

Chapter 12: The Statistical Definition of Entropy

The process of folding a protein to produce the active conformation restricts torsions along the polypeptide backbone and side chain torsions for amino acids that are buried in the interior of the protein. Calculate the conformational entropy of the side chain of the amino acid valine.

Energy is transferred from a hotter to a colder body. This energy transfer corresponds to a dispersal of energy. These generalizations come from observations of simple physical phenomena, such as the heating of a pot of water on a stove. Energy dispersal is a part of our common experience. All processes occur subject to the First Law of thermodynamics. Internal energy is conserved for all processes spontaneous or not, reversible or irreversible. So the First Law doesn't tell us the direction of spontaneous change. Can we develop a thermodynamic state function that predicts the spontaneous direction of chemical change based on energy dispersal? The first step is to quantitatively characterize energy dispersal. We will then use this indicator of energy dispersal to predict the spontaneous direction for all physical processes and the position of equilibrium. For the example of a bouncing ball, Section 10.1, potential energy is converted into random thermal kinetic energy with each bounce. This conversion suggests that statistical considerations may be central in understanding energy dispersal. The intriguing question is: does random chance determine the outcome of chemical reactions?

You may have too deterministic an interpretation of chemical reactions to appreciate that random chance plays a role in the outcome of chemical processes. The laws of statistics may seem to be too devoid of physical relevance to be the predominant factor in determining the extent of energy dispersal. But **statistical mechanics** is the discipline within chemistry that applies the laws of probability to determine the spontaneous direction of chemical reactions and the position of equilibrium.

12.1 Thermodynamic Properties are Average Values

The laws of probability determine the likely outcomes of coin tosses and throwing dice. According to the laws of chance, on average a coin toss yields heads 50% of the time and tails 50% of the time. This expectation is more closely met as the number of trials increases. The larger the number of trials, the simpler the behavior is to predict. Thermodynamics applies to macroscopic systems that contain large numbers of molecules. The laws of probability allow our description of systems to be greatly simplified; we only need to predict the average behavior.

Thermodynamic variables result from averages over all the degrees of freedom of a system as a function of time. For example, the pressure of a gas is the time averaged force of collisions of molecules with the walls of the container per unit area, Figure 7.5.1. The degrees of freedom of a system include translation, rotation, vibration, and electronic energies. A given system at some time may have more rotational energy than vibrational energy, and at some later time more vibrational energy than rotational. However, the time average energy for a system at equilibrium gives just one parameter, the internal energy. In thermodynamics, just a few variables are necessary to specify the state of the system, for example the temperature, volume, and molar amounts. Much specific information is lost when averaging over the variables for all the degrees of freedom to give just a few thermodynamic variables. However, the loss of information is compensated by a gain of generality, simplicity, and predictive power.

Following the time dependence of a system is difficult, because there are so many variables. An easier approach is to consider a large number of identical systems, each with the same number of molecules, N , and the same volume, V . A group of identical systems is called an **ensemble**. Gibbs suggested that averaging a property over all the systems in an ensemble at one instant in time is equivalent to averaging the property of a single system over time. Such an average is called an **ensemble average**.

Conceptually, the systems in the ensemble are held in an insulated constant volume container, so that the ensemble is isolated from the surroundings, Figure 12.1.1. The total energy of the ensemble is constant. If the systems in the ensemble are allowed to come into thermal contact, they will exchange energy and on average all the systems in the ensemble will have the same average energy and temperature. This kind of ensemble is called a **canonical ensemble**. The number of systems in the ensemble is given the symbol \mathcal{N} and the total energy of the ensemble is \mathcal{E} . Properties of the ensemble as a whole are written in “script,” (\mathcal{N} and \mathcal{E}), and properties of a single system are in the normal font (N , E). Even though the average energy of each system is constant, there are still fluctuations in the energy of each system. The energy in each degree of freedom also varies with time. At a particular instant one system may have excess rotational energy and another may have excess vibrational energy. Ensemble averages are much easier to calculate than time averages, because the ensemble average can be accomplished using the laws of statistics. The average behavior of the system is determined by the probability of occurrence of each possible energy state of the system. Ensemble averages are valid in the limit of large numbers of systems, $\mathcal{N} \rightarrow \infty$. We will discuss the validity of ensemble averages more completely in the later chapter on the foundations of statistical mechanics and Sec. 12.7. However, keep in mind that ensemble averages are just an easy way to find the average energy of a single system. In essence we just average many simultaneous experiments on a set of identical systems instead of repeating experiments on the same system and doing the average over time. The ensemble represents a system for averaging purposes. We will use a “{“ to indicate this representation.

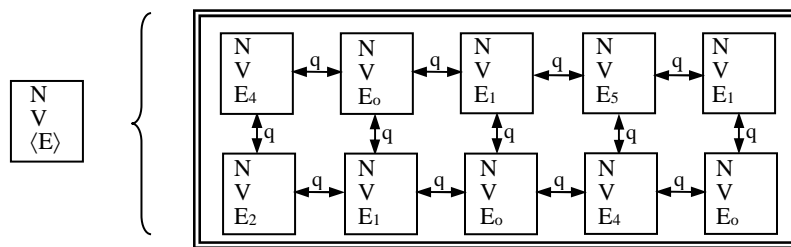


Figure 12.1.1: An isolated group of identical systems that are in thermal contact is called a canonical ensemble. Meaningful thermodynamic averages are taken over the systems in the ensemble in the limit $\mathcal{N} \rightarrow \infty$. The ensemble represents the system for averaging purposes.

The average energy of the system is given by the ensemble average. The ensemble average is particularly easy to calculate. Even though the energy of each system in the ensemble may be different, on average the energy of a system must be the total energy of the ensemble divided by the number of systems in the ensemble. For \mathcal{N} systems in the ensemble with total energy \mathcal{E} , the average energy of a system is just:

$$\langle E \rangle = \frac{\mathcal{E}}{\mathcal{N}} \quad (\text{canonical ensemble}) \quad 12.1.1$$

where the average is indicated using braces $\langle E \rangle$. The central hypothesis of statistical mechanics is that the ensemble average energy is equal to the thermodynamic internal energy:

$$U - U(0) = \langle E \rangle \quad 12.1.2$$

The inclusion of the $U(0)$ term allows for differences in reference points. As noted in Chapter 8, only differences in the internal energy may be measured and not absolute values, so an arbitrary reference is chosen. Eq. 12.1.2 can be read as: the internal energy of the system above the reference point is given by the ensemble average energy over all the degrees of freedom. As we stated in Section 7.8, the internal energy is the sum of the average kinetic and potential energy over all the atoms and molecules that make up the system. We now know how that averaging is done. The next step is to use the concept of the ensemble and averaging to develop a measure of energy dispersal. Determining energy dispersal and the spontaneity of physical processes flows naturally once we accept the notion that the laws of random chance determine the probability of occurrence of the different energy states of a system.

12.2 Energy Dispersal is Measured by Changes in Entropy

The dispersal of energy is determined by finding the **number of ways of arranging the energy states** for a given amount of energy and the energy states of the ensemble. The number of ways of arranging the energy states is given the symbol \mathcal{W} . Consider a group of three systems, $\mathcal{N} = 3$, with minimum energy, Figure 12.1.1 top-left. All of the systems are in the lowest possible energy state giving $\mathcal{E} = 0$. There is only one way of arranging the energy states of the ensemble, all systems are in the lowest energy state, giving $\mathcal{W} = 1$. For ease of discussion we will assume that the energy states are equally spaced, as is the case for harmonic molecular vibrations, and that energy is available in packets equal to the energy difference between the states. Transfer one packet of energy to the ensemble from the surroundings giving $\mathcal{E} = 1\epsilon$, Figure 12.2.1 top-right. Each of the systems can in turn share that packet, which results in three ways of arranging the available energy among the systems, Abc , aBc , and abC , giving $\mathcal{W} = 3$ and average energy $\frac{1}{3}\epsilon$. Each individual energy configuration of the ensemble is called a **microstate**. The ensemble can be in only one of the microstates at any time. Fluctuations cause the ensemble to jump from one microstate to another. The fluctuations are caused by collisions.

Transfer one additional packet of energy to the ensemble. There are now two packets of energy available, $\mathcal{E} = 2\epsilon$, and six total ways of arranging the energy among the three systems. There are now six microstates consistent with $\mathcal{E} = 2\epsilon$ and average energy $\frac{2}{3}\epsilon$. The transfer of energy to the ensemble increases the number of ways that the available energy can be arranged among the systems. The energy is dispersed among the available energy states. Transfers of energy into the ensemble correspond to heat transfers. Heat transfer increases the number of ways of arranging the energy states of the ensemble, because there is more energy available.

Now consider increasing the number of systems in the ensemble. With $\mathcal{N} = 4$, and minimum energy, there is still only one way of arranging the energy states and $\mathcal{W} = 1$. After transfer of one packet of energy, $\mathcal{E} = 1\epsilon$, there are four ways of arranging the energy states: $Abcd$, $aBcd$, $abCd$, and $abcd$, or $\mathcal{W} = 4$. The change in energy dispersal for the transfer of one packet of energy is greater with four systems than with three since there are more available energy states. \mathcal{W} increases with the number of systems in the ensemble, for a fixed amount of energy.

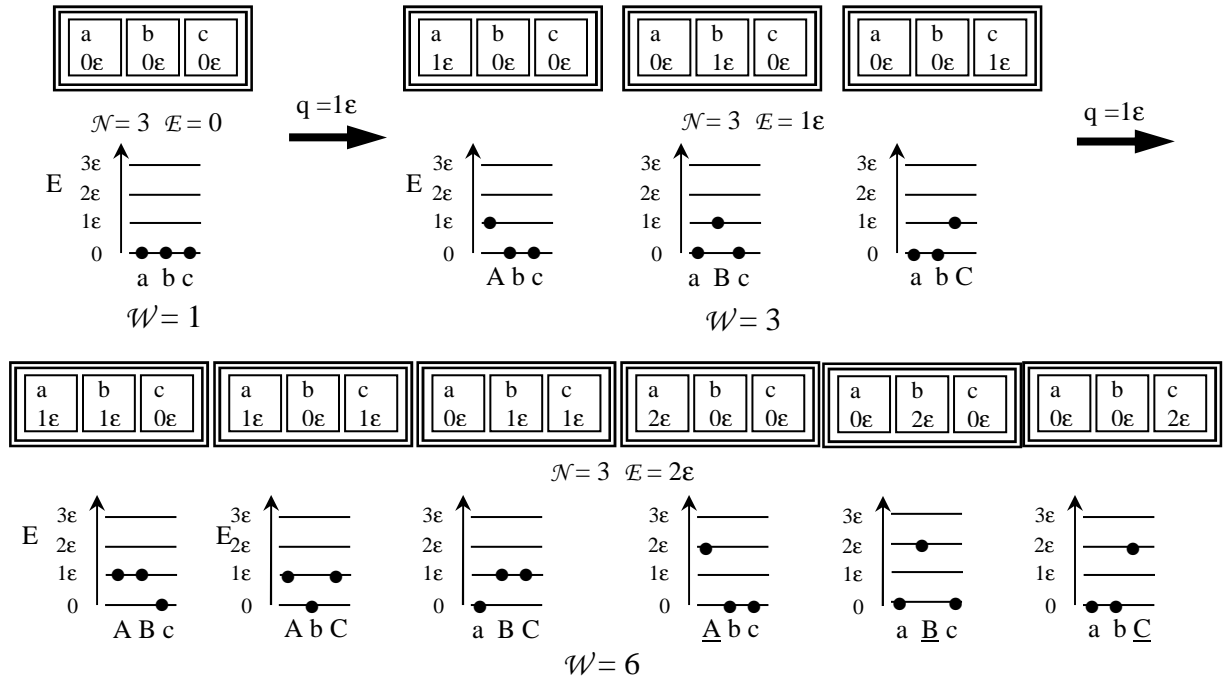


Figure 12.2.1: There is only one way of arranging the energy states for minimum energy. After transfer of one packet of energy there are three ways of arranging the energy states. For two packets of energy there are six total ways of arranging the energy states. A lower case letter represents a system with 0 packets of energy, a Capital letter represents a system with 1ε, and an underlined Capital letter 2ε.

Extensive properties increase with the size of a system. We conclude that the thermodynamic state function that corresponds to energy dispersal should be an extensive property that is increased by thermal energy transfer. The thermodynamic property that characterizes energy dispersal is called the **entropy**.

The statistical definition of the entropy of a system in the ensemble is:

$$S \equiv \frac{k}{\mathcal{N}} \ln \mathcal{W}_{\max} \tag{ensemble} \quad 12.2.1$$

where \mathcal{W}_{\max} is the maximum number of ways of arranging the energy states of the ensemble, k is Boltzmann’s constant, and \mathcal{N} is the number of systems in the ensemble. The maximum in \mathcal{W} is necessary to ensure that the system is at equilibrium and correspondingly that entropy is a state function. Boltzmann’s constant is just the gas constant on a per molecule basis instead of per mole, $k = R/N_A = 1.38066 \times 10^{-23} \text{ J K}^{-1}$, where N_A is Avogadro’s number. The units of entropy are then J K^{-1} . In Eq. 12.2.1, division by \mathcal{N} determines the ensemble average, just as we did for the ensemble averaged energy, Eq. 12.1.1. As the number of ways of arranging the states of the system increases, the entropy increases. \mathcal{W} increases through transfers of heat and increases in the number of systems and the volume. This statistical definition of entropy, which is often called the Boltzmann entropy, was developed by Ludwig Boltzmann and first written in this form by Max Planck in 1901.¹⁻³ Entropy is a measure of the statistical likelihood of a given configuration of energy states of the system.

Instead of an ensemble average, we often consider just one system and find the number of ways of arranging the energy states of the molecules within a single system, W_{\max} . The Boltzmann formula is then just $S = k \ln W_{\max}$ for a single system. The practical application of the Boltzmann formula requires that we develop an easy method for evaluating \mathcal{W} using standard statistical arguments.

Averages are Calculated Using Distribution Numbers: Enumerating all the individual configurations of the available energy, as we did in Figure 12.2.1, becomes tedious even for small systems. To determine thermodynamic averages, we don't need to know all the individual configurations; we just need to know the number of systems in each energy state. We don't care which specific system has a given amount of energy, we only need to know the number of systems that have each possible amount of energy. The number of systems in each energy state is specified using a set of **distribution numbers**, and all averages can be calculated using the distribution numbers. For a simple example, assume that the available energy levels are $E_i = 0, 1, 2, 3, 4, \text{ or } 5$ packets of energy. We determine the number of packets of energy in each system in the ensemble. Let the measurements for the energy in each system be:

4, 1, 2, 4, 4, 1, 5 packets

for $\mathcal{N} = 7$ total systems. The average energy is then $(4+1+2+4+4+1+5)/7 = 3.0$. However, we can do the averaging in a different order. We can count the number of times each possible result is found; these counts are the distribution numbers. In our example data set, the value 0 is not found, the value 1 occurs twice, the value 2 occurs once, the value 3 is not found, the value 4 occurs three times, and the value 5 occurs once. The set of distribution numbers, $\{n_i\}$ is then $(0, 2, 1, 0, 3, 1)$. The sum of the distribution numbers must equal the total number of trials:

$$\sum_i n_i = \mathcal{N} \quad 12.2.2$$

where the sum extends over all possible energy states, i . The average energy is then calculated using the distribution numbers and all the possible experimental values, E_i :

$$\langle E \rangle = \frac{1}{\mathcal{N}} \sum_i n_i E_i \quad 12.2.3$$

For our example:

$$\langle E \rangle = \frac{1}{\mathcal{N}} \sum_i n_i E_i = \frac{1}{\mathcal{N}} (n_0 E_0 + n_1 E_1 + n_2 E_2 + n_3 E_3 + n_4 E_4 + n_5 E_5)$$

$$\langle E \rangle = \frac{1}{\mathcal{N}} \sum_i n_i E_i = \frac{0(0) + 2(1) + 1(2) + 0(3) + 3(4) + 1(5)}{7} = 3.0 \quad 12.2.4$$

The distribution numbers can also be related to the probability of occurrence of each possible value of the experiment. The probability of occurrence, p_i , of value E_i is given by:

$$p_i = \frac{n_i}{\mathcal{N}} \quad 12.2.5$$

which is the ensemble average of the occupation of state i . The average value of the energy is given by substituting this last equation for each term in the sum of Eq. 12.2.3:

$$\langle E \rangle = U - U(0) = \sum_i p_i E_i \quad \backslash \quad 12.2.6$$

which is subject to normalization of the probability:

$$\sum_{i=0}^{\mathcal{N}} p_i = \frac{1}{\mathcal{N}} \sum_i n_i = 1 \quad 12.2.7$$

which is obtained by dividing Eq. 12.2.2 by \mathcal{N} . In other words, the sum of the probabilities must equal one. The advantage of using distribution numbers is that the calculation of \mathcal{W} becomes straightforward without explicitly specifying the individual microstates.

The number of ways of arranging the energy states, \mathcal{W} , is given by a standard statistical formula:

$$\mathcal{W} = \frac{\mathcal{N}!}{n_0! n_1! n_2! \dots} \quad (\text{distinguishable}) \quad 12.2.8$$

where \mathcal{N} is the total number of distinguishable systems in the ensemble, and n_0, n_1, \dots are the distribution numbers for the energy states. \mathcal{W} is the number of microstates for the given set of distribution numbers. The “!” indicates the factorial of the number. For example, the factorial of the number 5 is defined as $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1$, while $0! = 1$. Non-chemical examples of the use of this formula are given in the addendum. \mathcal{W} is often called the **statistical weight**, or **degeneracy**, of a given set of distribution numbers. Degenerate states have the same energy. A set of distribution numbers with a larger statistical weight is more probable because the set of distribution numbers has more ways of occurring.

We can verify Eq. 12.2.8 using some examples. Consider an ensemble consisting of three identical systems, $\mathcal{N} = 3$. Assume that three packets of energy are available to distribute among the three systems; the total energy of the ensemble is $\mathcal{E} = 3\epsilon$. The average energy of each system in the ensemble is then $\langle E \rangle = \mathcal{E}/\mathcal{N} = 1\epsilon$, Figure 12.2.2.

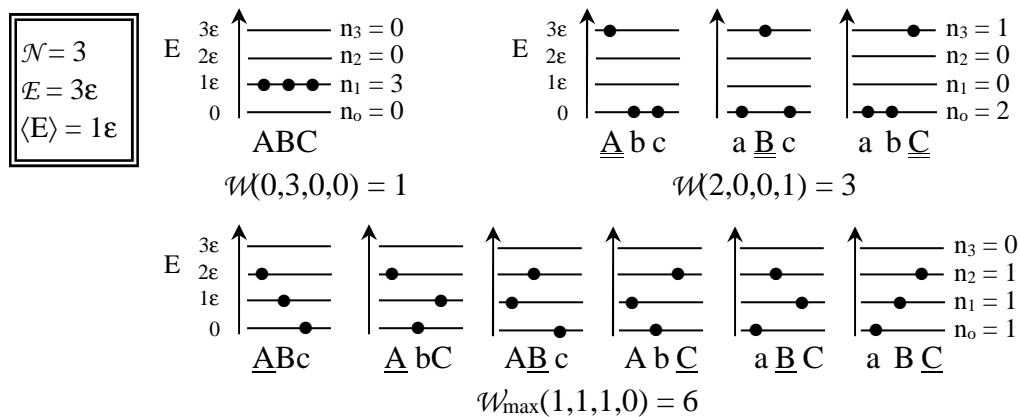


Figure 12.2.2: The number of ways of arranging the energy states, \mathcal{W} , for three systems with three packets of energy. \mathcal{W} is determined by the distribution numbers for the energy states.

One way to distribute the available total energy is for each system to hold one packet of energy, which corresponds to the distribution numbers (0,3,0,0). There is only one way of arranging the states of this “equal sharing” distribution. The complete set configurations for the two additional sets of distribution numbers, (2,0,0,1) and (1,1,1,0), are diagrammed in Figure 12.2.2. The number of ways of arranging the energy states can also be calculated using Eq. 12.2.8 as:

$$\mathcal{W}(2,0,0,1) = \frac{3!}{2! 0! 0! 1!} = 3 \quad \text{and} \quad \mathcal{W}(1,1,1,0) = \frac{3!}{1! 1! 1! 0!} = 6 \quad 12.2.9$$

which give the same result as the exhaustive configurations in the figure.

Microstates Have Equal a priori Probabilities: A fundamental assumption is that each microstate has the same probability of occurrence; in other words, each microstate has an equal *a priori* probability. A set of distribution numbers corresponds to a **macrostate**. For example, from Figure 12.2.2, the set of distribution numbers (2,0,0,1) corresponds to one macrostate that has $\mathcal{W} = 3$ corresponding microstates. While each microstate is equally probable; the value of \mathcal{W} shows that some macrostates are more probable than others. A simple example of this conclusion is shown by coin tosses. The possible outcomes of two successive coin tosses are HH, HT, TH, and TT, with H = heads and T = tails. Note that HH is just as probable as TH. The individual results HH and TH have an equal *a priori* probability. However, there are two ways of obtaining one H and one T, which are HT and TH. So the probability of observing one H and one T, in any order, is twice as probable as observing two successive H tosses. In other words, if the distribution numbers are arranged by the number of heads and tails (H,T), the macrostate (1,1) is twice as probable as (2,0):

Outcomes:	HH	HT	TH	TT
		⏟		
Distribution numbers:	(2,0)	(1,1)		(0,2)
Statistical weight:	$\mathcal{W} = 1$	$\mathcal{W} = 2$		$\mathcal{W} = 1$

Greater energy dispersal results from a macrostate with larger \mathcal{W} . The **most probable distribution** is the distribution that has the largest number of ways of arranging the energy states, \mathcal{W}_{\max} .

Statistical Weights Multiply for Composite Systems: In thermodynamics we often discuss processes that result from two systems coming into contact to form a composite system. In particular, energy dispersal is the result of spontaneous heat transfer between two systems that come into contact. Statistical weights and probabilities multiply when, with no other changes, a composite is formed. Consider two systems that contain N_1 and N_2 molecules in volumes V_1 and V_2 , respectively. System 1 has a statistical weight of \mathcal{W}_1 and system 2 has a statistical weight of \mathcal{W}_2 . When the two systems are considered as a composite, with no other changes, the statistical weight is the product of the two separate weights:

$$\mathcal{W} = \mathcal{W}_1 \mathcal{W}_2 \quad 12.2.10$$

Figure 12.2.3 shows the result for two identical systems, $N_1 = N_2$, $V_1 = V_2$, $\langle E \rangle_1 = \langle E \rangle_2 = \frac{1}{3} \epsilon$. Ensembles were chosen with $\mathcal{N} = 3$ to make it easy to list all the microstates.

In this example we chose same size systems for convenience in counting the number of microstates; in general the systems can be different and Eq. 12.2.10 remains applicable. Now that

we know that statistical weights combine by multiplication, we can determine the changes that occur upon heat transfer.

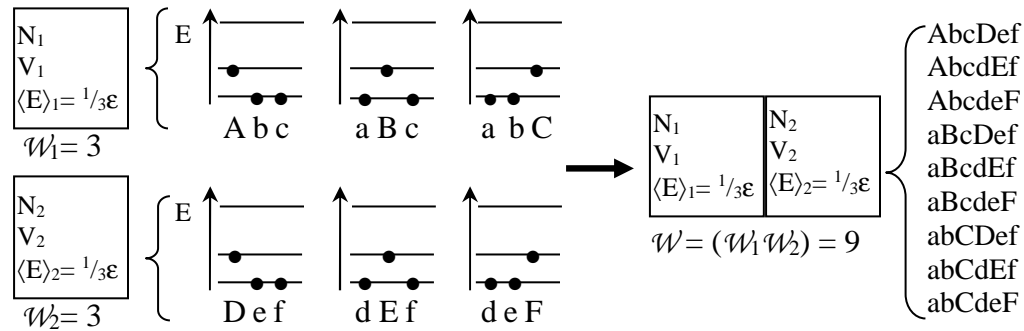


Figure 12.2.3: Statistical weights multiply upon forming a composite system, with no other changes.

Heat is Transferred from a Hotter to a Colder Body: Our experience tells us that energy in the form of heat is transferred from hot objects to cold objects. We can prove that this observation results from simple statistical considerations. Consider two systems, system 1 with a large initial energy and system 2 with a minimal initial energy, Figure 12.2.4.

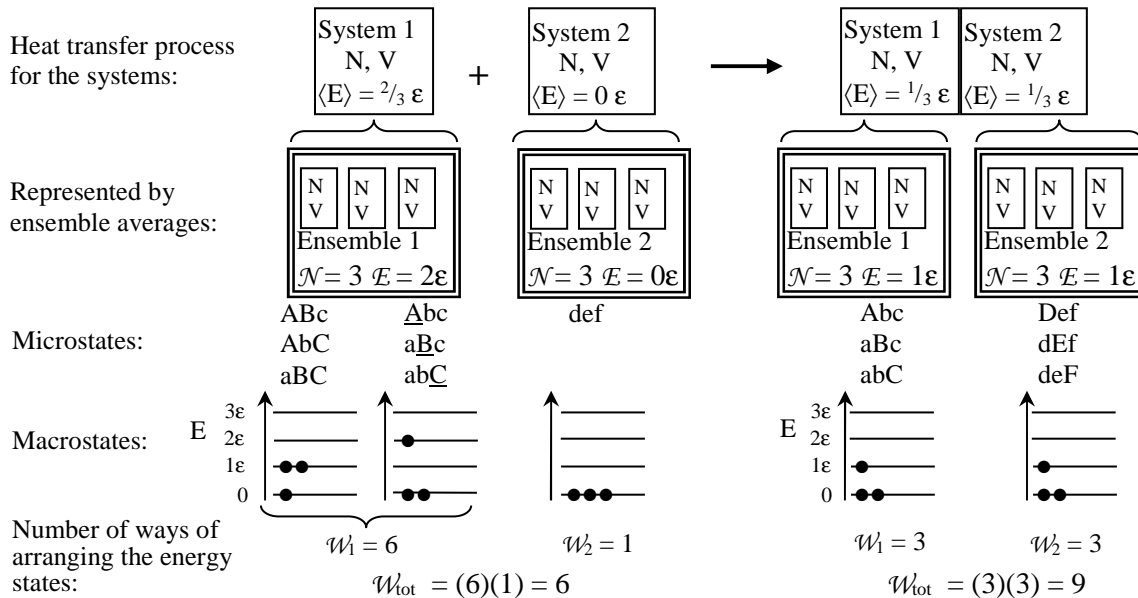


Figure 12.2.4: Energy is always transferred from a hotter object to a colder object because the energy transfer increases the number of ways of arranging the energy states for the composite of the two systems.

We assume that the two systems are at constant volume and isolated from the surroundings so that the only process that can occur is the transfer of energy between the two systems. For illustrative purposes, to make the state counting easy, we assume the systems are identical, with

the same number of atoms and volume. We represent each system with an ensemble containing three duplicate systems for averaging purposes. The statistical weight representing system 1 is $\mathcal{W}_1 = 6$ and for system 2 is $\mathcal{W}_2 = 1$ for a total of 6 ways of arranging the energy states. The two systems are brought into contact. One packet of energy is transferred from ensemble 1 to ensemble 2. Energy is conserved in the process, but is the energy transfer spontaneous? After the transfer, system 1 has a statistical weight of 3 and system 2 has a statistical weight of 3. The final total number of ways of arranging the energy states of the composite of the two systems is then $(3)(3) = 9$. The final state has a higher probability of occurrence than the initial state, so the transfer is spontaneous. The heat transfer takes place because the final state is more probable than the initial state. The laws of probability govern energy transfer.

The two systems will exchange additional packets of energy if the result gives a greater \mathcal{W} . The exchange of energy continues until the maximum number of ways of arranging the states of the composite has been attained, \mathcal{W}_{\max} . The resulting state is the most probable state. Once the most probable state has been reached, no further spontaneous processes will occur. If there is no further impetus for change, then the composite is at equilibrium. The equilibrium state is the most probable state. Any additional transfer of energy between the two systems results in smaller \mathcal{W} . For example in Figure 12.2.4 no further transfer of energy will occur, because the transfer of another packet will decrease \mathcal{W} . The configuration shown is at equilibrium, $\mathcal{W} = \mathcal{W}_{\max}$.

Our example is artificial in several ways. First, we kept the amount of energy and the number of systems in the ensemble small to make it easy to count the number of microstates. For real heat transfer processes, the energy and the number of systems would be much larger. The result would be a vastly greater increase in \mathcal{W} for the equilibrium state. Second, because we chose identical systems, the equilibrium state corresponded to equal average energy for the two systems, 0.333ε . For systems of different sizes, the average energy of the two systems will not be equal at equilibrium.

The Equilibrium State is the Most Probable State: Thermodynamic properties like internal energy and pressure are averages over all possible macrostates. However, the most probable macrostate plays a predominant role in determining thermodynamic averages. As the number of systems increases, the difference between \mathcal{W} for the most probable macrostate and the next most probable macrostate increases rapidly. For $\mathcal{N} = 3$ this difference is $6 - 3 = 3$, Figure 12.2.1. For $\mathcal{N} = 6$ this difference is $120 - 90 = 30$, Figure 12.2.5.

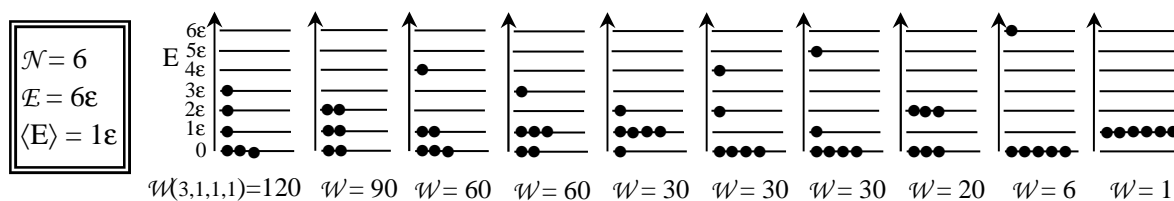


Figure 12.2.5: The number of ways of arranging the energy states, \mathcal{W} , for six systems with six packets of energy. Only the macrostates are diagrammed.

Thermodynamically meaningful averaging corresponds to very large \mathcal{N} , on the order of Avogadro's number, $\sim 10^{23}$. The limit as $\mathcal{N} \rightarrow \infty$ is called the **thermodynamic limit**. For such very large numbers of systems, the most probable distribution is the overwhelmingly

predominant distribution.⁴ The most probable distribution is characterized by \mathcal{W}_{\max} ways of arranging the energy states, which is an astronomically large number. In fact, with negligible error, we can use only the most probable distribution to calculate thermodynamic averages. Figure 12.2.6 summarizes the flow of concepts: the time average is replaced by the ensemble average, the ensemble average is dominated by the most probable distribution. The most probable distribution is the Boltzmann distribution. When the system is at equilibrium the state of the system is given by the most probable distribution. If the system is not at equilibrium, it is not in its most probable state.

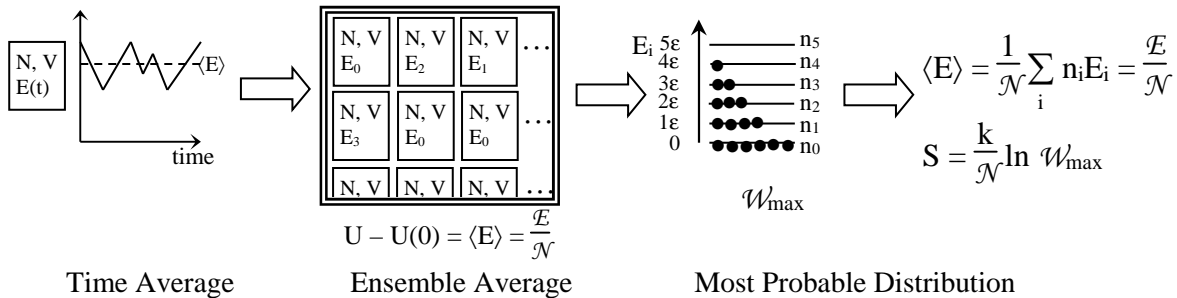


Figure 12.2.6: Thermodynamic averages are taken over the most probable distribution. The most probable distribution is the Boltzmann distribution.

The equilibrium state of the system is the most probable state as determined by the laws of random chance.⁵ It is by random chance that the result of many coin tosses gives 50% heads and 50% tails, because this result is the most probable outcome. A thermodynamic system evolves to equilibrium precisely because the equilibrium state is the most likely state. In fact the equilibrium state is overwhelmingly likely. There are so many more ways for the equilibrium state to occur than any non-equilibrium state that we don't expect any non-equilibrium state to occur, barring kinetic roadblocks. Any shift away from equilibrium takes the system to a much less probable state, $\mathcal{W} \ll \mathcal{W}_{\max}$. A shift away from equilibrium is very improbable, and is only a temporary, thermodynamically unobservable, small fluctuation in the system.^{4,5} Random chance provides a common sense explanation for the equilibrium state. We can also use a corresponding argument to provide a statistical interpretation of reversible processes.

We saw in Sections 1.2 and 7.4 that a reversible process occurs when a given large change is made through a sequence of small essentially reversible steps. The process occurs as a constraint, such as a moving piston, is displaced in many small steps, Figure 7.4.3. A reversible process occurs through a sequence of equilibrium or near-equilibrium states. Therefore, a reversible process takes place through a series of intermediate states with each step attaining the maximum number of ways of arranging the energy states of the system. That is, a reversible process occurs with $\mathcal{W} = \mathcal{W}_{\max}$ through each intermediate step, subject to the constraints placed on the system. An irreversible process occurs when $\mathcal{W} < \mathcal{W}_{\max}$ through at least some portion of the overall process. Of course, \mathcal{W} is strictly always less than \mathcal{W}_{\max} for an irreversible process, since \mathcal{W}_{\max} is the maximum possible value for the statistical weight. Any other \mathcal{W} must be smaller.

Notice that our specific process, Figure 12.2.4, was specified to be for the transfer of energy between two systems that are isolated from their surroundings. The requirement for isolation is an important restriction that must not be forgotten.

Entropy Always Increases for a Spontaneous Process in an Isolated System: A spontaneous process occurs when the system undergoes a change that increases \mathcal{W} . The system reaches equilibrium when it reaches the most probable state with $\mathcal{W} = \mathcal{W}_{\max}$. A reversible process occurs through a sequence of steps with $\mathcal{W} = \mathcal{W}_{\max}$. The definition of the entropy as $S \equiv k/\mathcal{N} \ln \mathcal{W}_{\max}$ then allows us to predict the direction for spontaneous change. Since \mathcal{W} always increases for a spontaneous process in an isolated system, the entropy also increases for a spontaneous process in an isolated system. Since the entropy is defined in terms of \mathcal{W}_{\max} , when we evaluate the change in entropy for a process, the initial and final states must be at equilibrium and the process must occur along a reversible path. For initial state 1 and final state 2:

$$\Delta S = \frac{k}{\mathcal{N}} (\ln \mathcal{W}_{\max,2} - \ln \mathcal{W}_{\max,1}) \quad 12.2.11$$

If $\mathcal{W}_{\max,2} > \mathcal{W}_{\max,1}$ the process is spontaneous. If $\mathcal{W}_{\max,2} = \mathcal{W}_{\max,1}$, then no spontaneous process will occur and the system is at equilibrium. Combining the logarithmic terms then gives:

$$\Delta S = \frac{k}{\mathcal{N}} \ln \left(\frac{\mathcal{W}_{\max,2}}{\mathcal{W}_{\max,1}} \right) \geq 0 \quad > \text{for spontaneous} \quad = \text{equilibrium} \quad (\text{isolated}) \quad 12.2.12$$

This result is a statement of the Second Law of Thermodynamics and can be phrased as “entropy always increases for a spontaneous process in an isolated system.” This result is entirely reconciled in terms of the probability of occurrence of the macrostates of the ensemble and the definition of entropy, Eq. 12.2.1. However, we have only considered heat transfer. Using Eq. 12.2.12 as the motivation for the Second Law is specific to processes at constant volume and no other forms of work. In the next chapter we show that the Second Law holds for any process in an isolated system, which extends these powerful ideas to chemical reactions.

12.3 Entropy is an Extensive State Function

In Eq. 12.2.1 we simply stated the statistical definition of entropy. In this section we prove that the logarithmic form of Eq. 12.2.1 is the only possible functional form for the dependence of entropy on \mathcal{W}_{\max} . We first focus on the unique property of logarithmic functions.

How can we relate the number of ways of arranging the energy states of the system to thermodynamic properties? Statistical weights and probabilities multiply, Eq. 12.2.10, but extensive thermodynamic state functions add. Consider again the two systems in Figure 12.2.4. The extensive thermodynamic variable that is an indicator of these statistical probabilities must be additive, $S = S_1 + S_2$, where S_1 is a function of \mathcal{W}_1 for system 1, S_2 is a function of \mathcal{W}_2 for system 2, and S is the result for the composite system:

$$S(\mathcal{W}) \equiv S_1(\mathcal{W}_1) + S_2(\mathcal{W}_2) \quad 12.3.1$$

The statistical weight for the composite system is given by Eq. 12.2.10. Substituting $\mathcal{W} = \mathcal{W}_1 \mathcal{W}_2$ into Eq. 12.3.1 gives:

$$S(\mathcal{W}) = S(\mathcal{W}_1 \mathcal{W}_2) = S_1(\mathcal{W}_1) + S_2(\mathcal{W}_2) \quad 12.3.2$$

The logarithmic function has this property; $\ln(ab) = \ln a + \ln b$. Logarithmic functions convert products into sums. We can conjecture that $S \equiv \xi \ln \mathcal{W}$, with ξ a constant. Substituting $S_i \equiv \xi \ln \mathcal{W}_i$ into Eq. 12.3.1 for the composite system and the individual systems gives:

$$S(\mathcal{W}) = S_1 + S_2 = \kappa \ln \mathcal{W}_1 + \kappa \ln \mathcal{W}_2 = \kappa \ln(\mathcal{W}_1 \mathcal{W}_2) = S(\mathcal{W}_1 \mathcal{W}_2) \quad 12.3.3$$

as required by Eq. 12.3.2. Now that we have shown that logarithmic functions convert products of statistical weights into sums, can we prove our conjecture that $S \equiv \kappa \ln \mathcal{W}$? We need to prove that a logarithmic function is the only possible functional form that satisfies Eq. 12.3.2.⁶

Take a step back for a moment and assume that we don't know the functional dependence of entropy on \mathcal{W} . To find the functional dependence of the entropy on \mathcal{W} , find the differential of the total entropy for a composite system starting from Eq. 12.3.1:

$$dS(\mathcal{W}) = dS_1(\mathcal{W}_1) + dS_2(\mathcal{W}_2) \quad 12.3.4$$

Expressing the differentials in terms of \mathcal{W} , \mathcal{W}_1 , and \mathcal{W}_2 gives:

$$\frac{dS(\mathcal{W})}{d\mathcal{W}} d\mathcal{W} = \frac{dS_1(\mathcal{W}_1)}{d\mathcal{W}_1} d\mathcal{W}_1 + \frac{dS_2(\mathcal{W}_2)}{d\mathcal{W}_2} d\mathcal{W}_2 \quad 12.3.5$$

The derivatives are regular one-dimensional derivatives because S is a function of \mathcal{W} only, S_1 is a function of \mathcal{W}_1 only, and S_2 is a function of \mathcal{W}_2 only. However, from $\mathcal{W} = \mathcal{W}_1 \mathcal{W}_2$ and the product rule, $d\mathcal{W} = d(\mathcal{W}_1 \mathcal{W}_2) = \mathcal{W}_1 d\mathcal{W}_2 + \mathcal{W}_2 d\mathcal{W}_1$. Substitution for $d\mathcal{W}$ into the left-hand side of Eq. 12.3.5 results in:

$$\frac{dS(\mathcal{W})}{d\mathcal{W}} (\mathcal{W}_1 d\mathcal{W}_2 + \mathcal{W}_2 d\mathcal{W}_1) = \frac{dS_1(\mathcal{W}_1)}{d\mathcal{W}_1} d\mathcal{W}_1 + \frac{dS_2(\mathcal{W}_2)}{d\mathcal{W}_2} d\mathcal{W}_2 \quad 12.3.6$$

The coefficients of $d\mathcal{W}_1$ and $d\mathcal{W}_2$ on both sides of this last equation must be equal, since \mathcal{W}_1 and \mathcal{W}_2 are for separate systems that are independent of each other:

$$\frac{dS(\mathcal{W})}{d\mathcal{W}} \mathcal{W}_2 d\mathcal{W}_1 = \frac{dS_1(\mathcal{W}_1)}{d\mathcal{W}_1} d\mathcal{W}_1 \quad \text{and} \quad \frac{dS(\mathcal{W})}{d\mathcal{W}} \mathcal{W}_1 d\mathcal{W}_2 = \frac{dS_2(\mathcal{W}_2)}{d\mathcal{W}_2} d\mathcal{W}_2 \quad 12.3.7$$

Dividing the first equation by $\mathcal{W}_2 d\mathcal{W}_1$ and the second by $\mathcal{W}_1 d\mathcal{W}_2$ gives:

$$\frac{dS(\mathcal{W})}{d\mathcal{W}} = \frac{1}{\mathcal{W}_2} \frac{dS_1(\mathcal{W}_1)}{d\mathcal{W}_1} \quad \text{and} \quad \frac{dS(\mathcal{W})}{d\mathcal{W}} = \frac{1}{\mathcal{W}_1} \frac{dS_2(\mathcal{W}_2)}{d\mathcal{W}_2} \quad 12.3.8$$

Setting these two equations equal to each other and cross multiplying results in:

$$\mathcal{W}_1 \frac{dS_1(\mathcal{W}_1)}{d\mathcal{W}_1} = \mathcal{W}_2 \frac{dS_2(\mathcal{W}_2)}{d\mathcal{W}_2} \quad 12.3.9$$

However, \mathcal{W}_1 and \mathcal{W}_2 are independent of each other; we placed no restrictions on the number of molecules and the volume of each of these separate systems. This last equality can hold for any arbitrary values of \mathcal{W}_1 and \mathcal{W}_2 only if each equation is separately equal to a constant:

$$\mathcal{W}_1 \frac{dS_1(\mathcal{W}_1)}{d\mathcal{W}_1} = \kappa \quad \text{and} \quad \mathcal{W}_2 \frac{dS_2(\mathcal{W}_2)}{d\mathcal{W}_2} = \kappa \quad 12.3.10$$

We can then solve either of these last two equivalent equations for dS . Solving for dS_1 from the first equation and integrating gives:

$$dS_1 = \frac{\kappa}{\mathcal{W}_1} d\mathcal{W}_1 \quad \text{and} \quad S_1 = \int \frac{\kappa}{\mathcal{W}_1} d\mathcal{W}_1 = \kappa \ln \mathcal{W}_1 + \text{cst} \quad 12.3.11$$

which is Boltzmann's entropy to within an additive constant. These last equations hold for any system, so we can drop the subscripts. The constant is the value of S at absolute zero in temperature, S_0 , when \mathcal{W} is at its minimum value of $\mathcal{W} = 1$ (usually). In general then the entropy is defined as:

$$S \equiv \kappa \ln \mathcal{W}_{\max} + S_0 \quad 12.3.12$$

noting that \mathcal{W} is a maximum for a reversible process. The integral in Eq. 12.3.11 shows that the logarithmic function is the only function that satisfies the additivity of entropy for composite systems as expressed by Eq 12.3.2. We have not proved that the constant κ in Eqs. 12.3.10-12.3.12 is given by Boltzmann's constant k for a single system or k/\mathcal{N} for an ensemble. We will see that $\kappa = k/\mathcal{N}$ is required for agreement between the statistical definition of entropy and the thermodynamic definition of entropy. The Boltzmann entropy is then seen to follow directly from the supposition that processes are governed purely by statistical rules for the distribution of energy among the available energy states.

12.4 Larger Number of Ways of Arranging the Microstates Gives Larger Probability

The definition of entropy based on \mathcal{W}_{\max} , Eq. 12.2.1, is useful for understanding the fundamental issues linking probability theory and energy transfer. However, evaluating \mathcal{W}_{\max} for practical chemical processes is difficult. In addition, neither Boltzmann nor Gibbs used this formula to develop the concept of entropy. Instead, the Gibbs and original Boltzmann definitions used the probability of occurrence of a given energy state as the fundamental statistical measure. Luckily the probability of occurrence of a macrostate is directly proportional to the number of ways of arranging the particular macrostate. The more ways of arriving at a given set of distribution numbers the more probable the state. The more probable the state the more often it occurs, since there are many ways of arriving at that state. How can we recast the entropy of a system directly in terms of probability?

Entropy and Probability; $S = -k \sum p_i \ln p_i$: The number of ways of arranging the energy states of the system is given by Eq. 12.2.5. The logarithm of \mathcal{W} is:

$$\ln \mathcal{W} = \ln \mathcal{N}! - \sum_i \ln n_i! \quad 12.4.1$$

There is a very useful formula for approximating the factorial of very large numbers called Sterling's formula, which we derive in the addendum. Sterling's formulas for $N!$ are:

$$N! = \left(\frac{N}{e}\right)^N \quad \text{and} \quad \ln N! = N \ln N - N \quad (N \rightarrow \infty) \quad 12.4.2$$

Using Sterling's approximation for the factorials in Eq. 12.4.1 gives:

$$\ln \mathcal{W} = \mathcal{N} \ln \mathcal{N} - \mathcal{N} - \sum_i (n_i \ln n_i - n_i) \quad 12.4.3$$

Summations can be done in any order; summations are associative, $\Sigma(a_i + b_i) = \Sigma a_i + \Sigma b_i$.

Carrying the summation sign through the terms in parentheses gives:

$$\ln \mathcal{W} = \mathcal{N} \ln \mathcal{N} - \mathcal{N} - \sum_i n_i \ln n_i + \sum_i n_i \quad 12.4.4$$

Using Eq. 12.2.2, the \mathcal{N} and $\sum n_i$ cancel:

$$\ln \mathcal{W} = \mathcal{N} \ln \mathcal{N} - \sum_i n_i \ln n_i = \sum_i n_i \ln \mathcal{N} - \sum_i n_i \ln n_i \quad 12.4.5$$

Combining sums and distributing out the common factor of n_i , $\sum a_i b_i + \sum a_i c_i = \sum a_i(b_i + c_i)$, gives:

$$\ln \mathcal{W} = - \sum_i n_i (\ln n_i - \ln \mathcal{N}) \quad 12.4.6$$

The difference in the logs is the log of the ratio; $\ln a - \ln b = \ln(a/b)$:

$$\ln \mathcal{W} = - \sum_i n_i \ln \frac{n_i}{\mathcal{N}} \quad 12.4.7$$

Substitution of $\ln \mathcal{W}$ from Eq. 12.4.7, with the most probable distribution, into Eq. 12.2.1 for the entropy gives:

$$S = -k \sum_i \frac{n_i}{\mathcal{N}} \ln \frac{n_i}{\mathcal{N}} \quad 12.4.8$$

But n_i/\mathcal{N} is the probability of finding a system in state i , $p_i = n_i/\mathcal{N}$:

$$S = -k \sum_i p_i \ln p_i \quad 12.4.9$$

This last equation is an alternate form for the definition of entropy, in terms of the probability of occurrence of each energy state of the ensemble. Does this last equation have the correct behavior? When all systems in the ensemble are in the lowest energy state, $p_0 = 1$, and the entropy is zero. There is no dispersal of energy, since all systems are in the lowest state. As energy is transferred into the system, more states are accessible and there are more terms in the sum, increasing the entropy, Example 12.4.1. Eq. 12.4.9 is closely related to the Boltzmann H-function, which Boltzmann originally proposed for the definition of entropy.¹ For practical applications, Eq. 12.4.9 is particularly useful for calculating the residual entropy of crystals at absolute zero and the conformational entropy in studies of protein folding and binding.

Example 12.4.1: Internal Energy and Entropy from Distribution Numbers

The occupations for the energy states for 15 systems are given for three different cases in Figure 12.4.1. Determine the internal energy, in multiples of ϵ above the reference energy, and entropy for each distribution.

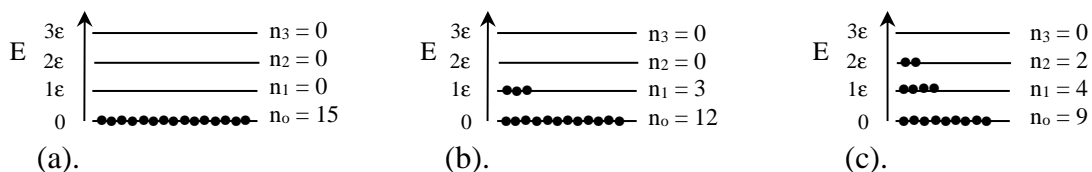


Figure 12.4.1: The occupations of the energy states for 15 systems at three different temperatures.

Answer: (a). The probabilities for each energy state are: $p_0 = 15/15$, $p_1 = 0$, $p_2 = 0$, $p_3 = 0$. The set of distribution numbers is (15,0,0,0) giving $\mathcal{W}(15,0,0,0) = 1$ and $U - U(0) = \langle E \rangle = 0$. The entropy using Eq. 12.4.9 is then:

$$S = -k(1 \ln 1) = 0$$

(b). The probabilities for each energy state are: $p_0 = 12/15$, $p_1 = 3/15$, $p_2 = 0$, $p_3 = 0$. The set of distribution numbers is (12,3,0,0) giving $\mathcal{W}(12,3,0,0) = 455$. The internal energy, using Eq. 12.2.6, and the entropy are:

$$U - U(0) = \langle E \rangle = \sum_i p_i E_i = (12/15) 0\epsilon + (3/15) 1\epsilon = 3/15 \epsilon$$

$$S = -k(12/15 \ln 12/15 + 3/15 \ln 3/15) = 6.9 \times 10^{-24} \text{ J K}^{-1}$$

(c). The probabilities for each energy state are: $p_0 = 9/15$, $p_1 = 4/15$, $p_2 = 2/15$, $p_3 = 0$. The set of distribution numbers is (9,4,2,0) giving $\mathcal{W}(9,4,2,0) = 75075$. The internal energy, and the entropy are:

$$U - U(0) = \langle E \rangle = \sum_i p_i E_i = (9/15) 0\epsilon + (4/15) 1\epsilon + (2/15) 2\epsilon = 8/15 \epsilon$$

$$S = -k(9/15 \ln 9/15 + 4/15 \ln 4/15 + 2/15 \ln 2/15) = 1.3 \times 10^{-23} \text{ J K}^{-1}$$

This example is schematic; Eqs. 12.2.1 and 12.4.9 are valid only in the limit of large occupations. The entropies are small because there are so few systems.

Residual Entropy at Absolute Zero Temperature: Eq. 12.3.12 gives the entropy of a system referenced to the entropy of the system at absolute zero. In Section 13.4 we will discuss the Third Law of Thermodynamics in more detail. The Third Law states that the absolute entropy of a pure, perfect crystalline solid is zero at absolute zero. In other words $S_0 = 0$ in Eq. 12.3.12 for most substances. However, some substances are difficult to prepare as perfect crystalline solids. When a substance freezes at the melting point, the available thermal energy, RT , can be greater than the intermolecular forces that determine the orientation of the molecules in the crystal lattice. Two examples are $\text{C}\equiv\text{O}$ and $\text{N}=\text{N}=\text{O}$. Crystals of solid $\text{C}\equiv\text{O}$ are experimentally found to have molecules in random orientations, rather than perfect alignment, Figure 12.4.3. When the crystals of the substance are cooled to absolute zero, this orientational randomness is “locked in,” because the kinetics of reorientation are too slow. The entropy caused by this randomness is called the **residual entropy** of the substance. The definitions of entropy in Eqs. 12.2.1 and 12.4.9 can alternatively be used to determine the residual entropy in crystalline substances.^{7,8}



Figure 12.4.3: The residual entropy for CO is the result of orientational randomness.

For a molecule like $\text{C}\equiv\text{O}$, there are two distinguishable ways of packing the molecules into the crystal lattice: $\text{C}\equiv\text{O C}\equiv\text{O}$ and $\text{C}\equiv\text{O O}\equiv\text{C}$. Let the number of ways of arranging a molecule in the lattice be w , for a single molecule. If there are N molecules in the system, there are w^N total ways of arranging the molecules in the crystal. The residual entropy is then given by

$$S_o = k \ln W_{\max} = k \ln w^N = N k \ln w = nR \ln w \quad (\text{molecular probabilities}) \quad 12.4.10$$

since Nk can be rewritten as $Nk = (N/N_A)(N_A k) = nR$. For $\text{C}\equiv\text{O}$ with $w = 2$ gives $S_o = 5.76 \text{ J K}^{-1} \text{ mol}^{-1}$. The R appears in this last equation because we are using w for a molecule instead of W for a full system; we multiply by Avogadro's number to put the entropy on a per mole basis.

Alternatively, Eq. 12.4.9 can also be used as a basis for the calculation of the residual entropy. If the intermolecular forces are weak compared to RT , then the probability of placing a molecule in the lattice as $\text{C}\equiv\text{O}$ and $\text{O}\equiv\text{C}$ are equal, $p(\text{C}\equiv\text{O}) = p(\text{O}\equiv\text{C}) = 1/2$ and Eq. 12.4.9 gives:

$$S_o = -R \sum_i p_i \ln p_i = -R (1/2 \ln 1/2 + 1/2 \ln 1/2) = R \ln 2 = 5.76 \text{ J K}^{-1} \text{ mol}^{-1} \quad 12.4.11$$

Once again, the R appears instead of k to put the entropy on a per mole basis. The advantage of Eq. 12.4.9 or 12.4.11 is that the probability based equation is useful if there is some partial alignment at the melting point; partial alignment occurs when the intermolecular forces are comparable to RT .

The Residual Entropy Can also be Calculated as the Entropy of Mixing: Another method for calculating the residual entropy is to find the entropy change as the lattice is formed from a mixture of molecules with the allowed orientations.⁷ The entropy of mixing is also a useful concept for understanding mixtures of gases and liquid solutions. Consider a crystalline lattice that can have molecules in two different orientations. For convenience, we diagram the two orientations as dark and light squares, Figure 12.4.4. The entropy of mixing is the difference in entropy of the initial state with unmixed orientations and the final state with each lattice site occupied by either orientation at random.

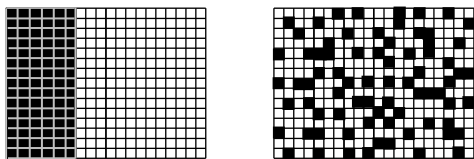


Figure 12.4.4: Entropy of mixing of n_1 molecules of orientation 1 with n_2 molecules of orientation 2. The number of ways of arranging the initial unmixed state is $W_{\text{unmixed}} = 1$. The distribution on the right is only one of a very large number of equally probable random configurational microstates.

Assume that there are n_1 molecules of orientation 1 with n_2 molecules of orientation 2. The number of ways of arranging the initial unmixed state is one, $W_{\text{unmixed}} = 1$. There is only one way of arranging the blocks of unmixed orientations. Assume the orientation of the molecule in a given cell is independent of the orientations of the molecules in the adjoining cells. The number of ways of arranging the random distribution of orientations in the mixed state is:

$$W_{\max} = \frac{N!}{n_1! n_2!} \quad \text{with } N = n_1 + n_2 \quad \text{(independent) 12.4.12}$$

and the entropy of mixing is given as:

$$\Delta_{\text{mix}}S = S_{\text{mixed}} - S_{\text{unmixed}} = k \ln \frac{W_{\text{mixed}}}{W_{\text{unmixed}}} = k \ln \left(\frac{N!}{n_1! n_2!} \right) \quad \text{(independent) 12.4.13}$$

Using Sterling's approximation for the factorials gives:

$$\Delta_{\text{mix}}S = k (\ln N! - \ln n_1! - \ln n_2!) = k (N \ln N - N - n_1 \ln n_1 + n_1 - n_2 \ln n_2 + n_2) \quad 12.4.14$$

Noting the cancellations using $N = n_1 + n_2$ and substituting for N in the $N \ln N$ term gives:

$$\Delta_{\text{mix}}S = k (N \ln N - n_1 \ln n_1 - n_2 \ln n_2) = k [(n_1 + n_2) \ln N - n_1 \ln n_1 - n_2 \ln n_2] \quad 12.4.15$$

Collecting terms in n_1 and n_2 results in:

$$\Delta_{\text{mix}}S = -k \left(n_1 \ln \frac{n_1}{N} + n_2 \ln \frac{n_2}{N} \right) \quad \text{(independent) 12.4.16}$$

Multiplying and dividing each term by N gives:

$$\Delta_{\text{mix}}S = -Nk \left(\frac{n_1}{N} \ln \frac{n_1}{N} + \frac{n_2}{N} \ln \frac{n_2}{N} \right) \quad \text{(independent) 12.4.17}$$

Note that $Nk = (N/N_A)(N_Ak) = nR$. The number ratios are the probabilities of the occurrence of the given orientation; that is $p_1 = n_1/N$ and $p_2 = n_2/N$:

$$\Delta_{\text{mix}}S = -nR (p_1 \ln p_1 + p_2 \ln p_2) \quad \text{(independent) 12.4.18}$$

This equation corresponds to Eq. 12.4.9, but it is written in terms of the probabilities of individual molecular states, instead of states of the ensemble. Alternatively the probabilities for each orientation are equal to the mole fractions, $x_1 = n_1/N$ and $x_2 = n_2/N$:

$$\Delta_{\text{mix}}S = -nR (x_1 \ln x_1 + x_2 \ln x_2) \quad \text{(independent) 12.4.19}$$

Applying Eq. 12.4.19 to the residual entropy of $\text{C}\equiv\text{O}$, if the two orientations are equally probable the mole fractions of left and right-pointing molecules is equal with $x_1 = x_2 = 1/2$, which gives the same result as Eq. 12.4.11, $S_o = \Delta_{\text{mix}}S$.

The entropy of mixing given by Eq. 12.4.19 is also applicable to mixtures of ideal gases. The requirement for independent occupations can also be satisfied by real gases and solutions that have equal forces between all the molecules. For mixtures with more than two constituents, Eq. 12.4.19 generalizes to a sum over all components:

$$\Delta_{\text{mix}}S = -nR \sum_i x_i \ln x_i \quad \text{(equal forces) 12.4.20}$$

Relating Sums Over Molecular States to Sums Over System States in an Ensemble: The expression for the entropy in Eq. 12.4.9 requires the probability of occupation for each energy state of the system. The Boltzmann distribution for the canonical ensemble allows us to calculate these probabilities. We will derive the Boltzmann distribution for the canonical ensemble in Sec. 12.5. The result is that the probability of occurrence, p_i , of a system with energy E_i is given by:

$$p_i = \frac{e^{-E_i/kT}}{Q} \quad \text{with} \quad Q \equiv \sum_i e^{-E_i/kT} \quad (\text{canonical ensemble}) \quad 12.4.21$$

The ensemble partition function, Q , is the sum of the Boltzmann weighting factors over all system energies. Q is the normalization constant for the probability distribution. Notice the correspondence to the Boltzmann distribution over molecular energy states, Eq. 8.9.5. The Boltzmann distribution over molecular states is derived from the canonical ensemble Boltzmann distribution. The first step in determining the relationship between molecular and system properties is to assume that the number of molecules in each system and the number of systems in the ensemble are so large that the most probable distribution is the overwhelmingly predominant distribution. Each system in the ensemble, for the vast majority of the time, is in its most probable distribution over molecular states and has the ensemble average energy.⁹ The number of ways of arranging the states of the ensemble is then the number of ways of arranging the states of the composite of \mathcal{N} identical distinguishable systems, $\mathcal{W}_{\max} = W_{\max}^{\mathcal{N}}$, where W_{\max} is the number of ways of arranging the states of a single system. The ensemble average of the entropy, Eq. 12.2.1, reduces to a function of W_{\max} for a single system:

$$S = \frac{k}{\mathcal{N}} \ln \mathcal{W}_{\max} = \frac{k}{\mathcal{N}} \ln W_{\max}^{\mathcal{N}} = k \ln W_{\max} \quad (\text{system}) \quad 12.4.22$$

as we stated without proof in conjunction with Eq. 12.2.1. (The systems in the ensemble are distinguishable, because each system is considered as having a fixed position within the ensemble.) All we are saying is that the canonical ensemble represents the properties of a single system for averaging purposes, and for a single system $S = k \ln W_{\max}$.

For a system of N molecules, Eqs. 12.4.10, 12.4.11, 12.4.18, and 12.4.19 show examples of how to relate properties of an individual molecule to W_{\max} . In Chapt. 32, we will show in detail that the properties of the system as a whole can be related to the properties of individual molecules and then the properties of the individual molecules can be related to the separate degrees of freedom of each molecule. Anticipating these results, we simply stated Eq. 8.9.5; for an ideal gas the Boltzmann distribution applies to individual molecular energy states and the corresponding partition function is the sum over the states of single molecules. We can use the Boltzmann distribution to describe the occupation of system energy states in an ensemble and the occupation of molecular energy states in a system. To make the distinction between single molecule properties and system properties, we denote individual molecule energy states as ϵ_i and the molecular partition function as lower case “ q ,” while the system energy states are given as E_i and the ensemble partition function as Q . The relationships between system and molecular states are summarized in Eqs. 12.4.23° and 12.4.24°, for an ideal gas:

<p>Canonical Ensemble, general: Sum over system energy states, E_i</p> $p_i = \frac{e^{-E_i/kT}}{Q} \quad Q \equiv \sum_i e^{-E_i/kT}$ $S = \frac{k}{\mathcal{N}} \ln \mathcal{W}_{\max} = -k \sum_i p_i \ln p_i$		<p>Single System, ideal, internal degrees of freedom: Sum over distinguishable single molecule states, ϵ_i</p> $p_i = \frac{e^{-\epsilon_i/kT}}{q} \quad q \equiv \sum_i e^{-\epsilon_i/kT} \quad (\text{internal}) \quad 12.4.23^\circ$ $S = nR \ln w_{\max} = -nR \sum_i p_i \ln p_i \quad (\text{internal}) \quad 12.4.24^\circ$
--	--	---

where w_{\max} is the maximum number of ways of arranging the states for a single molecule. The last equation for the entropy makes some assumptions about the distinguishability of the molecules in the system. We will focus on this issue in Chapt. 32. For now, note that Eq. 12.4.24° holds for distinguishable molecules or in general for internal degrees of freedom. The details need not concern us for now. Using the Boltzmann distribution to explore the contributions of individual molecular degrees of freedom to the entropy is instructive as we consider the underlying concepts in chemical equilibrium.

Energy is Dispersed by Increasing the Number of Accessible Degrees of Freedom: All molecular degrees of freedom contribute to the internal energy and entropy. The energies of translation, rotation, vibration, and electronic degrees of freedom add to give the energy of individual molecules. The energies of all the molecules in a system add to give the overall internal energy of the system. For real gases and condensed phases the total energy of a system also includes intermolecular interactions. A specific degree of freedom sometimes plays a predominant role in chemical processes. We can determine the influence of specific degrees of freedom if intermolecular interactions and interactions between the degrees of freedom are negligible.¹⁰ The energy of the system is then the sum of single molecule energies, and the energy of a single molecule is then the sum of the individual contributions of the degrees of freedom, Eq. 8.7.1:

$$E_i(\text{system}) = \sum_{\substack{k \\ \text{all molecules}}}^N \epsilon_k \quad \text{with} \quad \epsilon_k = \epsilon_{k,\text{trans}} + \epsilon_{k,\text{rot}} + \epsilon_{k,\text{vib}} + \epsilon_{k,\text{elec}} \quad (\text{ideal gas}) \quad 12.4.25^\circ$$

The contribution of an individual internal degree of freedom to the entropy of the system is calculated using Eqs. 12.4.23° and 12.4.24° for a single system.

Determining the number of internal degrees of freedom for a system can be a useful way of estimating the entropy of a substance. Non-linear molecules have three rotational degrees of freedom, while linear molecules have only two. Non-linear molecules, therefore, typically have a greater contribution to the rotational entropy of the system than linear molecules (Sec. 8.8). The contributions of individual degrees of freedom to the entropy are discussed further in Sec. 10.3.

Only Low Frequency Vibrations Contribute to the Entropy: In Sec. 8.9 we pointed out that at 298.2 K only vibrations with wave numbers less than about 500 cm^{-1} contribute to the internal energy of a substance, Figure 8.9.2. The vibrational internal energy and entropy at 298.2 K are listed in Table 12.4.1 for different frequency vibrations.

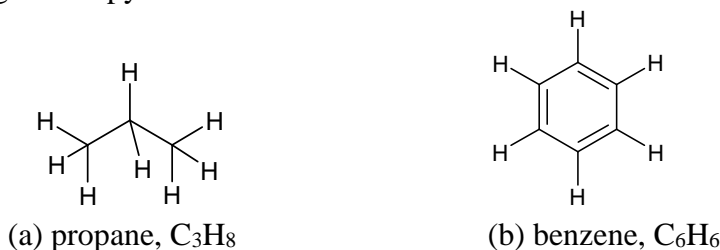
Table 12.4.1: Vibrational contribution at 298.2 K to the internal energy and entropy of an ideal gas.

Frequency (cm^{-1})	$U - U(0)$ (kJ mol^{-1})	S_{298} ($\text{J K}^{-1}\text{mol}^{-1}$)
1000	0.10	0.39
800	0.21	0.87
500	0.59	2.75
200	1.47	8.92
100	1.93	14.45
10	2.42	33.52

Counting low frequency vibrations is a useful way to estimate the entropy of a substance. In summary then, additional rotational and low energy vibrational degrees of freedom provide more accessible energy states for the system and therefore provide for greater energy dispersal.

Example 12.4.2: Vibrational Contribution to Absolute Entropy

Consider the gas phase entropy of propane and benzene, C_3H_8 and C_6H_6 . Which substance has the larger entropy?



Answer: Propane has two low frequency torsional vibrations, so the contribution from the vibrational degrees of freedom to the overall entropy is greater for propane. A normal mode analysis is not required for our purposes. However, the results are instructive. The low frequency vibrations are listed in Table 12.4.2.

Table 12.4.2: Low frequency vibrations for propane and benzene from molecular mechanics (MMFF), semi-empirical (AM1), a moderate level *ab initio* molecular orbital method (B3LYP/6-31G*), and experiment.^{11,12} Doubly degenerate vibrations are marked as *x2*. Torsions are in *italics*.

	Propane (cm^{-1})			Benzene (cm^{-1})			
	MMFF	AM1	Exp.	MMFF	AM1	B3LYP	Exp.
236	<i>174</i>	221	202	431 <i>x2</i>	371 <i>x2</i>	415 <i>x2</i>	405 <i>x2</i>
283	203	272	283	534 <i>x2</i>	617	622 <i>x2</i>	606 <i>x2</i>
433	413	366	375	677	648 <i>x2</i>	693	671
828	805	759	748	699	748	718	849

The experimental entropy of propane is slightly larger than benzene, $S^\circ(\text{propane,g}) = 269.91 \text{ J K}^{-1} \text{ mol}^{-1}$ versus $S^\circ(\text{benzene,g}) = 269.31 \text{ J K}^{-1} \text{ mol}^{-1}$, even though benzene has more atoms and a higher molar mass. The low frequency torsional vibrations of propane provide more accessible energy states than the corresponding bending vibrations in benzene.

Conformational Entropy: One important degree of freedom especially for studies of protein folding and binding interactions is the torsional degree of freedom.^{10,13} Torsions are low frequency vibrations that result from changes in dihedral angles. Figure 8.7.10 defines the dihedral angles along the backbone of a protein. The transition from an unfolded, random coil geometry to the folded active conformation of a protein requires restricting most of the torsional angles along the backbone. The folding process, viewed from the point of view of the backbone

conformation, is an entropically unfavorable process. Considering binding interactions, many enzyme substrates and pharmaceutically active enzyme inhibitors have torsional flexibility. Often when substrates or drugs bind to a receptor site of an enzyme, the torsional degrees of freedom are restricted. The “freezing out” of torsional degrees of freedom in the formation of molecular complexes is entropically very unfavorable. Consequently, medicinal chemists take torsional flexibility into account when designing new drugs. Designing drugs with limited conformational flexibility is one way to enhance the binding of drugs to their intended receptor sites. This process is called rigidification. One example is the enhanced binding of atrophine to the receptor site for acetylcholine in acetylcholine esterase, Figure 12.4.2:¹⁴

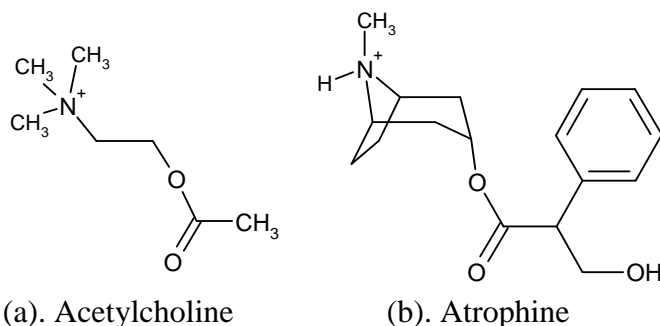


Figure 12.4.2: The decreased torsional flexibility attained by incorporating important bonds in a ring decreases the conformational entropy penalty for binding.

For illustrative purposes, we will focus on a very simple system, butane, $\text{CH}_3\text{CH}_2\text{CH}_2\text{CH}_3$. Butane has three torsional degrees of freedom. The barrier to internal rotation for methyl groups is so small that the $-\text{CH}_3$ groups are always in rapid motion, down to temperatures near absolute zero. The central $\text{C}-\text{C}$ torsion, however, is sensitive to the environment of the molecule. A plot of the torsional energy as a function of the $\text{C}-\text{C}-\text{C}-\text{C}$ dihedral angle is shown in Figure 12.4.2.

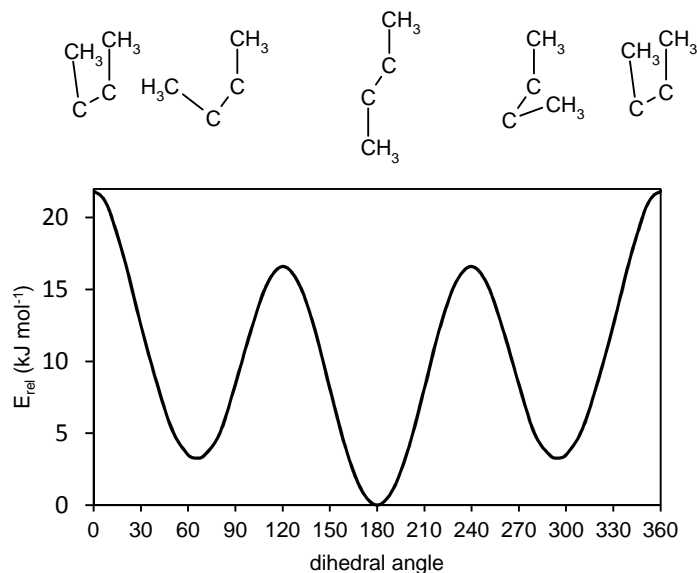


Figure 12.4.2: Dihedral angle dependence of the steric energy for butane using the MMFF forcefield. The *anti*-configuration is more stable than the *gauche* by 3.26 kJ mol^{-1} .

From a molecular mechanics perspective, the higher energy of the *gauche* compared to the *anti*-conformations is caused by the contribution of a one-fold torsional potential, Eq. 8.7.14, and Van der Waals interactions between the terminal methyl groups. The MMFF torsional potential for C–C–C–C is given by:

$$E_{\text{tor}} = 0.051 (1 + \cos \phi) - 0.341 (1 + \cos 2\phi) + 0.166 (1 + \cos 3\phi) \quad 12.4.16$$

The reference energy is taken as the *anti*-conformation. The two *gauche* states have the same energy. States with the same energy are said to be degenerate. Eq. 12.4.24 then gives the conformational entropy based on the energy difference between the *gauche* and *anti*-states.¹³

Example 12.4.3: Conformational Entropy

Calculate the conformational entropy for the torsional degree of freedom for butane. The energy difference between the *gauche* and *anti*-conformers in butane is 3.26 kJ mol⁻¹ using molecular mechanics and the MMFF forcefield.

Answer: The plan is to calculate the Boltzmann weighting factors for the conformational states and the sum of the weighting factors to find the partition function. The partition function is the probability normalization constant. The probabilities of occurrence of each of the energy states is then calculated and the sum of (p ln p) is then found over each state.

The Boltzmann weighting factor for the *anti*-conformer is 1, since we chose a reference energy of zero for the *anti*-conformer and $e^0 = 1$. The Boltzmann weighting factor for the *gauche*-conformer is:

$$e^{-\varepsilon_{\text{gauche}}/RT} = e^{-3.26 \times 10^3 \text{ J} / (8.314 \text{ J K}^{-1} \text{ mol}^{-1} \times 298.2 \text{ K})} = 0.2681$$

The partition function, Eq. 8.9.7, is the normalization for the probability distribution:

$$q = \sum_i e^{-\varepsilon_i/kT} = 1 + 0.2681 + 0.2681 = 1.5362$$

There is one *anti*-state and two degenerate *gauche*-states. The probabilities for each *gauche* and *anti*-energy state are then:

$$p(\text{gauche}) = e^{-\varepsilon_{\text{gauche}}/RT} / q = 0.2681 / 1.5362 = 0.1745$$

$$p(\text{anti}) = e^{-\varepsilon_{\text{anti}}/RT} / q = 1 / q = 1 / 1.5362 = 0.6510$$

The conformational entropy as given by Eq. 12.4.9 is for a sum over all the energies for a system. In this example, we are dealing with individual molecule energies. To convert to a per mole basis starting with the properties of a single molecule, Eq. 12.4.9 is multiplied by Avogadro's number and the sum is over the energy states for a single molecule, Eq. 12.4.24:

$$S = -N_A k \sum_i p_i \ln p_i = -R \sum_i p_i \ln p_i \quad (\text{molecular probabilities}) \quad 12.4.17$$

$$S = -8.314 \text{ J K}^{-1} \text{ mol}^{-1} [0.6510 \ln(0.6510) + 0.1745 \ln(0.1745) + 0.1745 \ln(0.1745)] \\ = 7.39 \text{ J K}^{-1} \text{ mol}^{-1}$$

When butane is bound in a restricted environment, most of this entropy is lost. For comparison, the difference in entropy between butane and *cis*-2-butene is $9.29 \text{ J K}^{-1} \text{ mol}^{-1}$. *cis*-2-Butene lacks the internal C–C–C torsion. Biochemists and medicinal chemists often use an average value of $19. \text{ J K}^{-1} \text{ mol}^{-1}$ for the entropy lost for each free internal rotation.¹⁵ Example 10.3.1 is also appropriate for your reading at this point.

The examples of conformational entropy and residual entropy show the intrinsic usefulness and power of the concept of entropy and the Boltzmann distribution. We introduced the Boltzmann distribution without proof in Chapter 8. Now we have the tools to derive this most central and important result.

12.5 The Boltzmann Distribution Gives the Equilibrium State

The equilibrium distribution of the available energy among the energy states of the ensemble is given by the most probable distribution, Figure 12.2.6. What is the most probable distribution? Are there some general principles that we can glean from the most probable distribution? The most probable distribution maximizes the entropy by maximizing \mathcal{W} or $\sum p_i \ln p_i$. To find the most probable distribution, we vary the probabilities of the occupations of the energy states, Figure 12.5.1. However, as we adjust the occupations we need to keep the number of systems in the ensemble and the total ensemble energy constant, since the ensemble is isolated from the surroundings.

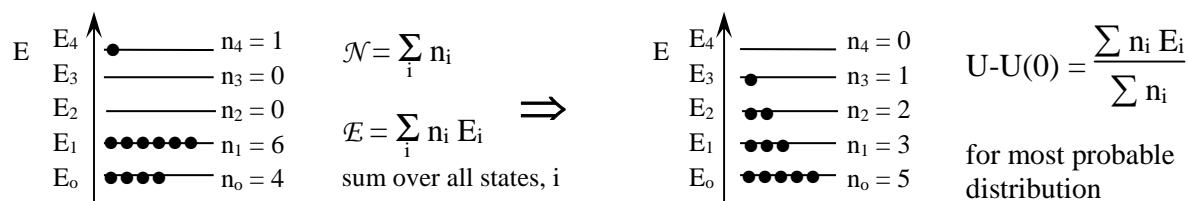


Figure 12.5.1: To find the most probable distribution, the probabilities of occurrence of the energy states are varied, while keeping the number of systems in the ensemble and the total ensemble energy constant.

States with High Energy Are Less Probable than States with Low Energy: To maximize the entropy we find the differential of the entropy with respect to changes in the probabilities and set the result equal to zero. For a general process, the change in entropy using Eq. 12.4.9 and the product rule is:

$$dS = -k \sum_i (p_i d \ln p_i + \ln p_i dp_i) \quad 12.5.1$$

Using the identity, $d \ln p_i = 1/p_i dp_i$, and separating the sums gives:

$$dS = -k \sum_i (1 + \ln p_i) dp_i = -k \sum_i dp_i - k \sum_i \ln p_i dp_i \quad 12.5.2$$

Note that $\sum p_i = 1$, by normalization, so that $d(\sum p_i) = \sum dp_i = 0$. In other words, if the probability of occurrence of one energy state increases, then the probability of some other state must decrease since the sum of the probabilities is always constant at 1. The first term on the right in Eq. 12.5.2 is then zero leaving:

$$dS = -k \sum_i \ln p_i dp_i \quad 12.5.3$$

The most probable distribution maximizes S:

$$dS = -k \sum_i \ln p_i dp_i = 0 \quad (\text{most probable}) \quad 12.5.4$$

However, the changes in the probabilities, as we just noted, are dependent on each other. The probabilities must sum to one and the total energy of the ensemble must be constant, since the ensemble is isolated: $\sum p_i = 1$ and $\sum E_i p_i = \mathcal{E}$. Correspondingly, we must keep $\sum dp_i = 0$ and $\sum E_i dp_i = 0$ as we adjust the probabilities of each state to find the maximum entropy. These two conditions are called constraints, and this type of problem is called a constrained maximization. To be as general as possible, note that constant values multiplying the constraints also give a zero result:

$$\alpha \sum_i dp_i = 0 \quad \text{and} \quad \beta \sum_i E_i dp_i = 0 \quad (\text{constraints}) \quad 12.5.5$$

Since both of these equations are equal to zero, adding Eqs. 12.5.4 and 12.5.5 still gives zero, which maximizes the entropy subject to the constraints:

$$dS = -k \sum_i \ln p_i dp_i + \alpha \sum_i dp_i + \beta \sum_i E_i dp_i = 0 \quad (\text{most probable}) \quad 12.5.6$$

The trick of adding in the constraints allows us to treat the changes in probabilities, dp_i , as independent variables. This constrained maximization method was developed by Joseph Louis Lagrange in 1804. The Lagrange method is discussed further in the addendum. The constants α and β are called **Lagrange multipliers** and the method is often called the “method of undetermined multipliers.” The important task we now face is to determine the value and meaning of α and β . Note that α is associated with the probability normalization and β is associated with conservation of the total ensemble energy, which in turn fixes the average energy of a system in the ensemble.

Combining sums and distributing out the common factor of dp_i in Eq. 12.5.6 gives:

$$dS = -k \sum_i (\ln p_i + \alpha + \beta E_i) dp_i = 0 \quad (\text{most probable}) \quad 12.5.7$$

Since the dp_i are now independent and can vary from 0 to 1, the only way to ensure that the sum in the last equation always gives zero is if the coefficient of each term of dp_i is separately equal to zero:

$$\ln p_i + \alpha + \beta E_i = 0 \quad (\text{most probable}) \quad 12.5.8$$

We can solve this last equation to find the most probable distribution:

$$p_i = e^{-\alpha} e^{-\beta E_i} \quad (\text{most probable}) \quad 12.5.9$$

The probability of occurrence of a given energy state is an exponentially decreasing function of the energy. Low energy states are more probable than high energy states; there are few high energy systems. The $e^{-\beta E_i}$ term is called the **Boltzmann weighting factor** for the energy state. We can find the value of $e^{-\alpha}$ through normalization using Eq. 12.2.7:

$$\sum_i p_i = \sum_i e^{-\alpha} e^{-\beta E_i} = e^{-\alpha} \sum_i e^{-\beta E_i} = 1 \quad 12.5.10$$

Solving for $e^{-\alpha}$ gives:

$$e^{-\alpha} = \frac{1}{\sum_i e^{-\beta E_i}} \quad 12.5.11$$

This factor is just the normalization for the probability distribution. Substitution of this last factor into Eq. 12.5.9 gives:

$$p_i = \frac{e^{-\beta E_i}}{\sum_i e^{-\beta E_i}} \quad (\text{most probable}) \quad 12.5.12$$

The sum in the denominator is called the partition function and is often given the symbol Q :

$$Q \equiv \sum_i e^{-\beta E_i} \quad (\text{canonical ensemble}) \quad 12.5.13$$

Substitution of this last definition into Eq. 12.5.12 gives the working formula:

$$p_i = \frac{e^{-\beta E_i}}{Q} \quad (\text{most probable}) \quad 12.5.14$$

The partition function, Q , is the sum of the Boltzmann weighing factors over all the accessible energy states of the ensemble. We will have much more to say about Q in the foundations of statistical mechanics chapter. However, for now, we will just consider Q as the normalization ensuring that $\sum p_i = 1$. We still need to find the value for β .

$\beta = 1/kT$: With the most probable distribution in hand we can now calculate the entropy change and average energy change for a process. The Boltzmann distribution applies when the system is at equilibrium. Taking the logarithm of the Boltzmann distribution, Eq. 12.5.14, gives:

$$\ln p_i = -\beta E_i - \ln Q \quad (\text{equilibrium}) \quad 12.5.15$$

Substitution of this last equation into the change in entropy, Eq. 12.5.3, gives:

$$dS = -k \sum_i (-\beta E_i - \ln Q) dp_i \quad (\text{equilibrium, cst.V}) \quad 12.5.16$$

The constant volume constraint results because a system in a canonical ensemble is at fixed volume, giving fixed energy states. Distributing through the summation gives:

$$dS = k\beta \sum_i E_i dp_i + k \ln Q \sum_i dp_i \quad (\text{equilibrium, cst.V}) \quad 12.5.17$$

Note once again that $\sum dp_i = 0$, giving the last term in Eq. 12.5.17 as zero, which leaves:

$$dS = k\beta \sum_i E_i dp_i \quad (\text{equilibrium, cst.V}) \quad 12.5.18$$

Notice that the summation in this last equation is the change in the ensemble averaged energy. The fundamental basis for the connection between statistical mechanics and thermodynamics is the equality of the internal energy and the ensemble averaged energy, $U - U(0) = \langle E \rangle$, Eq. 12.1.2. For a change during a process then $dU = d\langle E \rangle$. The energy states for a system are functions of the volume, but for a constant volume process the E_i are fixed. We can find the change in average energy for a constant volume process by taking the differential of the average energy, Eq. 12.2.6, for fixed E_i :

$$dU = d\langle E \rangle = \sum_i E_i dp_i \quad (\text{cst. V}) \quad 12.5.19$$

At constant volume the change in internal energy is given solely by changes in the populations of the energy states. Dividing Eq. 12.5.18 by 12.5.19, which is taken at constant volume, then gives one of the most fundamental results in thermodynamics:

$$\left(\frac{\partial S}{\partial U} \right)_V = k\beta \quad (\text{equilibrium, cst.V}) \quad 12.5.20$$

where k is Boltzmann's constant. This last equation is the key to finding the meaning of β . First notice the calculations of the change in entropy, Eq. 12.5.18, and the change in internal energy, Eq. 12.5.19, require detailed information about the energy states and the distribution of the energy among those states. However, the ratio of the changes always results in a number, $k\beta$. Remember that β is the Lagrange multiplier that is related to maintaining a fixed average energy for the system. Our examples, Figure 12.2.1 and 12.3.4, have shown that the entropy increases as energy is transferred into the system giving $(\partial S / \partial U) > 0$, which makes β a positive number. Summarizing then, β is an important characteristic of the system at equilibrium that is positive and characterizes the average energy. We are getting closer to the meaning of β .

The ideas that we have developed in this chapter are based on the everyday observation that heat is transferred from a hotter to a colder body. Eq. 12.5.20 characterizes thermal energy transfer. Consider energy transfer between two constant volume systems 1 and 2. Eq. 12.5.20 applies to both systems separately and solving for the entropy changes gives:

$$dS_1 = k\beta_1 dU_1 \quad \text{and} \quad dS_2 = k\beta_2 dU_2 \quad (\text{cst. V}) \quad 12.5.21$$

The total change in entropy for the transfer is the sum of both systems:

$$dS_{\text{tot}} = dS_1 + dS_2 = k\beta_1 dU_1 + k\beta_2 dU_2 \quad (\text{cst. V}) \quad 12.5.22$$

dS_{tot} is positive for a spontaneous process, as we determined in Sec. 12.2. The transfer of energy between the two systems is equal in magnitude and opposite in sign, $dU_1 = -dU_2$. Substituting for dU_2 in the last equation and assuming a spontaneous process gives:

$$dS_{\text{tot}} = k(\beta_1 - \beta_2) dU_1 > 0 \quad (\text{spontaneous, cst.V}) \quad 12.5.23$$

Now consider three cases for energy transfer; we need to choose the signs of $(\beta_1 - \beta_2)$ and dU_1 in Eq. 12.5.23 to give an overall positive entropy change:

1. If $\beta_1 < \beta_2$ then $(\beta_1 - \beta_2) < 0$ and we must have $dU_1 < 0$ to give $dS_{\text{tot}} > 0$. For $dU_1 < 0$ energy is transferred from system 1 to system 2. System 1 must be the hotter system.
2. If $\beta_1 > \beta_2$ then $(\beta_1 - \beta_2) > 0$ and we must have $dU_1 > 0$ to give $dS_{\text{tot}} > 0$. For $dU_1 > 0$ energy is transferred from system 2 to system 1. System 2 must be the hotter system.
3. If $\beta_1 = \beta_2$ then $dS_{\text{tot}} = 0$, there will be no energy transfer and the two systems must be at equilibrium.

We find that β is an indicator of the direction of energy flow and equilibrium. We can reconcile the three relationships if we set:

$$\beta = \frac{1}{kT} \quad 12.5.24$$

Then, for the three cases above, respectively:

1. If $\beta_1 < \beta_2$ then $(1/T_1 - 1/T_2) < 0$ giving $T_1 > T_2$ making system 1 the hotter system,
2. If $\beta_1 > \beta_2$ then $(1/T_1 - 1/T_2) > 0$ giving $T_1 < T_2$ making system 2 the hotter system,
3. If $\beta_1 = \beta_2$ then $T_1 = T_2$ and the systems are at the same temperature,

which agree with the previous conclusions. Energy is transferred from the hotter to the colder body for cases 1 and 2. Case 3 is particularly interesting. In Section 7.7 we introduced temperature as a measure of equilibrium, based on the Zeroth Law of thermodynamics.¹⁶ Two systems at equilibrium have the same temperature. Case 3 shows that the statistical and thermodynamic conceptions of temperature are consistent. The result that $\beta = 1/kT$ is consistent with our everyday conception of hotter and colder. Substituting Eq. 12.5.24 into Eq. 12.5.20 gives the most fundamental definition of temperature:¹⁶

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T} \quad (\text{equilibrium, cst. V}) \quad 12.5.25$$

We reached the same conclusion on purely thermodynamic grounds in Sec. 10.2, using the same process. You may want to read in Sec. 10.2 on the “Thermodynamic Definition of Temperature” to help reconcile this definition of temperature with your own personal experience. Temperature is an indicator of equilibrium. Systems at equilibrium have a Boltzmann distribution among their energy states characterized by the same temperature. We can summarize our results, written now in terms of the temperature, by substituting $\beta = 1/kT$ into Eqs. 12.5.12, 12.5.14, and 12.4.18. The equilibrium probability distribution for the energy states of the ensemble is:

$$p_i = \frac{e^{-E_i/kT}}{\sum_i e^{-E_i/kT}} \quad (\text{equilibrium}) \quad 12.5.26$$

The Boltzmann distribution written in terms of the partition function is given by:

$$p_i = \frac{e^{-E_i/kT}}{Q} \quad \text{with } Q \equiv \sum_i e^{-E_i/kT} \quad (\text{equilibrium}) \quad 12.5.27$$

The entropy change for a process is given by Eq. 12.5.18:

$$dS = \frac{1}{T} \sum_i E_i dp_i \quad (\text{equilibrium, cst.V}) \quad 12.5.28$$

These three equations with Eq. 12.2.6 and 12.3.9 provide a fundamental description of a system. Together these equations are used to predict the spontaneous direction for chemical reactions and the position of equilibrium.

Example 12.5.1: Determination of Temperature from the Boltzmann Distribution

The temperature is a single thermodynamic parameter that characterizes the distribution of systems among the allowed energy states. Determine the temperature for the energy distribution given in Figure 12.4.1c. Assume the energy state spacing is for a harmonic oscillator with frequency $200. \text{ cm}^{-1}$.

Answer: Taking the log of Eq. 12.4.27 gives a function that is linear in the energy states for the system:

$$\ln p_i = -\frac{E_i}{kT} - \ln Q \quad 12.5.29$$

The slope of a linear curve fit of $\ln p_i$ versus E_i gives slope = $-1/kT$. The units for the energy are joules. If the energy is converted to kJ mol^{-1} then the slope becomes slope = $-1/RT$. The energy state spacing is the fundamental vibration frequency (which you can read directly from the corresponding peak in the infrared spectrum) in kJ mol^{-1} , Eq. 8.9.2:

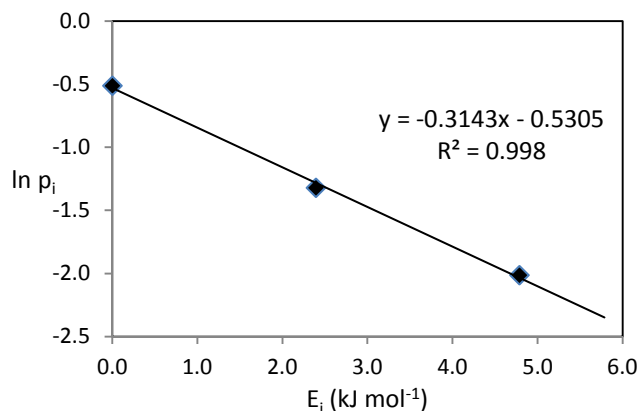
$$\begin{aligned} \varepsilon = \Delta E &= \frac{hc}{\lambda} = N_A hc \tilde{\nu} \\ &= 6.022 \times 10^{23} \text{ mol}^{-1} (6.626 \times 10^{-34} \text{ J s}) (2.998 \times 10^8 \text{ m s}^{-1}) (200. \text{ cm}^{-1}) (100 \text{ cm/1 m}) \\ &= 2393 \text{ J mol}^{-1} = 2.393 \text{ kJ mol}^{-1} \end{aligned}$$

The corresponding data and plot are given below.

i	E_i (kJ mol^{-1})	n_i	p_i	$\ln p_i$
0	0.000	9	0.600	-0.511
1	2.393	4	0.267	-1.322
2	4.785	2	0.133	-2.015

$$\begin{aligned} \text{slope} &= -1/RT = -0.3143 \text{ kJ}^{-1} \text{ mol} \\ RT &= 3.182 \text{ kJ mol}^{-1} \end{aligned}$$

$$T = \frac{3.182 \text{ kJ mol}^{-1}}{8.314 \times 10^{-3} \text{ kJ K}^{-1} \text{ mol}^{-1}} = 383. \text{ K}$$



12.6 Entropy and Heat Transfer

The entropy is defined through the statistical weights or the probability of occupation of the energy states of a system, Eqs. 12.2.1 and 12.4.9. These definitions are crucial in developing an understanding of the concept of entropy. However, these formulations are often cumbersome to apply in practical situations. A more easily applied definition of entropy in terms of thermodynamic variables is necessary. We began our discussion of entropy in Secs. 10.1 and 12.2 by focusing on the simple idea that heat is transferred from hotter to colder objects. That observation was then shown to be the result of the maximization of the number of accessible energy states for the two objects. However, we did not yet make a direct connection to heat transfer. The statistical definition of heat transfer is the link to the thermodynamic definition of entropy. Heat transfer results in a change of average energy and entropy for a system.

In general, a change in average energy for a process can result from changes in the occupations of the energy states, dn_i , or a change in the energy states themselves, dE_i . Taking the derivative of Eq. 12.2.3 for a general process:

$$d\langle E \rangle = \frac{1}{\mathcal{N}} \left(\sum_i E_i dn_i + \sum_i n_i dE_i \right) \quad 12.6.1$$

We can make this expression look a bit more familiar by realizing that work results from a change in volume of the system. Consider a gas confined in a container of volume V . The energy states for translation depend on the volume of the container:

$$dE_i = \left(\frac{\partial E_i}{\partial V} \right)_{n_i} dV \quad 12.6.2$$

Substitution of this explicit volume dependence back into Eq. 12.6.1 gives:¹⁷

$$d\langle E \rangle = \frac{1}{\mathcal{N}} \left(\underbrace{\sum_i E_i dn_i}_{\text{heat}} + \underbrace{\sum_i n_i \left(\frac{\partial E_i}{\partial V} \right)_{n_i} dV}_{\text{work}} \right) \quad 12.6.3$$

This last equation is now similar to the thermodynamic formula for the change in internal energy for a reversible process, $dU = dq_{\text{rev}} - P dV$. The first term in Eq. 12.5.3 gives the heat transfer for a reversible process while the second term gives the work. We can also relate the heat transfer to the change in probability of occupation of the energy states using the definition of the probability from Eq. 12.2.5:

$$dq_{\text{rev}} = \frac{1}{\mathcal{N}} \sum_i E_i dn_i = \sum_i E_i dp_i \quad (\text{reversible}) \quad 12.6.4$$

The transfer of heat into a system at constant volume results in a change in the occupations of the energy states, as we diagrammed in Figure 12.2.1.

The Thermodynamic Definition of Entropy: For a general process, the change in entropy is given by Eq. 12.5.28. Substituting the statistical definition of heat transfer, Eq. 12.6.4, into Eq. 12.5.28 gives:

$$dS = \frac{1}{T} \sum_i E_i dp_i = \frac{dq_{\text{rev}}}{T} \quad (\text{reversible}) \quad 12.6.5$$

This last equality gives the thermodynamic definition of the entropy change for the system. This definition is central and fundamental for the development of the thermodynamic theory of spontaneous processes and equilibrium. The result shows the direct relationship between heat transfer and entropy. As discussed in Sec. 12.2 and Eq. 12.5.4, the entropy is determined by \mathcal{W}_{\max} , which corresponds to a reversible process, so the heat transfer must also be reversible. The implications of the thermodynamic definition of entropy are developed in the next four chapters.

12.7 When Are Entropy and Disorder Related?

In general entropy and disorder are not directly related.¹⁸⁻²² Disorder is often used as a way of explaining the meaning of entropy; however, this relationship is not valid in many cases. The source of the confusion is primarily centered on the issue of indistinguishability, which we will discuss in Chapt. 32.²² One problem is that the scientific definition of disorder is not well established. Consider the two systems in Figure 12.7.1; both have $20 \times 15 = 300$ tiles. Which system is more disordered? There is only one way of arranging the tiles to give the exact configuration shown for each system, $W = 1$. Both of the systems in the figure have the same number of ways of occurrence, so they are both equally probable as single configurations. The number of ways of arranging the tiles for a random distribution is $W = 2^{300}$. The most probable distribution is half black and half white, $W_{\text{most probable}} = 300! / (150! 150!) = 2.04 \times 10^{90}$, where the black and white tiles can occur in any position. However, the exact configurations in the figure are both equally probable and so are equally “disordered.”²¹ It is not uncommon for the scientific applications of everyday concepts to be more restricted than their everyday usage. Disorder is not a scientifically precise concept.

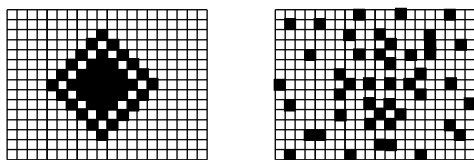


Figure 12.7.1: Each configuration has an equal probability of occurrence, $W = 1$ for both. Each system has the same amount of “disorder” as compared to all the ways of arranging the states of the system with random dark and light squares: $W = 2^{300}$.

Is disorder a useful analogy for entropy, even if the concept of disorder is only loosely associated with measurable thermodynamic properties? Consider the gas phase absolute entropy of propane and benzene, C_3H_8 and C_6H_6 , Example 12.4.1. Which substance has the larger entropy? Which substance has the greater disorder? Using disorder as an analogy for entropy is not helpful in determining the relative entropy of substances. In general you will find few analogies in physical chemistry texts. Analogies can often be misleading. However, while the concept of disorder is not particularly useful, configurational entropy does play an important role in the change of entropy for chemical processes.^{10,17}

The conformational entropy of molecular torsions and the residual entropy at absolute zero for $C \equiv O$ are examples in which configurational randomness plays an important role in chemical processes. Three things are necessary to associate configurational randomness with entropy:

1. The configurational randomness must be associated with the dispersal of energy.
2. The configurational randomness must lead to distinguishable states.
3. The system must have some mechanism for randomly accessing each possible microstate of the system.

Condition 1 is met if the configurational randomness results in the dispersal of energy in the underlying systems or molecules. For example, when considering the residual entropy at absolute zero of $\text{C}\equiv\text{O}$, the random orientation of the $\text{C}\equiv\text{O}$ molecules is associated with the electronic and vibrational energy of each molecule in the crystal lattice. Specifically, configurational randomness and entropy are related if the number of ways of arranging the states of the system can be factored into two terms, one relating to positional or orientational randomness and the other term the remaining degrees of freedom of the system. W can be factored if the remaining degrees of freedom are the same for each configuration of the molecule or system.^{10,17,23,24}

$$S = k \ln(W_{\text{config}} W'_{\text{sys}}) \quad (\text{equivalent configurational subsystems}) \quad 12.7.1$$

where W_{config} is the number of ways of arranging the states of the system for the different spatial configurations and W'_{sys} is the number of ways of arranging the energy states for each configuration of the system. When finding differences in entropy for a process, the statistical weights for the underlying degrees of freedom cancel:

$$\Delta S = S_2 - S_1 = [k \ln(W_{\text{config},2} W'_{\text{sys}})] - [k \ln(W_{\text{config},1} W'_{\text{sys}})] = k \ln\left(\frac{W_{\text{config},2}}{W_{\text{config},1}}\right) \quad 12.7.2$$

giving the appearance that the configurational change alone is responsible for the entropy change. However, there would be no change in entropy were the total statistical weight not associated with energy dispersal, as originally given in Eq. 12.7.1.

Condition 2 will be a major topic in the later chapter on the foundations of statistical mechanics. In general, only distinguishable states contribute to the entropy of a system. For example, the two nitrogen atoms in N_2 are indistinguishable. To follow the positions of the atoms, label the two nitrogen atoms as $\text{N}'\equiv\text{N}''$. The configurational change $\text{N}'\equiv\text{N}'' \rightarrow \text{N}''\equiv\text{N}'$ is indistinguishable and therefore does not contribute to configurational entropy. Our development in this chapter has focused on the system as a whole by using the canonical ensemble. We will discuss how to relate system energies, partition functions, and entropies to molecular properties in the subsequent foundations chapter. Indistinguishability plays an important role in the relationship of molecular properties to ensemble properties.²²

Condition 3 is the result of the **ergodic hypothesis**. The time average of a system may be replaced by an ensemble average only if the ergodic hypothesis is satisfied. The ergodic hypothesis requires that each possible microstate of the system be accessed at random with equal *a priori* probability. The ergodic hypothesis is satisfied by most chemical systems, except for some magnetic interactions.

Energy is Dispersed by Increasing Spatial Dispersion: For an ideal gas the number of ways of arranging the states of the system can be factored into a term that depends on the volume, V , and a term that depends on the molecular translational kinetic energy and the internal degrees of freedom. The configurational entropy for the location of the ideal gas molecules can then be used to calculate the change in entropy for a process for a change in volume using Eq. 12.7.2. Consider the isothermal expansion of an ideal gas from volume V_1 to a larger volume V_2 . The

result of the change in configurational entropy is the dispersal of molecular kinetic energy and potential energy from a small volume to a large volume, which is a favorable process. Suppose that N molecules are confined in a piston by a stop, Figure 10.1.2. The piston is isolated from the surroundings. The initial pressure of the gas is P_1 and the initial volume is V_1 . The other side of the piston is at vacuum. The initial number of ways of arranging the states of a single molecule is proportional to the initial volume and for the N molecules. For the system of N molecules then $W_{\max,1} \propto V_1^N$. The stop is removed allowing the gas to expand to occupy a new volume V_2 . The number of ways of arranging the states for the final volume are $W_{\max,2} \propto V_2^N$. Using Eq. 12.7.2, the change in entropy is given by:

$$\Delta S = k \ln \left(\frac{W_{\text{config},2}}{W_{\text{config},1}} \right) = k \ln \left(\frac{V_2^N}{V_1^N} \right) = Nk \ln \left(\frac{V_2}{V_1} \right) = nR \ln \left(\frac{V_2}{V_1} \right) \quad 12.7.3$$

If $V_2 > V_1$ then $\Delta S > 0$, showing that the expansion of an ideal gas into a vacuum is spontaneous for an isolated system.

12.8 Summary – Looking Ahead

The development of the concept of entropy is often considered the single most important development in physical science in the 19th century. This single concept laid the foundation for all modern science. Entropy is the statistical and thermodynamic extensive state function that determines the spontaneity of all macroscopic physical processes. The equation $S = k \ln W$ is inscribed on Boltzmann's tomb in Vienna, highlighting the contribution of the concept of entropy and of Boltzmann to our understanding of the physical universe.

We can conveniently combine both the statistical and thermodynamic perspectives on entropy by considering Eqs. 12.2.1, 12.5.7, 12.5.28, and 12.6.5 together to gain greater insight into the meaning of entropy:

$$S = \frac{k}{\mathcal{N}} \ln \mathcal{W}_{\max} = -k \sum_i p_i \ln p_i \quad \text{and} \quad dS = \frac{dq_{\text{rev}}}{T} = -k \sum_i \ln p_i dp_i \quad 12.8.1$$

Energy in the form of heat is transferred from hotter to colder bodies, which increases the number of ways of arranging the states of the composite system. Heat transfer into a system at constant volume results in changes in the occupations of the energy states, increasing the number of ways of arranging the energy states of the system. The number of ways of arranging the states of the system is a measure of energy dispersal. Energy dispersal corresponds to an increase in the probability of occurrence of many different energy states of the system. The equilibrium state is the most probable state. The most probable state is characterized by a single parameter called the temperature. Bodies in contact at equilibrium have the same temperature. Entropy is the extensive state function that characterizes energy dispersal. Entropy always increases for a spontaneous process in an isolated system because processes progress from states of lower probability to states of higher probability.²⁵

In the next chapter, we explore the thermodynamic definition of entropy, $dS = dq_{\text{rev}}/T$. The thermodynamic definition of entropy is often easier to use in practical problems, especially large scale problems. However, we can freely switch between the statistical and thermodynamic viewpoints at any point to find the practical solution to a new problem. The statistical definition of entropy reinforces the fact that random chance plays a determining role in chemical equilibria.

12.9 Addendum: Probability, Sterling's Approximation, and Constrained Maximization

Probability theory is used to predict the likelihood of a given set of events. For example, if you are a gambler, what is the probability of being dealt a royal flush or of a coin landing heads-up four times in a row? Consider N possible single events. For example, for coin tosses there are two possible events: heads (H) and tails (T) and $N = 2$. For dice games, rolling a single die gives 6 possible events: 1, 2, 3, 4, 5, and 6 giving $N = 6$. For selecting playing cards there are 52 different cards or events, $N = 52$.

We will make three important assumptions to establish the probability of occurrence of various events; we will assume that each event is independent of any other event, the events are mutually exclusive, and each individual outcome has an equal probability.²⁵ The assumption of independence means that the results of any single event are uncorrelated with previous events. Events are mutually exclusive if the occurrence of event A means that event B cannot occur. A coin cannot land with both heads and tails showing; heads and tails are mutually exclusive. We also assume that each individual event has an equal *a priori* probability of occurrence. For coin tosses, heads is just as likely to occur as tails. We always promise to use "honest" dice. Let the number of ways for a given event to occur be n_A . Then the probability of occurrence for the event A is:

$$p_A = n_A/N \quad (\text{independent, mutually exclusive}) \quad 12.9.1$$

For example, for coin tosses let the event be the occurrence of heads. There is only one way for a coin to land heads-up, so $n_H = 1$ out of $N = 2$ possible events, giving the probability for a coin landing heads-up as $p_H = 1/2$. For throwing a die let the event be the occurrence of a 3. There is only one way for a single die to land with a 3 showing, $n_3 = 1$ out of $N = 6$ possible events, giving the probability for a die landing with 3 showing $p_3 = 1/6$. The assumption of equal *a priori* probability requires that the probability of a die landing with a 1 showing is also $p_1 = 1/6$. What is the probability of selecting a ♥ from a deck of cards? There are $n_♥ = 13$ hearts in every deck, so the probability is $p_♥ = n_♥/N = 13/52$.

Now what happens for multiple events? Multiple events can occur as repeated selections in time or space. Flipping a single coin repeatedly is an example of a repeated selection in time. Filling several boxes is a multiple event in space. Probabilities for individual events combine in two different ways to give the probability of a multiple event.²⁷

Probabilities Add for the Occurrence of Events A OR B: Consider first a single event that can happen in multiple ways. What is the probability of selecting a Queen from a deck of cards? The probability of selecting a Q♥, or a Q♦, or a Q♣, or a Q♠ is each individually $1/52$. The probability of drawing a Queen is the sum of the probabilities of drawing a Queen of any suit: $p_Q = p_{Q♥} + p_{Q♦} + p_{Q♣} + p_{Q♠} = 4/52$. The probability of drawing any Queen is four times as probable as selecting any single particular Queen. Now consider multiple events. What is the probability of rolling a 6 on a single die in two rolls? The result can occur on the first roll or the second roll. The probabilities again add; the probability on the first selection is $1/6$ and on the second selection is $1/6$ adding to give $2/6$ overall. Probabilities of single events add when the selection can be expressed as an "OR" combination of events.²⁵

Probabilities Multiply for the Occurrence of Events A AND B: Consider a multiple event in time. What is the probability of two successive heads-up coin tosses? The complete list of

possibilities for two successive coin tosses is: HH, HT, TH, and TT. The occurrence of two successive heads-up, HH, is only one of the four possible outcomes giving a probability of $p_{HH} = 1/4$. The probabilities for the single events multiply, because we require H on the first toss and H on the second toss. The probability of heads on the first toss is p_H and on the second toss is p_H , giving for H and H, $p_{HH} = p_H^2$. Probabilities of single events multiply when the selection can be expressed as an “AND” combination of events. The probability of three successive H’s, HHH, is one out of eight or $p_{HHH} = (1/2)^3$:

HHH HHT HTH THH HTT THT TTH TTT

Notice that the specific arrangement HTH has the same probability of occurrence as HHH, $p_{HTH} = (1/2)^3$. The specific arrangement TTH also has the same probability as HHH. These equal probabilities result since the individual events are independent and have equal *a priori* probability. Another example of independence comes from the consideration of the third throw of three. Consider three successive coin tosses. Assume that the first two tosses both gave heads, HH. What is the probability that the third throw is also H? You might be tempted to say that the likelihood of the third throw being H is small, since the first two throws were both H, but this assumption is not true. After two successive heads, the probability of heads on the third throw is still $1/2$. The individual events are independent. The coin has no way of knowing that the first two throws were both heads. This result is shown explicitly by noting in the list of possible results that HHH and HHT are equally probable.

Usually Order Doesn’t Matter: Usually the distribution of events is important, but the specific order of multiple events does not make a difference in a physical property. For example, consider the three coin tosses. Notice that if order doesn’t matter, HHT, HTH, and THH each have two H’s and one T. Taken together, the probability of HHT or HTH or THH is three times more probable than any specific arrangement when order is taken into account. Grouping the outcomes for three coin tosses then gives the probabilities, irrespective of order:

HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
	⏟			⏟			
p_H^3	$3(p_H^2 p_T)$			$3(p_H p_T^2)$			p_T^3
$W(3,0) = 1$	$W(2,1) = 3$			$W(1,2) = 3$			$W(0,3) = 1$

The statistical weights, W , multiply the probability products. The weights result from the addition of the probabilities of the specific outcomes, in answer to “OR” questions. For example, HHT or HTH or THH are equivalent outcomes for our purposes. The statistical weights are specific to the distribution of H and T outcomes. The number of H and T results are listed in the order, $W(n_H, n_T)$, as distribution numbers for our coin tossing example. Since $p_H = p_T$ the products of the individual event probabilities for each set of distribution numbers is the same, $p_H^3 = p_H^2 p_T$, however we find that the probability of 2H’s and 1T, in any order, is three times as probable as HHH or TTT. The statistical weights are central to the development of the concept of entropy. The explicit enumeration of every possible outcome for multiple events is difficult to do, especially for large numbers of events. We need to consider counting problems in general to find a way of calculating the statistical weights for large systems. Central to the concept of a statistical weight is the idea of a permutation.

Permutations are the Rearrangements of the Order of a Series of Events: We often care about the results of a series of events, but not the order in which particular events occur. The concept of permutations helps us to count the number of equivalent series of events. Consider selecting three balls to place in a given box. The balls are identical, but we will assume they are labeled so we can observe the order of the selection events. The balls may be selected in the order:

1,2,3 1,3,2 2,1,3 2,3,1 3,1,2 or 3,2,1

Each series gives the same result—three balls in the box. Each specific order is called a **permutation**; which is a specific arrangement of the order of a series of events. For three balls in one box there are six equivalent permutations. How can we calculate the number of permutations for a series of events in general? Consider our present example; for the first selection we have 3 balls to choose from, leaving two remaining balls. For the second selection we have 2 balls to choose from, leaving only 1 ball to choose for the third event. The total number of ways of choosing the balls is given by $3 \cdot 2 \cdot 1$ or $3!$. The $3! = 6$ different permutations are shown above. For N objects there are $N!$ permutations or ways that the objects may be chosen. For example, there are N objects to choose for the first selection, $(N-1)$ objects to choose for the second selection and so on till the last object: $N! = N(N-1)(N-2)\dots(1)$. But what if there are more balls than the number to be selected?

Given N Choose n : Consider the number of ways of filling a box with n balls chosen from N total balls. In other words, given N balls choose n . Try an example first with $N = 4$ and $n = 3$. We need to choose 3 balls; for the first selection we have $N = 4$ to choose from, for the second selection we have $N-1 = 3$ to choose from, and for the third we have $N-2 = 2$ to choose from. For example, the three chosen balls might be balls 1, 2, and 3 or any permutation of 1, 2, and 3:

1,2,3 1,3,2 2,1,3 2,3,1 3,1,2 3,2,1

The $3!$ total permutations are equivalent from the perspective of a unique selection of the 4 initial balls. Balls 1, 2, and 3 are selected each time, but we don't care about the order of selection. The other possibilities, without listing the corresponding permutations are:

1,2,4 and the five permutations
 1,3,4 and the five permutations
 2,3,4 and the five permutations

There are four unique selections of the 4 balls, choosing 3 at a time. How many unique series are there for a general case? Consider the sequence of choices from our example and the number of ways of making each choice for the three balls: $N(N-1)(N-2)$. In general the series continues for n selections:

ways of selecting n objects = $N(N-1)\dots(N-n+1)$

However, these selections include all the permutations. In general there are $n!$ permutations of the n selected objects. The number of ways of selecting n distinguishable objects from N , irrespective of the order of the chosen objects is:

$$C[N \text{ choose } n] = \frac{N(N-1)\dots(N-n+1)}{n!} \quad 12.9.1$$

We can check this formula against our example for $N = 4$ and $n = 3$. The last term in the numerator will be $(N-n+1) = (4 - 3 + 1) = 2$:

$$C[4 \text{ choose } 3] = \frac{4(3)(2)}{3!} = \frac{4 \cdot 3 \cdot 2}{3 \cdot 2 \cdot 1} = 4 \quad 12.9.2$$

which agrees with our specific enumeration. But, what if there is more than one box to put the balls in?

The Number of Ways of Arranging a Set of Distribution Numbers: Consider $N = 7$ balls selected at random to fill three boxes. Assume that $n_a = 3$ balls are in box a, $n_b = 2$ balls in box b, and then $n_c = 2$ balls are left for box c. How many different unique arrangements are there? For the first box, box a, there are N balls and we need to select $n_a = 3$. The number of ways to fill box a is just Eq. 12.7.1 for $[N \text{ choose } n_a]$:

$$W_a = C[N \text{ choose } n_a] = \frac{N(N-1)\dots(N-n_a+1)}{n_a!} \quad 12.9.3$$

There are $(N-n_a)$ balls remaining. For our example, $(N-n_a) = (7 - 3) = 4$ remaining balls to choose from. So for box b, we need to select $n_b = 2$. The numbers of ways is just $[4 \text{ choose } 2]$ or in general $[(N-n_a) \text{ choose } n_b]$:

$$W_b = C[(N-n_a) \text{ choose } n_b] = \frac{(N-n_a)(N-n_a-1)\dots(N-n_a-n_b+1)}{n_b!} \quad 12.9.4$$

Now there are $(N-n_a-n_b)$ balls remaining. For our example, $(N-n_a-n_b) = (7 - 3 - 2) = 2$ remaining balls to choose from. So for box c, we need to select $n_c = 2$. The numbers of ways is just $[2 \text{ choose } 2]$ or in general $[(N-n_a-n_b) \text{ choose } n_c]$:

$$W_c = C[(N-n_a-n_b) \text{ choose } n_c] = \frac{(N-n_a-n_b)(N-n_a-n_b-1)\dots(1)}{n_c!} \quad 12.9.5$$

The total number of ways of selecting balls for the boxes is the product of Eqs. 12.9.3, 12.9.4, and 12.9.5, $W_{\text{total}} = W_a W_b W_c$:

$$W_{(n_a, n_b, n_c)} = \frac{N(N-1)\dots(N-n_a+1)}{n_a!} \frac{(N-n_a)(N-n_a-1)\dots(N-n_a-n_b+1)}{n_b!} \frac{(N-n_a-n_b)(N-n_a-n_b-1)\dots(1)}{n_c!} \quad 12.9.6$$

where we listed the number of balls in each box as the argument for the number of ways of making the selections, W . You will note from our discussion in Sec. 12.2 that the values (n_a, n_b, n_c) are called the distribution numbers for the occupations of each box. Notice that the numerator of Eq. 12.9.6 is really just the product of the complete series of numbers from N down to 1; the numerator is just $N!$:

$$W_{(n_a, n_b, n_c)} = \frac{N!}{n_a! n_b! n_c!} \quad (\text{distinguishable}) \quad 12.9.7$$

which is the number of ways of selecting distinguishable objects subject to the given set of distribution numbers, but irrespective of order within each box. The factorials in the denominator are the number of permutations of the balls in each box, to ensure that the count is irrespective of the order in each box. For our specific example, with $N = 7$ and distribution numbers $(3, 2, 2)$,

$W = 7!/(3! 2! 2!) = 210$ ways of arranging the balls among the boxes with the given set of distribution numbers. The number of ways of selecting a set of objects and the number of ways of arranging the objects for a given set of distribution numbers are equivalent views of the same counting process. Eq. 12.9.7 can be applied to the distribution of energy among the systems in an ensemble. The balls in our example are the systems in the ensemble and the boxes are the different energy states of the ensemble. Finding the distribution that gives a maximum in W determines the most probable state and the corresponding entropy.

Sterling's Approximation for the Factorial of a Distribution Number is Valid for Large Systems: The enormous number of systems in an ensemble allows us to focus on the average properties for the ensemble instead of looking at the details of each system in the ensemble. The average behavior becomes the overwhelmingly predominant behavior in the thermodynamic limit. The thermodynamic limit also allows the factorials to be approximated with no significant error. The logarithm of $N!$ can be written as the sum:

$$\ln N! = \sum_{i=1}^N \ln i \quad \text{for example } \ln 3! = \sum_{i=1}^3 \ln i = \ln 1 + \ln 2 + \ln 3 = \ln(3 \cdot 2 \cdot 1) \quad 12.9.10$$

For very large N , we can approximate the summation as an integral and then using Table 1.4.1:

$$\ln N! = \sum_{i=1}^N \ln i \approx \int_1^N \ln i \, di = [i \ln i - i]_1^N = N \ln N - N + 1 \approx N \ln N - N \quad 12.9.11$$

which is the form of Sterling's approximation that we introduced in Eq. 12.4.2. An approximation that is more accurate for small numbers is:

$$\ln N! = N \ln N - N + \frac{1}{2} \ln 2\pi N \quad \text{or } N! = \sqrt{2\pi N} \left(\frac{N}{e}\right)^N \quad 12.9.12$$

However, for very large numbers, the $\frac{1}{2} \ln 2\pi N$ term is small compared to the first two terms and Eq. 12.9.11 is sufficient. Several homework problems explore the validity of these equations. The use of Sterling's approximation for the factorials in calculating W using Eq. 12.9.7 allows the calculation of W_{\max} .

The Method of Lagrange Multipliers is Used for Constrained Maximization: The maximization or minimization of a function that is subject to constraints is a common problem in many areas of science, mathematics, statistics, and maximizing profits in economics. We will do a simple mathematical example in this section to highlight the method. Consider the bowl shaped function in Figure 12.9.2. The functional form is:

$$f(x,y) = 1 - x^2 - y^2 \quad 12.9.12$$

The total differential of f as we vary x and y to find the maximum is:

$$\begin{aligned} df &= \left(\frac{\partial f}{\partial x}\right)_y dx + \left(\frac{\partial f}{\partial y}\right)_x dy \\ df &= -2x dx - 2y dy = 0 \quad \text{(maximum)} \quad 12.9.13 \end{aligned}$$

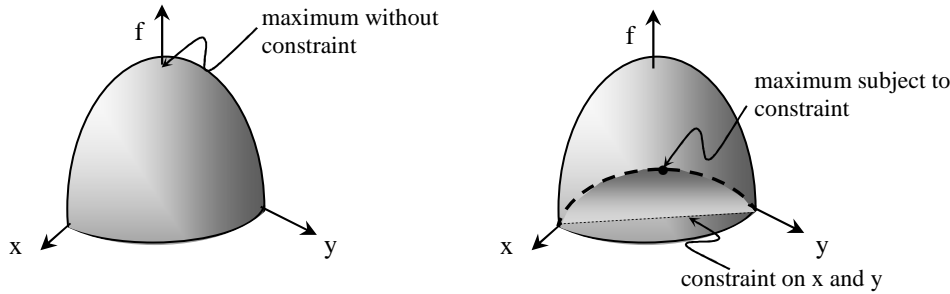


Figure 12.9.2: Constrained maximization of $f(x,y) = 1 - x^2 - y^2$.

The unconstrained maximum for this function is calculated by setting df equal to 0, giving $x = 0$ and $y = 0$ for the maximum, $f_{\max} = 1$. However, what happens if we have a constraint? Assume that x and y are constrained to lie along the straight line $y = 1 - x$, as shown in Figure 12.9.2. The constraint function, c , is then obtained by rearranging the constraint to give an equation equal to zero:

$$c = 1 - x - y = 0 \quad (\text{constraint}) \quad 12.9.14$$

As x and y are changed to find the maximum, the differential of the constraint gives:

$$dc = -dx - dy = 0 \quad (\text{constraint}) \quad 12.9.15$$

This equation can be multiplied by a constant, λ , and still give zero; the constant is the Lagrange multiplier:

$$\lambda(-dx - dy) = 0 \quad (\text{constraint}) \quad 12.9.15$$

Adding Eq. 12.9.13 and Eq. 12.9.15 still gives zero:

$$-2x dx - 2y dy + \lambda(-dx - dy) = 0 \quad (\text{constrained}) \quad 12.9.16$$

or collecting terms in dx and dy :

$$(-2x - \lambda) dx + (-2y - \lambda) dy = 0 \quad (\text{constrained}) \quad 12.9.17$$

We can now treat dx and dy as independent of each other. The assumption of independence is important. The reason is that dx can vary over its full range without regard to the changes in y and dy can vary over its full range without regard to the changes in x . When dx and dy are independent variables, the only way for Eq. 12.9.17 to always equal zero is if the coefficients of dx and dy are both always equal to zero:

$$(-2x - \lambda) = 0 \quad \text{and} \quad (-2y - \lambda) = 0 \quad (\text{constrained}) \quad 12.9.18$$

We now have two equations in two unknowns, which solve to give $x = y$. The maximum in our function occurs when $x = y$. We can solve for the values of x and y for the maximum by using the original constraint equation. Setting $x = y$ in Eq. 12.9.14 gives $x = 1/2$ and then $y = 1/2$. The value of f at the constrained maximum is then obtained from Eq. 12.9.12:

$$f(1/2, 1/2) = 1 - (1/2)^2 - (1/2)^2 = 1/2 \quad (\text{constrained}) \quad 12.9.19$$

as shown in the figure. Problem 23 uses the Lagrange method for an example from economics.

Chapter Summary

1. Energy is dispersed from a hotter to a colder body.
2. Thermodynamic variables are time averages over all degrees of freedom of a system.
3. Much specific information is lost when averaging over the variables for all the degrees of freedom to give just a few thermodynamic variables.
4. A group of \mathcal{N} identical systems is called an ensemble. An ensemble average is equivalent to a time average in the limit $\mathcal{N} \rightarrow \infty$, and subject to the ergotic hypothesis.
5. The ensemble is isolated from the surroundings; the total energy of the ensemble is constant. In a canonical ensemble the systems are in thermal contact, exchange energy, and have the same average energy and temperature.
6. The ensemble represents a single system for averaging purposes.
7. The fundamental postulate of statistical mechanics is: for \mathcal{N} systems in an ensemble with total energy \mathcal{E} , the ensemble average energy is equal to the internal energy of the system:

$$U - U(0) = \langle E \rangle = \frac{\mathcal{E}}{\mathcal{N}}$$

8. The dispersal of energy is determined by finding the number of ways of arranging the energy states for the ensemble, \mathcal{W} .
9. Each individual configuration of the systems among the energy states is called a **microstate**.
10. The statistical definition of the entropy for a system in the ensemble is:

$$S = \frac{k}{\mathcal{N}} \ln \mathcal{W}_{\max} \quad (\text{ensemble}) \qquad S = k \ln W_{\max} \quad (\text{single system})$$

where \mathcal{W}_{\max} is the maximum number of ways of arranging the energy states in the ensemble and W_{\max} is the number of ways of arranging the energy states within a single system.

11. The number of systems in each energy state is specified by a set of distribution numbers, $\{n_i\}$, and averages are calculated using the distribution numbers. The average energy is:

$$U - U(0) = \langle E \rangle = \frac{1}{\mathcal{N}} \sum_i n_i E_i \qquad \text{with normalization} \quad \sum_i n_i = \mathcal{N}$$

12. The probability of occurrence, p_j , of energy state E_j is given by:

$$p_j = \frac{n_j}{\mathcal{N}} \quad \text{and then} \quad U - U(0) = \langle E \rangle = \sum_i p_i E_i \quad \text{with normalization:} \quad \sum_{i=0}^{\mathcal{N}} p_i = 1$$

13. The number of ways of arranging the energy states for distinguishable systems is:

$$\mathcal{W}(n_0, n_1, n_2, \dots) = \frac{\mathcal{N}!}{n_0! n_1! n_2! \dots}$$

14. A fundamental assumption is that each microstate has equal *a priori* probability.
15. A set of distribution numbers corresponds to a **macrostate**. Greater energy dispersal results from a macrostate with larger \mathcal{W} . Thermodynamic properties are averages over all accessible macrostates.
16. Statistical weights and probabilities multiply when a composite is formed, with no other changes: $\mathcal{W} = \mathcal{W}_1 \mathcal{W}_2$.

17. The most probable distribution is the distribution that has the largest number of ways of arranging the energy states, \mathcal{W}_{\max} . For thermodynamically meaningful numbers of systems, the most probable distribution is the overwhelmingly predominant distribution. Thermodynamic averages are determined by the most probable distribution.

18. A spontaneous process occurs when the system undergoes a change that increases \mathcal{W} . The equilibrium state is the most probable state, with $\mathcal{W} = \mathcal{W}_{\max}$. A reversible process occurs through a sequence of steps with $\mathcal{W} = \mathcal{W}_{\max}$.

19. Entropy always increases for a spontaneous process in an isolated system.

20. A logarithmic function is the only possible form for the dependence of entropy on \mathcal{W}_{\max} , since entropy is an extensive state function.

21. Given the probability of occurrence of each energy state of the ensemble:

$$S = -k \sum_i p_i \ln p_i$$

22. The molar residual entropy at absolute zero is given by:

$$S_o = R \ln w \quad S_o = -R \sum_i p_i \ln p_i \quad S_o = \Delta_{\text{mix}} S = -R \sum_i p_i \ln p_i$$

23. The ideal entropy of mixing for an ideal gas or a mixture of molecules with equal intermolecular forces is $\Delta_{\text{mix}} S = -nR \sum x_i \ln x_i$.

24. The probability of occurrence of a system in the ensemble with energy E_i is given by:

$$p_i = \frac{e^{-E_i/kT}}{Q} \quad \text{with} \quad Q \equiv \sum_i e^{-E_i/kT} \quad (\text{canonical ensemble})$$

25. For systems with negligible interactions between distinguishable molecules or for internal degrees of freedom, the entropy in terms of the number of ways of arranging the states of the constituent molecules is $S = nR \ln w_{\max}$, and the sum over molecular states: $S = -nR \sum_i p_i \ln p_i$.

26. Energy is dispersed and entropy is increased by increasing the number of accessible degrees of freedom. Vibrations with frequencies less than about 500 cm^{-1} contribute to the internal energy and entropy, at room temperature.

27. The change in entropy for a process is: $dS = -k \sum \ln p_i dp_i$

28. The most probable state is characterized by a single thermodynamic parameter called the temperature. Bodies at equilibrium have the same temperature. The most fundamental definition of temperature is:

$$\left(\frac{\partial S}{\partial U} \right)_V \equiv \frac{1}{T} \quad \text{giving} \quad \beta = 1/kT$$

29. At constant V: $dU = d\langle E \rangle = \sum_i E_i dp_i$ and $dS = \frac{1}{T} \sum_i E_i dp_i$

30. For a reversible process: $dq_{\text{rev}} = \sum_i E_i dp_i$ and $dS = \frac{dq_{\text{rev}}}{T}$

31. To associate configurational randomness with entropy: the configurational randomness must be associated with the dispersal of energy, the configurational randomness must lead to

distinguishable states, and the system must have some mechanism for randomly accessing each possible microstate of the system.

32. The time average of a system may be replaced by an ensemble average only if the ergodic hypothesis is satisfied; each possible microstate of the system must be accessed at random with equal *a priori* probability.

33. Energy is dispersed by increasing spatial dispersion; the change in entropy for the isothermal expansion of an ideal gas is: $\Delta S = nR \ln(V_2/V_1)$.

34. Probabilities add for the occurrence of events A or B. Probabilities multiply for the occurrence of events A and B.

35. For N objects there are N! permutations or ways the objects may be chosen.

36. The number of ways of selecting n distinguishable objects from N, irrespective of the order:

$$C[N \text{ choose } n] = \frac{N(N-1)\dots(N-n+1)}{n!} = \frac{N!}{n!(n-r)!}$$

Literature Cited:

1. L. Boltzmann, "Über die Beziehung zwischen dem zweiten Hauptsatzes der mechanischen Wärmetheorie und der Wahrshcheinlichkeitsrechnung respektive den Sätzen über das Wärmegleichgewicht. *Wien. Ber.* 76, 373-435.
2. M. Planck, "Entropie und Temperatur strahlen der Wärme," *Ann. der Phys.*, **1900**, 1, 719-737.
3. M. Planck, "On the Law of Distribution of Energy in the Normal Spectrum," *Annalen der Physik*, **1901**, 4, 553.
4. L. K. Nash, *Elements of Statistical Thermodynamics*, 2nd. Ed., Addison-Wesley, Reading, Mass, 1974. Reprinted by Dover, Mineola, NY, 2006. pp. 19-26.
5. A. Ben-Naim, *Entropy Demystified: The Second Law Reduced to Plain Common Sense with Seven Simulated Games*, World Scientific, Singapore, 2008. pp. 143-145.
6. K. Denbigh, *The Principles of Chemical Equilibrium with Applications in Chemistry and Chemical Engineering*, 4th ed., Cambridge University Press, Cambridge, England, 1981. p. 344.
7. E. I. Kozliak, "Consistent Application of the Boltzmann Distribution to Residual Entropy in Crystals," *J. Chem. Educ.*, **2007**, 84, 493-498.
8. F. L. Lambert, "Configurational Entropy Revisited," *J. Chem. Educ.*, **2007**, 84, 1548-1550.
9. K. Huang, *Statistical Mechanics*, Wiley, New York, NY, 1965. pp. 159-162.
10. M. Karplus, J. N. Kushick, "Method for Estimating the Configurational Entropy of Macromolecules," *Macromolecules*, **1981**, 14, 325-332.
11. K. S. Pitzer, "The Molecular Structure and Thermodynamics of Propane: The Vibration Frequencies, Barrier to internal rotation, Entropy, and Heat Capacity," *J. Chem. Phys.*, **1944**, 12(7), 310-314.
12. D. H. Whiffen, "The Force Field, Vibration Frequencies, Normal Co-Ordinates, Infra-Red, and Raman Intensities for Benzene," *Phil. Trans. Roy. Soc. London Ser. A: Math. and Phys. Sci.*, **1955**, 248(942), 131-154.
13. K. H. Lee, D. Xie, E. Freire, L. M. Amzel, "Estimation of Changes in Side Chain Configurational Entropy in Binding and Folding: General Methods and Application to Helix Formation," *Proteins: Struct. Func. and Genetics*, **1994**, 20, 68-84.
14. G. L. Patrick, *An Introduction to Medicinal Chemistry*, Oxford University Press, Oxford, England, 1995. p. 224.

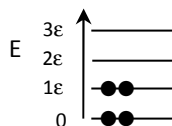
15. M. I. Page, W. P. Jencks, "Entropic Contributions to Rate Accelerations in Enzymic and Intramolecular Reactions and the Chelate Effect," *Proc. Nat. Acad. Sci. USA*, **1971**, 68(8), 1678-1683.
16. R. S. Berry, S. A. Rice, J. Ross, *Physical Chemistry*, Oxford University Press, New York, NY, 2000. pp. 359-361, 431-442.
17. K. Denbigh, *The Principles of Chemical Equilibrium with Applications in Chemistry and Chemical Engineering*, 4th ed., Cambridge University Press, Cambridge, England, 1981. Sec. 11.12, pp. 350-356.
18. J. P. Lowe, "Entropy: Conceptual Disorder," *J. Chem. Ed.*, 65(5), **1988**, 403-406.
19. H. S. Leff, "Thermodynamic Entropy: the Spreading and Sharing of Energy," *Am. J. Phys.*, **1996**, 64, 1261-1271.
20. H. S. Leff, "Entropy: Its Language, and Interpretation," *Found. Phys.*, **2007**, 37, 1744-1766. Entropy 2008, 10 283
21. A. Ben-Naim, *Entropy Demystified: The Second Law Reduced to Plain Common Sense with Seven Simulated Games*, World Scientific, Singapore, 2008. pp. 196-199.
22. J. N. Spencer, J. P. Lowe, "Entropy: The Effects of Distinguishability," *J. Chem. Ed.*, **2003**, 80(12), 1417-1424.
23. F. Hynne, "Understanding Entropy with the Boltzmann Formula," *Chemical Educator*, **2004**, 9, 74-79.
24. F. Hynne, "From Microstates to Thermodynamics," *Chemical Educator*, **2004**, 9, 262-269.
25. K. A. Dill, S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*, Garland Science, New York, NY, 2003. Chaps. 1-7.

Further Reading

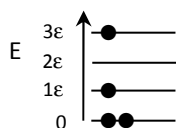
- K. A. Dill, S. Bromberg, *Molecular Driving Forces: Statistical Thermodynamics in Chemistry and Biology*, Garland Science, New York, NY, 2003. Chaps. 1-7.
- K. Denbigh, *The Principles of Chemical Equilibrium: with Applications in Chemistry and Chemical Engineering*, 4th Ed., Cambridge University Press, Cambridge, England, 1981. Chapter 11.
- J. S. Winn, *Physical Chemistry*, Harper Collins, New York, NY, 1995. Chapt. 22.
- A. Ben-Naim, *Entropy Demystified: The Second Law Reduced to Plain Common Sense with Seven Simulated Games*, World Scientific, Singapore, 2008.

Problems: The Statistical Definition of Entropy

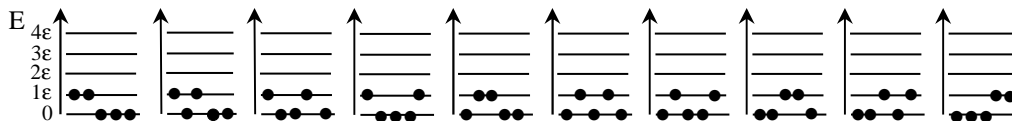
1. Calculate the number of ways of arranging the energy states and find all the microstates consistent with the distribution numbers (2,2,0,0). Find the average energy for the distribution. (Assume distinguishable systems.)



2. Calculate the number of ways of arranging the energy states and find all the microstates consistent with the distribution numbers (2,1,0,1). Find the average energy for the distribution. (Assume distinguishable systems.)



3. Given the following microstates:



- Find the set of distribution numbers.
- Specify the macrostate.
- Find the number of ways of arranging the energy states for the system for the set of distribution numbers. Is the given set of microstates complete?
- Find the statistical weight for the macrostate.
- Find the degeneracy for the macrostate.
- Find $\langle E \rangle$ and $U - U(0)$.
- Find the probability of occurrence of the first microstate, within the given macrostate.
- Find another macrostate with the same number of systems and the same energy. Which is more probable, the given macrostate or the new macrostate?
- What is the most probable distribution and \mathcal{W}_{\max} subject to $\mathcal{N} = 5$ and $\mathcal{E} = 2\epsilon$?
- What is the equilibrium distribution?

4. Assume a system has equally spaced energy states with spacing ϵ . (a) Find \mathcal{N} , \mathcal{E} , $\langle E \rangle$, and $U - U(0)$ for the distributions (26,14,9,5,3,2,1) and (25,16,8,5,3,2,1). (b) Which distribution is the more probable macrostate? (c) Find the difference between the two macrostates in the number of ways of arranging the energy states for the system. (d) One of the two is the most probable distribution. Which macrostate corresponds to the equilibrium state?

5. Calculate the residual entropy for a crystalline solid like $N=N=O$ assuming the energy difference for the two crystalline alignments is $0.300 \text{ kJ mol}^{-1}$. Assume that the distribution of alignments is “frozen in” at the melting point. Assume the melting point is -90.8°C . Compare the result to Eq. 12.4.10. Why is there a difference?

6. The goal of this problem is to help you become more comfortable with partition functions. Consider the residual entropy of $N=N=O$. Define the lowest energy alignment as having energy ϵ_0 and the higher energy alignment at ϵ_1 . A reasonable way to assess the degree of alignment is to calculate the fraction of molecules in the low energy state, f_0 , and the fraction of the molecules in the high energy state, f_1 :

$$f_0 = \frac{n_0}{n_0 + n_1} \quad f_1 = \frac{n_1}{n_0 + n_1}$$

where n_0 is the number of molecules in the low energy alignment and n_1 is the number of molecules in the high energy alignment.¹ Of course, $f_0 + f_1 = 1$. (a). The number of molecules in a specific energy state is proportional to the Boltzmann weighting factor, $n_i \propto e^{-\epsilon_i/kT}$. Use this proportionality to find the fractions f_0 and f_1 . (b). Alternatively, the probability of finding a molecule in a specific energy state, ϵ_i , is given by Eqs. 12.4.13 (Eq. 8.9.5). Show that your answers to part (a) are consistent with Eqs. 12.4.13.

7. Calculate the number of ways of arranging 10 distinguishable balls in three boxes with 3 in the first box, 5 in the second box, and 2 in the third box.

8. (a). Calculate the number of ways of arranging 3 distinguishable balls among 3 boxes with 2 balls in the first box, 1 ball in the second box, and 0 balls in the third box. (b). Calculate the number of ways of arranging the energy states of the system with 3 molecules given that 2 molecules are in the first energy level, 1 molecule is in the second energy level, and 0 molecules are in the third energy level. Draw the arrangements.

9. Find the set of distribution numbers that maximizes the number of arrangements for 3 balls in 3 boxes.

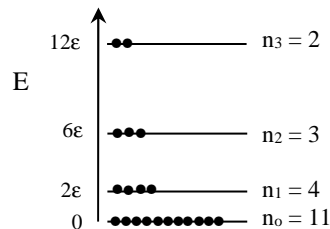
10. (a). Starting with $N! \approx (N/e)^N$ show that $\ln N! \approx N \ln N - N$. (b). Starting with $N! \approx \sqrt{2\pi N} (N/e)^N$ show that $\ln N! = N \ln N - N + \frac{1}{2} \ln 2\pi N$. (c) Compare the exact value of $\ln N!$ and the two approximations for the largest number your calculator can use.

11. Show that the percent error using Sterling's approximation for $\ln(N!)$ decreases with increasing N . (Excel has a larger range for valid N than most calculators.)

12. In Problem 4 the most probable distribution was determined, choosing from (26,14,9,5,3,2,1) and (25,16,8,5,3,2,1). Verify that the most probable distribution is a Boltzmann distribution.

13. The fundamental vibration frequency for I_2 is 214.50 cm^{-1} . Assume the vibrational states are equally spaced with spacing 214.50 cm^{-1} . Iodine vapor is held in an oven at elevated temperature. The relative occupations of the vibrational states were found to be 1.000 : 0.467 : 0.222 : 0.100. Calculate the temperature.

14. Is the following system at thermal equilibrium? Give the approximate temperature, assuming the unit of energy, ϵ , is 10.0 cm^{-1} .



15. The conformational entropy for butane was determined in Example 12.4.3 using the *gauche-anti*-energy difference from molecular mechanics. Use a molecular orbital calculation to estimate the energy difference and determine the corresponding conformational entropy. How sensitive is the conformational entropy to the value of the energy difference? Your instructor will assign a molecular orbital method from the following list depending on the resources available: AM1, PM3, HF 3-21G(*), HF 6-31G*, B3LYP/6-31G*, MP2/6-311G*//HF 6-31G* (single point energy at MP2/6-311G* for the geometry calculation at HF 6-31G*).

16. Determine the conformational entropy for 1,2-dichlorobutane. Your instructor will assign a molecular mechanics or molecular orbital method from the following list depending on available resources: MMFF, PM3, HF 3-21G(*), HF 6-31G*, B3LYP/6-31G*, MP2/6-311G*//HF 6-31G* (single point energy at MP2/6-311G* for the geometry calculation at HF 6-31G*).

17. The process of folding a protein to produce the active conformation restricts torsions along the polypeptide backbone and side chain torsions for amino acids that are buried in the interior of the protein. Calculate the conformational entropy of the side chain of the amino acid valine at 298.2 K. Use molecular mechanics with the MMFF force field in the gas phase for the zwitterionic form to determine the low energy side chain conformations and the corresponding energies.

18. Assume that the *gauche*-energy states for a 1,2-disubstituted ethane, X-CH₂-CH₂-Y, are at energy ϵ above the *anti*-state. The *anti*-state is set at the reference state. Show that the conformational entropy for the C-C bond in disubstituted ethane compounds is given by:

$$S = -\frac{R}{(1 + 2 e^{-\epsilon/RT})} \left[\ln\left(\frac{1}{1 + 2 e^{-\epsilon/RT}}\right) + 2 e^{-\epsilon/RT} \ln\left(\frac{e^{-\epsilon/RT}}{1 + 2 e^{-\epsilon/RT}}\right) \right]$$

19. Show that the maximum conformational entropy for freely rotating sp₃ hybridized bonds is given by $S = R \ln 3$. In other words, assume that the energy differences between the three conformational states is much less than RT. You also need to assume that the three conformations are distinguishable, as in the central butane dihedral or the side chain of valine but not -CH₃ torsions.

20. Calculate the conformational entropy for the C(sp₃)-C(sp₂) bond torsion leading to the phenyl ring in the side chain of the amino acid phenylalanine. Use molecular mechanics with the MMFF force field for gas phase energies. To obtain values that are appropriate for protein folding studies, build a protein in the alpha-helical form with 11 residues: five alanines followed by phenylalanine followed by five alanines. Acetylate the N-terminus and amidate the C-terminus to help stabilize the alpha-helix. Once the lowest energy structure is found, fix (or freeze) all of the atoms except those in the phenyl side chain. Then determine the low energy conformers as you rotate around the C_α-C_β(sp₃)-C_{ring}(sp₂)-C_{ring} dihedral. You will find four low energy conformers. However, the conformers are in two equivalent pairs. The conformers in each pair differ by rotation of the phenyl ring by 180°. The phenyl ring is symmetrical with respect to rotation by 180°, so the conformational states that differ by 180° are indistinguishable. The counting of states for calculation of the entropy is over distinguishable states. To correct for symmetry, then, calculate the entropy by summing over only the two unique, distinguishable states. The structure and the required dihedral are illustrated below, Figure P12.1.

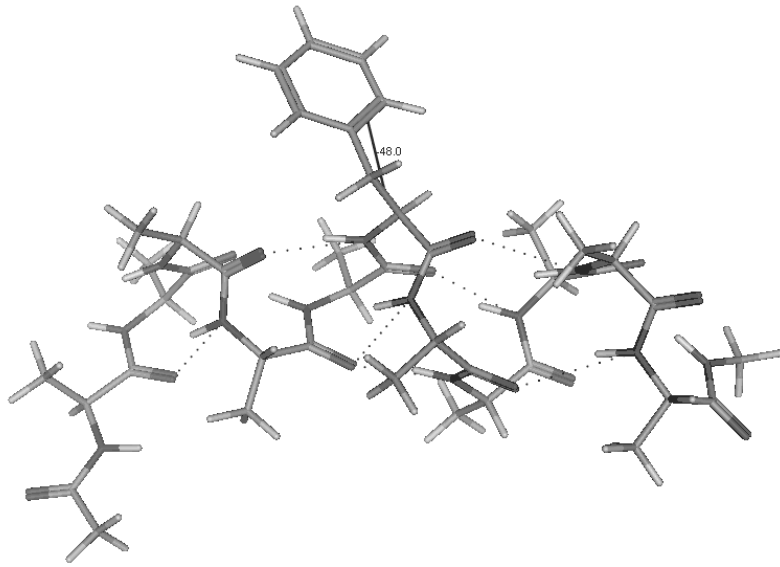


Figure P12.1: An alpha-helical 11-mer with phenylalanine in the center. The required dihedral is depicted with the angle of -48.0° , giving the lowest energy conformer. The dotted lines show the hydrogen-bonding pattern. The distortions in the hydrogen-bonding pattern are caused by the shortness of the peptide. Your model may differ in the conformations of the terminal residues.

21. The Boltzmann distribution is often derived directly by maximizing \mathcal{W} instead of maximizing the entropy using Eq. 12.4.9. In this problem, we derive the Boltzmann distribution in several steps directly from the statistical weights. (a) Starting with Eq. 12.4.1, show that without constraints:

$$d(\ln \mathcal{W}) = \sum \left(\frac{\partial \ln \mathcal{W}}{\partial n_i} \right) dn_i$$

(b). Add in the constraints to give:

$$0 = \sum \left(\frac{\partial \ln \mathcal{W}}{\partial n_i} \right) dn_i + \alpha \sum dn_i - \beta \sum E_i dn_i$$

(c). Show that the constrained maximization results in:

$$\left(\frac{\partial \ln \mathcal{W}}{\partial n_i} \right) + \alpha - \beta E_i = 0$$

(d). Note that the summation variable in Eq. 12.4.5 is an arbitrary index. We can also write:

$$\ln \mathcal{W} = \mathcal{N} \ln \mathcal{N} - \sum_j n_j \ln n_j$$

Show that the derivative of $\ln \mathcal{W}$ with respect to n_i while holding all the other distribution numbers constant gives just one term, which is:

$$\left(\frac{\partial \ln \mathcal{W}}{\partial n_i} \right) = -(\ln n_i + 1) \approx -\ln n_i$$

(e). Substitute this last result into the result from part (c) and solve for n_i to find:

$$n_i = e^{\alpha - \beta E_i} = e^{\alpha} e^{-\beta E_i}$$

(f). Use normalization to find the Boltzmann distribution:

$$\frac{n_i}{\mathcal{N}} = \frac{e^{-\beta E_i}}{Q}$$

22. Consider the bowl shaped function, $f(x,y) = 1 - x^2 - y^2$. Maximize the function subject to the constraint $y = 0.5$ using the Lagrange method of undetermined multipliers.

23. A scientific instrument company produces two different widgets. Let the number of widgets produced by the factory per day of the two different widgets be n_1 and n_2 , respectively. The profit obtained by selling type-one widgets, P_1 , and type-two widgets, P_2 , is given as:

$$P_1 = 40 n_1 - n_1^2 \qquad P_2 = 20 n_2 - 0.5 n_2^2$$

The negative terms in the profit equations result because as the production increases, the cost of labor increases (extra people need to be hired) and the marketing costs increase. The factory can make at most 25 widgets per day. Find the optimal level of production for the two widgets to maximize the overall profit. Compare the constrained result to the unconstrained result assuming the factory can produce any number of widgets per day.

24. Thermodynamic state functions can be written directly in terms of the partition function, Q , which adds to the importance of this central concept. Using Eqs. 12.4.9, 12.4.12, 12.2.6, and 12.1.2, show that the entropy can be written as:

$$S = k \ln Q + \frac{U - U(0)}{T}$$

25. What is the probability of selecting an Ace in 10 total cards? To avoid statistical complications, assume that after each selection the card is returned to the deck, so that each selection is made from a full deck of 52 cards.

26. The next five problems concern the relationship between statistical weights and the probability of occurrence of a particular set of events. The number of ways of selecting n objects from N , which we called $C[N \text{ choose } n]$, is also called the binomial coefficient and given the symbol $\binom{N}{n}$. Using Eq. 12.9.1 show that the binomial coefficient can be expressed as:

$$\binom{N}{n} = C[N \text{ choose } n] = \frac{N!}{n!(N-n)!}$$

27. In the last problem we showed that the binomial coefficient $\binom{N}{n}$ can be defined as:

$$\binom{N}{n} = C[N \text{ choose } n] = \frac{N!}{n!(N-n)!}$$

The binomial coefficient $\binom{N}{n}$ is the numerical coefficient for the n^{th} term in the N^{th} -order polynomial $(1 + x)^N$. For example:

$$\begin{aligned}(1 + x)^3 &= (1 + 2x + x^2)(1 + x) \\ &= 1 + 3x + 3x^2 + x^3 \\ &= \binom{3}{0} 1 + \binom{3}{1} x + \binom{3}{2} x^2 + \binom{3}{3} x^3\end{aligned}$$

Verify the corresponding result for $(1 + x)^4$.

28. Assume that 3 distinguishable balls are selected at random for placement into two boxes. The volume of box 1 is V_1 and the volume of box 2 is V_2 . The probability of a single ball landing in box 1 is proportional to its volume, $p_1 = V_1/(V_1 + V_2)$. The probability of a single ball landing in box 2 is proportional to its volume, $p_2 = V_2/(V_1 + V_2)$. What is the probability that all 3 balls land in the first box? There is only one way for all 3 balls to land in box 1 giving the probability as $p(3,0) = p_1^3$. There is only one way for all 3 balls to land in box 2 giving $p(0,3) = p_2^3$. Find the probability of 2 balls landing in box 1 and the remaining ball landing in box 2. Relate the results to the statistical weight $W(2,1)$.

29. Show that for a two-category problem with N distinguishable objects the binomial coefficient and statistical weight are related by:

$$W(n_1, n_2) = \binom{N}{n_1}$$

An example is the previous two-box problem. The result also holds for any molecular system that has only two energy levels. Use the result of the last problem as a specific example.

30. Assume that 3 distinguishable balls are selected at random for placement into two boxes. The volume of box 1 is V_1 and the volume of box 2 is V_2 . The probability of a single ball landing in box 1 is proportional to its volume, $p_1 = V_1/(V_1 + V_2)$. The probability of a single ball landing in box 2 is proportional to its volume, $p_2 = V_2/(V_1 + V_2)$. There is only one way for all 3 balls to land in box 1 giving the probability as $p(3,0) = p_1^3$. There is only one way for all 3 balls to land in box 2 giving $p(0,3) = p_2^3$. The probability of 2 balls landing in box 1 and the remaining ball landing in box 2 is $p(2,1) = 3(p_1^2 p_2)$, because there are 3 ways of arranging the set of distribution numbers. Likewise $p(1,2) = 3(p_1 p_2^2)$. Show the relationship of the probabilities $p(3,0)$, $p(2,1)$, $p(1,2)$, and $p(0,3)$ to the terms in the expansion of the polynomial $(p_1 + p_2)^3$.

Literature Cited:

1. E. I. Kozliak, "Consistent Application of the Boltzmann Distribution to Residual Entropy in Crystals," *J. Chem. Educ.*, **2007**, *84*, 493-498.
2. W. J. Hehre, *A Guide to Molecular Mechanics and Quantum Chemical Calculations*, Wavefunction, Inc., Irvine, CA, 2003. p. 276.