SIMPLE APPLICATIONS OF STATISTICAL MECHANICS

Andrew M. Steane

Exeter College and Department of Atomic and Laser Physics, University of Oxford.

February 24, 2025

In this note we will assume the following formulae, whose derivation is presented in STATISTICAL THERMAL PHYSICS: BASIC IDEAS:

$$\begin{aligned}
&\text{The Boltzmann distribution} & p_i &= \frac{1}{Z} e^{-\beta\epsilon_i} && (1)\\
&\text{The partition function} & Z &= \sum_i e^{-\beta\epsilon_i} && (2)\\
& & &= \mathrm{Tr}(e^{-\beta H}) && (3)\\
&\text{Internal energy} & U &= -\frac{\partial \ln Z}{\partial \beta} = -\frac{1}{Z}\frac{\partial Z}{\partial \beta} && (4)\\
& & &= k_\mathrm{B} T^2 \frac{\partial \ln Z}{\partial T} && (5)\\
&\text{Temperature} & T &= \frac{1}{k_\mathrm{B}\beta}, \qquad \beta = \frac{1}{k_\mathrm{B} T} && (6)\\
&\text{Helmholtz function} & F &= -k_\mathrm{B} T \ln Z && (7)\\
&\text{Entropy} & S &= -\frac{\partial F}{\partial T} && (8)\\
& & &= \frac{U}{T} + k_\mathrm{B} \ln Z && (9)\\
& & &= k_\mathrm{B}\left(\ln Z - \beta \frac{\partial \ln Z}{\partial \beta}\right) && (10)\\
&\text{Heat capacity} & C &= T\frac{\partial S}{\partial T} = \frac{\partial U}{\partial T} && (11)\\
&\text{Energy variance} & \Delta U^2 &= k_\mathrm{B} T^2 C && (12)
\end{aligned}$$

All the formulae apply to a system of any size (large or small, many parts or just one) in a condition of thermal equilibrium. The quantity U here is the average internal energy (averaged over time) and F is the average Helmholtz function. The latter may also be called *free energy.* In the expressions for the Boltzmann distribution and partition function, i labels individual quantum states (microstates) of the system in question.

The partial derivatives which appear in eqns (4), (8), (10), (11) are all at constant *energy levels.* That is, the derivative concerns a change where all the energy levels and quantum states of the system are unchanged, but energy may be entering or leaving the system in the form of heat.

In the case of a composite system made of N weakly interacting and distinguishable parts, one finds

$$Z = Z_1^N \tag{13}$$

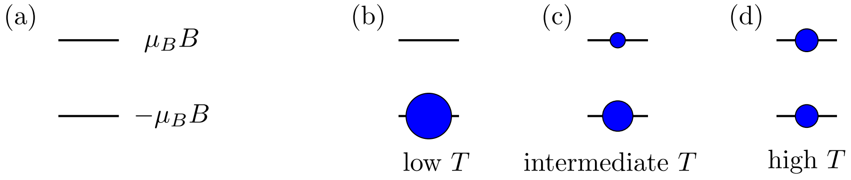

Figure 1: (a): A system with two quantum states of energies $\pm\mu_b B$. (b),(c),(d): the distribution of probability at three temperatures.

where Z_1 is the partition function of one of the parts. This relationship is very useful in practice and we will make much use of it. Here the term *distinguishable* is a subtle quantum-mechanical concept involving the difference between bosons and fermions. In order to be distinguishable it is sufficient if two parts are located at different places such that their position wavefunctions do not overlap. This is the case, for example, for atoms or molecules in a solid in ordinary circumstances, but not for atoms or molecules in a gas. For a gas all the formulae (1)–(12) remain valid but (13) does not.

1 Two-state system

When a particle (such as a single atom or a molecule) having magnetic dipole moment $\boldsymbol{\mu}$ interacts with an applied magnetic field $\mathbf{B}$, the interaction energy is given by

$$\hat{H} = -\hat{\boldsymbol{\mu}} \cdot \mathbf{B}. \tag{14}$$

It usually happens that the magnetic dipole moment is proportional to the internal angular momentum $\mathbf{J}$ of the particle:

$$\hat{\boldsymbol{\mu}} = -g_j \mu_{\mathrm{B}} (\hat{\mathbf{J}}/\hbar) \tag{15}$$

where g_j is a dimensionless constant which depends on the type of particle, and μ_{B} is a standard unit of magnetic dipole moment, called the *Bohr magneton*:

$$\mu_{\mathrm{B}} = \frac{e\hbar}{2m_{\mathrm{e}}} \tag{16}$$

in which m_{e} is the mass of an electron. Hence we find

$$\hat{H} = g_j \mu_{\mathrm{B}} B (\hat{J}_z/\hbar) \tag{17}$$

where we have taken the z axis along the direction of the magnetic field. For this Hamiltonian the energy eigenstates are the eigenstates of the operator $\hat{J}_z$. In the simplest case, such as a single electron, or any atom with total angular momentum one-half, there are just two states, with energies $\pm\mu_{\mathrm{B}} B$ (c.f. Fig. 1).

Using the standard formulae (2)–(11) one finds

$$\begin{aligned}
Z_1 &= 2\cosh x, && (18)\\
F_1 &= -k_{\mathrm{B}} T \ln(2\cosh x), && (19)\\
U_1 &= -\mu_{\mathrm{B}} B \tanh x, && (20)\\
C_B &= k_{\mathrm{B}} (x/\cosh x)^2, && (21)\\
S_1 &= k_{\mathrm{B}} (\ln 2 + \ln(\cosh x) - x\tanh x). && (22)
\end{aligned}$$

where $x = \beta\mu_{\mathrm{B}} B = \mu_{\mathrm{B}} B / k_{\mathrm{B}} T$ and the subscript 1 reminds us we have treated a single particle. The subscript on C_B is to remind us that this is the heat capacity at constant magnetic field.

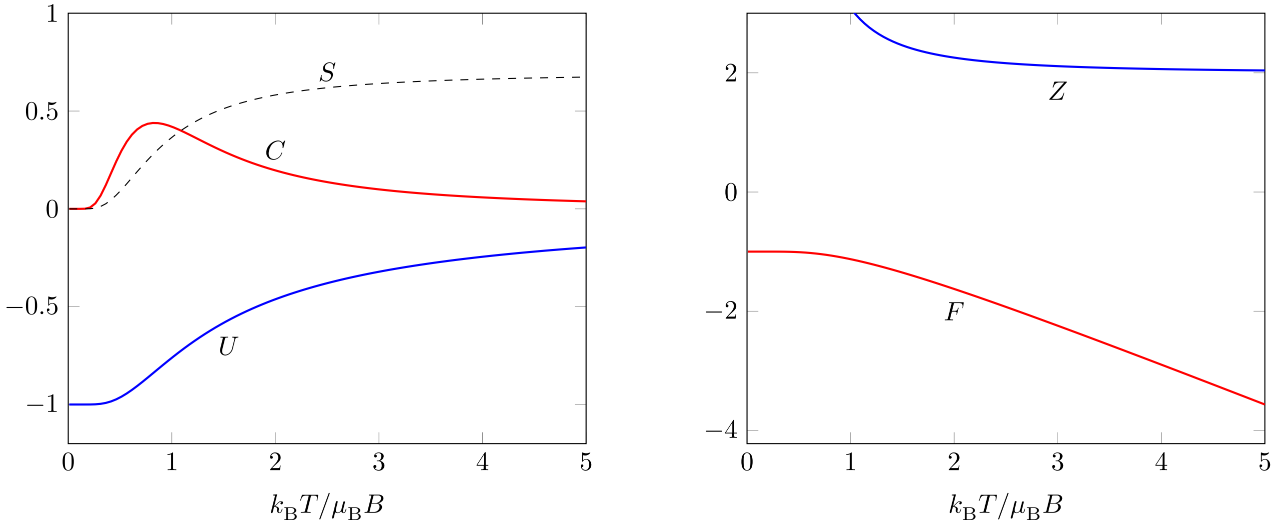

Figure 2: (a) Entropy, heat capacity and internal energy for the two-state system, as a function of temperature. S and C are shown in units of $k_{\rm B}$, U is in units of $\mu_{\rm B}B$. (b) Partition function and Helmholtz function. Z is dimensionless, F is shown in units of $\mu_{\rm B}B$. The behaviour of F can be understood from plot (a), by noting that $F = U - TS$.

It is instructive to examine all these results in detail: see Fig. 2. At low temperature, $U \to -\mu_{\rm B}B$. This is because the dipole is then in its ground state. At high temperature $U \to 0$ because then the dipole is equally likely to be in either state, so the average energy is zero. C is non-zero whenever U is changing with temperature. The peak value of C is of the order of $k_{\rm B}$ (the reader can find the precise number.) Note also that both C and the entropy S tend to zero at low temperature. At high temperature S tends to $k_{\rm B}\ln 2$. This is an example of Boltzmann's formula $S = k_{\rm B}\ln W$ because the dipole is then equally likely to be in either state and there are two states.

Notice also that the temperature where the heat capacity has a peak is such that $k_{\rm B}T$ is of the order of $\mu_{\rm B}B$. It follows that if one were to measure the heat capacity as a function of temperature at a known magnetic field, one could use the experimental results to deduce the value of $\mu_{\rm B}$. More generally, for other examples of two-state systems, one would thus determine the *energy gap*: the difference between the ground state energy and the excited state energy.

Figure 2 also shows the partition function and the Helmholtz function. For any system we can write

$$\begin{aligned} Z &= e^{-\beta\epsilon_0} + g_1 e^{-\beta\epsilon_1} + \cdots && (23)\\ &= e^{-\beta\epsilon_0}\left(1 + g_1 e^{-\beta\Delta E} + \cdots\right) && (24)\end{aligned}$$

where

$$\Delta E = \epsilon_1 - \epsilon_0 \qquad (25)$$

is the **energy gap**, ϵ_0 is the ground state energy and ϵ_1 is the lowest excited state energy, g_1 is the degeneracy of that energy level (i.e. how many states have energy ϵ_1) and we have assumed the ground state is non-degenerate (i.e. only one state has energy ϵ_0). At low temperature $\beta \to \infty$ and therefore Z tends to $+\infty$ if ϵ_0 is negative (as here) and to $-\infty$ if ϵ_0 is positive. At a temperature where the thermal energy $k_{\rm B}T$ is equal to the energy gap ΔE, Z is of order 1. At high temperature $\beta \to 0$ and then Z tends to the total number of states, i.e. 2 in the case of the two-state system.

All these observations for the two-state system have natural generalisations to systems having a larger number of states. In this way the two-state system gives a useful introduction to many aspects of thermal physics, especially the way macroscopic heat capacity measurements can be used to infer information about the microscopic quantum structure of the system.

Let us now consider a *spin-half paramagnetic solid*. This is a solid containing a large number N of magnetic particles of the type under discussion, weakly interacting. The system has a total dipole moment which we shall call $\mathbf{m}$. It is given by

$$\mathbf{m} = N\bar{\boldsymbol{\mu}}. \tag{26}$$

The fundamental thermodynamic relation for this system is

$$\mathrm{d}U = T\mathrm{d}S - \mathbf{m}\cdot\mathrm{d}\mathbf{B} \tag{27}$$

where $U = NU_1$ and $S = NS_1$. The Helmholtz function $F = U - TS$ then satisfies

$$\begin{aligned} \mathrm{d}F &= -S\mathrm{d}T - \mathbf{m}\cdot\mathrm{d}\mathbf{B} \\ &= -S\mathrm{d}T - m\mathrm{d}B. \end{aligned} \tag{28}$$

It follows that

$$\begin{aligned} m &= -\left.\frac{\partial F}{\partial B}\right|_T = -N\left.\frac{\partial F_1}{\partial B}\right|_T && (29)\\ &= N\mu_\mathrm{B}\tanh x && (30)\\ &= N\mu_\mathrm{B}\tanh(\mu_\mathrm{B}B/k_\mathrm{B}T). && (31) \end{aligned}$$

This is called the *equation of state*: it relates the thermodynamic state variables m and B to the temperature T. Observe that at low temperatures or high magnetic field m tends to $N\mu_\mathrm{B}$: the dipoles all line up with the field. At high temperatures or low magnetic field, $m \to 0$: the dipoles are randomly oriented. At temperatures small compared to $\mu_\mathrm{B}B/k_\mathrm{B}$ the relationship between m and B is linear, and we have

$$m \simeq \left(\frac{N\mu_\mathrm{B}^2}{k_\mathrm{B}}\right)\frac{B}{T}. \tag{32}$$

This is often written

$$M = \chi H \tag{33}$$

where the magnetisation $M = m/V$ is the dipole moment per unit volume, the field strength $H = B/\mu_0$ (with μ_0 the permeability of free space)[1] and

$$\chi = \left(\frac{N\mu_0\mu_\mathrm{B}^2}{Vk_\mathrm{B}}\right)\frac{1}{T}. \tag{34}$$

This χ is called the *magnetic susceptibility*. It is dimensionless, so one can infer that the quantity in the bracket in (34) has the dimensions of a temperature (which you may like to check). The general form $\chi \propto 1/T$ is known as **Curie's law** and it describes many magnetic systems quite well at intermediate temperatures.

A further noteworthy aspect of the two-state system is that the heat capacity at constant magnetization is zero. This means that if, in some process, the magnetization does not change, then neither does the entropy. One can deduce this from the fact that both the entropy (22) and the average dipole moment (30) are functions of a single variable, and it is the same variable for both of them: x. Therefore, for this system,

$$\text{constant } m \Leftrightarrow \text{constant } x \Leftrightarrow \text{constant } S \tag{35}$$

Another useful way to deduce this is to note that when there are only two states the entropy can be written as a function of a single probability (since the other is given by normalization) and so can the average dipole moment.

Let us revisit now the thermodynamic relation (27) and compare it with the hamiltonian (14). From the latter we have

$$\mathrm{d}\langle\hat{H}\rangle = -\mathbf{B}\cdot\mathrm{d}\langle\hat{\boldsymbol{\mu}}\rangle - \langle\hat{\boldsymbol{\mu}}\rangle\cdot\mathrm{d}\mathbf{B} \tag{36}$$

---

[1] Note that we here defined H to be the field strength associated with B in the absence of magnetization; this is valid when $\chi \ll 1$.

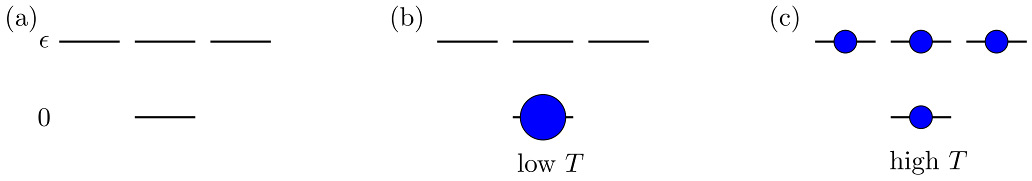

Figure 3: (a): A system with a single ground state and g excited states all of the same energy, shown for the example case $g = 3$. (b),(c): the distribution of probability at low and high temperatures.

and therefore

$$\mathrm{d}U = -\mathbf{B}\cdot\mathrm{d}\mathbf{m} - \mathbf{m}\cdot\mathrm{d}\mathbf{B}. \tag{37}$$

By comparing this with (27) we can deduce

$$T\mathrm{d}S = -\mathbf{B}\cdot\mathrm{d}\mathbf{m} \tag{38}$$

(which the reader can confirm by examining (22) and (30)). This is another way to see that entropy and magnetisation are intimately related for paramagnetic systems. It also confirms that $-\mathbf{m}\cdot\mathrm{d}\mathbf{B}$ is the correct expression for magnetic work.

2 Two energy levels with degeneracy

Our next interesting system is shown in Fig. 3. It has two energy levels and $g+1$ energy eigenstates: the ground state (a single state) of energy zero, and g states each having energy ϵ. One finds (exercise for the reader)

$$\begin{aligned} Z &= 1 + g e^{-x} & (39)\\ U &= \frac{g\epsilon}{g+e^{x}} & (40)\\ C &= k_{\mathrm{B}}\frac{g x^{2} e^{x}}{(g+e^{x})^{2}} & (41)\end{aligned}$$

where $x = \beta\epsilon = \epsilon/k_{\mathrm{B}}T$. What do you expect for the value of U at low and high temperature? Is this born out by (40)? What do you expect for the values of Z and S at high temperature? Is this born out by (9)?

Fig. 4 shows the heat capacity of this system for three example values of g, as a function of x. Since the curves are all different, it follows that one can, in principle, determine both the energy gap ϵ and the degeneracy g by measuring the heat capacity. Also, Fig. 4(b) illustrates an interesting universality: at low temperatures (so high x), the heat capacity tends to

Low temperature heat capacity

$$C \to k_{\mathrm{B}} g x^{2} e^{-x} \qquad \text{at low } T \text{ (high } x = \epsilon/k_{\mathrm{B}}T) \tag{42}$$

and therefore $C(x)/g$ *has the same functional form, independent of* g. All systems with an energy gap (i.e. a non-negligible difference in energy between the ground state and first excited state) will be described, at low enough temperature, by just the ground and first excited energies, because the higher states will not be populated. Therefore the behaviour shown in Fig. 4(b) and (42) will be observed *for any system at all* having a non-negligible energy gap, if g is the number of states at or very near the first excited state energy.

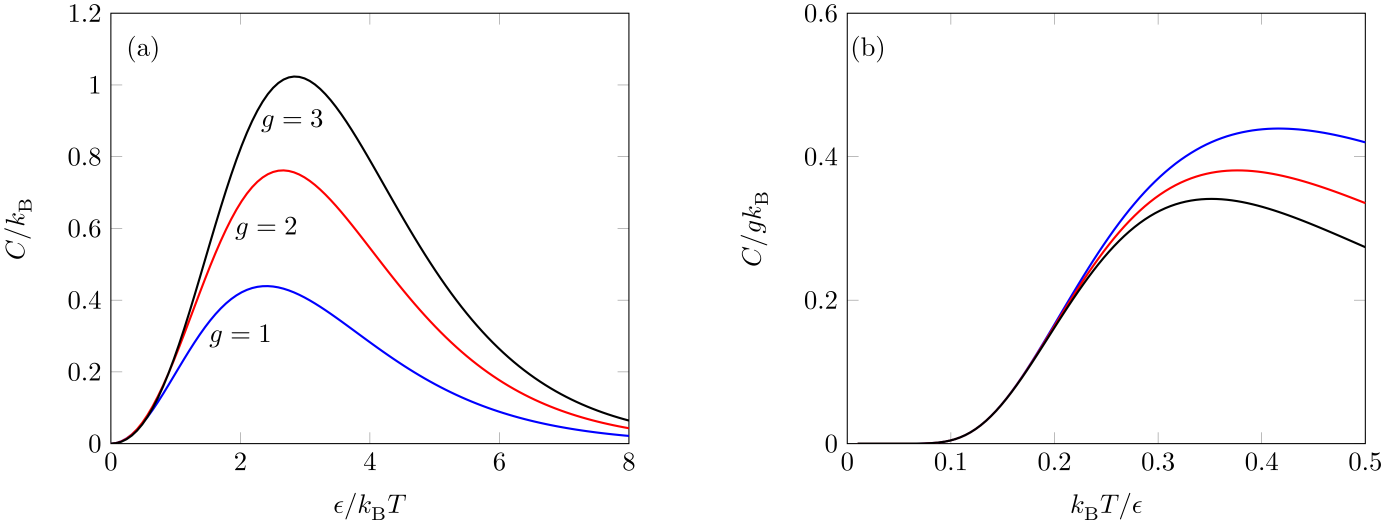

Figure 4: (a) Heat capacity of the system shown in Fig. 3 for three example values of g. (b) The same heat capacity, but now in units of gk_B, and plotted as a function of temperature not inverse temperature, in order to bring out the low-temperature behaviour.

3 Simple harmonic oscillator

The one-dimensional simple harmonic oscillator has Hamiltonian

$$\hat{H} = \frac{\hat{p}^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2. \tag{43}$$

The energy eigenvalues are $(n+1/2)\hbar\omega$ for non-negative integer values of n, and there is one state at each energy—see Fig. 5. One finds (exercise for the reader)

$$\begin{aligned} Z &= \sum_{n=0}^{\infty} e^{-\beta(n+1/2)\hbar\omega} = e^{-\beta\hbar\omega/2}\frac{1}{1-e^{-\beta\hbar\omega}}, && (44)\\ U &= \frac{\hbar\omega}{2} + \frac{\hbar\omega}{e^{\beta\hbar\omega}-1} && (45)\\ C &= k_\mathrm{B}\frac{x^2e^x}{(e^x-1)^2} && (46) \end{aligned}$$

where $x = \beta\hbar\omega$. At low temperature the heat capacity shows the universal behaviour already remarked in connection with (11). At high temperature the heat capacity tends to k_B. Note that this result is independent of ω. It is the same for all harmonic oscillators! This is an example of the *equipartition theorem* which is presented in another note (title: DERIVATION OF THE EQUIPARTITION THEOREM).

The energy U includes a constant part $\hbar\omega/2$ (the ground state energy) and a temperature-dependent part. The constant part is sometimes called *zero-point energy.* One can always set it to zero by shifting the zero of potential energy, so that all the energy levels move down by $\hbar\omega/2$. Whether or not one does this, the heat capacity only depends on the temperature-dependent part in any case.

The average degree of excitation can be obtained from

$$\bar{n} = \frac{U-(1/2)\hbar\omega}{\hbar\omega} = \frac{1}{e^{\beta\hbar\omega}-1}. \tag{47}$$

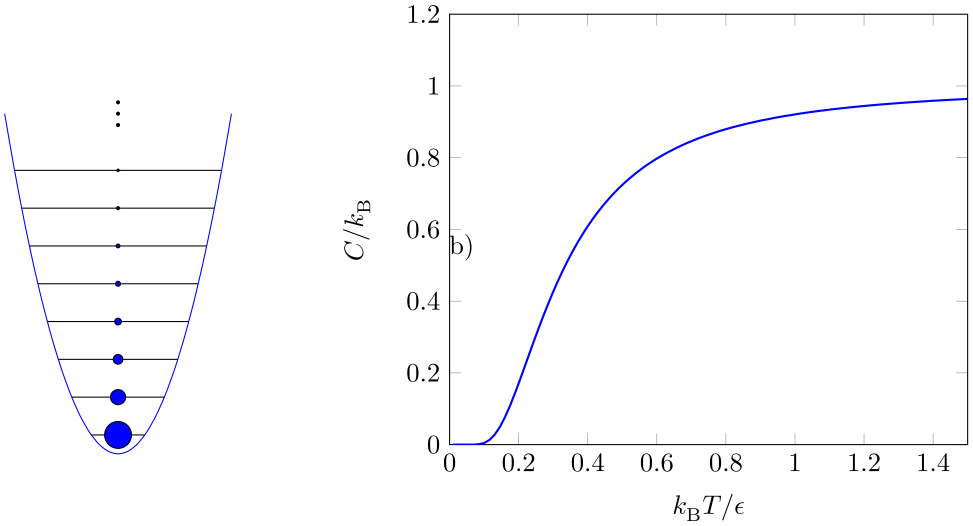

Figure 5: (a) Energy level diagram for a one-dimensional simple harmonic oscillator, with an example probability distribution. (b) Heat capacity.

This is also called the *mean vibrational quantum number*. At high temperature it evaluates to

$$\bar{n} \to \frac{k_\mathrm{B}T}{\hbar\omega} \tag{48}$$

(and Z tends to this same value).

4 Multiple harmonic oscillators

A set of N (weakly interacting, distinguishable) one-dimensional harmonic oscillators all of the same frequency is easily treated: one simply gets $U = NU_1$, $S = NS_1$, etc. An example occurs when a particle is situated in a harmonic potential well in 2 or 3 dimensions. For the 3-dimensional case, for example, the Hamiltonian is

$$\begin{aligned} \hat{H} &= \frac{\hat{p}_x^2}{2m} + \frac{\hat{p}_y^2}{2m} + \frac{\hat{p}_z^2}{2m} + \frac{1}{2}m\omega^2\left(\hat{x}^2 + \hat{y}^2 + \hat{z}^2\right) && (49) \\ &= \left[\frac{\hat{p}_x^2}{2m} + \frac{1}{2}m\omega^2\hat{x}^2\right] + \left[\frac{\hat{p}_y^2}{2m} + \frac{1}{2}m\omega^2\hat{y}^2\right] + \left[\frac{\hat{p}_z^2}{2m} + \frac{1}{2}m\omega^2\hat{z}^2\right] && (50) \end{aligned}$$

where we treated a spherically symmetric well, and the second version brings out the fact that $\hat{H}$ is a sum of three terms, each of which describes a one-dimensional harmonic oscillator. *It follows that the symmetric three-dimensional harmonic oscillator has precisely the same thermal behaviour as a set of 3 one-dimensional harmonic oscillators*, as long as some small asymmetry exists, sufficient to allow the three motions to exchange energy but without significantly affecting the energy levels.

The quantum states of the three-dimensional oscillator can be labelled by three non-negative integers n_x, n_y, n_z, such that the energy of the quantum state specified by $\{n_x, n_y, n_z\}$ is

$$(n_x + 1/2)\hbar\omega + (n_y + 1/2)\hbar\omega + (n_z + 1/2)\hbar\omega = (n + 3/2)\hbar\omega \tag{51}$$

where

$$n = n_x + n_y + n_z \tag{52}$$

Note, there is one quantum state for each $\{n_x, n_y, n_z\}$ and this means there is more than one quantum state for each n when $n > 0$. The reader is invited to confirm to their own satisfaction that the number of states at any given n is given by the triangular numbers $(1, 3, 6, 10, \cdots)$. Therefore the degeneracy of the energy level specified by n is

$$g_n = \frac{(n+1)(n+2)}{2}. \tag{53}$$

It follows that the partition function can be written

$$Z = \sum_i e^{-\beta\epsilon_i} = \sum_{n=0}^{\infty} \frac{(n+1)(n+2)}{2} e^{-\beta(n+3/2)\hbar\omega} \tag{54}$$

where the first version has been given in order to be clear that this is the standard sum over states, and the second version shows how that sum can be written by gathering the terms at each energy.

Eqn (54) is correct but it is not the most convenient way to find the partition function for this system. For didn't we just say that the system can be treated as a set of 3 one-dimensional oscillators, whose partition function we already know? We have displayed (54) merely in order to mention the following mathematical fact:

$$\sum_{n=0}^{\infty} \frac{(n+1)(n+2)}{2} r^n = \left(\sum_{n=0}^{\infty} r^n \right)^3 \tag{55}$$

By setting $r = \exp(-\beta\hbar\omega)$, and using $Z = Z_1^3$ for the composite system, this is another way to prove the equivalence of the 3d harmonic oscillator to a set of three 1d harmonic oscillators.

Let us now comment on the thermodynamic relation

$$\mathrm{d}U = T\mathrm{d}S + đW \tag{56}$$

where the second term is the work supplied. How is work done on a collection of harmonic oscillators? To answer this question, observe that in an **adiabatic** (i.e. constant entropy) process we have $đW = \mathrm{d}U$. What does a constant entropy process look like in a statistical treatment? Recall the definition of entropy: constant entropy means constant probabilities. An adiabatic process is one which shifts the energies of the quantum states, and may also gradually change the wavefunctions themselves, but without changing the distribution of probability over the states. In the present case this will happen if the potential well is gradually compressed or expanded, causing the oscillation frequency ω to change, without allowing heat to enter or leave. To calculate the effect of such a change the easy way is to start from

$$U = \sum_i \epsilon_i p_i \tag{57}$$

and then one obtains

$$\left.\frac{\partial U}{\partial \omega}\right|_S = \sum_i \frac{\partial \epsilon_i}{\partial \omega} p_i. \tag{58}$$

Whether for a single or multiple harmonic oscillators, $\partial\epsilon_i/\partial\omega = \epsilon_i/\omega$ so we deduce

$$\left.\frac{\partial U}{\partial \omega}\right|_S = \frac{U}{\omega}. \tag{59}$$

It is easy to generalise this:

If the energy levels of some system are all proportional to a parameter X, independent of their occupation probabilities, then

$$\left.\frac{\partial U}{\partial X}\right|_S = \frac{U}{X} \tag{60}$$

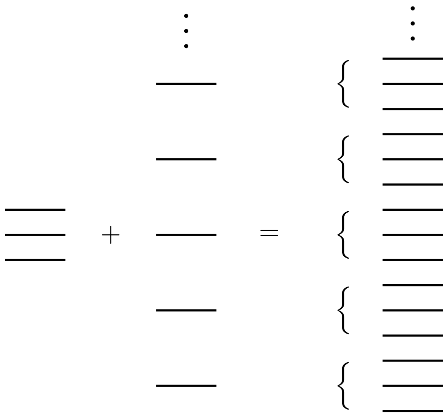

Figure 6: Equivalence between two different systems. On the left we show the energy-level structure of a magnetic system and a harmonic oscillator. For a system with both properties we must add the energies and then we get the energy-level structure shown on the right. The latter is the same as that of a single harmonic oscillator. This leads to results such as (65).

5 Magnetic system with $J > 1/2$

We return now to the magnetic behaviour described by (14) and 15, but now with angular momentum quantum number J larger than 1/2. In general J can take either a half-integer or an integer value, depending on the type of atom or molecule or whatever. For a given J the eigenvalues of the operator $\hat{J}_z$ are given by

$$M_J \hbar, \qquad M_J \in \{-J, -J+1, \cdots, J-1, J\}. \tag{61}$$

For example, for $J = 1$, the possible values of M_J are $\{-1, 0, 1\}$. For $J = 3/2$ the possible values of M_J are $\{-3/2, -1/2, 1/2, 3/2\}$. For any given J there are, therefore, $(2J + 1)$ possible values of M_J. Since the magnetic energy is proportional to M_J this leads to $(2J + 1)$ energy levels, arranged as an evenly-spaced ladder. The partition function is therefore

$$Z = \sum_{M=-J}^{J} e^{-Mx} = \frac{\sinh((J + 1/2)x)}{\sinh(x/2)} \tag{62}$$

where $x = \beta g_j \mu_B B$. The task to deduce internal energy, entropy, heat capacity and average magnetic dipole moment is left to the reader.

There are two interesting connections between the magnetic system and the harmonic oscillator:

1. When the temperature is low enough that almost all the probability is in the lowest $(2J + 1)$ states of a given harmonic oscillator, then the behaviour of the oscillator and that of a magnetic system with the same energy level spacing agree. For example, they will give the same heat capacity, as a function of temperature, at low temperature.
2. Let $\epsilon = g_j \mu_B B$ be the energy level spacing of a magnetic system. Then a system which is both magnetic and simple harmonic (e.g. a magnetic particle vibrating in a well, in a uniform **B** field) with frequency $\omega = (2J+1)\epsilon/\hbar$ will behave in the same way as a non-magnetic harmonic oscillator of frequency $\omega = \epsilon/\hbar$.

The first of these should be obvious. The second can be shown by inspecting Fig. 6. From the equality of the energy level diagrams, one finds immediately that

$$Z_{\text{shm}}(\epsilon) = Z_{\text{mag}}(J, \epsilon)\, Z_{\text{shm}}\left((2J + 1)\epsilon\right) \tag{63}$$

hence

$$\ln Z_{\text{mag}} = \ln\left[Z_{\text{shm}}(\epsilon)\right] - \ln\left[Z_{\text{shm}}((2J+1)\epsilon)\right]. \tag{64}$$

It follows that one can obtain the magnetic heat capacity, for example, by using

$$C_{\text{mag}} = C_{\text{shm}}(\epsilon) - C_{\text{shm}}((2J+1)\epsilon). \tag{65}$$

The purpose of this observation is not to suggest this method is any easier, from a mathematical point of view, but to offer further intuition about the behaviour of the heat capacity. By inspecting Fig. 5 and Eqn (65) we can deduce that the heat capacity will have a single broad peak, rising to about k_B at $T \sim 0.3\epsilon/k_B$ and falling back to zero at $T \sim (2J+1)0.3\epsilon/k_B$.

6 Rigid Rotator

A rigid rotator is a rigid body that can rotate about one or more axes. We will treat the case where the body is free to rotate in three dimensions. In classical mechanics the rotational energy (i.e. the kinetic energy associated with rotation) is given by

$$H = \frac{L_1^2}{2I_1} + \frac{L_2^2}{2I_2} + \frac{L_3^2}{2I_3} \tag{66}$$

where I_1, I_2, I_3 are the three principle moments of inertia, and $\mathbf{L}_1, \mathbf{L}_2, \mathbf{L}_3$ are the components of angular momentum about the principle axes.

The quantum mechanics of rotation is considerably complicated by the fact that the three components of the angular momentum operator do not commute. For this reason we will not treat a general rotator here, but restrict attention to two cases: a body with an axis of rotational symmetry (e.g. a cylinder or a dumbbell), and a body with no such axis but with all three moments of inertia the same, $I_1 = I_2 = I_3$ (e.g. a cube, a tetrahedron). A diatomic molecule is very well approximated by the first case; a small molecule with no axis of rotational symmetry (such as CH_4 or H_2O) is reasonably well approximated by the second case. One finds that for both these cases the energy levels are given by

$$E_L = \frac{L(L+1)\hbar^2}{2I} \tag{67}$$

where I is the moment of inertia, the allowed values of L are $L = 0, 1, 2, \cdots$ (with an exception to be described shortly) and the degeneracies are given by

$$g = \left\{ \begin{array}{cc} (2L+1) & \text{with axis of rotational symmetry} \\ (2L+1)^2 & \text{no axis of rotational symmetry} \end{array} \right. \tag{68}$$

The first case (a system with a symmetry axis) is treated in introductory quantum mechanics texts; it is connected with Laplace's equation and the spherical harmonic functions $Y_m^l(\beta, \phi)$. The second case is not normally treated in introductory texts. It is instructive to note that although the two cases have the same Hamiltonian and energy levels, they have different sets of quantum states, leading to different degeneracies, with more states in the second case. It is because in the second case there are three angles (e.g. the Euler angles) required to specify the orientation in space of the body, whereas in the first case only two are needed because of the axial symmetry.

We mentioned that there was an exception to the rule $L = 0, 1, 2, \cdots$. This concerns the case where the rotator is made of identical particles, such as, for example, the molecule H_2 or N_2. This introduces a restriction on the values of L such that in some cases only even values are allowed, in some cases only odd; the discussion of this is postponed to section 6.2. (This issues does not arise in the treatment of the electron wavefunctions in the hydrogen atom because the electron is not identical to the proton.)

When we apply the theory of the rigid rotator to molecules, it is the mass distribution not the charge distribution that concerns us. The mass is almost entirely concentrated in the atomic nuclei, and the radius of a nucleus is about 10^5

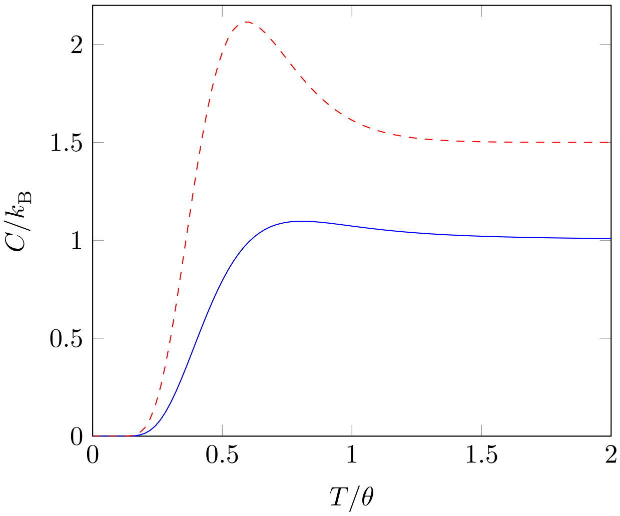

Figure 7: Heat capacity of a 3-dimensional rigid rotator. Full blue curve: with axial symmetry (e.g. a heteronuclear diatomic molecule); dashed red curve: no axis of rotational symmetry and a single moment of inertia (e.g. CH_4).

times smaller than the separation between the nuclei, therefore a diatomic molecule behaves like two point-like masses on the ends of a stick, not two overlapping spheres. The moment of inertia for rotations about any axis perpendicular to the line between the nuclei is accurately given by

$$I = \frac{m_1 m_2}{m_1 + m_2} d^2 \tag{69}$$

where d is the separation of the nuclei.

The partition function of the 3d rigid rotator with axial symmetry and non-identical particles is

$$Z = \sum_{L=0}^{\infty} (2L+1) e^{-L(L+1)x/2} \; = \; 1 + 3e^{-x} + 5e^{-3x} + 7e^{-6x} \cdots \tag{70}$$

where $x = \beta\hbar^2/I$. We defined x this way (rather than using $\beta\hbar^2/2I$ for example) because then $x = \beta\Delta E$ where ΔE is the energy gap. At low temperature the heat capacity is given by the universal equation (11) with $g = 3$.

There is no simple formula for Z valid at all temperatures, but the high temperature behaviour can be obtained as follows. When $k_B T \gg \hbar^2/I$ we have $x \ll 1$ and consequently a large number of terms contribute significantly to the sum forming Z. In this situation the sum is well approximated by an integral:

$$\begin{aligned} Z &= \sum_{L=0}^{\infty} (2L+1) e^{-L(L+1)x/2} \, \delta L \\ &\simeq \int_{L=0}^{\infty} (2L+1) e^{-L(L+1)x/2} \, \mathrm{d}L \\ &= \frac{2}{x} = \frac{T}{\theta} \end{aligned} \tag{71}$$

where we used $\delta L = 1$ in the first line, in order to make the transition to an integral clear, and in the last step we introduced

$$\theta = \frac{\hbar^2}{2Ik_B}. \tag{72}$$

This θ is a characteristic temperature; it gives the value of temperature at which rotational effects become significant.

You can easily show (exercise) that for any partition function having the functional form of a power-law dependence on T:

$$Z(T) = AT^n \tag{73}$$

where A and n do not depend on T, one finds

$$U = nk_{\mathrm{B}}T \tag{74}$$

and therefore $C = nk_{\mathrm{B}}$. In the present case, $n = 1$ so the heat capacity is $C = k_{\mathrm{B}}$ at high temperature. Fig. 7 shows the complete behaviour: C is very small at low T, as we have come to expect, then rises and slightly overshoots its high-temperature value, coming close to the latter at $T \geq 2\theta$.

Also shown on Fig. 7 is the heat capacity of the rigid rotator with no axis of rotational symmetry, for which there are more states at each energy, with the result that C is larger. You can show (exercise) that in the high temperature limit we now have $Z \propto T^{3/2}$ so $C \to (3/2)k_{\mathrm{B}}$.

6.1 Rotation with magnetism: Landau levels

A system with both paramagnetic and rotational energy will have energy levels given by

$$E(L, M_J) = g\mu_{\mathrm{B}}BM_J + L(L+1)\frac{\hbar^2}{2I} \tag{75}$$

An interesting case arises when the magnetic quantum number M_J is none other than the quantum number M_L associated with the rotation. In this case the $(2L+1)$ quantum states associated with a given value of L are precisely the ones referred to in the magnetic part, with energies $g\mu_{\mathrm{B}}BM_L$. This can happen when the rotation and the magnetism are two aspects of a single phenomenon. One then finds that when the magnetic energy levels exactly fill the gaps between the rotational energy levels, the result is a uniform ladder of energy levels with one quantum state at each energy: see Fig. **??**. Therefore in this situation the system is equivalent to a one-dimensional simple harmonic oscillator.

A physical situation related to this, but somewhat different, is the *cyclotron motion* of electrons in a conductor subject to a magnetic field. Equation (75) does not apply directly to cyclotron motion, because (75) treats a case where the moment of inertia is independent of L. To get a physical insight into cyclotron motion, let us develop an expression for the Hamiltonian using classical physics. We have

$$H = \frac{p^2}{2m} - \boldsymbol{\mu} \cdot \mathbf{B} \tag{76}$$

where the magnetic dipole $\boldsymbol{\mu}$ is owing to circular motion of the charged particle. Its magnitude μ is equal to the current in the loop multiplied by the area of the loop:

$$\mu = q\frac{v}{2\pi r}\pi r^2 = \frac{1}{2}q\omega r^2 \tag{77}$$

where r is the radius of the current loop, v the speed of the particle, q is its charge, and $\omega = v/r$. From Newton's second law we have $mv^2/r = qvB$ and therefore $mv = qBr$. Hence the **cyclotron frequency** is

$$\omega = qB/m. \tag{78}$$

It follows that

$$\mu B = \frac{1}{2}m\omega^2 r^2 \tag{79}$$

so the Hamiltonian can be written

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 r^2. \tag{80}$$

The sign is positive because the dipole moment and the B field are anti-aligned (which the reader is invited to confirm). This is the Hamiltonian of simple harmonic motion at frequency ω. In the quantum treatment a similar conclusion holds and one finds equally-spaced energy levels called *Landau levels*.

6.1.1 A more precise analysis

The above treatment using (76) gives a good insight but for a more complete analysis one should set out from

$$\hat{H} = (\hat{\mathbf{p}} - q\hat{\mathbf{A}})^2/2m \tag{81}$$

where $\mathbf{A}$ is the magnetic vector potential and $\hat{\mathbf{p}}$ is the canonical momentum operator. A vector potential $\mathbf{A} = (0, Bx, 0)$ corresponds to a $\mathbf{B}$ field along z. One then has

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{(\hat{p}_y - qBx)^2}{2m} + \frac{\hat{p}_z^2}{2m}. \tag{82}$$

The operators $\hat{p}_y$ and $\hat{p}_z$ commute with this. If we now adopt the basis of eigenstates of $\hat{p}_y$ for the y motion, we can replace $\hat{p}_y$ by its eigenvalue $\hbar k_y$ and then

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{1}{2}m\omega^2\left(x - \frac{\hbar k_y}{m\omega}\right)^2 + \frac{\hat{p}_z^2}{2m} \tag{83}$$

with $\omega = qB/m$ as in (78). This is the Hamiltonian for free motion in the z direction combined with simple harmonic motion in the x direction with an offset which depends on k_y. The energy levels will be $(n+1/2)\hbar\omega + (\hbar k_z)^2/2m$, with a degeneracy given by the number of k_y values. That number is finite if the motion is confined to a finite region, and infinite otherwise.

Eqn (83) treats the x-motion and y-motion differently, which should surprise us since the physical scenario is symmetric about z. However the eigenstates of p_y form a complete basis for motion in the y direction, so we can, in principle, restore the symmetry by a judicious grouping of eigenstates at each energy. Alternatively, one can set up the calculation more symmetrically by adopting $\mathbf{A} = (-By, Bx, 0)/2$ at the start. (The fact that the same $\mathbf{B}$ field can be obtained from two different vector potentials is an example of gauge invariance.)

6.2 Identical nuclei and the saga of hydrogen

Our treatment of rotation, when applied to diatomic molecules, has been restricted so far to molecules such as NO and HCl where the two nuclei are different (they are called heteronuclear). In the case of homonuclear molecules such as H_2 and N_2 another consideration comes into play, going by the technical name *exchange symmetry*. The issue is that when two particles are strictly indistinguishable then, according to quantum theory, the joint states available to them are restricted by an *exchange symmetry requirement*. The restriction can be expressed like this:

$$\text{identical bosons:} \qquad |\psi(1,2)\rangle = |\psi(2,1)\rangle \tag{84}$$

$$\text{identical fermions:} \qquad |\psi(1,2)\rangle = -|\psi(2,1)\rangle \tag{85}$$

Here *bosons* and *fermions* are the two possible types of particle, and the notation $|\psi(2,1)\rangle$ refers to a quantum state in which the first particle is labelled '1' and the second particle is labelled '2'. The requirement shows that when the labels are swapped, the state is either unaffected (bosons) or simply changes sign (fermions). In practice the main result of this requirement is that there are fewer quantum states available to pairs of identical particles than there are to pairs of otherwise similar but non-identical particles.

This affects the rigid rotator because a half-rotation of a diatomic molecule is like swapping the nuclei. To deduce the outcome one must study what happens to the quantum states. One finds:

even L: spatial state unchanged under a reflection through the origin
odd L: spatial state changes sign under a reflection through the origin

To complete the analysis we need to note also the *spin state*. The total quantum state $|\psi(1,2)\rangle$ can often be written as a product of a spatial part and a spin part. The exchange symmetry requirement is then

$$\text{bosons:} \quad |\psi(1,2)\rangle \;=\; \begin{cases} \text{either} & |\text{sym}\rangle_{\text{spatial}} \otimes |\text{sym}\rangle_{\text{spin}} \\ \text{or} & |\text{anti-sym}\rangle_{\text{spatial}} \otimes |\text{anti-sym}\rangle_{\text{spin}} \end{cases} \tag{86}$$

$$\text{fermions:} \quad |\psi(1,2)\rangle \;=\; \begin{cases} \text{either} & |\text{sym}\rangle_{\text{spatial}} \otimes |\text{anti-sym}\rangle_{\text{spin}} \\ \text{or} & |\text{anti-sym}\rangle_{\text{spatial}} \otimes |\text{sym}\rangle_{\text{spin}} \end{cases} \tag{87}$$

where *sym* and *anti-sym* refer to symmetric and anti-symmetric under exchange of labels. This is beginning to look complicated, but it is not. A simple example will help. We consider the case of hydrogen, where the two nuclei are simply protons. Each is a fermion with spin half. There is one antisymmetric joint spin state, called the singlet, and there are three symmetric spin states, together called a triplet. Therefore the possible states of rotation and spin are:

$$|\text{even } L, \text{ spin singlet}\rangle \qquad \text{or } |\text{odd } L, \text{ spin triplet}\rangle \tag{88}$$

The spin singlet case is called *parahydrogen*; the spin triplet case is called *orthohydrogen.* I don't find the names to be intuitive, but it is good to have names because they invite us to consider hydrogen molecules to come in these two versions. We can then say that in a case where the molecules are all of one type, and no nuclear spin state changes are occurring, then the rotational partition function is one of

$$Z_{\text{para}} = \sum_{\text{even } L} (2L+1) e^{-L(L+1)x/2} \tag{89}$$

$$Z_{\text{ortho}} = 3 \sum_{\text{odd } L} (2L+1) e^{-L(L+1)x/2} \tag{90}$$

in which the factor 3 accounts for the three different nuclear spin states for the triplet. Notice that the ground state for orthohydrogen occurs at $L = 1$ and therefore it has an energy substantially higher than that of the ground state of parahydrogen.

The reason this subsection has the title 'saga of hydrogen' is that it took much work and many attempts in the early twentieth century to make sense of the experimental data. This is because there are not just two possibilities but four! The multitude arises because changes of nuclear spin state are quite unlikely when two hydrogen molecules collide, so the system does not readily relax to an overall thermal equilibrium by that process. If the system started out as parahydrogen, then that is what it remains. But over long enough times, and especially in the presence of a catalyst such as activated charcoal, transitions between para and ortho can occur, and then the system reaches its true equilibrium. The four main possibilities are

1. Parahydrogen with no catalyst: $Z_1 = Z_{\text{para}}$.
2. Orthohydrogen with no catalyst: $Z_1 = Z_{\text{ortho}}$.
3. In the presence of a catalyst, or over very long times, equilibrium: $Z_1 = Z_{\text{para}} + Z_{\text{ortho}}$
4. Parahydrogen and orthohydrogen in a mixture in the proportions 1:3 (and those proportions staying fixed, i.e. no nuclear spin flips): $Z_1 = Z_{\text{para}} Z_{\text{ortho}}^3$, with $Z = Z_1^{N/4}$ and consequently $C = (C_{\text{para}} + 3C_{\text{ortho}})/4$

This leads to the four curves shown in Fig. 8.

A similar analysis will apply to any diatomic molecule with identical spin-half nuclei, but hydrogen is the only one where the behaviour is easily observed, because for other gases the rotational temperature is below the boiling point (at atmospheric pressure) so the low-temperature rotational behaviour is taking place in liquid state. For nuclear-spin zero (e.g. O_2), the nuclei are bosons and there is only one spin state, which is symmetric, so there is only one case: a sum over even L as in Z_{para}. For nuclear-spin one the nuclei are bosons again but now there are more joint spin states (some symmetric, some anti-symmetric) and consequently a greater range of possibilities.

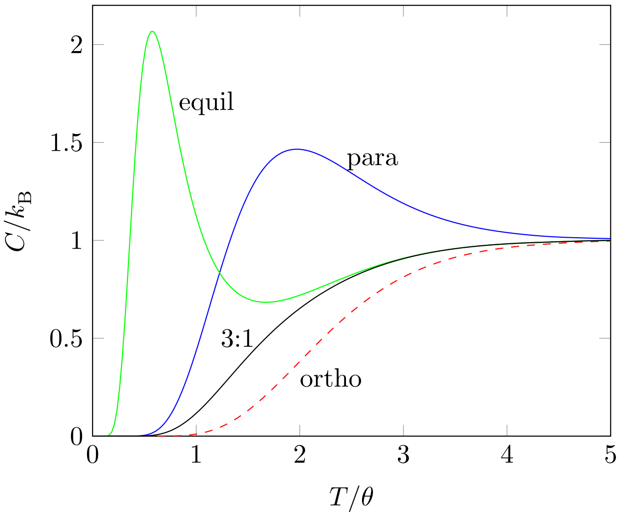

Figure 8: Rotational heat capacity of hydrogen (or, more generally, a diatomic molecule with identical spin-half nuclei). The 3:1 mixture is the case observed in normal circumstances.

7 Exercises

1. Confirm that (20) and (40) agree when $g = 1$ (caution: pay attention to the different definitions of x).
2. Taking the heat capacity C_B from (21), evaluate $\int_0^\infty (C_B/T)\,\mathrm{d}T$ and comment.
3. Prove that the entropy always go to zero at low temperature if the ground state is non-degenerate.
4. A certain system has 20 energy levels, all non-degenerate, distributed randomly between 0 and 0.01 eV, with the lowest two at 0 and 10^{-3} eV. Sketch the heat capacity C as a function temperature. Also sketch the function C/T as a function of temperature. What is the area under the second curve?