

CHAPTER 6

INFORMATION OF EVENTS WITH DISCRETE OUTCOMES: METHODS OF PHYSICS

ORIENTATION

It is important to understand the manner in which the concept of information is applied in physics, because it is closely aligned with the manner in which we shall be applying information in the analysis of sensory systems. In a way, it is not proper to speak of “the applications of information theory in physics,” because the statistical entropy concept was introduced into physics by Boltzmann more than half a century before Wiener and Shannon’s treatises and the coining of the term “information theory.” Physicist Leon Brillouin wrote two books dealing with the information concept in science (1962 and 1964). The earlier of the two books is the one more commonly cited, however, the later of the two books I find the more useful.

This chapter is designed to give the gist of how the information idea is used in physics. In the pages to follow, the first and second laws of thermodynamics will be presented as they were understood in the nineteenth century. It will be shown how Carnot engines operating under constraint of the second law gave rise to Clausius’ inequality, and how this inequality can be used to introduce the concept of *physical* or *thermodynamic entropy*. Change in physical entropy will be shown to be calculable from macroscopic measurements made in the laboratory. Physical entropy will then be expressed from the perspective of statistical mechanics – that is, in terms of the probabilistic behavior of molecules – and linked with *information theoretical entropy*. The interdiction of “It can be shown that” (Chapter 4, note 1) will be suspended here as we try to compress a rather lengthy subject into a few pages. The reader is encouraged to capture the flow of ideas rather than the details of the proofs. I acknowledge my supreme debt in preparing this chapter to F. W. Sears’ text on thermodynamics and statistical mechanics (1953). It was from Professor Sears’ book that I first learned thermodynamics as a student; and it is in this book that I search again and again for inherent jewels and errant joules. I also recommend the texts by Zemansky (1943) and Zemansky and Dittman (1981).

THE FIRST LAW OF THERMODYNAMICS

The development of thermodynamics preceded the establishment of the atomic theory of matter, so that the formulations of the first and second laws of thermodynamics in the nineteenth century were made largely while matter and energy were regarded as continua. The first law of thermodynamics is an expression of the principle of conservation of energy. A quantity called the *internal energy* of a system was formulated, which we shall represent by U . Any change in U is produced by the difference between the heat, Q , flowing *into* the system and work, \mathcal{W} , done *by* the system. That is,

$$dU = d'Q - d'\mathcal{W}. \quad (6.1)$$

The primes indicate that $d'Q$ and $d'\mathcal{W}$ are not exact differentials; that is, no meaning can be attached to terms such as “the heat in the system.” The difference between the two inexact differentials $d'Q$ and $d'\mathcal{W}$ is equal to the exact differential, dU . We say that the internal energy is a determinant of the *state* of the system.

Dividing Equation (6.1) by the number of moles in the system, we can write

$$du = d'q - d'w. \quad (6.2)$$

Internal energy, then, reflects a balance between heat flow into, and work done by the system. The first law is, in effect, an accounting principle for energy.

The first law of thermodynamics is certainly adequate, in itself, to account for many physical processes. For example, suppose that our system consists of two metal blocks at different temperatures that are in contact with each other, but insulated thermally from their environment. Since the system is thermally insulated, $d'Q = 0$; and if we permit the system to do no work, $d'W = 0$. By the first law, $dU = 0$. It is common experience that heat will flow from the block at the higher temperature to the block at the lower temperature until both blocks reach the same temperature. The first law simply acknowledges the legitimacy of the redistribution of heat between the blocks while conserving internal energy.

Suppose now that we *begin* with the two blocks at the same temperature and in contact with each other. Is it possible that heat will flow from one block to the other, leaving the first block at a lower temperature and the second at a higher temperature? Since no heat leaves the system and no work is done, dU remains equal to zero, and so the first law does not prohibit such occurrences. However, experience tells us that this phenomenon never occurs. From this example and similar considerations, it is recognized that the first law alone is not adequate to account for all thermodynamic phenomena. Another state variable like the internal energy is required.

THE SECOND LAW OF THERMODYNAMICS

Hence, one expression of the *second* law of thermodynamics, due to Clausius, is the following:

*No process is possible whose sole result is the removal of heat from a reservoir at one temperature (lower) and the absorption of an equal quantity of heat at a higher temperature.*¹

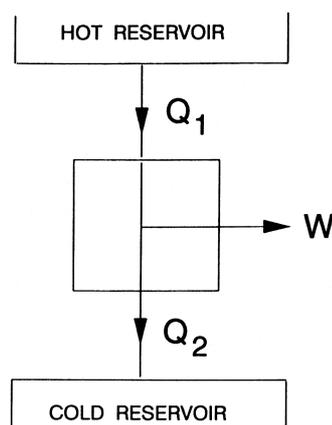
THE CARNOT CYCLE / ENGINE

Clausius was aided by the work of Sadi Carnot (1824; please see Carnot, 1977), who abstracted the properties of heat engines. Carnot understood even before the formulation of the first law that just as mechanical work is done by water falling from a higher to a lower level, so, too, mechanical work can be done when heat flows from a heat reservoir at a higher temperature to a heat reservoir at a lower temperature. He also understood that the heat engine is best analyzed with reference to a cyclical process where a gas or some other “working substance” is passed through a sequence of compressions and expansions. The heat engine is known as a *Carnot engine*, and the associated cycle as the *Carnot cycle*.

Let the heat issuing from the hot reservoir be Q_1 , that entering the cold reservoir be Q_2 , and the work done *by* the working substance (let's call it the *system*) be W . These processes are illustrated in Figure 6.1. The Carnot engine can operate in reverse, so that heat Q_2 may be absorbed from the cooler reservoir into the system, coupled with work, W , done by the system, and heat Q_1 released at the hotter reservoir. When operating in the latter fashion, the system is known as a Carnot refrigerator. The algebraic signs of Q_1 , Q_2 and W are taken with respect to the system; that is, $Q_1 > 0$ when heat flows from the hotter reservoir to the system, $Q_2 < 0$ when heat flows from the system to the colder reservoir, and $W > 0$ when the system does positive work on the outside world. The Carnot engine is a reversible² engine.

The exchanges of energy depicted in Figure 6.1 occurred during a Carnot cycle, after which the internal energy of the system was left unchanged. Therefore, from the first law as expressed by Equation (6.1), the changes in heat minus the change in work must equal zero, or

$$Q_1 + Q_2 - W = 0,$$



CARNOT ENGINE

Figure 6.1 Carnot Engine. Heat Q_1 flows from a heat reservoir at a higher temperature to a system (working substance). Heat Q_2 flows from the system to a heat reservoir at a lower temperature. Work \mathcal{W} is done by the system. Carnot refrigerator. Heat Q_2 is drawn from the cooler reservoir, work \mathcal{W} is done on the system. Heat Q_1 is delivered to the warmer reservoir.

or

$$\mathcal{W} = Q_1 + Q_2 . \quad (6.3)$$

These results are independent of the working substance, which is the substance undergoing the Carnot cycle.

Lord Kelvin then utilized the ratio of Q_1 to Q_2 to *define* the temperature of the reservoirs:

$$\frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0 . \quad (6.4)$$

It is possible to show that when the working substance is an ideal gas, Equation (6.4) can be derived using the ideal gas law, Equation (1.1), and the first law of thermodynamics, Equation (6.3). The Kelvin temperature is numerically equal to the absolute temperature.

We note that the second law of thermodynamics has not yet been brought to bear on the Carnot engine / refrigerator.

THE CLAUSIUS INEQUALITY

Armed now with a verbal statement of the second law, and Equation (6.4) relating heat flows from two heat reservoirs and the respective temperatures of these reservoirs, we can appreciate the development of Clausius' inequality.

We consider any arbitrarily chosen process in which a system is carried through a closed cycle (involving expansions, contractions, heat loss, heat gain, positive work done by the system, positive work done on the system), such that its end state is identical with its initial state (pressure, volume and temperature of a fixed quantity of matter returned to their original values). Then it can be shown with the aid of Carnot engines (that is, the first law of thermodynamics), using Equation (6.4) and using the second law of thermodynamics that

$$\sum_{\substack{\text{all} \\ \text{reservoirs}}} \frac{Q_i}{T_i} \leq 0 , \quad (6.5)$$

where T_i are the temperatures of the reservoirs and Q_i are the heat flows as seen by the system (that is,

Q_i is positive when heat proceeds from the heat reservoir toward the system). When only infinitesimally small quantities of heat are exchanged with infinitely many heat reservoirs, we may pass to the integral form

$$\int \frac{d'Q}{T} \leq 0. \quad (6.6)$$

Both (6.5) and (6.6) express the *Clausius inequality*, which is valid whether the cyclic process is reversible or irreversible.² The simplest example to illustrate (6.5) is the flow of, say, Q joules from a higher temperature reservoir, T_{high} , to a lower temperature reservoir, T_{low} . Then the summation on the left-hand side of (6.5) is just

$$\frac{Q}{T_{\text{high}}} + \frac{-Q}{T_{\text{low}}} = Q \left(\frac{1}{T_{\text{high}}} - \frac{1}{T_{\text{low}}} \right) \text{ joule / deg,}$$

which must be less than zero as demanded by the right-hand side of (6.5).

PHYSICAL ENTROPY

The term *entropy* was introduced by Clausius in 1854 to mean a transformation (Greek, *trepein*, to turn). We are now in position to understand what this term means and how it may be calculated.

While the Clausius inequality was derived for a cyclic system that was either reversible or irreversible, we now consider just a reversible system. Suppose our system cycles first forward and then backward returning to its starting point. Let $d'Q_f$ be the heat flowing into the system in its forward cycle, and $d'Q_b$ the heat in its reverse cycle. Then because of the reversible nature of the system

$$d'Q_f = -d'Q_b \quad (6.7)$$

and the temperature, T , of the heat reservoir is equal to that of the system. Using the contour integral \oint to represent integration around a cycle, we have from (6.6)

$$\oint \frac{d'Q_f}{T} \leq 0$$

and

$$\oint \frac{d'Q_b}{T} \leq 0. \quad (6.8)$$

Introducing (6.7) into (6.8) we see that the only way (6.8) can be true is if the equality sign holds, so that

$$\oint_{\text{rev}} \frac{d'Q}{T} = 0, \quad (6.9)$$

where *rev* indicates that the process is reversible. We are following the argument of Sears (1953) very closely.

We can now show that $d'Q/T$ is an exact differential. Consider an arbitrary closed path for the reversible cycle, represented schematically in Figure 6.2. A and B are any two points on the contour; path I leads from A to B and path II from B to A . Then

$$\oint_{\text{rev}} \frac{d'Q}{T} = \int_A^B \frac{d'Q}{T} + \int_B^A \frac{d'Q}{T} = 0. \quad (6.10)$$

along path I along path II

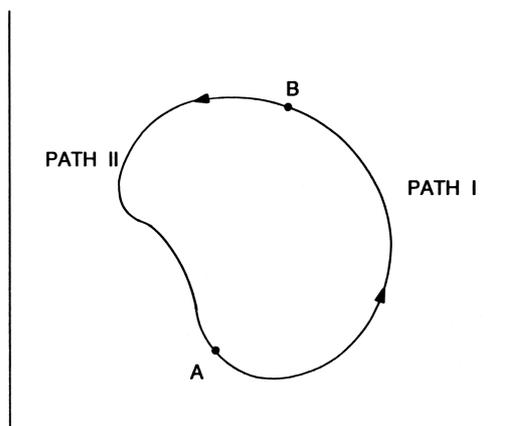


Figure 6.2 Reversible cycle

If path II were traversed in the opposite direction then, as a consequence of the reversibility of the reaction,

$$\int_B^A \frac{d'Q}{T} \underset{\text{along path II}}{=} - \int_A^B \frac{d'Q}{T} \underset{\text{along path II}}{=} . \quad (6.11)$$

Inserting this result into Equation (6.10), we have

$$\int_A^B \frac{d'Q}{T} \underset{\text{along path I}}{=} \int_A^B \frac{d'Q}{T} \underset{\text{along path II}}{=} . \quad (6.12)$$

That is, the integral between two points A and B has the same value along all reversible paths. Therefore, $d'Q/T$ is an exact differential and may be represented by dS :

$$\int_A^B \frac{d'Q}{T} = \int_A^B dS = S_B - S_A . \quad (6.13)$$

That is, the value of the integral depends not on the path, but only on its end points. S is called the *physical or thermodynamic entropy of the system*.

In this chapter we must distinguish explicitly between “physical” and “informational” entropy.

The second law of thermodynamics for a reversible process, therefore, permits the mathematical statement

$$dS = \frac{d'Q}{T} . \quad (6.14)$$

Specific entropy, s , can be obtained by dividing S by the number of moles in the system:

$$ds = d'q/T \quad (6.15)$$

(cf. Equation (6.2)). The units of entropy are, of course, joule. deg⁻¹ or joule. deg⁻¹. mole⁻¹.

THE PRINCIPLE OF INCREASE OF PHYSICAL ENTROPY

In Figure 6.3 is depicted a cyclical process wherein an isolated system proceeds from A to B irreversibly by some natural process, and is then returned from B to A reversibly. The cycle as a whole

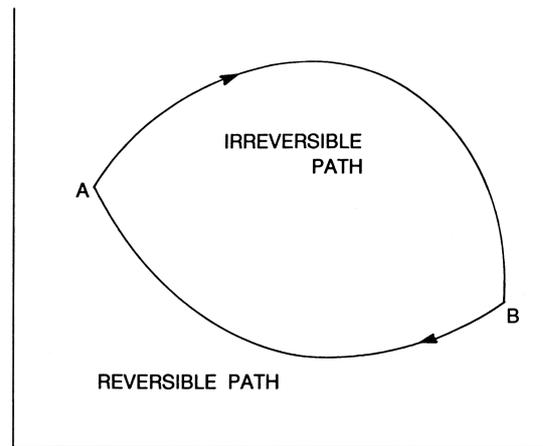


Figure 6.3

is irreversible. By the Clausius inequality (6.6),

$$\oint \frac{d'Q}{T} = \int_A^B \underbrace{\frac{d'Q}{T}}_{\text{irreversible path}} + \int_B^A \underbrace{\frac{d'Q}{T}}_{\text{reversible path}} < 0.$$

The quantity $d'Q/T$ in the first integral is not an exact differential because the process is irreversible, but $d'Q$ in the first integral is equal to zero because the system is isolated and does not exchange heat with its surroundings. The first integral is, therefore, equal to zero. The second integral does represent a reversible process, so its value is equal to $S_A - S_B$. Therefore,

$$\begin{aligned} S_A - S_B &< 0 \\ S_B &> S_A. \end{aligned} \quad (6.16)$$

The thermodynamic entropy of the system in state B is greater than its entropy in state A . Since the process was arbitrary, we learn as a consequence of the second law of thermodynamics that *the thermodynamic entropy of an isolated system increases in every natural (irreversible) process*.

THE LABORATORY MEASURE OF PHYSICAL ENTROPY CHANGE

Change in physical entropy is a measurable, nearly-palpable quantity, that is measured in joules/degree or joules/degree-mole. For example, a change in phase (solid to liquid or liquid to gas) is carried out at a constant temperature, while the phases remain in equilibrium with each other. When the phase change is carried out reversibly, $d'Q$, the heat absorbed, is the latent heat of fusion or vaporization, and the temperature will be the melting or boiling point respectively. To take a specific instance, consider the melting of ice. The latent heat of fusion of water is equal to 79.70 cal/g at one atmosphere. The entropy change as ice melts can be found as follows:

$$\begin{aligned} \text{Heat change } \Delta q &= (79.70 \text{ cal/g}) (4.18 \text{ joule/cal}) (18 \text{ g/mole}) \\ &= 600. \times 10^1 \text{ joule/mole} \end{aligned}$$

$$\text{Ice melts at } 0^\circ \text{ C} = 273^\circ \text{ K}$$

$$\text{Entropy change } \Delta s = \Delta q / T = 6000 / 273 = 22.0 \text{ joule/deg. mole.}$$

As another example, if a solid is heated from temperature T_1 to T_2 ,

$$dq = C dT,$$

where C is specific heat per mole. Then, calculating on the basis of a reversible reaction,

$$ds = dq / T = C dT / T$$

$$\Delta s = \int_{T_1}^{T_2} C dT / T = C \ln(T_2 / T_1) \text{ joule/deg.mole.}$$

Et cetera.

SUMMARY OF SECTION ON CLASSICAL THERMODYNAMICS

This is not a treatise on thermodynamics. The purpose of the preceding pages is to demonstrate (i) that changes in thermodynamic entropy are calculated from standard laboratory measurements of macroscopic variables made by calorimeters and thermometers; and (ii) that thermodynamic entropy was defined before the atomic theory of matter was established, and so calculations of entropy change can be made without reference to atoms and molecules.

As we proceed, it will be seen that direct comparison can be made between the thermodynamic variable, S , and the perceptual variable, F , that has been defined in Chapter 2.

STATISTICAL MECHANICS: μ -SPACE AND PHASE POINTS

The atomic theory of matter, conceived in the West in the fifth century BCE in Greece by Democritus and Leukippus, matured some 2300 years later in the nineteenth century. Physicists and chemists, employing the principles of *kinetic theory*, sought to translate the laws of macroscopic physics and chemistry into terms of atoms and molecules. Using a gas as a model, it is not difficult to see how the macroscopic variable, pressure, can be interpreted as many molecules colliding with a surface and changing their direction of motion. The rate of change of momentum of the molecules as they strike the surface constitutes a force, and the force per unit area is measured macroscopically as pressure. Temperature can be interpreted as a measure of the average kinetic energy of molecules comprising the gas. Physical entropy, however, is somewhat more difficult to express in terms of the properties of gas molecules. This task was achieved by the Austrian physicist, Ludwig Boltzmann, using the methods of what we now call *statistical mechanics*. The term was first coined by Gibbs in 1901.

Statistical mechanics concerns itself not with the motions and collisions of individual molecules, but with the properties of large assemblies of molecules, which it treats probabilistically. In order to do so, it is necessary to utilize spaces consisting of more than the usual three dimensions. Let's take a simple example from everyday life. Suppose that a helicopter is located 100 metres above a heliport, and is proceeding at constant altitude in a northwesterly direction at 80 kilometers per hour. The helicopter's 3 spatial coordinates are given by (i) its altitude and (ii) the longitude and (iii) latitude of the heliport; and its 3 coordinates in *velocity space* are given by (i) its vertical speed (0) and its horizontal speed in the directions (ii) north and (iii) west. Therefore, we see that even if the orientation of the helicopter (which direction it is facing and its inclination to the horizontal) is ignored, it still requires six coordinates, three in ordinary space and three in velocity space, to define its position and motion completely. This six-dimensional space is needed to define the position and velocity of a molecule, irrespective of its orientation. Three coordinates, x , y , z , define a molecule's position in configuration space, and three coordinates, v_x , v_y , v_z , define its position in velocity space. One speaks of the hyperspace needed to define the position and motion of a molecule as a *phase space*, and the particular six-dimensional space described above as a type of μ -space ("μ" Greek *mu*, for *molecule*).³ One also speaks of the coordinates of a *representative point* (x , y , z , v_x , v_y , v_z) of a type of molecule in μ -space.⁴ Such a representative point is called a *phase point*.

We can now subdivide μ -space into small, 6-dimensional elements of volume, $\delta v_\mu = \delta x \delta y \delta z \delta v_x \delta v_y \delta v_z$, which we call *cells* in μ -space. We can number the cells 1, 2, 3, ..., i , ..., and let N_i be the number of phase points in the i^{th} cell. The density of points in the i^{th} cell is, then, $N_i / \delta v_\mu$.

MACROSTATES AND MICROSTATES

We imagine our gas confined to a rigid container whose walls do not permit the exchange of either molecules or energy. We deal with the equilibrium situation.

Specification of the which cell is occupied by each molecule in our system would completely define a *microstate* of the system. For example, molecule 1 is in cell 1507, molecule 2 is in cell 2,345,678, ..., would define a specific microstate. However, clearly the macroscopic or observable properties of the gas do not permit the specification of a unique microstate. For example, the density of the gas will depend only on the total number of molecules contained in one cubic meter of gas, and will tell us little about which individual molecule occupies which cell in μ -space. The macroscopic properties of the gas will, in general, depend only on the numbers, N_i , the number of phase points that lie in each cell in μ -space. Complete specification of the N_i will define a *macrostate* of the system.

Clearly, many microstates make up one macrostate. For example, we could take one molecule from N_1 and place it in N_2 , and one from N_2 and place it in N_1 , producing a new microstate but leaving the macrostate unchanged. The number of microstates that make up a given macrostate is called the *thermodynamic probability* of that macrostate, and is commonly represented by W . It is not difficult to compute W given the values of N_i . To take a simple example, suppose that there are only 5 cells in μ -space: $N_1 = 2$, $N_2 = 1$, $N_3 = 1$, $N_4 = 0$, $N_5 = 0$. That is, the first cell contains two phase points, the second and third cells each contain a single point, while the fourth and fifth cells contain no points at all. In this example we shall have only four phase points which can be designated a , b , c , and d . The only possible distributions of the 4 phase points in the five cells are shown in Table 6.1. From combinatorial theory, the value of W is given by

$$W = \frac{4!}{2!1!1!0!0!},$$

where $0!$ can be taken equal to one (Refer to a standard text on probability such as Freund and Walpole). That is, $W = 12$ microstates per macrostate.⁵ Laboratory observation cannot distinguish one microstate from another. We can say that the uncertainty pertaining to a macrostate is a function of the number of microstates associated with it.

In general, if there are N representative points, the number of microstates is given by

$$W = \frac{N!}{N_1!N_2!N_3! \dots} \quad (6.17)$$

where

$$N = \sum_i N_i. \quad (6.18)$$

Table 6.1 Microstates Corresponding to a Single Macrostate

Cell	Microstates											
	1	2	3	4	5	6	7	8	9	10	11	12
N_1	ab	ac	ad	bc	bd	cd	ab	ac	ad	bc	bd	cd
N_2	c	b	b	a	a	a	d	d	c	d	c	b
N_3	d	d	c	d	c	b	c	b	b	a	a	a
N_4	–	–	–	–	–	–	–	–	–	–	–	–
N_5	–	–	–	–	–	–	–	–	–	–	–	–

There are 4 phase point: a , b , c , d . A total of 12 microstates correspond to the single macrostate where 2 particles are contained within phase cell N_1 , 1 particle within phase cells N_2 and N_3 , and 0 particles within the two remaining phase cells N_4 and N_5 .

Recall that we have, by assumption, a system with constant total energy. Moreover, we shall confine the system to have a constant volume, and to be in a state of equilibrium. The totality of microstates whose total energy is so constrained comprise what Gibbs has termed a *microcanonical ensemble*. Desloge (1966) offers a particularly clear description of microstates and ensembles.

THE BASIC ASSUMPTION OF STATISTICAL MECHANICS

All microstates of a system that have the same energy are equally probable.

This statement has been called (for example, Jackson) the basic assumption of statistical mechanics. With reference, then, to our microcanonical ensemble, all microstates expressed by Equation (6.17) are equally probable. Therefore (Equation (4.4)), the *information theoretical entropy* of the macrostate is given by

$$H_I = \ln W. \quad (6.19)$$

H_I encodes the uncertainty of the human perceiver about which particular microstate gives rise the a specified macrostate.

Boltzmann then postulated that *physical entropy*, S , will be proportional to (what we now call) *information theoretical entropy*, H_I . That is,⁶

$$S = k_B H_I, \quad (6.20)$$

or expressed alternatively,

$$S = k_B \ln W. \quad (6.21)$$

This remarkable equation, that equates the macroscopic, observable variable, S , to $\ln W$, the uncertainty about which molecular microstate makes up a macrostate, is inscribed on Boltzmann's gravestone, although he apparently never wrote it in this form. The k_B is known as *Boltzmann's constant*. Not surprisingly, not all physicists are happy about the interpretation of $\ln W$ as a human uncertainty, because it seems to be injecting a subjective quantity into Equation (6.21), which is an equation of physics and, therefore, an objective entity completely free from human influence. To avoid such quasi-subjectivity, $\ln W$ is simply taken to be the natural log of the thermodynamic probability, W . Period.

We have seen before how a macroscopic variable can be expressed as a measure of uncertainty, or information theoretical entropy,

$$F = k H_I. \quad (2.6)$$

We are now beginning to fill in the history of Equation (2.6). Let us continue with the development of statistical mechanics in order to appreciate the utility of Equations (6.20) and (6.21).

We shall need to use Stirling's approximation for evaluating $\ln x!$, where x is a large, positive integer. Stated here without proof,

$$\ln x! = x \ln x - x. \quad (6.22)$$

The proof is not difficult, and the reader is referred to the texts. Taking logarithms of both sides of Equation (6.17),

$$\ln W = \ln N! - \sum_i \ln N_i!$$

Applying Stirling's approximation,

$$\begin{aligned} \ln W &= (N \ln N - N) - \left(\sum_i N_i \ln N_i - \sum_i N_i \right) \\ &= N \ln N - \sum_i N_i \ln N_i \end{aligned} \quad (6.23)$$

using Equation (6.18).

We recall that the system is in a state of equilibrium, which we might designate by a maximum value of the thermodynamic probability, W . That is, the change in W at equilibrium produced by the aggregate of all the small changes in the occupancies N_i resulting from motions of the phase points in μ -space will equal zero. Expressing this variation mathematically using Equation (6.23),

$$\delta \ln W = \delta(N \ln N) - \delta \sum_i N_i \ln N_i = 0 .$$

That is,

$$0 - \sum_i \ln N_i \delta N_i - \sum_i N_i \delta \ln N_i = 0 .$$

But

$$\sum_i N_i \delta \ln N_i = \sum_i (N_i / N_i) \delta N_i = \sum_i \delta N_i = 0 ,$$

since the total number of particles does not change. Therefore (dropping the subscript under the summation sign for simplicity),

$$\sum \ln N_i \delta N_i = 0 . \quad (6.24)$$

This equation, then, expresses the maximum value of W at equilibrium. We introduce, now, the constraints imposed by the microcanonical ensemble. First, the total number of molecules, N , remains constant. Taking the variation of N ,

$$\delta N = \sum \delta N_i = 0 . \quad (6.25)$$

Second, the total energy of the system remains constant. If the energy of the i^{th} cell is designated by w_i , the total internal energy will be given by

$$U = \sum w_i N_i . \quad (6.26)$$

The variation of the internal energy is also equal to zero, so that

$$\delta U = \sum w_i \delta N_i = 0 . \quad (6.27)$$

Introducing three Lagrangian multipliers of undetermined value, we combine Equations (6.24), (6.25) and (6.27) to give

$$\sum (\ln N_i + \ln A + \beta w_i) \delta N_i = 0 . \quad (6.28)$$

Since the δN_i are independent, each coefficient is independent, so that

$$\ln N_i + \ln A + \beta w_i = 0$$

or

$$N_i = (1/A) \exp(-\beta w_i) , \quad (6.29)$$

which is the celebrated Maxwell-Boltzmann energy distribution.

$$\sum N_i = N = (1/A) \sum \exp(-\beta w_i) = Z / A .$$

That is,

$$Z = \sum \exp(-\beta w_i) \quad (6.30)$$

where Z , which is known as the *partition function*, stands for the German word *Zustandssumme*, usually translated as “sum of states.” Since

$$1/A = N / Z ,$$

Table 6.2 Illustrating that Different Macrostates Are Compatible with the Same Total Energy.

Cell	Energy	Configuration (macrostate)	
		Configuration A	Configuration B
N_1	1	2	3
N_2	2	1	0
N_3	3	1	0
N_4	4	0	1
N_5	5	0	0
Total energy =		7	7
Number of microstates =		$\frac{4!}{2!1!1!0!0!} = 12$	$\frac{4!}{3!0!0!1!0!} = 4$

Note. Total energy for configuration A is obtained from (2 phase points)(energy = 1) + (1 phase point)(energy = 2) ... = 7 units of energy. Similarly, the total energy for configuration B is 7 units. Although both configurations of 4 phase points have total energy equal to 7, configuration A is the more spread out or disordered and, therefore, has the greater number of microstates or (“thermodynamic probability”).

(6.29) becomes

$$N_i = \frac{N}{Z} \exp(-\beta w_i) \quad (6.31)$$

giving the number of phase points in the i^{th} cell in the state of maximum thermodynamic probability.

We are almost home.

Substituting the value of N_i from (6.31) into (6.23),

$$\begin{aligned} \ln W &= N \ln N - \sum N_i (\ln N - \ln Z - \beta w_i) \\ &= N \ln N - \ln N \sum N_i + \ln Z \sum N_i + \beta \sum w_i N_i . \end{aligned}$$

Now

$$\sum N_i = N \quad \text{and} \quad \sum w_i N_i = U ,$$

so that

$$\ln W = N \ln N - N \ln N + N \ln Z + \beta U .$$

Since

$$S = k_B \ln W , \quad (6.21)$$

$$S = k_B N \ln Z + k_B \beta U . \quad (6.32)$$

Using simple thermodynamic arguments, it can be shown that $k\beta$ is equal to the reciprocal of temperature, so that

$$S = Nk_B \ln Z + U/T . \quad (6.33)$$

Using the partition function, we have been able to “identify” the thermodynamic entropy, S , with a statistical function of molecular behavior in a gas. Again, using Equations (6.26) with (6.30) and (6.31), the internal energy can also be identified with statistical properties of molecules:

$$U = Nk_B T^2 \frac{d(\ln Z)}{dT} . \quad (6.34)$$

The relationship $S = k_B \ln W$ has been the key that permitted the mapping of macroscopic thermodynamics onto the microscopic function of (invisible) molecules in a gas.

The value of W is the maximum value consistent with constant energy, volume and number of particles. To continue the example of Table 6.1, suppose that cell N_1 contained phase points with energy $w_1 = 1$, N_2 with energy $w_2 = 2$, N_3 with energy $w_3 = 3$, N_4 with energy 4, and N_5 with energy 5. Then the microstate with two phase points in cell 1, one point in cell 2, one in cell 3, and none in cells 4 and 5, has a total energy of $(2)(1) + (1)(2) + (1)(3) + (0)(4) + (0)(5) = 7$. W for this configuration was shown to be equal to 12. However another way of obtaining a macrostate with total energy 7 would be $N_1 = 3, N_2 = 0, N_3 = 0, N_4 = 1$, and $N_5 = 0$ (see Table 6.2). W for this new configuration would be equal to

$$\frac{4!}{3!0!0!1!0!} = 4.$$

That is, the configuration (2, 1, 1, 0, 0) has a greater value of W than the configuration (3, 0, 0, 1, 0), and would, therefore, be closer to the maximum value of W for the given total energy of 7 units. The reader might like to check this configuration for maximum W approximately, using N_i -values from Equation (6.29) with $\beta = 1$.

PHYSICAL ENTROPY AND DISORDER

The quantity W permits us to relate physical entropy to “disorder,” which is its usual interpretation in the lay literature. The more phase cells occupied by the gas in configuration space (x - y - z -space), the greater is the physical entropy, S . Dissemination of phase points throughout a larger volume signifies larger W , more disorder, and greater physical entropy.

We can illustrate this idea by calculating the change in thermodynamic entropy in the isothermal, reversible expansion of one mole of an ideal gas from volume V_1 to volume V_2 . If we allow that the internal energy of an ideal gas is independent of its volume, then from the first law of thermodynamics, Equation (6.1),

$$d'Q = d'W = p dV.$$

From the ideal gas law, Equation (1.1), $p dV = (RT/V) dV$.

From the second law of thermodynamics,

$$dS = \frac{d'Q}{T} = \frac{p dV}{T} = \frac{R}{V} dV,$$

so that

$$\Delta S = \int_{V_1}^{V_2} \frac{R}{V} dV = R \ln \frac{V_2}{V_1}. \quad (6.35)$$

Now let us calculate the same quantity, ΔS , from statistical mechanics. Suppose that V_1 comprises n_1 cells of volume δv so that $V_1 = n_1 \delta v$. Similarly, $V_2 = n_2 \delta v$. There are n_1 ways of putting one molecule into V_1 , n_1^2 ways of putting two molecules in V_1 , ..., $n_1^{N_o}$ ways of putting Avogadro's number, N_o , molecules into V_1 . Therefore, $W_1 = n_1^{N_o}$. Similarly, $W_2 = n_2^{N_o}$. Since $S = k_B \ln W$,

$$\begin{aligned} \Delta S &= k_B \ln \Delta W = k_B \ln n_2^{N_o} - k_B \ln n_1^{N_o} \\ &= k_B N_o \ln(n_2/n_1) \\ &= R \ln(n_2/n_1) \end{aligned}$$

since $R = k_B N_o$ (where R is the gas constant, k_B is Boltzmann's constant, and N_o is Avogadro's number). That is,

$$\Delta S = R \ln \left(\frac{V_2/\delta v}{V_1/\delta v} \right)$$

or

$$\Delta S = R \ln \frac{V_2}{V_1} \quad (6.36)$$

exactly as before (6.35).

Thus, we can see clearly how the critical equation, $S = k_B \ln W$, permits the mapping of macroscopic thermodynamics onto the properties of molecules.

APPLICATIONS OF ENTROPY CHANGE IN CHEMISTRY

We recall from inequality (6.16) that the physical entropy of an isolated system increases in every natural (irreversible) process. We now understand with reference to the molecular representation of entropy that the increase in thermodynamic entropy can be interpreted as a progression from a *more*-ordered to a *less*-ordered state. In chemistry, this principle of increasing physical entropy is applied to determine whether a system will change spontaneously. If a change, such as that produced by chemical reaction, will result in an increase in physical entropy, the reaction is likely to go forward (see, for example, Hargreaves and Socrates (1973)).

THE PURPOSE OF THIS CHAPTER

In this chapter we have made a survey of the concept of physical entropy as it was born into classical thermodynamics, and as it was translated, largely through the work of Boltzmann, into statistical mechanics. Physical entropy was seen, through the $S = k_B \ln W = k_B H_I$ relationship, to be proportional to the information theoretical entropy, in the equilibrium state. Through the use of $S \propto H_I$, classical macroscopic thermodynamics became interpretable completely in terms of the properties of molecules. The mapping process was not completed during Boltzmann's lifetime.

We learned from Chapter 5 that the information concept, as it has been applied in communications science and psychology and biochemistry, is non-unique and extrinsic to the system. By contrast, the use of the information measure in physics is both unique and intrinsic. The measure

$$H_I = \ln W \quad (6.19)$$

is a unique measure of uncertainty. No other measure is acceptable. No other measure will permit the mapping of thermodynamics onto statistical mechanics. That is, if $H_I = \ln W$ is correct, our calculations of physical events will be correct; but if $H_I = \ln W$ were incorrect, we would get wrong answers. Our calculations of the thermodynamic entropy, internal energy, free energy, etc., would be wrong. $H_I = \ln W$ is "intrinsic" in the sense that this equation enters physics as a natural law. It is an integral part of physics.

There is yet another way of viewing the two statements,

$$\mathcal{J}(X|Y) = H(X) - H(X, Y) \quad (4.21)$$

and

$$H_I = \ln W. \quad (6.19)$$

Both statements, the first about $\mathcal{J}(X|Y)$ and the second about H_I , might be called *metastatements* (Greek, *meta*: after). That is, Equation (4.21) arose in the case of a channel transmitting a message, or making a *statement*. The statement is arbitrary; for example, "The weather is fine today." Then Equation (4.21) is a metastatement because it talks *about* the statement. It tells us how much information is transmitted by the statement without actually telling us what the statement is. Similarly, $H_I = \ln W$ is a metastatement, giving us the uncertainty that any particular microstate is associated with a macrostate. The corresponding statement would be "We have microstate *bc, a, d, -, -*" (see Table 6.1). However, the statement is hidden from us. Instead, embedded as an integral part of our physical law is

the metastatement “The uncertainty about which particular microstate is present is the maximum value of W consistent with constraints ...”

The contrasts and comparisons made above are important, because when I propose Equation (2.6)⁷

$$F = kH_I,$$

I suggest its use in the physicist’s sense rather than in the communication scientist’s sense. That is, I propose H_I as a *unique* measure, *intrinsic* to the physiological function of an organism. Moreover, by its use I propose that sensory messages that are relayed neuronally to the brain are, in reality, metastatements about the so-called external world. That is, the brain will receive only messages that express the state of uncertainty of the peripheral sensory receptors about the world external, but will never receive messages that directly describe the state of this world. For example, the conscious brain will not receive a message about the brightness of a light, but only about how uncertain the eye is about how bright the light is (Norwich, 1977).

These ideas now set the stage for development of the perceptual theory. There is, however, one more topic we must study by way of background before developing the perceptual model: the informational entropy of continuous distributions.

NOTES

1. Clausius’ statement of the second law of thermodynamics can be shown to be equivalent to Kelvin’s statement of the second law (Sears, 1953)

No process is possible whose sole result is the abstraction of heat from a single reservoir and the performance of an equivalent amount of work. That is, heat cannot be totally converted into work.

2. A reversible process is one carried out with the system effectively in equilibrium with its environment: a succession of equilibrium states. Pressure, temperature and density of each portion of the system remain uniform.

3. μ -space may have more than 6 dimensions if one allows for the internal configuration of the molecule.

4. Often, the coordinates are x, y, z , and p_x, p_y, p_z , where the latter 3 coordinates are the momenta of the particle, mv_x, mv_y, mv_z .

5. Notice that we do not consider permutations within a cell; representative points ab and ba within a phase cell are considered identical.

6. Equation (6.20), as we have written it, is, of course, an anachronism. H_I was not “born,” so to speak, until the time of Wiener and Shannon. Boltzmann did, indeed, have an H -function, but we shall not introduce it at this time.

7. The “ k ” in this equation is a constant, but not Boltzmann’s constant.

REFERENCES

- Brillouin, L. 1962. *Science and Information Theory*. 2nd edition. Academic Press, New York.
- Brillouin, L. 1964. *Scientific Uncertainty and Information*. Academic Press, New York.
- Carnot, S. 1977. *Reflections on the Motive Power of Fire* (E. Medonza, ed.). Peter Smith, Gloucester, M.A.
- Desloge, E. A. 1966. *Statistical Physics*. Holt, Rinehart and Winston, New York.
- Freund, J.E. and Walpole, R.E. 1980. *Mathematical Statistics*. 3rd edition. Prentice-Hall, Englewood Cliffs, N.J.
- Hargreaves, G., and Socrates, G. 1973. *Elementary Chemical Thermodynamics*. Butterworths, London.
- Jackson, E.A. 1968. *Equilibrium Statistical Mechanics*. Prentice-Hall, Englewood Cliffs N.J.
- Norwich, K.H. 1977. On the information received by sensory receptors. *Bulletin of Mathematical Biology*, **39**, 453-461.
- Sears, F.W. 1953. *An Introduction to Thermodynamics, the Kinetic Theory of Gases, and Statistical Mechanics*. 2nd ed. Addison-Wesley, Reading, MA.
- Zemansky, M.W. 1943. *Heat and Thermodynamics*. McGraw-Hill, New York.
- Zemansky, M.W. and Dittman. 1981. *Heat and Thermodynamics: An Intermediate Textbook*, 6th edition. McGraw-Hill, N.Y.