilon_3/kT} + e^{-\epsilon_4/kT} \dots} \quad \text{Eq. 11}$$

Equation 11 is now an expression for the thermal energy, E , that has the form dx/x , and we know from calculus that dx/x equals $d \ln x$,

$$\frac{dx}{x} = d \ln x$$

We can therefore rewrite Equation 11 as Equation 12.

$$E = N \left(kT^2 \frac{\partial \ln Q}{\partial T} \right)_V \quad \text{Eq. 12}$$

or, for one mole of particles, when N equals Avogadro's number and $Nk = R$

$$E = \left(RT^2 \frac{\partial \ln Q}{\partial T} \right)_V \quad \text{Eq. 12a}$$

We see that the thermal energy, E , depends upon the temperature, T , and the partition function, Q , which depends upon the temperature and the (constant) spacing of the energy levels.

This interdependence can be seen more clearly by dividing both sides of Equation 12 by RT to get Equation 13.

$$\frac{E}{RT} = T \frac{\partial \ln Q}{\partial T} \quad \text{Eq. 13}$$

In Equation 13, E , the thermal energy per mole, is expressed as dimensionless multiples of RT , energy per mole. The right hand side of equation 13 is also dimensionless, as it is T times the rate of change of the dimensionless number $\ln Q$, with T .

A closer look at Equation 13

Equation 13 indicates that for a unit change in energy a cooler system (T is smaller) will have a larger rate of change of $\ln Q$ with T ($\partial \ln Q / \partial T$ is larger).

Or, saying it the other way around, Equation 13 indicates that if the same transfer of energy is made when the system is at a higher temperature, when T is larger, the rate of change of $\ln Q$ with a change in temperature, $\partial \ln Q / \partial T$, will be smaller.

The reason that two bricks come to the same temperature

Imagine that two identical bricks, one hot and one cold, are wrapped up together and then isolated from the rest of the world. We know that they will come to the same temperature. How do we account for this outcome?

Energy in the form of heat will move from the hot brick to the cold brick. The energy, E , lost by the hot brick will be exactly equal to the energy gained by the cold brick. Energy is conserved.

At the same time that the energy is transferred the multiplicity, Q , of the hot brick will decrease, and the multiplicity of the cold brick will increase. Unlike the energies, however, the changes in multiplicity will not be the same. Multiplicity, or Q , or W , or S , entropy, is not conserved. For the reasons outlined just above the decrease in multiplicity of the hotter brick as it cools will be less than the increase of multiplicity of the colder brick as it warms. The flow of energy and the changes in multiplicity will continue until the opposing rates of transfer of multiplicity become equal; energy will flow until the temperatures of the bricks become the same; until thermal equilibrium has been reached.

While there will be no net change in energy there will be a net increase in multiplicity. Energy is conserved; entropy increases.

Energy is transferred to maximize multiplicity

“The tendency for heat to flow is not always a tendency to equalize energy. It is a tendency to maximize multiplicity.” (Dill and Bromberg, page 46)

“The tendency toward maximum multiplicity is a tendency toward equal temperatures, not [a tendency toward] equal energies. (Dill and Bromberg, page 46).

Dill and Bromberg summarize the first and second laws in this way. “The First Law says that heat is a form of energy exchange, and that the sum of heat plus work, called the internal energy, is a quantity that obeys a conservation law. The First Law is a bookkeeping tool. It catalogs the balance of heat and work. It doesn’t tell us why heat flows. The Second Law says that systems tend toward their states of maximum multiplicity. Heat flows to maximize multiplicity.” (Dill and Bromberg, page 47).

The tale of two bricks is often told as a story of equalization of two temperatures, but it is better told as the maximization of multiplicity. The maximization story should be preferred because it provides the *reason* for equalization of temperatures. It is a *more probable* state for the temperatures to be equal. It is an *overwhelmingly more probable* state for the temperatures to be equal. It is actually *inevitable* that the temperatures become equal. It is a *certainty* that the temperatures will become equal (within the ability to detect any difference).

More examples of maximizing multiplicity

The flow of heat from a hot object to a cold object is only one example of a change that is driven by an increase in multiplicity. Two other examples include the expansion of one gas, and the mixing of two gases.

The expansion of a gas

The volume of a gas is its constraint. Relieve the constraint and the gas will expand. The gas will expand because the spacing of the translational

energy levels will thereby become less, and energy levels that were not previously “accessible” become “accessible”. The multiplicity of the system increases. Thus what we call “pressure” is an indication of the tendency for a gas to increase its multiplicity by changing to a state of higher probability.

As Dill and Bromberg say it “If the degree of freedom of the system is its volume, the particles will spread out into the largest possible volume to maximize the multiplicity of the system.” (Dill and Bromberg, page 33). They also provide a simple example that is based on a “lattice model”, page 332. The math is of the form $W = M!/N!(1-N)!$.

Changes in w , q , and E during expansion of a gas

Depending upon the conditions of the expansion of a gas there can be different changes in w , q , and E during the expansion.

A “free expansion”

If the expansion takes place “freely” into a vacuum against zero resistance no work is done; $w = 0$. If the gas is ideal no attractive forces will have been overcome during the expansion; $q = 0$. Thus in a free expansion of an ideal gas there will be no change in the internal energy: $w + q = 0$. The temperature will remain constant during the free expansion of an ideal gas. The only thing that “happens” is that the multiplicity of the system increases. That is “why” it happens.

PV work during an expansion

If a gas expands against a resisting pressure, work will be done by the gas; $w \neq 0$. If the expansion is done “slowly”, “reversibly”, “in a quasi-static manner”, so that a Boltzmann distribution of the molecules over energy levels is maintained, or nearly maintained, throughout the expansion, w can be calculated. If the expansion is reversible and *isothermal*, $+q$ will equal $-w$, and E will be unchanged. If the expansion is reversible and *adiabatic*, q will

equal 0, and the internal energy, E , will be diminished by w . E will equal $-w$.

The mixing of two gases

According to Dill and Bromberg, page 33, “While the *pressure* is a force for expansion, the *chemical potential* is a force for mixing, as when a dye [tea color] diffuses in water. The chemical potential describes the tendencies of substances to dissolve or to partition between different phases of matter or to transform into different chemical species.” Here, too, they provide a simple lattice model example for the increase in multiplicity that accompanies mixing. The math is of the form $W = W_1 \cdot W_2$.

The melting of ice

When we drop two or three ice cubes into a glass of “iced tea” it doesn’t take long for the cubes to melt. The “flow” of heat is from the liquid (tea) into the solid (ice), and the ice turns into water (melts). The water molecules of the ice, which are held rather firmly in place, become water molecules of water, which are more free to move about in the liquid phase than in the solid phase. The melted state has a much greater multiplicity than does the frozen state. As long as thermal energy is available the ice will melt.

A summary

While the first law provides a limit to what *can* happen (energy must be conserved) it is only the second law that provides an indication of what *will* happen (entropy will increase). The “problem” with entropy was that while a change in entropy could be *measured* (the integral of dq/T) a change in entropy could not be *predicted* without statistical mechanics. Concepts such as “disorder” or “randomness” were not well defined.

There also were phenomena that defied explanation, such as the free expansion of an ideal gas into a vacuum. For this spontaneous process both w and q are zero, and there is no change in E or temperature. Yet something

has happened, and since it happened there must have been an increase in entropy. Now, with the help of statistical thermodynamics, we understand that when the constraint on volume is removed a macrostate of greater multiplicity is reached as the motions of the particles allow them to move into a larger volume where their translational energy levels are more closely spaced.

It now appears that every change in entropy can be accounted for by a change in the number of microstates, which is conveniently indicated by the change in the partition function, Q .

Finally, as we said earlier, changes in all the thermodynamic variables can be calculated by use of the partition function, Q .

A deep thought

Counting is the essence of statistical thermodynamics, which implies that there is something to count. In statistical thermodynamics the items counted are individual particles with particular energies. “There is, essentially, only one problem in statistical thermodynamics: the *distribution* of a given amount of energy E over N identical systems.” (S. p. 1) This implies that both matter and energy have a discontinuous, or particle, nature so that there will be countable combinations of matter and energy. Perhaps the existence of entropy is all the evidence we need for the particle nature of both matter and energy.

Notes

1. When there are fractions of *exactly* the same size (when there are two or more levels with *exactly* the same energy; when there are *degeneracies*) , these identical fractions are often combined and the total of the combined fractions can be more than 1.

Additional reading in the Journal of Chemical Education

The Journal of Chemical Education's on-line index indicates that more than 100 papers have been published by This Journal in which the word "Entropy" appears in the title. I mention here, in chronological order, some of those papers that I found to be most helpful.

1. Lowe, J. P. **Entropy: Conceptual Disorder**, *J. Chem. Educ.* **1988**, *65*, 403-406. This is the transcript of a conversation between a student, S, and a professor, P. The conversation starts like this. "*Student*: Good morning, Professor. I have a few questions about entropy. Would this be a good time to discuss them with you? *Professor*: By all means! Come in, come in. Entropy is one of my most favorite subjects." The student has found that disorder and the number of microstates are not always related. A brilliant conversation.
2. McDowell, S. A. C. **A Simple Derivation of the Boltzmann Distribution** *J. Chem. Educ.* **1999**, *76*, 1393-1394. This turns out to be the approach used by Henry Bent, as the subsequent Letter indicates.
3. Lambert, F. L. **Disorder - A Cracked Crutch for Supporting Entropy Discussions** *J. Chem. Educ.* **2002**, *79*, 187-192. Let me quote from this paper: "The important point here is that *this kind of mistake is fundamentally due to a focus on disorder rather than on the correct cause of entropy change*, [which is] *energy flow toward dispersal*." Italics in the original. Then "Entropy is not disorder. Entropy is not a measure of disorder or chaos. Entropy is not a driving force. Energy's diffusion, dissipation, or dispersion in a final state compared to an initial state is the driving force in chemistry." The examples make the case. The author states in footnote 11 that "It was Planck who, in December 1900, found that a $k, R/N$, united Clausius' S with Boltzmann's logarithm

of the probability of a state.”

4. Lambert, F. L. **Entropy is Simple, Qualitatively** *J. Chem. Educ.* **2002**, *79*, 1241-1246. Again, “...the reason for transformation in classical thermodynamics is energy’s dissipating, dispersing, spreading out from where it is confined in a small space to a larger volume whenever it is not restricted from doing so. Entropy is the measure or the index of energy dispersal at a temperature T.” We might say, correspondingly, that the reason for transformation in statistical thermodynamics is the availability of a macrostate with a larger number of microstates that have come into existence through the removal of constraints. Remember: when the macrostates are *distinguishable* the corresponding sets of microstates are *different*.

5. Spencer, J. N., and Lowe, J. P. **Entropy: The Effects of Distinguishability**, *J. Chem. Educ.* **2003**, *80*, 1417-1424. This is a sequel to the conversation between the Professor and the Student. The Student has now become The Young Thermo Teacher, and his mentor is the Old Thermo Teacher.

The conversation starts in this way: Young Thermo Teacher: “Good Morning, Professor. This is my first year of teaching physical chemistry, and I hope that you might share with me some of the insights you have gained from teaching courses in thermodynamics over your long career.”

Old Thermo Teacher: Ah, yes, thermodynamics, that most subtle of subjects. I’ve been teaching it for 35 years, and each year I become aware of new ways to understand its relationships and better ways to try to explain them.” Again we are privileged to have available a transcript of this conversation. One thing we learn from this conversation is that there are *three* factors that affect entropy change. The first is the amount of

energy available. The second is the number of *energy levels* available to the particles, and the third, which is new to us, is the extent to which the particles are distinguishable. A second thing we learn from this conversation is that there can be two ways to think about an entropy change for which the symbols and mathematical operations are identical, but for which the meanings of the symbols are different: “OTT: Right. The formula is always the same, but the T, V, and N factors change with choice of route and show what the explanation for ΔS is for each choice.”

6. Gislason, E. A. and Craig, N. C. **Why Do Two Objects of Different Temperatures Come to a Common Intermediate Temperature When Put in Contact? Entropy is Maximized** *J. Chem. Educ.* **2006**, *83*, 885-889.

The authors answer the question using heat capacities. Part of their conclusion is “... for a spontaneous infinitesimal process the second law requires that the warmer object lose thermal energy to the cooler one, thereby lowering the temperature of the warmer object and raising that of the cooler one.” The next paragraph states “This result for the direction of transfer of thermal energy is a direct consequence of the inverse temperature dependence in $dS = Dq_{\text{rev}}/T$. For a given heat transfer, Dq_{rev} , the object at higher temperature undergoes a smaller change in entropy than the one at lower temperature. In other words, the warmer object decreases in entropy at the expense of the greater entropy increase of the cooler object.”

According to statistical thermodynamics we would say that the loss of *multiplicity* of the warmer object as it cools is more than compensated by the gain in *multiplicity* of the cooler object as it warms. The multiplicity of the system, the two objects together, will be maximized at an intermediate

temperature. This is the temperature at which there will be no net change in multiplicity, and no net heat transfer. This is the equilibrium temperature.

7. Lambert, F. L. **Configurational Entropy Revisited** *J. Chem. Educ.* **2007**, *84*, 1548-1550. This paper holds that there is only one kind of entropy, and that the entropy is determined by the number of microstates that correspond to a macrostate. If the system has “motional energy”, which will be conserved (first law), and if constraints are removed, the motional energy of the system will allow the system to evolve to a new, equilibrium macrostate that has a greater number of microstates (second law). The reason this happens is that the number of microstates for the new, equilibrium, macrostate will be overwhelmingly larger than the number of microstates for any other experimentally detectable macrostate, and the macrostate with the greatest number of microstates prevails at equilibrium. Admittedly, there are shortcuts for calculating differences in values of state functions, but we need to remember that it is only our minds that are taking the shortcuts, not the molecules.

Additional Reading

These are the authoritative sources that I found to be most helpful.

1. Andrews, Frank C., *Equilibrium Statistical Mechanics*, John Wiley & Sons, Inc., 1963

I found Chapter 1, Introduction, Chapter 2, Thermodynamics versus Statistical Mechanics, and Chapter 3, The Problem of Statistical Mechanics, pages 3-7, and The Ensemble in Statistical Mechanics, pages 18 and 19, to be of help in seeing the Big Picture of statistical mechanics. However, I found that the little pictures of many of the other books were needed in order to bring the Big Picture into focus. After you see the Big

Picture you can then consider *The Mechanical View of Thermodynamics*, pages 182-186.

2. Ben-Naim, Arieh, *Entropy Demystified*, World Scientific, Singapore, 2008

The subtitle to this book is “The Second Law Reduced to Plain Common Sense”. Ben-Naim also explains what it is that, like entropy, will always increase. It is the probability of the equilibrium state, and when something is sufficiently probable you might well take it as certain. Ben-Naim speaks of “dim states” and “specific states”, as in “Each of these dim states contains many, many specific states ...”. I believe that we would call his “specific states” microstates, but I am less sure that I understand the “dim states”. I did appreciate a paragraph that he quoted from George Gamov’s book *Mr. Tompkins in Wonderland*: “The liquid in the glass was covered with violently bursting bubbles, and a thin cloud of steam was rising slowly toward the ceiling. It was particularly odd, however, that the [tea] was boiling only in a comparatively small area around the ice cube. The rest of the [tea] was quite cold”.

3. Bent, Henry Henry A., *The Second Law. An Introduction to Classical and Statistical Thermodynamics*, Oxford University Press, 1965.

Again, a helpful view of the Big Picture is to be found in the Introduction and Synopsis, pages 1-5.

Part III An Introduction to Statistical Thermodynamics contains: Chapter 20, entitled Boltzmann’s Relation; Entropy and Ω , Chapter 21, called Statistical Mechanics of Very Small Systems; and Chapter 23, describing Boltzmann’s Factor and Partition Functions, and General Formulas for E and S. Here you will find closeups of many of the little pictures that give you the details of the Big Picture.

4. Davies, William G., *Introduction to Chemical Thermodynamics*, W. B.

Saunders, Philadelphia, 1972.

Chapter 2, pages 23-49, contains sections called A Simple Example, Permutations and Factorials, A Formula for **W**, The Boltzmann Distribution Law, and Some Features of a Boltzmann Distribution.

Here you will become familiar with permutations and combinations, you will see that there are always more permutations than combinations, and you will learn that in statistical mechanics it is only combinations that count (that should be counted).

5. Dill, Ken A., and Bromberg, Sarina, *Molecular Driving Forces. Statistical Driving Forces in Chemistry and Biology*, Garland Science, 2003.

This book at, 3 1/2 pounds and 666, pages has it all. It starts with Principles of Probability in Chapter 1, then goes to Heat, Work, and Energy in Chapter 3, Entropy & the Boltzmann Distribution Law in Chapter 6, Thermodynamic Driving Forces in Chapter 7, and the Boltzmann Distribution Law in Chapter 10, and we are only 1/4 of the way through the book. We skipped Extremum Principles, Series and Approximations, and Multivariate Calculus. If you really want to understand statistical thermodynamics, how it provides a driving force, and its applications in chemistry and biology, this is your book. It is clearly written, and beautifully illustrated with a multitude of helpful diagrams and graphs.

You should set aside a year or so to read it.

6. Laidler, Keith J. *The World of Physical Chemistry*, Oxford University Press, Oxford, 1993. The sections entitled The statistical basis of the second law, pages 154 through 160, and Statistical mechanics, pages 161 through 164 provide interesting and useful historical and philosophical insights.

7. McQuarrie, Donald A., *Statistical Mechanics*, University Science Books, Sausalito, CA 2000. I found Chapter 2 of this book, The Canonical Ensemble, to be of great help in understanding the book *Statistical Thermodynamics*, by Erwin Schrödinger.
8. Nash, Leonard K. *Chemthermo: A Statistical Approach to Classical Chemical Thermodynamics*, Addison-Wesley Publishing Company, Reading, MA, 1971
9. Nash, Leonard K. *Elements of Statistical Thermodynamics*, second edition, Dover Publications, Inc., Mineola, New York, 2006
10. Schrödinger, Erwin, *Statistical Thermodynamics*, Cambridge University Press, 1948. This book, of which I do not claim to have read more than 14 pages, also says it all, but at only 6 ounces and 88 pages. It is a beautiful book in the sense that the drawings of Picasso are beautiful: so much message; so little ink. Schrödinger chooses to follow the method of Gibbs: “It has a particular beauty of its own, is applicable quite generally to every physical system, and has some advantages to be mentioned forthwith.”

Literature Cited

Additional reading in the Journal of Chemical Education

1. Lowe, J. P. *J. Chem. Educ.* **1988**, *65*, 403.
2. McDowell, S. A. C. *J. Chem. Educ.* **1999**, *76*, 1393.
3. Lambert, F. L. *J. Chem. Educ.* **2002**, *79*, 187.
4. Lambert, F. L. *J. Chem. Educ.* **2002**, *79*, 1241.
5. Spencer, J. N., and Lowe, J. P. *J. Chem. Educ.* **2003**, *80*, 1417.
6. Gislason, E. A. and Craig, N. C. *J. Chem. Educ.* **2006**, *83*, 885.
7. Lambert, F. L. *J. Chem. Educ.* **2007**, *84*, 1548.

Additional reading

1. Andrews, Frank C., *Equilibrium Statistical Mechanics*, John Wiley & Sons, Inc., 1963
2. Ben-Naim, Arieh, *Entropy Demystified*, World Scientific, Singapore, 2008
3. Bent, Henry A., *The Second Law. An Introduction to Classical and Statistical Thermodynamics*, Oxford University Press, 1965.
4. Davies, William G., *Introduction to Chemical Thermodynamics*, W. B. Saunders, Philadelphia, 1972.
5. Dill, Ken A., and Bromberg, Sarina, *Molecular Driving Forces. Statistical Driving Forces in Chemistry and Biology*, Garland Science, 2003.
6. Laidler, Keith J. *The World of Physical Chemistry*, Oxford University Press, Oxford, 1993.
7. McQuarrie, Donald A., *Statistical Mechanics*, University Science Books, Sausalito, CA 2000.
8. Nash, Leonard K. *Chemthermo: A Statistical Approach to Classical Chemical Thermodynamics*, Addison-Wesley Publishing Company, Reading, MA, 1971
9. Nash, Leonard K. *Elements of Statistical Thermodynamics*, second edition, Dover Publications, Inc., Mineola, New York, 2006
10. Schrödinger, Erwin, *Statistical Thermodynamics*, Cambridge University Press, 1948.

Table 2

An illustration of the relationships between energy levels, the absolute temperature, T , the Boltzmann distribution, and the partition function, Q .

state	ϵ/kT	$1/e^{\epsilon/kT}$	number
fourteenth excited	14	0.000 001	0
thirteenth excited	13	0.000 002	0
twelfth excited	12	0.000 006	0
eleventh excited	11	0.000 017	0
tenth excited	10	0.000 045	0
ninth excited	9	0.000 123	1
eighth excited	8	0.000 335	2
seventh excited	7	0.000 911	6
sixth excited	6	0.002 478	16
fifth excited	5	0.006 737	43
fourth excited	4	0.018 315	116
third excited	3	0.049 787	315
second excited	2	0.135 335	855
first excited	1	0.367 879	2325
ground state	0	1.000 000	6321 = n_0
		sum = $Q = 1.581\ 98$	sum = $N = 10,000$