

5.3: Heat Capacity of an Ideal Gas

5.3.1 The equipartition theorem

Equipartition theorem. (For classical systems.) Suppose the Hamiltonian $H(\Gamma)$ decouples into one piece involving a single phase space variable—call it H_1 , plus another piece which involves all the other phase space variables—call it $H_2(\Gamma_2)$. Suppose further that the energy depends quadratically upon this single phase space variable, and that this variable may take on values from $-\infty$ to $+\infty$. Then, in classical statistical mechanics, the mean contribution to the energy due to that single variable is

$$\langle H_1 \rangle = \frac{1}{2} k_B T. \quad (5.3.1)$$

Notice how general this theorem is. The remaining piece of the Hamiltonian, $H_2(\Gamma_2)$, might decouple further or it might not. The phase space variable entering H_1 might be a momentum,

$$H_1(p) = \frac{p^2}{2m}, \quad (5.3.2)$$

or an angular momentum,

$$H_1(\ell) = \frac{\ell^2}{2I}, \quad (5.3.3)$$

or even a position coordinate, as in the simple harmonic oscillator energy

$$H_1(x) = \frac{1}{2} k x^2. \quad (5.3.4)$$

Furthermore, in all these circumstances, the mean energy is *independent* of the particular parameters m or I or k . . . it depends only upon the temperature. This explains the origin of the name “equipartition”: the mean translational energy due to motion in the x direction is equal to the mean rotational energy due to the change of θ , and this holds true even if the gas is a mixture of molecules with different masses and different moments of inertia. The energy is equally partitioned among all these different ways of holding energy.

Proof. We will write

$$H_1(p) = ap^2, \quad (5.3.5)$$

although the variable might not be a linear momentum. Then the average of H_1 is

$$\langle H_1 \rangle = \frac{\int d\Gamma H_1 e^{-\beta H(\Gamma)}}{\int d\Gamma e^{-\beta H(\Gamma)}} = \frac{\int_{-\infty}^{+\infty} dp H_1 e^{-\beta H_1} \int d\Gamma_2 e^{-\beta H_2(\Gamma_2)}}{\int_{-\infty}^{+\infty} dp e^{-\beta H_1} \int d\Gamma_2 e^{-\beta H_2(\Gamma_2)}}. \quad (5.3.6)$$

Clearly, the integrals over Γ_2 cancel in this last expression. (This explains why the form of H_2 is irrelevant to the theorem.) We are left with

$$\langle H_1 \rangle = \frac{\int_{-\infty}^{+\infty} dp a p^2 e^{-\beta a p^2}}{\int_{-\infty}^{+\infty} dp e^{-\beta a p^2}}. \quad (5.3.7)$$

These two integrals could be evaluated in terms of Gamma functions (see appendix C), but they don’t need to be evaluated yet. Think for a moment about our “slick trick” of parametric differentiation. . . using it we can write

$$\langle H_1 \rangle = -\frac{d}{d\beta} \ln \left[\int_{-\infty}^{+\infty} dp e^{-\beta a p^2} \right]. \quad (5.3.8)$$

The integral that remains is of Gaussian character and we could evaluate it using the results of Appendix B. But before rushing in to integrate, let’s employ the substitution $u = \sqrt{\beta a} p$ to find

$$\langle H_1 \rangle = -\frac{d}{d\beta} \ln \left[\frac{1}{\sqrt{\beta a}} \int_{-\infty}^{+\infty} du e^{-u^2} \right] = -\frac{d}{d\beta} \left\{ \ln \left[\frac{1}{\sqrt{\beta}} \right] + \ln \left[\frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} du e^{-u^2} \right] \right\}. \quad (5.3.9)$$

This last expression shows that there's no need to evaluate the integral. Whatever its value is, it is some number, not a function of β , so when we take the derivative with respect to β the term involving that number will differentiate to zero. Similarly for the constant a , which explains why the equipartition result is independent of that prefactor. We are left with

$$\langle H_1 \rangle = -\frac{d}{d\beta} \ln \frac{1}{\sqrt{\beta}} = \frac{1}{2} \frac{d}{d\beta} \ln \beta = \frac{1}{2} \frac{1}{\beta} \quad (5.3.10)$$

or, finally, the desired equipartition result

$$\langle H_1 \rangle = \frac{1}{2} k_B T. \quad (5.3.11)$$

5.3.2 Applications of equipartition; Comparison with experiment

5.3.3 Crossover between classical and quantal behavior; Freeze out

At high temperatures, typical thermal energies are much greater than level spacings. Transitions from one level to another are very easy to do and the granular character of the quantized energies can be ignored. This is the classical limit, and equipartition holds!

At low temperatures, typical thermal energies are less than the level spacing between the ground state and the first excited state. There is so little thermal energy around that the molecule cannot even be excited out of its ground state. Virtually all the molecules are in their ground states, and the excited states might as well just not exist.

In classical mechanics, a diatomic molecule offered a small amount of rotational energy will accept that energy and rotate slowly. But in quantum mechanics, a diatomic molecule offered a small amount of rotational energy will reject that energy and remain in the ground state, because the energy offered is not enough to lift it into the first excited state.² The quantal diatomic molecule does not rotate at all at low temperatures, so it behaves exactly like a monatomic molecule with only center-of-mass degrees of freedom.

In short, we explain the high-temperature rotational specific heat ($c^{\text{rot}}_V = k_B$) through equipartition. We explain the low-temperature rotational specific heat (c^{rot}_V vanishes) through difficulty of promotion to the first quantal excited state. This fall-off of specific heat as the temperature is reduced is called “freeze out”.

The crossover between the high-temperature and low-temperature regimes occurs in the vicinity of a characteristic temperature θ at which the typical thermal energy is equal to energy separation between the ground state and the first excited state. If the energies of these two states are ϵ_0 and ϵ_1 respectively, then we define the characteristic crossover temperature through

$$k_B \theta \equiv \epsilon_1 - \epsilon_0 \quad (5.3.12)$$

5.2 Generalized equipartition theorem and the ultra-relativistic gas

- a. Suppose the Hamiltonian $H(\Gamma)$ decouples into two pieces

$$H(\Gamma) = a|p|^n + H_2(\Gamma_2) \quad (5.3.13)$$

where p is some phase space variable that may take on values from $-\infty$ to $+\infty$, and where Γ_2 represents all the phase space variables except for p . (Note that the absolute value $|p|$ is needed in order to avoid, for example, taking the square root of a negative number in the case $n = 1/2$.) Show that, in classical statistical mechanics, the mean contribution to the energy due to that single variable is

$$\langle a|p|^n \rangle = \frac{1}{n} k_B T \quad (5.3.14)$$

- b. In special relativity, the energy of a free (i.e. non-interacting) particle is given by

$$\sqrt{(mc^2)^2 + (pc)^2} \quad (5.3.15)$$

where c is the speed of light. As you know, when $v \ll c$ this gives the non-relativistic kinetic energy $\text{KE} \approx mc^2 + p^2/2m$. In the “ultra-relativistic” limit, where v is close to c , the energy is approximately pc . What is the heat capacity of a gas of non-interacting ultra-relativistic particles?

- c. Estimate the crossover temperature between the non-relativistic and ultra-relativistic regimes.

5.3 Another generalization of equipartition

Consider the same situation as the equipartition theorem in the text, but now suppose the single phase space variable takes on values from 0 to $+\infty$. What is the corresponding result for $\langle H_1 \rangle$?

5.4 Equipartition and the virial theorem

Look up the term “virial theorem” in a classical mechanics textbook. Is there any relation between the virial theorem of classical mechanics and the equipartition theorem of classical statistical mechanics?

²This paragraph is written in the “shorthand” language discussed on page 113, as if energy eigenstates were the only allowed quantal states.

The \mathcal{O} Notation

Approximations are an important part of physics, and an important part of approximation is to ensure their reliability and consistency. The \mathcal{O} notation (pronounced “the big-oh notation”) is an important and practical tool for making approximations reliable and consistent. The technique is best illustrated through an example. Suppose you desire an approximation for

$$f(x) = \frac{e^{-x}}{1-x} \quad (5.3.16)$$

valid for small values of x , that is, $x \ll 1$. You know that

$$e^{-x} = 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \dots \quad (5.3.17)$$

and that

$$\frac{1}{1-x} = 1 + x + x^2 + x^3 + \dots \quad (5.3.18)$$

so it seems that reasonable approximations are

$$e^{-x} \approx 1 - x \quad (5.3.19)$$

and

$$\frac{1}{1-x} \approx 1 + x \quad (5.3.20)$$

whence

$$\frac{e^{-x}}{1-x} \approx (1-x)(1+x) = 1 - x^2. \quad (5.3.21)$$

Let’s try out this approximation at $x_0 = 0.01$. A calculator shows that

$$\frac{e^{-x_0}}{1-x_0} = 1.0000503\dots \quad (5.3.22)$$

while the value for the approximation is

$$1 - x_0^2 = 0.9999000. \quad (5.3.23)$$

This is a very poor approximation indeed. . . the deviation from $f(0) = 1$ is even of the wrong sign!

Let’s do the problem over again, but this time keeping track of exactly how much we’ve thrown away while making each approximation. We write

$$e^{-x} = 1 - x + \frac{1}{2}x^2 - \frac{1}{6}x^3 + \dots \quad (5.3.24)$$

as

$$e^{-x} = 1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3), \quad (5.3.25)$$

where the notation $\mathcal{O}(x^3)$ stands for the small terms that we haven't bothered to write out explicitly. The symbol $\mathcal{O}(x^3)$ means “terms that are about the magnitude of x^3 , or smaller” and is pronounced “terms of order x^3 ”. The \mathcal{O} notation will allow us to make controlled approximations in which we keep track of exactly how good the approximation is.

Similarly, we write

$$\frac{1}{1-x} = 1 + x + x^2 + \mathcal{O}(x^3), \quad (5.3.26)$$

and find the product

$$f(x) = \left[1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)\right] \times [1 + x + x^2 + \mathcal{O}(x^3)] \quad (5.3.27)$$

$$= \left[1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)\right] \quad (5.3.28)$$

$$+ \left[1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)\right] x \quad (5.3.29)$$

$$+ \left[1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)\right] x^2 \quad (5.3.30)$$

$$+ \left[1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)\right] \mathcal{O}(x^3) \quad (5.3.31)$$

Note, however, that $x \times \frac{1}{2}x^2 = \mathcal{O}(x^3)$, and that $x^2 \times \mathcal{O}(x^3) = \mathcal{O}(x^3)$, and so forth, whence

$$f(x) = \left[1 - x + \frac{1}{2}x^2 + \mathcal{O}(x^3)\right] \quad (5.3.32)$$

$$+ [x - x^2 + \mathcal{O}(x^3)] \quad (5.3.33)$$

$$+ [x^2 + \mathcal{O}(x^3)] \quad (5.3.34)$$

$$+ \mathcal{O}(x^3) \quad (5.3.35)$$

$$= 1 + \frac{1}{2}x^2 + \mathcal{O}(x^3) \quad (5.3.36)$$

Thus we have the approximation

$$f(x) \approx 1 + \frac{1}{2}x^2 \quad (5.3.37)$$

Furthermore, we know that this approximation is accurate to terms of order $\mathcal{O}(x^2)$ (i.e. that the first neglected terms are of order $\mathcal{O}(x^3)$). Evaluating this approximation at $x_0 = 0.01$ gives

$$1 + \frac{1}{2}x_0^2 = 1.0000500, \quad (5.3.38)$$

far superior to our old approximation (5.35).

What went wrong on our first try? The $-x^2$ in approximation (5.35) is the same as the $-x^2$ on line (5.47). However, lines (5.46) and (5.48) demonstrate that there were other terms of about the same size (that is, other “terms of order x^2 ”) that we neglected in our first attempt.

The \mathcal{O} notation is superior to the “dot notation” (such as \cdots) in that dots stand for “a bunch of small terms”, but the dots don't tell you just how small they are. The symbol $\mathcal{O}(x^3)$ also stands for “a bunch of small terms”, but in addition it tells you precisely how small those terms are. The \mathcal{O} notation allows us to approximate in a consistent manner, unlike the uncontrolled approximations where we ignore a “small term” without knowing whether we have already retained terms that are even smaller.

This page titled [5.3: Heat Capacity of an Ideal Gas](#) is shared under a [CC BY-SA](#) license and was authored, remixed, and/or curated by [Daniel F. Styer](#).