1.D: Interpretation of Entropy on the Microscopic Scale - The Connection between Randomness and Entropy

1.D.1 Entropy Change in Mixing of Two Ideal Gases

Consider an insulated rigid container of gas separated into two halves by a heat conducting partition so the temperature of the gas in each part is the same. One side contains air, the other side another gas, say argon, both regarded as ideal gases. The mass of gas in each side is such that the pressure is also the same.

The entropy of this system is the sum of the entropies of the two parts: $S_{system} = S_{air} + S_{argon}$. Suppose the partition is taken away so the gases are free to diffuse throughout the volume. For an ideal gas, the energy is not a function of volume, and, for each gas, there is no change in temperature. (The energy of the overall system is unchanged, the two gases were at the same temperature initially, so the final temperature is the same as the initial temperature.) The entropy change of each gas is thus the same as that for a reversible isothermal expansion from the initial specific volume v_i to the final specific volume, v_f. For a mass m of ideal gas, the entropy change is $\Delta S = mR \ln(v_f / v_i)$. The entropy change of the system is

$$\Delta S_{system} = \Delta S_{air} + \Delta S_{argon} = m_{air}R_{air}\ln\left[(v_f)_{air}/(v_i)_{air}\right] + m_{argon}R_{argon}\ln\left[(v_f)_{argon}/(v_i)_{argon}\right]. \quad \text{(D.1.1)}$$

Equation (D.1.1) states that there is an entropy increase due to the increased volume that each gas is able to access.

Examining the mixing process on a molecular level gives additional insight. Suppose we were able to see the gas molecules in different colors, say the air molecules as white and the argon molecules as red. After we took the partition away, we would see white molecules start to move into the red region and, similarly, red molecules start to come into the white volume. As we watched, as the gases mixed, there would be more and more of the different color molecules in the regions that were initially all white and all red. If we moved further away so we could no longer pick out individual molecules, we would see the growth of pink regions spreading into the initially red and white areas. In the final state, we would expect a uniform pink gas to exist throughout the volume. There might be occasional small regions which were slightly more red or slightly more white, but these fluctuations would only last for a time on the order of several molecular collisions.

In terms of the overall spatial distribution of the molecules, we would say this final state was more random, more mixed, than the initial state in which the red and white molecules were confined to specific regions. Another way to say this is in terms of "disorder"; there is more disorder in the final state than in the initial state. One view of entropy is thus that increases in entropy are connected with increases in randomness or disorder. This link can be made rigorous and is extremely useful in describing systems on a microscopic basis. While we do not have scope to examine this topic in depth, the purpose of Section 1.D is to make plausible the link between disorder and entropy through a statistical definition of entropy.

1.D.2 Microscopic and Macroscopic Descriptions of a System

The **microscopic** description of a system is the complete description of each particle in this system. In the above example, the microscopic description of the gas would be the list of the state of each molecule: position and velocity in this problem. It would require a great deal of data for this description; there are roughly 10^{19} molecules in a cube of air one centimeter on a side at room temperature and pressure.   The **macroscopic** description, which is in terms of a few (two!) properties is thus far more accessible and useable for engineering applications, although it is restricted to equilibrium states.

To address the description of entropy on a microscopic level, we need to state some results concerning microscopic systems. These, and the computations and arguments below are taken almost entirely from the excellent discussion in Chapter 6 of *Engineering Thermodynamics* by Reynolds and Perkins (1977)[*].

For a given macroscopic system, there are many microscopic states.   A key idea from quantum mechanics is that the states of atoms, molecules, and entire systems are discretely quantized. This means that a system of particles under certain constraints, like being in a box of a specified size, or having a fixed total energy, can exist in a finite number of allowed microscopic states. This number can be very big, but it is finite.

The microstates of the system keep changing with time from one quantum state to another as molecules move and collide with one another. The probability for the system to be in a particular quantum state is defined by its quantum-state probability p_i. The set of the p_i is called the distribution of probability. The sum of the probabilities of all the allowed quantum states must be unity, hence for any time t,

$$\sum_i p_i = 1 \qquad\qquad\qquad (D.2.1)$$

When the system reaches equilibrium, the individual molecules still change from one quantum state to another. In equilibrium, however, the system state does not change with time; so the probabilities for the different quantum states are independent of time. This distribution is then called the equilibrium distribution, and the probability p_i can be viewed as the fraction of time a system spends in the ith quantum state. In what follows, we limit consideration to equilibrium states.

We can get back to macroscopic quantities from the microscopic description using the probability distribution. For instance, the macroscopic energy of the system would be the weighted average of the successive energies of the system (the energies of the quantum states); weighted by the relative time the system spends in the corresponding microstates. In terms of probabilities, the average energy, $\langle E \rangle$, is

$$\langle E \rangle = \sum_i p_i \, \varepsilon_i, \text{ where } \varepsilon_i \text{ is the energy of a quantum state.} \qquad (D.2.2)$$

---

[*] Reynolds, W.C., and Perkins, H.C., *Engineering Thermodynamics*, McGraw-Hill Book Co., 1977.

The probability distribution provides information on the randomness of the equilibrium quantum states. For example, suppose the system can only exist in three states (1, 2 and 3). If the distribution probability is

$$p_1=1, \qquad p_2=0, \qquad p_3=0$$

the system is in quantum state 1 and there is no randomness. If we were asked what quantum state the system is in, we would be able to say it is always in state 1. If the distribution were

$$p_1=0.3, \qquad p_2=0.2, \qquad p_3=0.5$$

or

$$p_1=0.5, \qquad p_2=0.2, \qquad p_3=0.3$$

the randomness would not be zero and would be equal in both cases. We would be more uncertain about the instantaneous quantum state than in the first situation.

Maximum randomness corresponds to the case where the three states are equally probable:

$$p_1=1/3, \qquad p_2=1/3, \qquad p_3=1/3$$

In this case, we can only guess the instantaneous state with 33 per cent probability.

1.D.3 A Statistical Definition of Entropy

The list of the p_i is a precise description of the randomness in the system, but the number of quantum states in almost any industrial system is so high this list is not useable. We thus look for a single quantity, which is a function of the p_i that gives an appropriate measure of the randomness of a system. As shown below, the entropy provides this measure.

There are several attributes that the sought for function should have. The first is that the average of the function over all of the microstates should have an extensive behavior, in other words the microscopic description of the entropy of a system C, composed of parts A and B is given by

$$S_C = S_A + S_B. \tag{D.3.1}$$

Second is that entropy should increase with randomness and should be largest for a given energy when all the quantum states are equiprobable.

The average of the function over all the microstates is defined by

$$S = \langle f \rangle = \sum_i p_i f(p_i) \tag{D.3.2}$$

where the function $f(p_i)$ is to be found. Suppose that system A has n microstates and system B has m microstates. The entropies of systems A, B, and C, are defined by

$$S_A = \sum_{i=1}^{n} p_i f(p_i) \tag{D.3.3a}$$

$$S_B = \sum_{j=1}^{m} p_j f(p_j) \tag{D.3.3b}$$

$$S_C = \sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij} f(p_{ij}) = \sum_{i=1}^{n} p_i f(p_i) \sum_{j=1}^{m} p_j f(p_j) \tag{D.3.3c}$$

In Equations (D.3.2 and D.3.3), the term p_{ij} means the probability of a microstate in which system A is in state i and system B is in state j.

For Equation (D.3.1) to hold,

$$S_C = \sum_{i=1}^{n} p_i f(p_i) \sum_{j=1}^{m} p_j f(p_j) = \sum_{i=1}^{n} p_i f(p_i) + \sum_{j=1}^{m} p_j f(p_j) = S_A + S_B. \tag{D.3.4}$$

The function f must be such that this is true regardless of the values of the probabilities p_i and p_j. This will occur if $f(\ ) = ln(\ )$ because $ln(a \cdot b) = ln(a) + ln(b)$. Making this substitution, the expression for S_c in Equation (D.3.4) becomes

$$\sum_{i=1}^{n}\sum_{j=1}^{m} p_i p_j \, ln(p_i) + \sum_{i=1}^{n}\sum_{j=1}^{m} p_i p_j \, ln(p_j). \tag{D.3.5a}$$

Rearranging the sums, D.3.5a becomes

$$\sum_{i=1}^{n}\left\{ p_i \, ln(p_i) \left[\sum_{j=1}^{m} p_j \right] \right\} + \sum_{j=1}^{m}\left\{ p_j \, ln(p_j) \left[\sum_{i=1}^{n} p_i \right] \right\}. \tag{D.3.5b}$$

Because

$$\sum_{i=1}^{n} p_i = \sum_{j=1}^{m} p_j = 1, \tag{D.3.6}$$

the left hand side of Equation (D.3.4) can be written as

$$S_C = \sum_{i=1}^{n} p_i \, ln(p_i) + \sum_{j=1}^{m} p_j \, ln(p_j). \tag{D.3.7}$$

This means that Equation (D.3.4) is satisfied for any p_i, p_j, n, m. Reynolds and Perkins show that the most general $f(p_i)$ is $f = C \, ln(p_i)$, where C is an arbitrary constant. Because the p_i are less than unity, the constant is chosen to be negative to make the entropy positive.

Based on the above a the statistical definition of entropy can be given as:

$$\boxed{S = -k \sum_j p_i \, ln p_i.} \tag{D.3.8}$$

The constant k is known as the Boltzmann constant,

$$k = 1.380 \times 10^{-23} \frac{J}{K}.$$ (D.3.9)

The value of k is (another wonderful result!) given by $k = \mathbf{R}/N_{Avogadro}$, where R is the universal gas constant, 8.3143 J/(mol-K) and $N_{Avogadro}$ is Avogadro's number, 6.02×10^{23} molecules per mol. Sometimes k is called the gas constant per molecule. With this value for $k,$ the statistical definition of entropy is identical with the macroscopic definition of entropy.

1.D.4 Connection between the Statistical Definition of Entropy and Randomness

We need now to examine the behavior of the statistical definition of entropy as regards randomness. Because a uniform probability distribution reflects the largest randomness, a system with n allowed states will have the greatest entropy when each state is equally likely. In this situation, the probabilities become

$$p_i = p = \frac{1}{\Omega},$$ (D.4.1)

where Ω is the total number of microstates. The entropy is thus

$$S = -k \sum_{i=1}^{\Omega} \frac{1}{\Omega} ln\left(\frac{1}{\Omega}\right) = -k\Omega \left[\frac{1}{\Omega} ln\left(\frac{1}{\Omega}\right)\right] = -k\, ln\left(\frac{1}{\Omega}\right)$$

$$\boxed{S = k\, ln\, \Omega.}$$ (D.4.2)

Equation (D.4.1) states that the larger the number of possible states the larger the entropy.

The behavior of the entropy stated in Equation (D.4.2) can be summarized as follows:
    a) S is maximum when Ω is maximum, which means many permitted quantum states, hence much randomness,
    b) S is minimum when Ω is minimum. In particular, for $\Omega = 1$, there is no randomness and $S = 0$.
These trends are in accord with our qualitative ideas concerning randomness. Equation (D.4.2) is carved on Boltzmann's tombstone (he died about a hundred years ago) in Vienna.

We can also examine the additive property of entropy with respect to probabilities. If we have two systems, A and B, which are viewed as a combined system, C, the quantum states for the combined system are the combinations of the quantum states from A and B. The quantum state where A is in its state x and B in its state y would have a probability $p_{A_x} \cdot p_{B_y}$ because the two probabilities are independent. The number of probabilities for the combined system, Ω_C, is thus defined by $\Omega_C = \Omega_A \cdot \Omega_B$. The entropy of the combined system is

$$S_C = k\,ln\left(\Omega_A \cdot \Omega_B\right) = k\,ln\,\Omega_A + k\,ln\,\Omega_B = S_A + S_B \tag{D.4.3}$$

Equation (D.4.2) is sometimes taken as the basic definition of entropy, but it should be remembered that it is only appropriate when each quantum state is equally likely. Equation (D.3.8) is more general and applies equally for equilibrium and non-equilibrium situations.

A simple numerical example shows trends in entropy changes and randomness for a system which can exist in three states. Consider the five probability distributions

i) $p_1 = 1.0$, $p_2 = 0$, $p_3 = 0$; $S = -k(1\,ln\,1 + 0\,ln\,0 + 0\,ln\,0) = 0$

ii) $p_1 = 0.8$, $p_2 = 0.2$, $p_3 = 0$; $S = -k\left[0.8\,ln(0.8) + 0.2\,ln(0.2) + 0\,ln(0)\right] = 0.500k$

iii) $p_1 = 0.8$, $p_2 = 0.1$, $p_3 = 0.1$; $S = -k\left[0.8\,ln(0.8) + 0.1\,ln(0.1) + 0.1\,ln(0.1)\right] = 0.639k$

iv) $p_1 = 0.5$, $p_2 = 0.3$, $p_3 = 0.2$; $S = -k\left[0.5\,ln(0.5) + 0.3\,ln(0.3) + 0.2\,ln(0.2)\right] = 1.03k$

v) $p_1 = 1/3$, $p_2 = 1/3$, $p_3 = 1/3$; $S = -3k\left[\dfrac{1}{3}ln\left(\dfrac{1}{3}\right)\right] = 1.099k$.

The first distribution has no randomness. For the second, we know that state 3 is never found. States (iii) and (iv) have progressively greater uncertainty about the distribution of states and thus higher randomness. State (v) has the greatest randomness and uncertainty and also the largest entropy.

1.D.5. *Numerical Example of the Approach to the Equilibrium Distribution*

Reynolds and Perkins give a numerical example which illustrates the above concepts and also the tendency of a closed isolated system to tend to equilibrium.

Reynolds and Perkins, Engineering Thermodynamics, McGraw-Hill, 1977. Sec. 6.7. pp.177-183.

1.D.6 Summary and Conclusions

a) Entropy as defined from a microscopic point of view is a measure of randomness in a system.

b) The entropy is related to the probabilities p_i of the individual quantum states of the system by

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

where k, the Boltzmann constant is given by $\mathbf{R} / N_{Avogadro}$.

c) For a system in which there are Ω quantum states, all of which are equally probable (for which the probability is $p_i = 1/\Omega$), the entropy is given by

$$S = k \, ln \, \Omega .$$

The more quantum states, the more the randomness and uncertainty that a system is in a particular quantum state.

d) From the statistical point of view there is a finite, but *exceedingly small* possibility that a system that is well mixed could suddenly "unmix" and that all the air molecules in the room could suddenly come to the front half of the room. The unlikelihood of this is well described by Denbigh [*Principles of Chemical Equilibrium*, 1981] in a discussion of the behavior of an isolated system:

> "In the case of systems containing an appreciable number of atoms, it becomes increasingly improbably that we shall ever observe the system in a non-uniform condition. For example, it is calculated that the probability of a relative change of density, $\Delta\rho/\rho$, of only 0.001% in 1 cm^3 of air is smaller than 10^{-10^8} and would not be observed in trillions of years. Thus, according to the statistical interpretation the discovery of an appreciable and spontaneous decrease in the entropy of an isolated system, if it is separated into two parts, is not impossible, but exceedingly improbable. We repeat , however, that it is an *absolute impossibility* to know when it will take place."

e) The definition of entropy in the form $S = -k \sum_i p_i \, lnp_i$ arises in other aerospace fields, notably that of information theory. In this context, the constant k is taken as unity and the entropy becomes a dimensionless measure of the uncertainty represented by a particular message. There is no underlying physical connection with thermodynamic entropy, but the underlying uncertainty concepts are the same.

f) The presentation of entropy in this subject is focused on the connection to macroscopic variables and behavior. These involve the definition of entropy given in Section 1.B of the notes and the physical link with lost work, neither of which makes any mention of molecular (microscopic) behavior. The approach in other sections of the notes is only connected to these macroscopic processes and does not rely at all upon the microscopic viewpoint. Exposure to the statistical definition of entropy, however, is helpful as another way not only to answer the question of "What is entropy?" but also to see the depth of this fundamental concept and the connection with other areas of technology.