

### 3: The Microscopic Response

#### Statistics for large numbers

In the previous section, we established that, among the extensive properties of macroscopic systems, there is one, called the entropy, that is not macroscopically observable. Notice that although the development did not depend in any way on any features of the microscopic structure of matter, it did seem to reflect hidden or averaged degrees of freedom within the system. Of course, we now live in a time when the microscopically discrete nature of matter has been deeply characterized and is taken completely for granted. It is clear that very many sets of the positions and velocities of the microscopic constituents of a system can correspond to a single macroscopic state of that system. Considering the very large number of particles involved ( $\sim 10^{23}$ ), constructing relationships between the numerous microstates in the microscopic picture and the macroscopic properties is inevitably a statistical enterprise.

#### Averages for energy and volume

Given the complexity, it helps to start with the energy and the volume, extensive state functions that we have more intuition about. Let's, once again, consider a system A, a system B, and the composite system AB comprising A and B as independent subsystems. It seems reasonable to suppose that the macroscopically observed energy and volume,  $E$  and  $V$ , of each of these systems will be the average over those of the distinguishable microstates of each system,

$$\begin{aligned} E_A &= \sum_i \varepsilon_i p_i & V_A &= \sum_i v_i p_i \\ E_B &= \sum_j \varepsilon_j p_j & V_B &= \sum_j v_j p_j \\ E_{AB} &= \sum_{ij} \varepsilon_{ij} p_{ij} & V_{AB} &= \sum_{ij} v_{ij} p_{ij} \end{aligned} \quad (16)$$

where  $p_i$  is the probability that the system A is in microstate  $i$  with energy  $\varepsilon_i$  and volume  $v_i$ ,  $p_j$  is the probability that the system B is in microstate  $j$  with energy  $\varepsilon_j$  and volume  $v_j$ , and  $p_{ij}$  is the probability that the system AB is in microstate  $ij$  (i.e., with subsystem A in state  $i$  and subsystem B in state  $j$ ). As always, energies and volumes are additive,

$$\varepsilon_{ij} = \varepsilon_i + \varepsilon_j \quad \text{and} \quad v_{ij} = v_i + v_j \quad (17)$$

and, since the two subsystems are independent of one another,

$$\sum_{ij} = \sum_i \sum_j \quad \text{and} \quad p_{ij} = p_i p_j \quad (18)$$

It follows that

$$\begin{aligned} E_{AB} &= \sum_{ij} \varepsilon_{ij} p_{ij} & V_{AB} &= \sum_{ij} v_{ij} p_{ij} \\ &= \sum_i \sum_j (\varepsilon_i + \varepsilon_j) p_i p_j & &= \sum_i \sum_j (v_i + v_j) p_i p_j \\ &= \sum_i \sum_j \varepsilon_i p_i p_j + \sum_i \sum_j \varepsilon_j p_i p_j & &= \sum_i \sum_j v_i p_i p_j + \sum_i \sum_j v_j p_i p_j \\ &= \left( \sum_i \varepsilon_i p_i \right) \left( \sum_j p_j \right) + \left( \sum_i p_i \right) \left( \sum_j \varepsilon_j p_j \right) & &= \left( \sum_i v_i p_i \right) \left( \sum_j p_j \right) + \left( \sum_i p_i \right) \left( \sum_j v_j p_j \right) \end{aligned} \quad (19)$$

And, since probabilities sum to 1 (by definition),

$$E_{AB} = \sum_i \varepsilon_i p_i + \sum_j \varepsilon_j p_j = E_A + E_B \quad V_{AB} = \sum_i v_i p_i + \sum_j v_j p_j = V_A + V_B \quad (20)$$

In other words, our interpretation of the macroscopic values of extensive variables as averages of microscopic ones is self-consistent. This can be generalized to the other macroscopic extensive variables,  $A$ ,  $L$ ,  $Q$ ,  $N_i$  and presumably  $S$ .

#### Averages for entropy

An analogous microscopic accounting for  $S$  would posit a feature  $\sigma$  of microstates, with the averages

$$\begin{aligned}
 S_A &= \sum_i \sigma_i p_i \\
 S_B &= \sum_j \sigma_j p_j \\
 S_{AB} &= \sum_{ij} \sigma_{ij} p_{ij}
 \end{aligned}
 \tag{21}$$

The question is what  $\sigma$  should be, given that it must have a value for each microstate that is distinct from those directly related to the macroscopic observables. The remaining possibility is that it is a function of the microstate's probability. But what sort of function would that be?

As the above analysis for  $E$  and  $V$  shows, for  $S$  to be extensive, i.e.,

$$S_{AB} = S_A + S_B \tag{22}$$

it is necessary that

$$\sigma_{ij}(p_{ij}) = \sigma_i(p_i) + \sigma_j(p_j) \tag{23}$$

which corresponds to the well-known property of the logarithm,

$$\ln(p_{ij}) = \ln(p_i p_j) = [\ln(p_i) + \ln(p_j)] \tag{24}$$

Since this relationship also holds when multiplied by any constant, we can now posit that

$$\begin{aligned}
 \sigma_{ij} &= -k \ln(p_{ij}) \\
 \sigma_i &= -k \ln(p_i) \\
 \sigma_j &= -k \ln(p_j)
 \end{aligned}
 \tag{25}$$

where  $k$  is a yet-to-be-determined constant, and find that, by substitution of Eq.25 into Eq. 21, a proper accounting of the macroscopic entropy  $S$  can be had from the averages

$$\begin{aligned}
 S_A &= -k \sum_i p_i \ln(p_i) \\
 S_B &= -k \sum_j p_j \ln(p_j) \\
 S_{AB} &= -k \sum_{ij} p_{ij} \ln(p_{ij})
 \end{aligned}
 \tag{26}$$

or, using angle bracket notation for averaging,

$$S = -k \langle \ln(p) \rangle \tag{27}$$

Notice that the values of  $\sigma$  and  $S$  are always positive if  $k$  is positive, because the probabilities are all  $\leq 1$ . Notice also that  $p \ln(p)$  has the peculiar property that it is zero for both  $p = 0$  and for  $p = 1$ , about which more below.

In some simple cases, all the microstates have the same probability. If there are  $W$  of them

$$p = 1/W \tag{28}$$

and

$$S = k \ln W \tag{29}$$

This is the expression that appears on Ludwig Boltzmann's grave stone.

The one remaining unknown is the value of the constant  $k$ , named after Boltzmann. Once again, we need consistency with macroscopic observations. As shown in Appendix G, this dictates that

$$k = \frac{R}{N_A} \tag{30}$$

where  $N_A$  is Avogadro's constant. In other words, Boltzmann's constant  $k$  and the gas constant  $R$  are the same constant in different units (per particle in  $k$  and per mole of particles in  $R$ ).

## Information, order and spread

What more can we say about the entropy from the microscopic vantage point?

As we have seen,  $S$  is always positive.

The lower limit,  $S = 0$ , corresponds to the special situation of certainty that the system exists in just one microstate, in other words that  $p = 1$  (so  $p \ln(p) = 0$ ) for one microstate and  $p = 0$  (so  $p \ln(p) = 0$ ) for all others.

On the other hand,  $S$  has its maximum value in the case of maximum uncertainty where the system is equally likely to be in any microstate, in other words where  $p$  is the same for all the microstates. This is the result  $S = k \ln W$  for a system with  $W$  equally probable states (i.e., Equation 29).

In general, increasing  $S$  is associated with decreasing microscopic information or order, and can be imagined as a measure of the "spread" of the system's visits among microstates.

## The probabilities

So far, we have seen that understanding the behavior of a system at the microscopic level amounts to enumerating all distinct microstates and determining their probabilities. However, where do the probabilities come from? Specifically, how do the properties of microstates figure into their probabilities of occurring?

Once again, agreement with macroscopic observations is essential. In particular, the probabilities have to comport with the extremum condition that describes equilibrium for each type of system. It is also necessary that the probabilities are normalized (i.e., sum to 1). This sort of constrained optimization is most readily accomplished using the [Method of Lagrange Multipliers](#) with the following results for the most common types of systems.

- In an isolated system (i.e., with rigid, insulating and impermeable walls, such that  $N, V$  and  $E$  are fixed), equilibrium corresponds to the maximum of

$$S = -k \sum_j p_j \ln(p_j) \quad (31)$$

and the corresponding probabilities are all equal to

$$p_j = \frac{1}{Z_{NVE}} \quad (32)$$

where

$$\begin{aligned} Z_{NVE} &= \sum_k 1 = W \\ &= \text{the "multiplicity" (i.e., the number of available microstates)} \end{aligned} \quad (33)$$

normalizes the probabilities (i.e., assures that they sum to 1). Notice that this result followed entirely from the necessary relationship to macroscopic observations. You may come across some presentations of statistical mechanics that avoid this reference to macroscopic observations by citing the equal probabilities of states of the same content, volume and energy as a postulate dubbed "The Principle of Equal A Priori Probabilities". However, it is far more satisfying to derive this result than to assume it.

- For an NVT system (i.e., with thermally conducting, but still rigid and impermeable walls), equilibrium corresponds to the minimum of

$$\begin{aligned} F &= E - TS \\ &= \sum_j \epsilon_j p_j + kT \sum_j p_j \ln(p_j) \end{aligned} \quad (34)$$

and the corresponding probabilities are equal to

$$p_j = \frac{\exp(-\epsilon_j/kT)}{Z_{NVT}} \quad (35)$$

where

$$Z_{NVT} = \sum_k \exp(-\varepsilon_k/kT) \quad (36)$$

normalizes the probabilities. In other words, in this non-isolated system, the temperature imposed by the surroundings influences the probability distribution and it does so in a fashion that satisfies the multiplication rule for a system composed of two independent subsystems (see Equation 18). The probability distribution in Equation 35 is known as "the Boltzmann distribution" and it is the basis for the Maxwell-Boltzmann distribution of molecular speeds in a gas.

- For an NPT system (i.e., with thermally conducting and flexible, but still impermeable walls), equilibrium corresponds to the minimum of

$$\begin{aligned} G &= E + PV - TS \\ &= \sum_j \varepsilon_j p_j + \sum_j P v_j p_j + kT \sum_j p_j \ln(p_j) \end{aligned} \quad (37)$$

and the corresponding probabilities are equal to

$$p_j = \frac{\exp(-(\varepsilon_j + P v_j)/kT)}{Z_{NPT}} \quad (38)$$

where

$$Z_{NPT} = \sum_k \exp(-(\varepsilon_k + P v_k)/kT) \quad (39)$$

normalizes the probabilities. In other words, in this non-isolated system, the probability distribution is influenced by the temperature and pressure imposed by the surroundings, in a fashion that satisfies the multiplication rule for a composite system comprising two independent subsystems (see Equation 18).

## The partition function

The symbol  $Z$  comes from the German "Zustandssumme" for "sum over states". Of course, the sum varies with the values of the applied conditions and, in this respect,  $Z$  is known as the partition function. As for any function of multiple variables, we can take partial derivatives and we find that those with respect to intensive variables provide averages of practical significance. In particular, for the NVT system

$$\frac{\partial Z_{NVT}}{\partial(1/T)} = -E/k \quad \text{or} \quad E = kT^2 \left[ \frac{\partial \ln(Z_{NVT})}{\partial(T)} \right] \quad (40)$$

and for the NPT system

$$\frac{\partial Z_{NPT}}{\partial(1/T)} = -(E + PV)/k \quad \text{or} \quad H = kT^2 \left[ \frac{\partial \ln(Z_{NPT})}{\partial(T)} \right] \quad (41)$$

and

$$\frac{\partial Z_{NPT}}{\partial(P)} = -V/kT \quad \text{or} \quad V = -kT \left[ \frac{\partial \ln(Z_{NPT})}{\partial(P)} \right] \quad (42)$$

From these results, other state functions can be obtained by using the relationships in Appendices E and F.

This page titled [3: The Microscopic Response](#) is shared under a [CC BY-NC-ND 4.0](#) license and was authored, remixed, and/or curated by [Judith Herzfeld](#).