

6

The Classical Ideal Gas: Energy Dependence of Entropy

You should call it entropy, for two reasons. In the first place your uncertainty function has been used in statistical mechanics under that name, so it already has a name. In the second place, and more important, no one really knows what entropy really is, so in a debate you will always have the advantage.

John von Neumann, suggesting to Claude Shannon a name for his new uncertainty function, as quoted in *Scientific American*, (1971) 225(3): 180.

This chapter provides the second half of the derivation of the entropy of the classical ideal gas, as outlined in Chapter 2 and begun in Chapter 4. In the present chapter we will calculate the probability distribution for the energy of each subsystem. The logarithm of that probability then gives the energy-dependent contribution to the entropy of the classical ideal gas. The total entropy is just the sum of the configurational entropy and the energy-dependent terms, as discussed in Section 4.1.

6.1 Distribution for the Energy between Two Subsystems

We again consider the composite system that consists of the two subsystems discussed in Chapter 2, and we will extend the discussion to $M \geq 2$ systems in Section 6.6. Subsystem j contains N_j particles and subsystem k contains N_k particles, with a total of $N_{T,jk} = N_j + N_k$ particles in the composite system. Since we are dealing with classical, non-interacting particles, the energy of each subsystem is given by

$$E_\alpha = \sum_{i=1}^{N_\alpha} \frac{|\vec{p}_{\alpha,i}|^2}{2m}, \quad (6.1)$$

where $\alpha = j$ or k and m is the mass of a single particle. The momentum of the i -th particle in subsystem α is $\vec{p}_{\alpha,i}$.

The composite system is perfectly isolated from the rest of the universe, and the total energy $E_{T,jk} = E_j + E_k$ is fixed. We are interested in the case in which the two subsystems can be brought into thermal contact to enable the subsystems to exchange energy, so that the composite system can come to thermal equilibrium.

A partition that is impervious to particles, but allows energy to be exchanged between subsystems, is called a ‘diathermal’ wall, in contrast to an adiabatic wall that prevents the exchange of either particles or energy. We wish to calculate the probability distribution of the energy between two subsystems separated by a diathermal wall.

To proceed with the entropy calculation, we must make an assumption about the probability distribution in momentum space. The simplest assumption is that the momentum distribution is constant, subject to the constraints on the energy. This is again the model probability that we introduced in Chapter 4.

With the assumption of uniform probability density in momentum space, we can calculate the probability distribution for the energy from that of the momenta, using the methods from Section 5.4. We use Dirac delta functions to select states for which system j has energy E_j and system k has energy E_k . Conservation of energy in the composite system of course requires that $E_k = E_{T,jk} - E_j$, but this form of writing the probability distribution will be useful in Section 6.2 when we separate the contributions of the two subsystems.

Since the notation can become unwieldy at this point, we use a compact notation for the integrals

$$\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} (\cdots) dp_{\alpha,1}^3 \cdots dp_{\alpha,N_j}^3 \equiv \int_{-\infty}^{\infty} (\cdots) dp_{\alpha}, \quad (6.2)$$

and write:

$$P(E_j, E_k) = \frac{\int_{-\infty}^{\infty} \delta\left(E_j - \sum_{i=1}^{N_j} \frac{|\vec{p}_{j,i}|^2}{2m}\right) dp_j \int_{-\infty}^{\infty} \delta\left(E_k - \sum_{\ell=1}^{N_k} \frac{|\vec{p}_{k,\ell}|^2}{2m}\right) dp_k}{\int_{-\infty}^{\infty} \delta\left(E_{T,jk} - \sum_{i=1}^{N_{T,jk}} \frac{|\vec{p}_i|^2}{2m}\right) dp}. \quad (6.3)$$

The integral over dp in the denominator goes over the momenta of all particles in systems j and k .

The denominator of eq. (6.3) is chosen to normalize the probability distribution,

$$\int_0^{\infty} P(E_j, E_{T,jk} - E_j) dE_j = 1. \quad (6.4)$$

By defining a function

$$\Omega_p(E_\alpha, N_\alpha) = \int_{-\infty}^{\infty} \delta\left(E_\alpha - \sum_{i=1}^{N_\alpha} \frac{|\vec{p}_{\alpha,i}|^2}{2m}\right) dp_\alpha, \quad (6.5)$$

we can write eq. (6.3) in an even more compact form that highlights its similarity to eq. (4.13)

$$P(E_j, E_k) = \frac{\Omega_p(E_j, N_j) \Omega_p(E_k, N_k)}{\Omega_p(E_{T,jk}, N_{T,jk})}. \quad (6.6)$$

6.2 Evaluation of Ω_p

To complete the derivation of $P(E_j, E_k)$, we need to evaluate the function $\Omega_p(E, N)$ in eq. (6.5), where E and N are generic variables that stand for E_j, E_k , or $E_{T,jk}$ and N_j, N_k , or $N_{T,jk}$. Fortunately, this can be done exactly by taking advantage of the symmetry of the integrand to transform the $3N$ -dimensional integral to a one-dimensional integral, and then relating the integral to a known function.

6.2.1 Exploiting the Spherical Symmetry of the Integrand

The delta function in eq. (6.5) makes the integrand vanish everywhere in momentum space, except on a sphere in momentum space defined by energy conservation

$$2mE = \sum_{i=1}^N |\vec{p}_i|^2. \quad (6.7)$$

The radius of this $3N$ -dimensional sphere is clearly $\sqrt{2mE}$.

Since the area of an n -dimensional sphere is given by $A = S_n r^{n-1}$, where S_n is a constant that depends on the dimension, n , the expression for the function can be immediately transformed from a $3N$ -dimensional integral to a one-dimensional integral,

$$\begin{aligned} \Omega_p(E, N) &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \delta\left(E - \sum_{i=1}^N \frac{|\vec{p}_i|^2}{2m}\right) dp_j \cdots dp_{3N} \\ &= \int_0^{\infty} S_n p^{3N-1} \delta\left(E - \frac{p^2}{2m}\right) dp. \end{aligned} \quad (6.8)$$

The new integration variable is just the radial distance in momentum space,

$$p^2 = \sum_{i=1}^N |\vec{p}_i|^2. \quad (6.9)$$

To evaluate the integral in eq. (6.8), transform the variable of integration by defining $x = p^2/2m$, which implies that $p^2 = 2mx$ and $p dp = m dx$. Inserting this in the expression for $\Omega_p(E, N)$, we can evaluate the integral

$$\begin{aligned} \Omega_p(E, N) &= S_n \int_0^\infty (2mx)^{(3N-1)/2} \delta(E-x) (2mx)^{-1/2} m dx \\ &= S_n m (2mE)^{3N/2-1}. \end{aligned} \quad (6.10)$$

The only thing remaining is to evaluate S_n , which is the surface area of an n -dimensional sphere of unit radius.

6.2.2 The Surface Area of an n -Dimensional Sphere

The volume of a sphere in n dimensions is given by $C_n r^n$, where C_n is a constant. By differentiation, the area of the sphere is $nC_n r^{n-1}$, so that $S_n = nC_n$.

To evaluate C_n we can use a trick involving Gaussian integrals. We begin by taking the n -th power of both sides of eq. (3.61):

$$\begin{aligned} \left[\int_{-\infty}^{\infty} e^{-x^2} dx \right]^n &= [\sqrt{\pi}]^n = \pi^{n/2} \\ &= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \exp \left(- \sum_{j=1}^n x_j^2 \right) dx_1 dx_2, \dots, dx_n \\ &= \int_0^\infty \exp(-r^2) S_n r^{n-1} dr \\ &= nC_n \int_0^\infty \exp(-r^2) r^{n-1} dr. \end{aligned} \quad (6.11)$$

We can transform the final integral in eq. (6.11) to a more convenient form by changing the integration variable to $t = r^2$,

$$\pi^{n/2} = nC_n \int_0^\infty \exp(-r^2) r^{n-1} dr$$

$$\begin{aligned}
&= \frac{1}{2} n C_n \int_0^\infty \exp(-r^2) r^{n-2} 2r dr \\
&= \frac{1}{2} n C_n \int_0^\infty e^{-t} t^{n/2-1} dt.
\end{aligned} \tag{6.12}$$

The integral in eq. (6.12) is well known; it is another representation of the factorial function

$$m! = \int_0^\infty e^{-t} t^m dt. \tag{6.13}$$

We can see why this is so by induction. Begin with $m = 0$,

$$\int_0^\infty e^{-t} t^0 dt = \int_0^\infty e^{-t} dt = 1 = 0!. \tag{6.14}$$

If eq. (6.13) is valid for m , then we can use integration by parts to prove that it is valid for $(m+1)$,

$$\begin{aligned}
\int_0^\infty e^{-t} t^{m+1} dt &= [e^{-t} (m+1) t^m]_0^\infty - (m+1) \int_0^\infty (-e^{-t}) t^m dt \\
&= (m+1) \int_0^\infty e^{-t} t^m dt \\
&= (m+1) m! \\
&= (m+1)!.
\end{aligned} \tag{6.15}$$

This confirms the validity of eq. (6.13) for all positive integers.

A convenient consequence of eq. (6.13) is that it provides an analytic continuation of $m!$ to non-integer values. It is traditional to call this extension of the concept of factorials a Gamma (Γ) function,

$$\Gamma(m+1) = m! = \int_0^\infty e^{-t} t^m dt. \tag{6.16}$$

For many purposes, $\Gamma(\cdot)$ is a very useful notation. However, the shift between m and $m+1$ can be a mental hazard. I will stay with the factorial notation, even for non-integer values.

Returning to eq. (6.12), we can now complete the derivation of C_n ,

$$\begin{aligned}\pi^{n/2} &= \frac{1}{2} n C_n \int_0^\infty e^{-t} t^{n/2-1} dt \\ &= C_n \left(\frac{n}{2}\right) (n/2 - 1)! \\ &= (n/2)! C_n.\end{aligned}\tag{6.17}$$

This gives us values for C_n and $S_n = n C_n$:

$$C_n = \frac{\pi^{n/2}}{(n/2)!}\tag{6.18}$$

$$S_n = n \frac{\pi^{n/2}}{(n/2)!}.\tag{6.19}$$

6.2.3 Exact Expression for Ω_p

Inserting eq. (6.19) into the expression for $\Omega_p(E, N)$ in eq. (6.10) we find

$$\Omega_p(E, N) = \frac{3N \pi^{3N/2}}{(3N/2)!} m (2mE)^{3N/2-1}.\tag{6.20}$$

An excellent approximation to the logarithm of $\Omega_p(E, N)$ can be found using Stirling's approximation. The errors are of order $\ln N/N$,

$$\ln \Omega_p(E, N) \approx N \left[\frac{3}{2} \ln \left(\frac{E}{N} \right) + X \right].\tag{6.21}$$

The constant X in eq. (6.21) can be calculated, but will not be needed at this point.

6.3 Distribution of Energy between Two Isolated Subsystems that were Previously in Equilibrium

As was the case for the probability distribution of particles, the probability distribution of the energy is exactly the same in equilibrium and for isolated systems that were previously in equilibrium. The act of separating the two systems has no effect on the probability distribution of the energy. Analogous to Subsection 4.3, when the two systems are allowed to exchange energy, the probability distribution for E_j is given by the product of the two factors.

6.4 Probability Distribution for Large N

Putting the functional form of Ω_p into eq. (6.6)—and ignoring constants for the time being—we find the energy-dependence of the probability distribution,

$$P(E_j, E_{T,jk} - E_j) = \frac{\Omega_p(E_j, N_j) \Omega_p(E_{T,jk} - E_j, N_k)}{\Omega_p(E_{T,jk}, N_{T,jk})} \propto (E_j)^{3N_j/2-1} (E_{T,jk} - E_j)^{3N_k/2-1}. \quad (6.22)$$

We are primarily interested in finding the average energy, $\langle E_j \rangle$. The energy E_j has a very narrow probability distribution, and the measured energy will almost always be equal to $\langle E_j \rangle$ within experimental error. We want to find the width of the distribution, δE_j , primarily to confirm that the distribution is narrow. As before, note that the neglect of the width of the distribution is a extremely good approximation, but an approximation nevertheless. We will return to this point in Chapter 21.

The average, $\langle E_j \rangle$, is most easily found by approximating it by the mode of the probability distribution, which is the location of the maximum probability. The mode is found by taking the derivative of the logarithm of the energy distribution and setting it equal to zero,

$$\frac{\partial}{\partial E_j} \ln P(E_j, E_{T,jk} - E_j) = \left(\frac{3N_j}{2} - 1 \right) \frac{1}{E_j} - \left(\frac{3N_k}{2} - 1 \right) \frac{1}{E_{T,jk} - E_j} = 0, \quad (6.23)$$

Solving this equation gives us the location of the maximum of the probability distribution, which is located at $\langle E_j \rangle$ to very high accuracy,

$$\langle E_j \rangle = E_{j,\max} = \left(\frac{3N_j - 2}{3N_{T,jk} - 4} \right) E_{T,jk}. \quad (6.24)$$

For a large number of particles, this becomes

$$E_{j,\max} = \left(\frac{N_j}{N_{T,jk}} \right) E_{T,jk}, \quad (6.25)$$

or

$$\frac{E_{j,\max}}{N_j} = \frac{E_{T,jk}}{N_{T,jk}} = \frac{\langle E_j \rangle}{N_j}, \quad (6.26)$$

with a relative error of the order of $N_{T,jk}^{-1}$. When $N_{T,jk} = 10^{20}$, or even $N_{T,jk} = 10^{12}$, this is certainly an excellent approximation.

Note that the energy per particle is the same in each subsystem. This is a special case of the equipartition theorem, which we will discuss in more detail when we come to the canonical ensemble in Chapter 19.

For large N_j , the probability distribution is very sharply peaked. It can be approximated by a Gaussian, and we can use the same methods as we did in Section 3.9 to determine the appropriate parameters. In particular, we can calculate the width of the probability distribution by taking the second derivative of the logarithm, which gives us the negative reciprocal of the variance:

$$\frac{\partial^2}{\partial E_j^2} \ln P(E_j, E_{T,jk} - E_j) = - \left(\frac{3N_j}{2} - 1 \right) \frac{1}{E_j^2} - \left(\frac{3N_k}{2} - 1 \right) \frac{1}{(E_{T,jk} - E_j)^2}. \quad (6.27)$$

Evaluating this at the maximum of the function, $E_{j,\max} = E_{T,jk} N_j / N_{T,jk}$, and assuming that we are dealing with large numbers of particles so that we can approximate $3N_j/2 - 1 \approx 3N_j/2$, this becomes

$$\begin{aligned} \frac{\partial^2}{\partial E_j^2} \ln P(E_j, E_{T,jk} - E_j) |_{E_j = E_{T,jk} N_j / N_{T,jk}} &= - \left(\frac{3N_j}{2} \right) \frac{N_{T,jk}^2}{E_{T,jk}^2 N_j^2} - \left(\frac{3N_k}{2} \right) \frac{N_{T,jk}^2}{E_{T,jk}^2 N_k^2} \\ &= - \frac{3N_{T,jk}^2}{2E_{T,jk}^2} \left(\frac{1}{N_j} + \frac{1}{N_k} \right) \\ &= - \frac{3N_j^2}{2E_j^2} \left(\frac{N_{T,jk}}{N_j N_k} \right) \\ &= -\sigma_{E_j}^{-2}. \end{aligned} \quad (6.28)$$

The variance of E_j is then:

$$\sigma_{E_j}^2 = \frac{2\langle E_j \rangle^2}{3N_j^2} \left(\frac{N_j N_k}{N_{T,jk}} \right) = \frac{\langle E_j \rangle^2}{N_{T,jk}} \left(\frac{2N_k}{3N_j} \right). \quad (6.29)$$

The width of the energy distribution is then given by the standard deviation $\delta E_j = \sigma_{E_j}$,

$$\sigma_{E_j} = \langle E_j \rangle \sqrt{\frac{1}{N_{T,jk}} \left(\frac{2N_k}{3N_j} \right)^{1/2}}$$

$$= \langle E_j \rangle \sqrt{\frac{1}{N_j} \left(\frac{2N_k}{3N_{T,jk}} \right)^{1/2}}. \quad (6.30)$$

From eq. (6.30), we see that since $E_j \propto N_j$, the width of the probability distribution increases with $\sqrt{N_j}$, while the relative width,

$$\frac{\delta E_j}{\langle E_j \rangle} = \frac{\sigma_{E_j}}{\langle E_j \rangle} = \sqrt{\frac{1}{N_j} \left(\frac{2N_k}{3N_{T,jk}} \right)^{1/2}} \quad (6.31)$$

decreases with $\sqrt{N_j}$.

The behavior of the width and relative width is analogous to the $\sqrt{N_j}$ behavior we saw in eq. (4.9) for the width of the probability distribution for N_j . As the size of the system increases, the width of the probability distribution for E_j increases, but the relative width decreases. For a system with 10^{20} particles, the typical relative deviation of E_j from its average value is of the order of 10^{-10} , which is a very small number.

This increase of the standard deviation δE_j with the size of the system has the same significance discussed in Section 4.5 for δN_j . The average value of the energy is a description of the subsystem that we find useful; it is not the same as the true energy of the subsystem at any given time.

For both the energy and the number of particles in a subsystem, we have found that the relative width of the probability distribution is very narrow. As long as the width of the probability distribution is narrower than the uncertainty in the experimental measurements, we can regard the predictions of probability theory (statistical mechanics) as giving effectively deterministic values.

It is important to note both that it is not necessary to take the ‘thermodynamic limit’ (infinite size) for thermodynamics to be valid, and that it is necessary for the system to be large enough so that the relative statistical fluctuations of measured quantities is negligible in comparison with the experimental error.

6.5 The Logarithm of the Probability Distribution and the Energy-Dependent Terms in the Entropy

In direct analogy to the derivation of eq. (4.17) for the configurational contributions in Chapter 4, we can find the contributions from the momentum degrees of freedom to the entropy (the energy-dependent term) by taking the logarithm of eq. (6.6).

In analogy to eq. (4.15), we can define a function to describe the contributions to the entropy of the classical ideal gas from the momentum degrees of freedom,

$$S_{p,\alpha} = k \ln \Omega_p(E_\alpha, N_\alpha) \quad (6.32)$$

where α refers to one of the subsystems, j or k , or the total system when the subscript T, jk is used. Taking the logarithm of eq. (6.22), the energy-dependent contributions to the entropy of the composite system are seen to be:

$$\begin{aligned} S_{p,tot}(E_j, N_j, E_k, N_k) &\equiv k \ln P(E_j, N_j, E_k, N_k) + S_p(E_{T,jk}, N_{T,jk}) \\ &= S_p(E_j, N_j) + S_p(E_k, N_k). \end{aligned} \quad (6.33)$$

$S_{p,tot}(E_j, N_j, E_k, N_k)$ is the energy-dependent contribution to the total entropy of the composite system, while $S_p(E_j, N_j)$ and $S_p(E_k, N_k)$ as the corresponding energy-dependent contributions to the entropies of the subsystems. Eq. (6.33) shows that the energy-dependent contributions to the entropy of an ideal gas are additive, as the configurational terms were shown to be in Section 4.7.

As was the case for the configurational entropy, the height of the probability distribution for E_j is proportional to $\sqrt{N_j}$ (or $\sqrt{N_k}$), so that in equilibrium

$$k \ln P(E_j, N_j, E_k, N_k) \propto \ln N_j \lll S_p(E_{T,jk}, N_{T,jk}) \propto N_{T,jk} \quad (6.34)$$

and the energy-dependent contributions to the entropy of the composite system in equilibrium is given by $S_p(E_{T,jk}, N_{T,jk})$,

$$S_{p,tot}(E_j, N_j, E_k, N_k) = S_p(E_j, N_j) + S_p(E_k, N_k) = S_p(E_{T,jk}, N_{T,jk}). \quad (6.35)$$

If E and N are generic variables standing for N_j or N_k , we can see from eq. (6.21), that the energy contributions of the entropy for large N are given by

$$S_p(E, N) = k \ln \Omega_p(E, N) \approx kN \left[\frac{3}{2} \ln \left(\frac{E}{N} \right) + X \right] \quad (6.36)$$

where the errors are of order $1/N$, and X is a constant that we will not need to evaluate at this point.

6.6 The Generalization to $M \geq 2$ systems

As in Section 4.8, the equation for the entropy can be generalized to an arbitrary number $M \geq 2$ systems. Eq. (6.6) simply becomes a product of M factors, one for each system,

$$P(\{E_k, N_k | k = 1, \dots, M\}) = \frac{\sum_{k=1}^M \Omega_p(E_k, N_k)}{\Omega_p(E_T, N_T)}, \quad (6.37)$$

where $\Omega_p(E_T, N_T)$ is a normalization constant. In this equation, $E_T = \sum_{k=1}^M E_k$ and $N_T = \sum_{k=1}^M N_k$ are constants. Eq. (6.37) is analogous to eq. (4.22), which was derived from the discrete probability distribution for the particle distribution.

Finally, in analogy to Chapter 4, we can define the energy contributions to the entropy as

$$S_p(E_k, N_k) = \ln \Omega_p(E_k, N_k) \quad (6.38)$$

so that

$$S_p(\{E_k, N_k | k = 1, \dots, M\}) = \sum_{k=1}^M S_p(E_k, N_k) - S_p(E_T, N_T) + C' \quad (6.39)$$

where E_T , N_T , and C' are constants, whose values have no physical consequences.

Eq. (6.39) shows that contributions to the entropy of an individual system, $S_p(E_k, N_k)$, are exactly the same as derived from considering just two systems, and, in fact, are independent of the total number of systems. The validity of this equation is not dependent on whether the systems are isolated or in equilibrium with each other.