Contents

List of Tables

List of Figures

Chapter 4

Statistical Ensembles

4.1 References

- F. Reif, *Fundamentals of Statistical and Thermal Physics* (McGraw-Hill, 1987)
  This has been perhaps the most popular undergraduate text since it first appeared in 1967, and with good reason.
- A. H. Carter, *Classical and Statistical Thermodynamics*
  (Benjamin Cummings, 2000)
  A very relaxed treatment appropriate for undergraduate physics majors.
- D. V. Schroeder, *An Introduction to Thermal Physics* (Addison-Wesley, 2000)
  This is the best undergraduate thermodynamics book I've come across, but only 40% of the book treats statistical mechanics.
- C. Kittel, *Elementary Statistical Physics* (Dover, 2004)
  Remarkably crisp, though dated, this text is organized as a series of brief discussions of key concepts and examples. Published by Dover, so you can't beat the price.
- M. Kardar, *Statistical Physics of Particles* (Cambridge, 2007)
  A superb modern text, with many insightful presentations of key concepts.
- M. Plischke and B. Bergersen, *Equilibrium Statistical Physics* ($3^{\rm rd}$ edition, World Scientific, 2006)
  An excellent graduate level text. Less insightful than Kardar but still a good modern treatment of the subject. Good discussion of mean field theory.
- E. M. Lifshitz and L. P. Pitaevskii, *Statistical Physics* (part I, $3^{\rm rd}$ edition, Pergamon, 1980)
  This is volume 5 in the famous Landau and Lifshitz *Course of Theoretical Physics*. Though dated, it still contains a wealth of information and physical insight.

d

4.2 Microcanonical Ensemble (μCE)

4.2.1 The microcanonical distribution function

We have seen how in an *ergodic* dynamical system, time averages can be replaced by phase space averages:

$$\text{ergodicity} \quad \Longleftrightarrow \quad \big\langle f(\boldsymbol{\varphi})\big\rangle_t = \big\langle f(\boldsymbol{\varphi})\big\rangle_S \quad , \tag{4.1}$$

where

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_t = \lim_{T\to\infty} \frac{1}{T}\int\limits_0^T dt\, f\big(\boldsymbol{\varphi}(t)\big) \quad . \tag{4.2}$$

and

$$\big\langle f(\boldsymbol{\varphi})\big\rangle_S = \int d\mu\, f(\boldsymbol{\varphi})\,\delta\big(E-\hat{H}(\boldsymbol{\varphi})\big) \Big/ \int d\mu\,\delta\big(E-\hat{H}(\boldsymbol{\varphi})\big) \quad . \tag{4.3}$$

Here $\hat{H}(\boldsymbol{\varphi}) = \hat{H}(\boldsymbol{q},\boldsymbol{p})$ is the Hamiltonian, and where $\delta(x)$ is the Dirac δ-function[1]. Thus, averages are taken over a constant energy hypersurface which is a subset of the entire phase space.

We've also seen how any phase space distribution $\varrho(\mathit{\Lambda}_1,\ldots,\mathit{\Lambda}_k)$ which is a function of conserved quantities $\mathit{\Lambda}_a(\boldsymbol{\varphi})$ is automatically a stationary (time-independent) solution to Liouville's equation. Note that the microcanonical distribution,

$$\varrho_E(\boldsymbol{\varphi}) = \delta\big(E-\hat{H}(\boldsymbol{\varphi})\big) \Big/ \int d\mu\,\delta\big(E-\hat{H}(\boldsymbol{\varphi})\big) \quad , \tag{4.4}$$

is of this form, since $\hat{H}(\boldsymbol{\varphi})$ is conserved by the dynamics. Linear and angular momentum conservation generally are broken by elastic scattering off the walls of the sample.

So averages in the microcanonical ensemble are computed by evaluating the ratio

$$\big\langle A\big\rangle = \frac{\mathsf{Tr}\; A\,\delta(E-\hat{H})}{\mathsf{Tr}\;\delta(E-\hat{H})} \quad , \tag{4.5}$$

where Tr means 'trace', which entails an integration over all phase space:

$$\mathsf{Tr}\; A(q,p) \equiv \frac{1}{N!}\prod_{i=1}^{N}\int \frac{d^dp_i\, d^dq_i}{(2\pi\hbar)^d}\, A(q,p) \quad . \tag{4.6}$$

Here N is the total number of particles and d is the dimension of physical space in which each particle moves. The factor of $1/N!$, which cancels in the ratio between numerator and denominator, is present for

[1] We write the Hamiltonian as $\hat{H}$ (classical or quantum) in order to distinguish it from magnetic field, H. In chapter 2, we used the symbol H for enthalpy and wrote the magnetic field as H, but since enthalpy doesn't get a mention in this chapter, we shift notation slightly.

indistinguishable particles[2]. The normalization factor $(2\pi\hbar)^{-Nd}$ renders the trace dimensionless. Again, this cancels between numerator and denominator. These factors may then seem arbitrary in the definition of the trace, but we'll see how they in fact are required from quantum mechanical considerations. So we now adopt the following metric for classical phase space integration:

$$d\mu = \frac{1}{N!}\prod_{i=1}^{N}\frac{d^dp_i\,d^dq_i}{(2\pi\hbar)^d} \quad . \tag{4.7}$$

4.2.2 Density of states

The denominator,

$$D(E) = \mathsf{Tr}\;\delta(E-\hat{H}) \quad , \tag{4.8}$$

is called the *density of states*. It has dimensions of inverse energy, such that

$$\begin{aligned} D(E)\,\Delta E &= \int_E^{E+\Delta E} dE' \int d\mu\;\delta(E'-\hat{H}) = \int\limits_{E<\hat{H}<E+\Delta E} d\mu \\ &= \#\text{ of states with energies between } E \text{ and } E+\Delta E \quad . \end{aligned} \tag{4.9}$$

Let us now compute $D(E)$ for the nonrelativistic ideal gas. The Hamiltonian is

$$\hat{H}(q,p) = \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m} \quad . \tag{4.10}$$

We assume that the gas is enclosed in a region of volume V, and we'll do a purely classical calculation, neglecting discreteness of its quantum spectrum. We must compute

$$D(E) = \frac{1}{N!}\int\prod_{i=1}^{N}\frac{d^dp_i\,d^dq_i}{(2\pi\hbar)^d}\;\delta\Big(E-\sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m}\Big) \quad . \tag{4.11}$$

We shall calculate $D(E)$ in two ways. The first method utilizes the *Laplace transform*, $Z(\beta)$:

$$Z(\beta) = \mathcal{L}\big[D(E)\big] \equiv \int_0^\infty dE\;e^{-\beta E}\,D(E) = \mathsf{Tr}\;e^{-\beta\hat{H}} \quad . \tag{4.12}$$

The inverse Laplace transform is then

$$D(E) = \mathcal{L}^{-1}\big[Z(\beta)\big] \equiv \int_{c-i\infty}^{c+i\infty}\frac{d\beta}{2\pi i}\;e^{\beta E}\,Z(\beta) \quad , \tag{4.13}$$

[2]More on this in chapter 5.

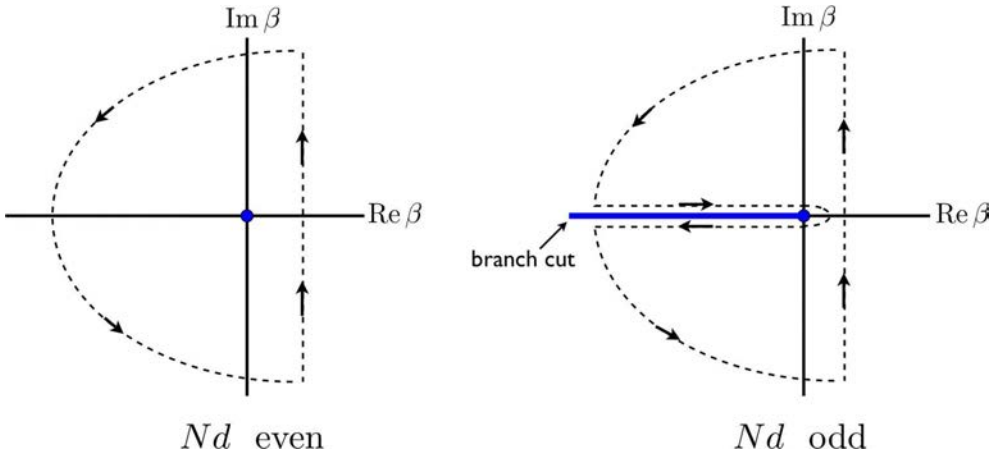

Figure 4.1: Complex integration contours $\mathcal{C}$ for inverse Laplace transform $\mathcal{L}^{-1}\big[Z(\beta)\big] = D(E)$. When the product dN is odd, there is a branch cut along the negative $\mathrm{Re}\,\beta$ axis.

where c is such that the integration contour is to the right of any singularities of $Z(\beta)$ in the complex β-plane. We then have

$$\begin{aligned} Z(\beta) &= \frac{1}{N!}\prod_{i=1}^{N}\int\frac{d^dx_i\,d^dp_i}{(2\pi\hbar)^d}\,e^{-\beta\boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\,e^{-\beta p^2/2m}\right)^{Nd} = \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\beta^{-Nd/2} \quad . \end{aligned} \tag{4.14}$$

The inverse Laplace transform is then

$$\begin{aligned} D(E) &= \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\oint_{\mathcal{C}}\frac{d\beta}{2\pi i}\,e^{\beta E}\,\beta^{-Nd/2} \\ &= \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\frac{E^{\frac{1}{2}Nd-1}}{\Gamma(Nd/2)} \quad , \end{aligned} \tag{4.15}$$

exactly as before. The integration contour for the inverse Laplace transform is extended in an infinite semicircle in the left half β-plane. When Nd is even, the function $\beta^{-Nd/2}$ has a simple pole of order $Nd/2$ at the origin. When Nd is odd, there is a branch cut extending along the negative $\mathrm{Re}\,\beta$ axis, and the integration contour must avoid the cut, as shown in fig. 4.1. One can check that this results in the same expression above, *i.e.* we may analytically continue from even values of Nd to all positive values of Nd.

For a general system, the Laplace transform, $Z(\beta) = \mathcal{L}\big[D(E)\big]$ also is called the *partition function*. We shall again meet up with $Z(\beta)$ when we discuss the ordinary canonical ensemble.

Our final result, then, is

$$D(E,V,N) = \frac{V^N}{N!}\left(\frac{m}{2\pi\hbar^2}\right)^{Nd/2}\frac{E^{\frac{1}{2}Nd-1}}{\Gamma(Nd/2)} \quad . \tag{4.16}$$

Here we have emphasized that the density of states is a function of E, V, and N. Using Stirling's approximation,

$$\ln N! = N\ln N - N + \tfrac{1}{2}\ln N + \tfrac{1}{2}\ln(2\pi) + \mathcal{O}\big(N^{-1}\big) \quad , \tag{4.17}$$

we may define the *statistical entropy*,

$$S(E,V,N) \equiv k_{\rm B}\ln D(E,V,N) = Nk_{\rm B}\,\phi\left(\frac{E}{N}\,,\,\frac{V}{N}\right) + \mathcal{O}(\ln N) \quad , \tag{4.18}$$

where

$$\phi\left(\frac{E}{N}\,,\,\frac{V}{N}\right) = \frac{d}{2}\,\ln\left(\frac{E}{N}\right) + \ln\left(\frac{V}{N}\right) + \frac{d}{2}\,\ln\left(\frac{m}{d\pi\hbar^2}\right) + \big(1+\tfrac{1}{2}d\big) \quad . \tag{4.19}$$

Recall $k_{\rm B} = 1.3806503\times10^{-16}\,{\rm erg/K}$ is Boltzmann's constant.

Second method

The second method invokes a mathematical trick. First, let's rescale $p_i^\alpha \equiv \sqrt{2mE}\,u_i^\alpha$. We then have

$$D(E) = \frac{V^N}{N!}\left(\frac{\sqrt{2mE}}{h}\right)^{Nd}\frac{1}{E}\int d^Mu\;\delta\big(u_1^2+u_2^2+\ldots+u_M^2-1\big) \quad . \tag{4.20}$$

Here we have written $\boldsymbol{u} = (u_1,u_2,\ldots,u_M)$ with $M = Nd$ as a M-dimensional vector. We've also used the rule $\delta(Ex) = E^{-1}\delta(x)$ for δ-functions. We can now write

$$d^Mu = u^{M-1}\,du\,d\Omega_M \quad , \tag{4.21}$$

where $d\Omega_M$ is the M-dimensional differential solid angle. We now have our answer:[3]

$$D(E) = \frac{V^N}{N!}\left(\frac{\sqrt{2m}}{h}\right)^{Nd}E^{\frac{1}{2}Nd-1}\cdot\tfrac{1}{2}\,\Omega_{Nd} \quad . \tag{4.22}$$

What remains is for us to compute Ω_M, the total solid angle in M dimensions. We do this by a nifty mathematical trick. Consider the integral

$$\begin{aligned}\mathcal{I}_M &= \int d^Mu\,e^{-u^2} = \Omega_M\int_0^\infty du\,u^{M-1}\,e^{-\boldsymbol{u}^2}\\ &= \tfrac{1}{2}\Omega_M\int_0^\infty ds\,s^{\frac{1}{2}M-1}\,e^{-s} = \tfrac{1}{2}\Omega_M\,\Gamma\big(\tfrac{1}{2}M\big) \quad ,\end{aligned} \tag{4.23}$$

where $s = u^2$, and where

$$\Gamma(z) = \int_0^\infty dt\,t^{z-1}\,e^{-t} \tag{4.24}$$

[3]The factor of $\frac{1}{2}$ preceding Ω_M in eqn. 4.22 appears because $\delta(u^2-1) = \frac{1}{2}\,\delta(u-1) + \frac{1}{2}\,\delta(u+1)$. Since $u = |\boldsymbol{u}| \ge 0$, the second term can be dropped.

is the Gamma function, which satisfies $z\,\Gamma(z) = \Gamma(z+1)$.[4] On the other hand, we can compute $\mathcal{I}_M$ in Cartesian coordinates, writing

$$\mathcal{I}_M = \left(\int_{-\infty}^{\infty} du_1\, e^{-u_1^2} \right)^M = \left(\sqrt{\pi}\right)^M \quad . \tag{4.25}$$

Therefore

$$\Omega_M = \frac{2\pi^{M/2}}{\Gamma(M/2)} \quad . \tag{4.26}$$

Thus we obtain $\Omega_2 = 2\pi$, $\Omega_3 = 4\pi$, $\Omega_4 = 2\pi^2$, *etc.*, the first two of which are familiar.

4.2.3 Arbitrariness in the definition of $S(E)$

Note that $D(E)$ has dimensions of inverse energy, so one might ask how we are to take the logarithm of a dimensionful quantity in eqn. 4.18. We must introduce an energy scale, such as ΔE in eqn. 4.9, and define $\tilde{D}(E;\Delta E) = D(E)\,\Delta E$ and $S(E;\Delta E) \equiv k_{\rm B} \ln \tilde{D}(E;\Delta E)$. The definition of statistical entropy then involves the arbitrary parameter ΔE, however this only affects $S(E)$ in an additive way. That is,

$$S(E,V,N;\Delta E_1) = S(E,V,N;\Delta E_2) + k_{\rm B} \ln\left(\frac{\Delta E_1}{\Delta E_2}\right) \quad . \tag{4.27}$$

Note that the difference between the two definitions of S depends only on the ratio $\Delta E_1/\Delta E_2$, and is independent of E, V, and N.

4.2.4 Ultra-relativistic ideal gas

Consider an ultrarelativistic ideal gas, with single particle dispersion $\varepsilon(p) = cp$. We then have

$$Z(\beta) = \frac{V^N}{N!}\frac{\Omega_d^N}{h^N d}\left(\int_0^{\infty} dp\, p^{d-1}\, e^{-\beta cp} \right)^N = \frac{V^N}{N!}\left(\frac{\Gamma(d)\,\Omega_d}{c^d\,h^d\,\beta^d}\right)^N \quad . \tag{4.28}$$

The statistical entropy is $S(E,V,N) = k_{\rm B} \ln D(E,V,N) = N k_{\rm B}\, \phi\big(\frac{E}{N}, \frac{V}{N}\big)$, with

$$\phi\left(\frac{E}{N}\,,\,\frac{V}{N}\right) = d\,\ln\left(\frac{E}{N}\right) + \ln\left(\frac{V}{N}\right) + \ln\left(\frac{\Omega_d\,\Gamma(d)}{(dhc)^d}\right) + (d+1) \tag{4.29}$$

4.2.5 Discrete systems

For classical systems where the energy levels are discrete, the states of the system $|\,\boldsymbol{\sigma}\,\rangle$ are labeled by a set of discrete quantities $\{\sigma_1, \sigma_2, \ldots\}$, where each variable σ_i takes discrete values. The number of ways

[4] Note that for integer argument, $\Gamma(k) = (k-1)!$

of configuring the system at fixed energy E is then

$$\Omega(E,N) = \sum_{\boldsymbol{\sigma}} \delta_{\hat{H}(\boldsymbol{\sigma}),E} \quad , \tag{4.30}$$

where the sum is over all possible configurations. Here N labels the total number of particles. For example, if we have N spin-$\frac{1}{2}$ particles on a lattice which are placed in a magnetic field H, so the individual particle energy is $\varepsilon_i = -\mu_0 H\sigma$, where $\sigma = \pm 1$, then in a configuration in which $N_\uparrow$ particles have $\sigma_i = +1$ and $N_\downarrow = N - N_\uparrow$ particles have $\sigma_i = -1$, the energy is $E = (N_\downarrow - N_\uparrow)\mu_0 H$. The number of configurations at fixed energy E is

$$\Omega(E,N) = \binom{N}{N_\uparrow} = \frac{N!}{(\frac{N}{2} - \frac{E}{2\mu_0 H})!\,(\frac{N}{2} + \frac{E}{2\mu_0 H})!} \quad , \tag{4.31}$$

since $N_{\uparrow/\downarrow} = \frac{N}{2} \mp \frac{E}{2\mu_0 H}$. The statistical entropy is $S(E,N) = k_{\rm B} \ln \Omega(E,N)$.

4.3 The Quantum Mechanical Trace

Thus far our understanding of ergodicity is rooted in the dynamics of classical mechanics. A Hamiltonian flow which is ergodic is one in which time averages can be replaced by phase space averages using the microcanonical ensemble. What happens, though, if our system is quantum mechanical, as all systems ultimately are?

4.3.1 The density matrix

First, let us consider that our system S will in general be in contact with a world W. We call the union of S and W the universe, $U = W \cup S$. Let $\big|\, N \,\big\rangle$ denote a quantum mechanical state of W, and let $\big|\, n \,\big\rangle$ denote a quantum mechanical state of S. Then the most general wavefunction we can write is of the form

$$\big|\, \Psi \,\big\rangle = \sum_{N,n} \Psi_{N,n} \,\big|\, N \,\big\rangle \otimes \big|\, n \,\big\rangle \quad . \tag{4.32}$$

Now let us compute the expectation value of some operator $\hat{\mathcal{A}}$ which acts as the identity within W, meaning $\big\langle\, N \,\big|\, \hat{\mathcal{A}} \,\big|\, N' \,\big\rangle = \hat{A}\,\delta_{NN'}$, where $\hat{A}$ is the 'reduced' operator which acts within S alone. We then have

$$\big\langle\, \Psi \,\big|\, \hat{\mathcal{A}} \,\big|\, \Psi \,\big\rangle = \sum_{N,N'} \sum_{n,n'} \Psi^*_{N,n}\, \Psi_{N',n'}\, \delta_{NN'} \,\big\langle\, n \,\big|\, \hat{A} \,\big|\, n' \,\big\rangle = \mathsf{Tr}\left(\hat{\varrho}\,\hat{A}\right) \quad , \tag{4.33}$$

where

$$\hat{\varrho} = \sum_N \sum_{n,n'} \Psi^*_{N,n}\, \Psi_{N,n'} \,\big|\, n' \,\big\rangle \,\big\langle\, n \,\big| \tag{4.34}$$

is the *density matrix*. The time-dependence of $\hat{\varrho}$ is easily found:

$$\hat{\varrho}(t) = \sum_N \sum_{n,n'} \Psi^*_{N,n}\, \Psi_{N,n'} \,\big|\, n'(t) \,\big\rangle \,\big\langle\, n(t) \,\big| = e^{-i\hat{H}t/\hbar}\; \hat{\varrho}\; e^{+i\hat{H}t/\hbar} \quad , \tag{4.35}$$

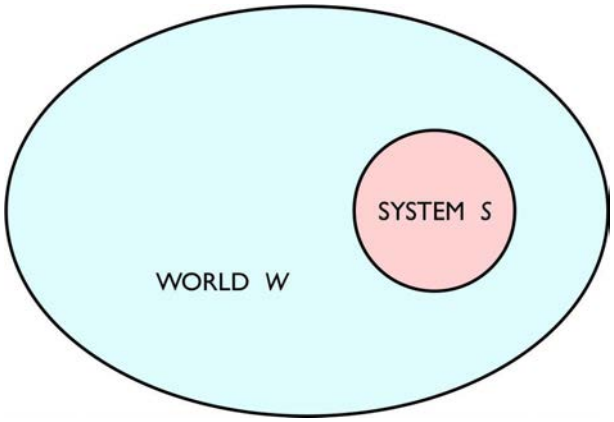

Figure 4.2: A system S in contact with a 'world' W. The union of the two, universe $U = W \cup S$, is said to be the 'universe'.

where $\hat{H}$ is the Hamiltonian for the system S. Thus, we find

$$i\hbar \frac{\partial \hat{\varrho}}{\partial t} = \big[\hat{H}, \hat{\varrho}\big] \quad . \tag{4.36}$$

Note that the density matrix evolves according to a slightly different equation than an operator in the Heisenberg picture, for which

$$\hat{A}(t) = e^{+iHt/\hbar}\, A\, e^{-i\hat{H}t/\hbar} \qquad \Longrightarrow \qquad i\hbar \frac{\partial \hat{A}}{\partial t} = \big[\hat{A}, \hat{H}\big] = -\big[\hat{H}, \hat{A}\big] \quad . \tag{4.37}$$

For Hamiltonian systems, we found that the phase space distribution $\varrho(q, p, t)$ evolved according to the Liouville equation, $i\,\partial\varrho/\partial t = L\,\varrho$, where the Liouvillian L is the differential operator

$$L = -i \sum_{j=1}^{Nd} \left\{ \frac{\partial \hat{H}}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial \hat{H}}{\partial q_j} \frac{\partial}{\partial p_j} \right\} \quad . \tag{4.38}$$

Accordingly, any distribution $\varrho(I_1, \ldots, I_k)$ which is a function of constants of the motion $I_a(q, p)$ is a stationary solution to the Liouville equation: $\partial_t\, \varrho(I_1, \ldots, I_k) = 0$. Similarly, any quantum mechanical density matrix which commutes with the Hamiltonian is a stationary solution to eqn. 4.36. The corresponding microcanonical distribution is $\hat{\varrho}_E = \delta\big(E - \hat{H}\big)$.

4.3.2 Averaging the DOS

If our quantum mechanical system is placed in a finite volume, the energy levels will be discrete, rather than continuous, and the density of states (DOS) will be of the form

$$D(E) = \mathsf{Tr}\; \delta\big(E - \hat{H}\big) = \sum_l \delta(E - E_l) \quad , \tag{4.39}$$

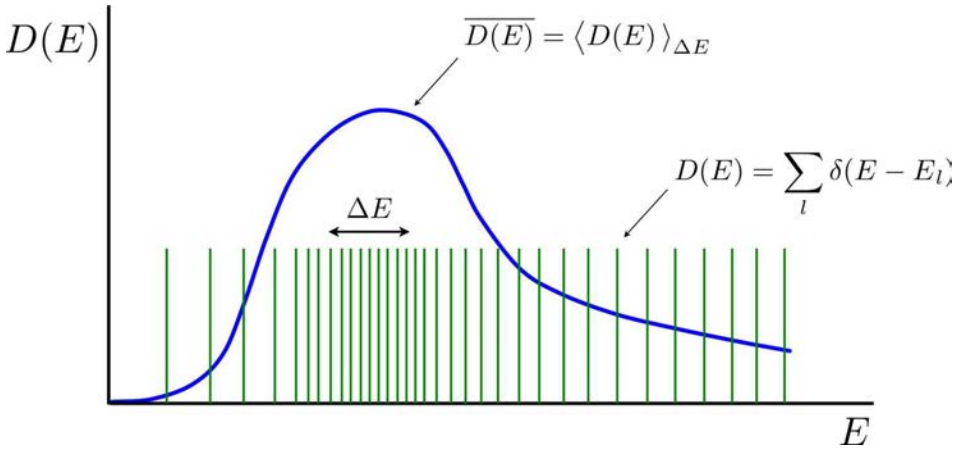

Figure 4.3: Averaging the quantum mechanical discrete density of states yields a continuous curve.

where $\{E_l\}$ are the eigenvalues of the Hamiltonian $\hat{H}$. In the thermodynamic limit, $V \to \infty$, and the discrete spectrum of kinetic energies remains discrete for all finite V but must approach the continuum result. To recover the continuum result, we average the DOS over a window of width ΔE:

$$\overline{D(E)} = \frac{1}{\Delta E} \int_E^{E+\Delta E} dE' \, D(E') \quad . \tag{4.40}$$

If we take the limit $\Delta E \to 0$ but with $\Delta E \gg \delta E$, where δE is the spacing between successive quantized levels, we recover a smooth function, as shown in fig. 4.3. We will in general drop the bar and refer to this function as $D(E)$. Note that $\delta E \sim 1/D(E) = e^{-N\phi(\varepsilon,v)}$ is (typically) *exponentially* small in the size of the system, hence if we took $\Delta E \propto V^{-1}$ which vanishes in the thermodynamic limit, there are still exponentially many energy levels within an interval of width ΔE.

4.3.3 Coherent states

The quantum-classical correspondence is elucidated with the use of *coherent states*. Recall that the one-dimensional harmonic oscillator Hamiltonian may be written

$$\hat{H}_0 = \frac{p^2}{2m} + \tfrac{1}{2} m \omega_0^2 \, q^2 = \hbar\omega_0 \left(a^\dagger a + \tfrac{1}{2} \right) \quad , \tag{4.41}$$

where a and $a^\dagger$ are *ladder operators* satisfying $\left[a, a^\dagger \right] = 1$, which can be taken to be

$$a = \ell \, \frac{\partial}{\partial q} + \frac{q}{2\ell} \qquad , \qquad a^\dagger = -\ell \, \frac{\partial}{\partial q} + \frac{q}{2\ell} \quad , \tag{4.42}$$

with $\ell = \sqrt{\hbar/2m\omega_0}$. Note that $q = \ell \, (a + a^\dagger)$ and $p = \frac{\hbar}{2i\ell} \, (a - a^\dagger)$.

The ground state satisfies $a \, \psi_0(q) = 0$, which yields

$$\psi_0(q) = (2\pi\ell^2)^{-1/4} \, e^{-q^2/4\ell^2} \quad . \tag{4.43}$$

The normalized *coherent state* $|\, z\,\rangle$ is defined as

$$|\, z\,\rangle = e^{-\frac{1}{2}|z|^2}\, e^{z a^\dagger} |\, 0\,\rangle = e^{-\frac{1}{2}|z|^2} \sum_{n=0}^\infty \frac{z^n}{\sqrt{n!}}\, |\, n\,\rangle \quad . \tag{4.44}$$

The overlap of coherent states is given by

$$\langle\, z_1 \,|\, z_2\,\rangle = e^{-\frac{1}{2}|z_1|^2}\, e^{-\frac{1}{2}|z_2|^2}\, e^{\bar z_1 z_2} \quad , \tag{4.45}$$

hence different coherent states are not orthogonal. Despite this nonorthogonality, the coherent states allow a simple resolution of the identity,

$$1 = \int \frac{d^2z}{2\pi i}\, |\, z\,\rangle\langle\, z\,| \quad ; \quad \frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re} z\, d\,\mathrm{Im} z}{\pi} \tag{4.46}$$

which is straightforward to establish.

To gain some physical intuition about the coherent states, define

$$z \equiv \frac{Q}{2\ell} + \frac{i\ell P}{\hbar} \tag{4.47}$$

and write $|\, z\,\rangle \equiv |\, Q, P\,\rangle$. One finds (*exercise!*)

$$\psi_{Q,P}(q) = \langle\, q \,|\, z\,\rangle = (2\pi\ell^2)^{-1/4}\, e^{-iPQ/2\hbar}\, e^{iPq/\hbar}\, e^{-(q-Q)^2/4\ell^2} \quad , \tag{4.48}$$

hence the coherent state $\psi_{Q,P}(q)$ is a wavepacket Gaussianly localized about $q = Q$, but oscillating with average momentum P.

For example, we can compute

$$\langle\, Q, P \,|\, q \,|\, Q, P\,\rangle = \langle\, z \,|\, \ell\,(a + a^\dagger) \,|\, z\,\rangle = 2\ell\, \mathrm{Re}\, z = Q \tag{4.49}$$

$$\langle\, Q, P \,|\, p \,|\, Q, P\,\rangle = \langle\, z \,|\, \frac{\hbar}{2i\ell}\,(a - a^\dagger) \,|\, z\,\rangle = \frac{\hbar}{\ell}\, \mathrm{Im}\, z = P \tag{4.50}$$

as well as

$$\langle\, Q, P \,|\, q^2 \,|\, Q, P\,\rangle = \langle\, z \,|\, \ell^2\,(a + a^\dagger)^2 \,|\, z\,\rangle = Q^2 + \ell^2 \tag{4.51}$$

$$\langle\, Q, P \,|\, p^2 \,|\, Q, P\,\rangle = -\langle\, z \,|\, \frac{\hbar^2}{4\ell^2}(a - a^\dagger)^2 \,|\, z\,\rangle = P^2 + \frac{\hbar^2}{4\ell^2} \quad . \tag{4.52}$$

Thus, the root mean square fluctuations in the coherent state $|\, Q, P\,\rangle$ are

$$\Delta q = \ell = \sqrt{\frac{\hbar}{2m\omega_0}} \qquad , \qquad \Delta p = \frac{\hbar}{2\ell} = \sqrt{\frac{m\hbar\omega_0}{2}} \quad , \tag{4.53}$$

and $\Delta q \cdot \Delta p = \frac{1}{2}\,\hbar$. Thus we learn that the coherent state $\psi_{Q,P}(q)$ is localized in phase space, *i.e.* in both position and momentum. If we have a general operator $\hat A(q,p)$, we can then write

$$\langle\, Q, P \,|\, \hat A(q,p) \,|\, Q, P\,\rangle = A(Q,P) + \mathcal{O}(\hbar) \quad , \tag{4.54}$$

where $A(Q,P)$ is formed from $\hat{A}(q,p)$ by replacing $q \to Q$ and $p \to P$.

Since

$$\frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re}z\; d\,\mathrm{Im}z}{\pi} = \frac{dQ\,dP}{2\pi\hbar} \quad , \tag{4.55}$$

we can write the trace using coherent states as

$$\mathrm{Tr}\ \hat{A} = \frac{1}{2\pi\hbar}\int\limits_{-\infty}^{\infty} dQ \int\limits_{-\infty}^{\infty} dP\, \langle\, Q,P \,|\, \hat{A} \,|\, Q,P \,\rangle \quad . \tag{4.56}$$

We now can understand the origin of the factor $2\pi\hbar$ in the denominator of each (q_i,p_i) integral over classical phase space in eqn. 4.6.

Note that ω_0 is arbitrary in our discussion. By increasing ω_0, the states become more localized in q and more plane wave like in p. However, so long as ω_0 is finite, the width of the coherent state in each direction is proportional to $\hbar^{1/2}$, and thus vanishes in the classical limit.

4.4 Thermal Equilibrium

4.4.1 Two systems in thermal contact

Consider two systems in thermal contact, as depicted in fig. 4.4. The two subsystems #1 and #2 are free to exchange energy, but their respective volumes and particle numbers remain fixed. We assume the contact is made over a surface, and that the energy associated with that surface is negligible when compared with the bulk energies E_1 and E_2. Let the total energy be $E = E_1 + E_2$. Then the density of states $D(E)$ for the combined system is

$$D(E) = \int dE_1\, D_1(E_1)\, D_2(E - E_1) \quad . \tag{4.57}$$

The probability density for system #1 to have energy E_1 is then

$$P_1(E_1) = \frac{D_1(E_1)\, D_2(E-E_1)}{D(E)} \quad . \tag{4.58}$$

Note that $P_1(E_1)$ is normalized: $\int dE_1\, P_1(E_1) = 1$. We now ask: what is the most probable value of E_1? We find out by differentiating $P_1(E_1)$ with respect to E_1 and setting the result to zero. This requires

$$\begin{aligned} 0 &= \frac{1}{P_1(E_1)}\frac{dP_1(E_1)}{dE_1} = \frac{\partial}{\partial E_1}\,\ln P_1(E_1) \\ &= \frac{\partial}{\partial E_1}\,\ln D_1(E_1) + \frac{\partial}{\partial E_1}\,\ln D_2(E-E_1) \quad . \end{aligned} \tag{4.59}$$

We conclude that the maximally likely partition of energy between systems #1 and #2 is realized when

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2} \quad . \tag{4.60}$$

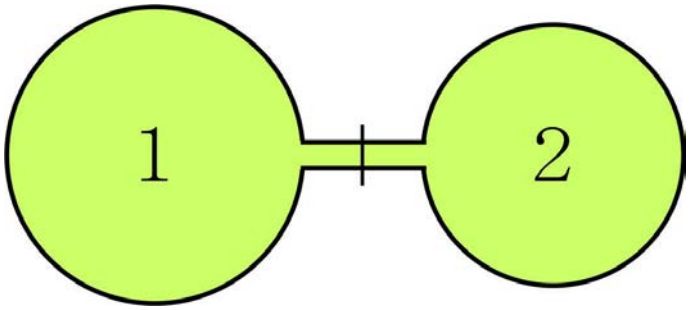

Figure 4.4: Two systems in thermal contact.

This guarantees that

$$S(E, E_1) = S_1(E_1) + S_2(E - E_1) \tag{4.61}$$

is a maximum with respect to the energy E_1, at fixed total energy E.

The *temperature* T is defined as

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N} \quad , \tag{4.62}$$

a result familiar from thermodynamics. The difference is now we have a more rigorous definition of the entropy. When the total entropy S is maximized, we have that $T_1 = T_2$. Once again, two systems in thermal contact and can exchange energy will in equilibrium have equal temperatures.

According to eqns. 4.19 and 4.29, the entropies of nonrelativistic and ultrarelativistic ideal gases in d space dimensions are given by

$$S_{\rm NR} = \tfrac{1}{2} N d\, k_{\rm B} \ln\left(\frac{E}{N}\right) + N k_{\rm B} \ln\left(\frac{V}{N}\right) + \text{const.} \tag{4.63}$$

$$S_{\rm UR} = N d\, k_{\rm B} \ln\left(\frac{E}{N}\right) + N k_{\rm B} \ln\left(\frac{V}{N}\right) + \text{const.} \quad . \tag{4.64}$$

Invoking eqn. 4.62, we then have $E_{\rm NR} = \frac{1}{2} N d\, k_{\rm B} T$ and $E_{\rm UR} = N d\, k_{\rm B} T$.

We saw that the probability distribution $P_1(E_1)$ is maximized when $T_1 = T_2$, but how sharp is the peak in the distribution? Let us write $E_1 = E_1^* + \Delta E_1$, where E_1^* is the solution to eqn. 4.59. We then have

$$\ln P_1(E_1^* + \Delta E_1) = \ln P_1(E_1^*) + \frac{1}{2k_{\rm B}} \frac{\partial^2 S_1}{\partial E_1^2}\bigg|_{E_1^*} (\Delta E_1)^2 + \frac{1}{2k_{\rm B}} \frac{\partial^2 S_2}{\partial E_2^2}\bigg|_{E_2^*} (\Delta E_1)^2 + \ldots \quad , \tag{4.65}$$

where $E_2^* = E - E_1^*$. We must now evaluate

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_{V,N} = -\frac{1}{T^2\, C_V} \quad , \tag{4.66}$$

where $C_V = \big(\partial E/\partial T\big)_{V,N}$ is the heat capacity. Thus,

$$P_1 = P_1^*\, e^{-(\Delta E_1)^2 / 2 k_{\rm B} T^2 \bar{C}_V} \quad , \tag{4.67}$$

where

$$\bar{C}_V = \frac{C_{V,1}\, C_{V,2}}{C_{V,1} + C_{V,2}} \quad . \tag{4.68}$$

The distribution is therefore a Gaussian, and the fluctuations in ΔE_1 can now be computed:

$$\big\langle (\Delta E_1)^2 \big\rangle = k_{\rm B} T^2\, \bar{C}_V \qquad \Longrightarrow \qquad (\Delta E_1)_{\rm RMS} = k_{\rm B} T \sqrt{\bar{C}_V / k_{\rm B}} \quad . \tag{4.69}$$

The individual heat capacities $C_{V,1}$ and $C_{V,2}$ scale with the volumes V_1 and V_2, respectively. If $V_2 \gg V_1$, then $C_{V,2} \gg C_{V,1}$, in which case $\bar{C}_V \approx C_{V,1}$. Therefore the RMS fluctuations in ΔE_1 are proportional to the *square root* of the system size, whereas E_1 itself is extensive. Thus, the ratio $(\Delta E_1)_{\rm RMS}/E_1 \propto V^{-1/2}$ scales as the inverse square root of the volume. The distribution $P_1(E_1)$ is thus *extremely sharp*.

4.4.2 Thermal, mechanical and chemical equilibrium

We have $dS\big|_{V,N} = \frac{1}{T}\, dE$, but in general $S = S(E, V, N)$. Equivalently, we may write $E = E(S, V, N)$. The full differential of $E(S, V, N)$ is then $dE = T\, dS - p\, dV + \mu\, dN$, with $T = \big(\frac{\partial E}{\partial S}\big)_{V,N}$ and $p = -\big(\frac{\partial E}{\partial V}\big)_{S,N}$ and $\mu = \big(\frac{\partial E}{\partial N}\big)_{S,V}$. As we shall discuss in more detail, p is the *pressure* and μ is the *chemical potential*. We may thus write the total differential dS as

$$dS = \frac{1}{T}\, dE + \frac{p}{T}\, dV - \frac{\mu}{T}\, dN \quad . \tag{4.70}$$

Employing the same reasoning as in the previous section, we conclude that entropy maximization for two systems in contact requires the following:

- If two systems can exchange energy, then $T_1 = T_2$. This is *thermal equilibrium*.
- If two systems can exchange volume, then $p_1/T_1 = p_2/T_2$. This is *mechanical equilibrium*.
- If two systems can exchange particle number, then $\mu_1/T_1 = \mu_2/T_2$. This is *chemical equilibrium*.

4.4.3 Gibbs-Duhem relation

The energy $E(S, V, N)$ is an extensive function of extensive variables, *i.e.* it is homogeneous of degree one in its arguments. Therefore $E(\lambda S, \lambda V, \lambda N) = \lambda E$, and taking the derivative with respect to λ yields

$$\begin{aligned} E &= S \left(\frac{\partial E}{\partial S}\right)_{V,N} + V \left(\frac{\partial E}{\partial V}\right)_{S,N} + N \left(\frac{\partial E}{\partial N}\right)_{S,V} \\ &= TS - pV + \mu N \quad . \end{aligned} \tag{4.71}$$

Taking the differential of each side, using the Leibniz rule on the RHS, and plugging in $dE = T\, dS - p\, dV + \mu\, dN$, we arrive at the *Gibbs-Duhem relation*[5],

$$S\, dT - V dp + N\, d\mu = 0 \quad . \tag{4.72}$$

[5] See §2.7.4.

This, in turn, says that any one of the intensive quantities (T, p, μ) can be written as a function of the other two, in the case of a single component system.

4.5 Ordinary Canonical Ensemble (OCE)

4.5.1 Canonical distribution and partition function

Consider a system S in contact with a world W, and let their union $U = W \cup S$ be called the 'universe'. The situation is depicted in fig. 4.2. The volume $V_{\rm S}$ and particle number $N_{\rm S}$ of the system are held fixed, but the energy is allowed to fluctuate by exchange with the world W. We are interested in the limit $N_{\rm S} \to \infty$, $N_{\rm W} \to \infty$, with $N_{\rm S} \ll N_{\rm W}$, with similar relations holding for the respective volumes and energies. We now ask what is the probability that S is in a state $|\, n \,\rangle$ with energy E_n. This is given by the ratio

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0} \frac{D_{\rm W}(E_{\rm U} - E_n)\,\Delta E}{D_{\rm U}(E_{\rm U})\,\Delta E} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.73}$$

Then

$$\begin{aligned} \ln P_n &= \ln D_{\rm W}(E_{\rm U} - E_n) - \ln D_{\rm U}(E_{\rm U}) \\ &= \ln D_{\rm W}(E_{\rm U}) - \ln D_{\rm U}(E_{\rm U}) - E_n \left.\frac{\partial \ln D_{\rm W}(E)}{\partial E}\right|_{E=E_{\rm U}} + \ldots \\ &\equiv -\alpha - \beta E_n \quad . \end{aligned} \tag{4.74}$$

The constant β is given by

$$\beta = \left.\frac{\partial \ln D_{\rm W}(E)}{\partial E}\right|_{E=E_{\rm U}} = \frac{1}{k_{\rm B} T} \quad . \tag{4.75}$$

Thus, we find $P_n = e^{-\alpha}\, e^{-\beta E_n}$. The constant α is fixed by the requirement that $\sum_n P_n = 1$:

$$P_n = \frac{1}{Z}\, e^{-\beta E_n} \qquad , \qquad Z(T, V, N) = \sum_n e^{-\beta E_n} = \mathsf{Tr}\; e^{-\beta \hat{H}} \quad . \tag{4.76}$$

We've already met $Z(\beta)$ in eqn. 4.12 – it is the Laplace transform of the density of states. It is also called the *partition function* of the system S. Quantum mechanically, we can write the ordinary canonical density matrix as

$$\hat{\varrho} = \frac{e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \quad , \tag{4.77}$$

which is known as the *Gibbs distribution*. Note that $\left[\hat{\varrho}, \hat{H}\right] = 0$, hence the ordinary canonical distribution is a stationary solution to the evolution equation for the density matrix. Note that the OCE is specified by three parameters: T, V, and N.

4.5.2 The difference between $P(E_n)$ and P_n

Let the total energy of the Universe be fixed at $E_{\rm U}$. The joint probability density $P(E_{\rm S}, E_{\rm W})$ for the system to have energy $E_{\rm S}$ and the world to have energy $E_{\rm W}$ is

$$P(E_{\rm S}, E_{\rm W}) = D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm W})\, \delta(E_{\rm U} - E_{\rm S} - E_{\rm W}) / D_{\rm U}(E_{\rm U}) \quad , \tag{4.78}$$

where

$$D_{\rm U}(E_{\rm U}) = \int\limits_{-\infty}^{\infty} dE_{\rm S}\; D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S}) \quad , \tag{4.79}$$

which ensures that $\int dE_{\rm S} \int dE_{\rm W}\, P(E_{\rm S}, E_{\rm W}) = 1$. The *probability density* $P(E_{\rm S})$ is defined such that $P(E_{\rm S})\, dE_{\rm S}$ is the (differential) probability for the system to have an energy in the range $[E_{\rm S}, E_{\rm S} + dE_{\rm S}]$. The units of $P(E_{\rm S})$ are E^{-1}. To obtain $P(E_{\rm S})$, we simply integrate the joint probability density $P(E_{\rm S}, E_{\rm W})$ over all possible values of $E_{\rm W}$, obtaining

$$P(E_{\rm S}) = \frac{D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S})}{D_{\rm U}(E_{\rm U})} \quad , \tag{4.80}$$

as we have in eqn. 4.73. Suppose we wish to know the probability P_n that the system is in a *particular* state $|\, n \,\rangle$ with energy E_n. Clearly

$$P_n = \lim_{\Delta E \to 0} \frac{\text{probability that } E_{\rm S} \in [E_n, E_n + \Delta E]}{\text{\# of S states with } E_{\rm S} \in [E_n, E_n + \Delta E]} = \frac{P(E_n)\, \Delta E}{D_{\rm S}(E_n)\, \Delta E} = \frac{D_{\rm W}(E_{\rm U} - E_n)}{D_{\rm U}(E_{\rm U})} \quad . \tag{4.81}$$

4.5.3 Additional remarks

The formula of eqn. 4.73 is quite general and holds in the case where $N_{\rm S}/N_{\rm W} = \mathcal{O}(1)$, so long as we are in the thermodynamic limit, where the energy associated with the interface between S and W may be neglected. In this case, however, one is not licensed to perform the subsequent Taylor expansion, and the distribution P_n is no longer of the Gibbs form. It is also valid for quantum systems[6], in which case we interpret $P_n = \langle n | \varrho_{\rm S} | n \rangle$ as a diagonal element of the density matrix $\varrho_{\rm S}$. The density of states functions may then be replaced by

$$\begin{aligned} D_{\rm W}(E_{\rm U} - E_n)\, \Delta E \to e^{S_{\rm W}(E_{\rm U} - E_n\,,\, \Delta E)} &\equiv \mathop{\rm Tra}_{\rm W} \int\limits_{E_{\rm U} - E_n}^{E_{\rm U} - E_n + \Delta E} dE\; \delta(E - \hat{H}_{\rm W}) \\ D_{\rm U}(E_{\rm U})\, \Delta E \to e^{S_{\rm U}(E_{\rm U}\,,\, \Delta E)} &\equiv \mathop{\rm Tra}_{\rm U} \int\limits_{E_{\rm U}}^{E_{\rm U} + \Delta E} dE\; \delta(E - \hat{H}_{\rm U}) \quad . \end{aligned} \tag{4.82}$$

The off-diagonal matrix elements of $\varrho_{\rm S}$ are negligible in the thermodynamic limit.

[6] See T.-C. Lu and T. Grover, arXiv 1709.08784.

4.5.4 Averages within the OCE

To compute averages within the OCE,

$$\langle \hat{A} \rangle = \mathsf{Tr}\left(\hat{\varrho}\,\hat{A}\right) = \frac{\sum_n \langle n|\hat{A}|n\rangle\, e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} \quad , \tag{4.83}$$

where we have conveniently taken the trace in a basis of energy eigenstates. In the classical limit, we have

$$\varrho(\boldsymbol{\varphi}) = \frac{1}{Z}\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \quad Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \int d\mu\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \tag{4.84}$$

with $d\mu = \frac{1}{N!}\prod_{j=1}^{N}(d^d q_j\, d^d p_j/h^d)$ for identical particles ('Maxwell-Boltzmann statistics'). Thus,

$$\langle A \rangle = \mathsf{Tr}\,(\varrho A) = \frac{\int d\mu\, A(\boldsymbol{\varphi})\, e^{-\beta \hat{H}(\boldsymbol{\varphi})}}{\int d\mu\, e^{-\beta \hat{H}(\boldsymbol{\varphi})}} \quad . \tag{4.85}$$

4.5.5 Entropy and free energy

The *Boltzmann entropy* is defined by

$$S = -k_{\rm B}\, \mathsf{Tr}\left(\hat{\varrho} \ln \hat{\varrho}\right) = -k_{\rm B} \sum_n P_n\, \ln P_n \quad . \tag{4.86}$$

The Boltzmann entropy and the statistical entropy $S = k_{\rm B} \ln D(E)$ are identical in the thermodynamic limit. We define the Helmholtz free energy $F(T,V,N)$ as

$$F(T,V,N) = -k_{\rm B} T \ln Z(T,V,N) \quad , \tag{4.87}$$

hence $P_n = e^{\beta F}\, e^{-\beta E_n}$ and $\ln P_n = \beta F - \beta E_n$. The entropy is then

$$S = -k_{\rm B} \sum_n P_n \left(\beta F - \beta E_n\right) = -\frac{F}{T} + \frac{\langle\, \hat{H}\,\rangle}{T} \quad , \tag{4.88}$$

which is to say $F = E - TS$, where

$$E = \sum_n P_n\, E_n = \frac{\mathsf{Tr}\; \hat{H}\, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \tag{4.89}$$

is the average energy. We also see that

$$Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \sum_n e^{-\beta E_n} \quad \Longrightarrow \quad E = \frac{\sum_n E_n\, e^{-\beta E_n}}{\sum_n e^{-\beta E_n}} = -\frac{\partial}{\partial \beta} \ln Z = \frac{\partial}{\partial \beta}(\beta F) \quad . \tag{4.90}$$

Thus, $F(T,V,N)$ is a Legendre transform of $E(S,V,N)$, with

$$dF = -S\, dT - p\, dV + \mu\, dN \quad , \tag{4.91}$$

which means

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \quad , \quad p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \quad , \quad \mu = +\left(\frac{\partial F}{\partial N}\right)_{T,V} \quad . \tag{4.92}$$

4.5.6 Fluctuations in the OCE

In the OCE, the energy is not fixed. It therefore fluctuates about its average value $E = \langle \hat{H} \rangle$. Note that

$$\begin{aligned} -\frac{\partial E}{\partial \beta} &= k_{\rm B} T^2 \, \frac{\partial E}{\partial T} = \frac{\partial^2 \ln Z}{\partial \beta^2} \\ &= \frac{\mathsf{Tr}\ \hat{H}^2 \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} - \left(\frac{\mathsf{Tr}\ \hat{H} \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} \right)^2 \\ &= \langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2 \quad . \end{aligned} \tag{4.93}$$

Thus, the heat capacity is related to the fluctuations in the energy, just as we saw at the end of §4.4:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{1}{k_{\rm B} T^2} \left(\langle \hat{H}^2 \rangle - \langle \hat{H} \rangle^2 \right) \tag{4.94}$$

For the nonrelativistic ideal gas, we found $C_V = \frac{d}{2} N k_{\rm B}$, hence the ratio of RMS fluctuations in the energy to the energy itself is

$$\frac{\sqrt{\langle \, (\Delta \hat{H})^2 \, \rangle}}{\langle \hat{H} \rangle} = \frac{\sqrt{k_{\rm B} T^2 \, C_V}}{\frac{d}{2} N k_{\rm B} T} = \sqrt{\frac{2}{N d}} \quad , \tag{4.95}$$

and the ratio of the RMS fluctuations to the mean value vanishes in the thermodynamic limit.

The full distribution function for the energy is

$$P(\mathcal{E}) = \big\langle \delta(\mathcal{E} - \hat{H}) \big\rangle = \frac{\mathsf{Tr}\ \delta(\mathcal{E} - \hat{H}) \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} = \frac{1}{Z} \, D(\mathcal{E}) \, e^{-\beta \mathcal{E}} \quad . \tag{4.96}$$

Thus,

$$P(\mathcal{E}) = \frac{e^{-\beta [\mathcal{E} - TS(\mathcal{E})]}}{\int d\mathcal{E}' \, e^{-\beta [\mathcal{E}' - TS(\mathcal{E}')]}} \quad , \tag{4.97}$$

where $S(\mathcal{E}) = k_{\rm B} \ln D(\mathcal{E})$ is the statistical entropy. Let's write $\mathcal{E} = E + \delta\mathcal{E}$, where E extremizes the combination $\mathcal{E} - T\,S(\mathcal{E})$, *i.e.* the solution to $T\,S'(E) = 1$, where the energy derivative of S is performed at fixed volume V and particle number N. We now expand $S(E + \delta\mathcal{E})$ to second order in $\delta\mathcal{E}$, obtaining

$$S(E + \delta\mathcal{E}) = S(E) + \frac{\delta\mathcal{E}}{T} - \frac{\big(\delta\mathcal{E}\big)^2}{2T^2 \, C_V} + \ldots \tag{4.98}$$

Recall that $S''(E) = \frac{\partial}{\partial E} \left(\frac{1}{T} \right) = -\frac{1}{T^2 C_V}$. Thus,

$$\mathcal{E} - T\,S(\mathcal{E}) = E - T\,S(E) + \frac{(\delta\mathcal{E})^2}{2T\,C_V} + \mathcal{O}\big((\delta\mathcal{E})^3\big) \quad . \tag{4.99}$$

Applying this to both numerator and denominator of eqn. 4.97, we obtain[7]

$$P(\mathcal{E}) = \mathcal{N} \, \exp\left[-\frac{(\delta\mathcal{E})^2}{2 k_{\rm B} T^2 \, C_V} \right] \quad , \tag{4.100}$$

[7] In applying eqn. 4.99 to the denominator of eqn. 4.97, we shift $\mathcal{E}'$ by E and integrate over the difference $\delta\mathcal{E}' \equiv \mathcal{E}' - E$, retaining terms up to quadratic order in $\delta\mathcal{E}'$ in the argument of the exponent.

where $\mathcal{N} = (2\pi k_{\rm B} T^2 C_V)^{-1/2}$ is a normalization constant which guarantees $\int d\mathcal{E}\, P(\mathcal{E}) = 1$. Once again, we see that the distribution is a Gaussian centered at $\langle \mathcal{E} \rangle = E$, and of width $(\Delta\mathcal{E})_{\rm RMS} = \sqrt{k_{\rm B} T^2\, C_V}$. This is a consequence of the Central Limit Theorem.

4.5.7 Thermodynamics revisited

The average energy within the OCE is

$$E = \sum_n E_n P_n \quad , \tag{4.101}$$

and therefore

$$dE = \sum_n E_n\, dP_n + \sum_n P_n\, dE_n = đQ - đW \quad , \tag{4.102}$$

where

$$đQ = \sum_n E_n\, dP_n \qquad , \qquad đW = -\sum_n P_n\, dE_n \quad . \tag{4.103}$$

Finally, from $P_n = Z^{-1}\, e^{-E_n/k_{\rm B}T}$, we can write

$$E_n = -k_{\rm B} T \ln Z - k_{\rm B} T \ln P_n \quad , \tag{4.104}$$

with which we obtain

$$\begin{aligned} đQ &= \sum_n E_n\, dP_n \\ &= -k_{\rm B} T \ln Z \sum_n dP_n - k_{\rm B} T \sum_n \ln P_n\, dP_n \\ &= T\, d\Big(- k_{\rm B} \sum_n P_n \ln P_n \Big) = T\, dS \quad . \end{aligned} \tag{4.105}$$

Note also that

$$\begin{aligned} đW &= -\sum_n P_n\, dE_n \\ &= -\sum_n P_n \left(\sum_i \frac{\partial E_n}{\partial X_i}\, dX_i \right) \\ &= -\sum_{n,i} P_n \left\langle\, n \,\middle|\, \frac{\partial \hat{H}}{\partial X_i} \,\middle|\, n \,\right\rangle dX_i \equiv \sum_i F_i\, dX_i \quad , \end{aligned} \tag{4.106}$$

so the generalized force F_i conjugate to the generalized displacement dX_i is

$$F_i = -\sum_n P_n \frac{\partial E_n}{\partial X_i} = -\left\langle \frac{\partial \hat{H}}{\partial X_i} \right\rangle \quad . \tag{4.107}$$

This is the force acting *on* the system[8]. In the chapter on thermodynamics, we defined the generalized force conjugate to X_i as $y_i \equiv -F_i$.

[8] In deriving eqn. 4.107, we have used the so-called Feynman-Hellman theorem of quantum mechanics: $d\langle n|\hat{H}|n\rangle = \langle n|\, d\hat{H}\, |n\rangle$, if $|n\rangle$ is an energy eigenstate.

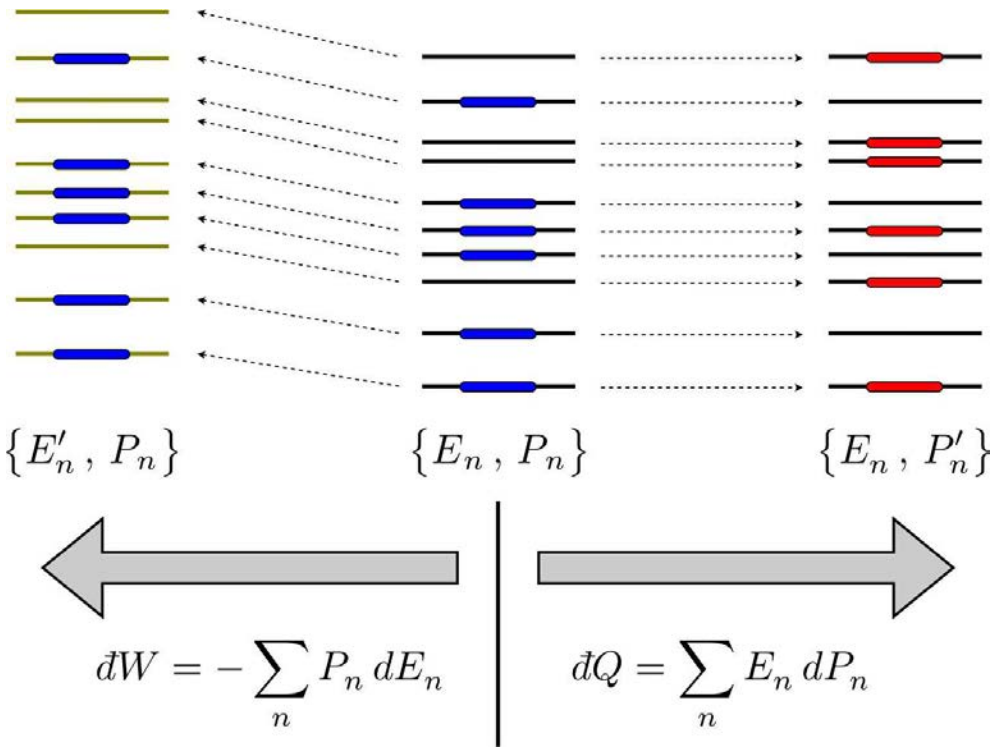

Figure 4.5: Microscopic, statistical interpretation of the First Law of Thermodynamics.

Thus we see from eqn. 4.102 that there are two ways that the average energy can change; these are depicted in the sketch of fig. 4.5. Starting from a set of energy levels $\{E_n\}$ and probabilities $\{P_n\}$, we can shift the energies to $\{E'_n\}$. The resulting change in energy $(\Delta E)_{\rm I} = -W$ is identified with the work done *on* the system. We could also modify the probabilities to $\{P'_n\}$ without changing the energies. The energy change in this case is the heat absorbed *by* the system: $(\Delta E)_{\rm II} = Q$. This provides us with a statistical and microscopic interpretation of the First Law of Thermodynamics.

4.5.8 Generalized susceptibilities

Suppose our Hamiltonian is of the form

$$\hat{H} = \hat{H}(\lambda) = \hat{H}_0 - \lambda\,\hat{Q} \quad , \tag{4.108}$$

where λ is an intensive parameter, such as magnetic field. Then $Z(\lambda) = \mathsf{Tr}\; e^{-\beta(\hat{H}_0 - \lambda\hat{Q})}$ and

$$\frac{1}{Z}\frac{\partial Z}{\partial \lambda} = \beta \cdot \frac{1}{Z}\,\mathsf{Tr}\left(\hat{Q}\,e^{-\beta\hat{H}(\lambda)}\right) = \beta\,\langle \hat{Q}\rangle \quad . \tag{4.109}$$

But then from $Z = e^{-\beta F}$ we have

$$Q(\lambda, T) = \langle\,\hat{Q}\,\rangle = -\left(\frac{\partial F}{\partial \lambda}\right)_T \quad . \tag{4.110}$$

Typically we will take Q to be an extensive quantity. We can now define the *susceptibility* χ as

$$\chi = \frac{1}{V}\frac{\partial Q}{\partial \lambda} = -\frac{1}{V}\frac{\partial^2 F}{\partial \lambda^2} \quad . \tag{4.111}$$

The volume factor in the denominator ensures that χ is intensive.

It is important to realize that we have assumed here that $\left[\hat{H}_0\,,\,\hat{Q}\right] = 0$, *i.e.* the 'bare' Hamiltonian $\hat{H}_0$ and the operator $\hat{Q}$ commute. If they do not commute, then the response functions must be computed within a proper quantum mechanical formalism, which we shall not discuss here.

Note also that we can imagine an entire family of observables $\left\{\hat{Q}_k\right\}$ satisfying $\left[\hat{Q}_k\,,\,\hat{Q}_{k'}\right] = 0$ and $\left[\hat{H}_0\,,\,\hat{Q}_k\right] = 0$, for all k and k'. Then for the Hamiltonian

$$\hat{H}(\vec{\lambda}) = \hat{H}_0 - \sum_k \lambda_k \hat{Q}_k \quad , \tag{4.112}$$

we have that

$$Q_k(\vec{\lambda}, T) = \langle\, \hat{Q}_k \,\rangle = -\left(\frac{\partial F}{\partial \lambda_k}\right)_{T,\, N_a,\, \lambda_{k'\neq k}} \tag{4.113}$$

and we may define an entire matrix of susceptibilities,

$$\chi_{kl} \equiv \frac{1}{V}\frac{\partial Q_k}{\partial \lambda_l} = -\frac{1}{V}\frac{\partial^2 F}{\partial \lambda_k\, \partial \lambda_l} \quad . \tag{4.114}$$

4.6 Grand Canonical Ensemble (GCE)

4.6.1 Grand canonical distribution and partition function

Consider once again the situation depicted in fig. 4.2, where a system S is in contact with a world W, their union $U = W \cup S$ being called the 'universe'. We assume that the system's volume $V_{\rm S}$ is fixed, but otherwise it is allowed to exchange energy and particle number with W. Hence, the system's energy $E_{\rm S}$ and particle number $N_{\rm S}$ will fluctuate. We ask what is the probability that S is in a state $|\, n \,\rangle$ with energy E_n and particle number N_n. This is given by the ratio

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0} \frac{D_{\rm W}(E_{\rm U} - E_n\,,\, N_{\rm U} - N_n)\,\Delta E}{D_{\rm U}(E_{\rm U}, N_{\rm U})\,\Delta E} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n \text{ and } N_{\rm S} = N_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.115}$$

Then

$$\begin{aligned}\ln P_n &= \ln D_{\rm W}(E_{\rm U} - E_n\,,\, N_{\rm U} - N_n) - \ln D_{\rm U}(E_{\rm U}, N_{\rm U}) \\ &= \ln D_{\rm W}(E_{\rm U}, N_{\rm U}) - \ln D_{\rm U}(E_{\rm U}, N_{\rm U}) \\ &\qquad - E_n \left.\frac{\partial \ln D_{\rm W}(E,N)}{\partial E}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} - N_n \left.\frac{\partial \ln D_{\rm W}(E,N)}{\partial N}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} + \ldots \\ &\equiv -\alpha - \beta E_n + \beta\mu N_n \quad .\end{aligned} \tag{4.116}$$

The constants β and μ are given by

$$\beta = \left.\frac{\partial \ln D_{\rm W}(E,N)}{\partial E}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} = \frac{1}{k_{\rm B}T} \tag{4.117}$$

$$\mu = -k_{\rm B}T \left.\frac{\partial \ln D_{\rm W}(E,N)}{\partial N}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} \quad . \tag{4.118}$$

The quantity μ has dimensions of energy and is called the *chemical potential*. *Nota bene:* Some texts define the 'grand canonical Hamiltonian' $\hat{K}$ as $\hat{K} \equiv \hat{H} - \mu\hat{N}$. Thus, $P_n = e^{-\alpha}\, e^{-\beta(E_n - \mu N_n)}$. Once again, the constant α is fixed by the requirement that $\sum_n P_n = 1$:

$$\begin{aligned}P_n &= \frac{1}{\Xi}\, e^{-\beta(E_n - \mu N_n)} \\ \Xi(\beta, V, \mu) &= \sum_n e^{-\beta(E_n - \mu N_n)} = \mathsf{Tr}\; e^{-\beta\hat{K}} \quad .\end{aligned} \tag{4.119}$$

Thus, the quantum mechanical *grand canonical density matrix* is given by

$$\hat{\varrho} = \frac{e^{-\beta\hat{K}}}{\mathsf{Tr}\; e^{-\beta\hat{K}}} \quad . \tag{4.120}$$

Note that $\left[\hat{\varrho}, \hat{K}\right] = 0$. The quantity $\Xi(T, V, \mu)$ is called the *grand partition function*. It stands in relation to a corresponding free energy in the usual way:

$$\Xi(T,V,\mu) \equiv e^{-\beta\Omega(T,V,\mu)} \qquad \Longleftrightarrow \qquad \Omega = -k_{\rm B}T\, \ln \Xi \quad , \tag{4.121}$$

where $\Omega(T, V, \mu)$ is the *grand potential*, also known as the *Landau free energy*. The dimensionless quantity $z \equiv e^{\beta\mu}$ is called the *fugacity*.

If $\left[\hat{H}, \hat{N}\right] = 0$, the grand potential may be expressed as a sum over contributions from each N sector, *viz.*

$$\Xi(T,V,\mu) = \sum_N e^{\beta\mu N}\, Z(T,V,N) \quad . \tag{4.122}$$

When there is more than one species, we have several chemical potentials $\{\mu_a\}$, and accordingly we define

$$\hat{K} = \hat{H} - \sum_a \mu_a \hat{N}_a \quad , \tag{4.123}$$

with $\Xi = \mathsf{Tr}\; e^{-\beta\hat{K}}$ as before.

4.6.2 Entropy and Gibbs-Duhem relation

In the GCE, the Boltzmann entropy is

$$\begin{aligned} S &= -k_{\rm B} \sum_n P_n \ln P_n \\ &= -k_{\rm B} \sum_n P_n \Big(\beta\Omega - \beta E_n + \beta\mu N_n \Big) \\ &= -\frac{\Omega}{T} + \frac{\langle \hat{H} \rangle}{T} - \frac{\mu \langle \hat{N} \rangle}{T} \quad , \end{aligned} \tag{4.124}$$

which says

$$\Omega = E - TS - \mu N \quad , \tag{4.125}$$

where

$$\begin{aligned} E &= \sum_n E_n \, P_n = \mathsf{Tr} \, \big(\hat{\varrho} \, \hat{H} \big) \\ N &= \sum_n N_n \, P_n = \mathsf{Tr} \, \big(\hat{\varrho} \, \hat{N} \big) \quad . \end{aligned} \tag{4.126}$$

Therefore, $\Omega(T, V, \mu)$ is a double Legendre transform of $E(S, V, N)$, with

$$d\Omega = -S \, dT - p \, dV - N \, d\mu \quad , \tag{4.127}$$

which entails

$$S = -\left(\frac{\partial \Omega}{\partial T} \right)_{V,\mu} \qquad , \qquad p = -\left(\frac{\partial \Omega}{\partial V} \right)_{T,\mu} \qquad , \qquad N = -\left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} \quad . \tag{4.128}$$

Since $\Omega(T, V, \mu)$ is an extensive quantity, we must be able to write $\Omega = V\omega(T, \mu)$. We identify the function $\omega(T, \mu)$ as the negative of the pressure:

$$\begin{aligned} \frac{\partial \Omega}{\partial V} &= -\frac{k_{\rm B} T}{\Xi} \left(\frac{\partial \Xi}{\partial V} \right)_{T,\mu} = \frac{1}{\Xi} \sum_n \frac{\partial E_n}{\partial V} \, e^{-\beta(E_n - \mu N_n)} \\ &= \left(\frac{\partial E}{\partial V} \right)_{T,\mu} = -p(T, \mu) \quad . \end{aligned} \tag{4.129}$$

Therefore, $\Omega = -pV$, and $p = p(T, \mu)$ is an *equation of state*. This is consistent with the result from thermodynamics that $G = E - TS + pV = \mu N$. Taking the differential, we recover the Gibbs-Duhem relation,

$$d\Omega = -S \, dT - p \, dV - N \, d\mu = -p \, dV - V \, dp \quad \Rightarrow \quad S \, dT - V \, dp + N \, d\mu = 0 \quad . \tag{4.130}$$

4.6.3 Generalized susceptibilities in the GCE

We can appropriate the results from §4.5.8 and apply them, *mutatis mutandis*, to the GCE. Suppose we have a family of observables $\{\hat{Q}_k\}$ satisfying $\big[\hat{Q}_k\,,\,\hat{Q}_{k'}\big]=0$ and $\big[\hat{H}_0\,,\,\hat{Q}_k\big]=0$ and $\big[\hat{N}_a\,,\,\hat{Q}_k\big]=0$ for all k, k', and a. Then for the grand canonical Hamiltonian

$$\hat{K}(\vec{\lambda}) = \hat{H}_0 - \sum_a \mu_a \hat{N}_a - \sum_k \lambda_k \hat{Q}_k \quad , \tag{4.131}$$

we have that

$$Q_k(\vec{\lambda},T) = \langle\,\hat{Q}_k\,\rangle = -\left(\frac{\partial\Omega}{\partial\lambda_k}\right)_{T,\mu_a,\lambda_{k'(\neq k)}} \tag{4.132}$$

and we may define the matrix of generalized susceptibilities,

$$\chi_{kl} = \frac{1}{V}\,\frac{\partial Q_k}{\partial\lambda_l} = -\frac{1}{V}\,\frac{\partial^2\Omega}{\partial\lambda_k\,\partial\lambda_l} \quad . \tag{4.133}$$

4.6.4 Fluctuations in the GCE

Both energy and particle number fluctuate in the GCE. Let us compute the fluctuations in particle number. We have

$$N = \langle\,\hat{N}\,\rangle = \frac{\mathsf{Tr}\;\hat{N}\,e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathsf{Tr}\;e^{-\beta(\hat{H}-\mu\hat{N})}} = \frac{1}{\beta}\,\frac{\partial}{\partial\mu}\,\ln\Xi \quad . \tag{4.134}$$

Therefore,

$$\begin{aligned}\frac{1}{\beta}\,\frac{\partial N}{\partial\mu} &= \frac{\mathsf{Tr}\;\hat{N}^2\,e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathsf{Tr}\;e^{-\beta(\hat{H}-\mu\hat{N})}} - \left(\frac{\mathsf{Tr}\;\hat{N}\,e^{-\beta(\hat{H}-\mu\hat{N})}}{\mathsf{Tr}\;e^{-\beta(\hat{H}-\mu\hat{N})}}\right)^2 \\ &= \big\langle\hat{N}^2\big\rangle - \big\langle\hat{N}\big\rangle^2 \quad .\end{aligned} \tag{4.135}$$

Note now that

$$\frac{\big\langle\hat{N}^2\big\rangle - \big\langle\hat{N}\big\rangle^2}{\big\langle\hat{N}\big\rangle^2} = \frac{k_{\rm B}T}{N^2}\left(\frac{\partial N}{\partial\mu}\right)_{T,V} = \frac{k_{\rm B}T}{V}\,\kappa_T \quad , \tag{4.136}$$

where κ_T is the isothermal compressibility. Note:

$$\begin{aligned}\left(\frac{\partial N}{\partial\mu}\right)_{T,V} &= \frac{\partial(N,T,V)}{\partial(\mu,T,V)} = -\frac{\partial(N,T,V)}{\partial(V,T,\mu)} \\ &= -\frac{\partial(N,T,V)}{\partial(N,T,p)}\cdot\frac{\partial(N,T,p)}{\partial(V,T,p)}\cdot\overbrace{\frac{\partial(V,T,p)}{\partial(N,T,\mu)}}^{1}\cdot\frac{\partial(N,T,\mu)}{\partial(V,T,\mu)} \\ &= -\frac{N^2}{V^2}\left(\frac{\partial V}{\partial p}\right)_{T,N} = \frac{N^2}{V}\,\kappa_T \quad .\end{aligned} \tag{4.137}$$

Thus,

$$\frac{(\Delta N)_{\rm RMS}}{N} = \sqrt{\frac{k_{\rm B}T\,\kappa_T}{V}} \quad , \tag{4.138}$$

which again scales as $V^{-1/2}$.

4.6.5 Gibbs ensemble

Let the system's particle number N be fixed, but let it exchange energy and volume with the world W. *Mutatis mutandis*, we have

$$P_n = \lim_{\Delta E\to 0}\lim_{\Delta V\to 0}\frac{D_{\rm W}(E_{\rm U}-E_n\,,\,V_{\rm U}-V_n)\,\Delta E\,\Delta V}{D_{\rm U}(E_{\rm U},V_{\rm U})\,\Delta E\,\Delta V}. \tag{4.139}$$

Then

$$\begin{aligned}\ln P_n &= \ln D_{\rm W}(E_{\rm U}-E_n\,,\,V_{\rm U}-V_n) - \ln D_{\rm U}(E_{\rm U},V_{\rm U}) \\ &= \ln D_{\rm W}(E_{\rm U},V_{\rm U}) - \ln D_{\rm U}(E_{\rm U},V_{\rm U}) \\ &\qquad - E_n\,\frac{\partial \ln D_{\rm W}(E,V)}{\partial E}\bigg|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} - V_n\,\frac{\partial \ln D_{\rm W}(E,V)}{\partial V}\bigg|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} + \ldots \\ &\equiv -\alpha - \beta E_n - \beta p\,V_n \quad .\end{aligned} \tag{4.140}$$

The constants β and p are given by

$$\beta = \frac{\partial \ln D_{\rm W}(E,V)}{\partial E}\bigg|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} = \frac{1}{k_{\rm B}T} \tag{4.141}$$

$$p = k_{\rm B}T\,\frac{\partial \ln D_{\rm W}(E,V)}{\partial V}\bigg|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} \quad . \tag{4.142}$$

The corresponding partition function is

$$Y(T,p,N) = {\rm Tr}\;e^{-\beta(\hat H+pV)} = \frac{1}{V_0}\int\limits_0^\infty dV\,e^{-\beta pV}\,Z(T,V,N) \equiv e^{-\beta G(T,p,N)} \quad , \tag{4.143}$$

where V_0 is a constant which has dimensions of volume. The factor V_0^{-1} in front of the integral renders Y dimensionless. Note that $G(V_0') = G(V_0) + k_{\rm B}T\ln(V_0'/V_0)$, so the difference is not extensive and can be neglected in the thermodynamic limit. In other words, it doesn't matter what constant we choose for V_0 since it contributes subextensively to G. Moreover, in computing averages, the constant V_0 divides out in the ratio of numerator and denominator. Like the Helmholtz free energy, the Gibbs free energy $G(T,p,N)$ is also a double Legendre transform of the energy $E(S,V,N)$, *viz.*

$$\begin{aligned}G &= E - TS + pV \\ dG &= -S\,dT + V\,dp + \mu\,dN \quad ,\end{aligned} \tag{4.144}$$

which entails

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N} \quad , \quad V = +\left(\frac{\partial G}{\partial p}\right)_{T,N} \quad , \quad \mu = +\left(\frac{\partial G}{\partial N}\right)_{T,p} \quad . \tag{4.145}$$

4.7 Statistical Ensembles from Maximum Entropy

The basic principle: maximize the entropy,

$$S = -k_{\rm B} \sum_n P_n \ln P_n \quad . \tag{4.146}$$

4.7.1 μCE

We maximize S subject to the single constraint

$$C = \sum_n P_n - 1 = 0 \quad . \tag{4.147}$$

We implement the constraint $C = 0$ with a Lagrange multiplier, $\bar{\lambda} \equiv k_{\rm B}\,\lambda$, writing

$$S^* = S - k_{\rm B}\,\lambda\,C \quad , \tag{4.148}$$

and freely extremizing over the distribution $\{P_n\}$ and the Lagrange multiplier λ. Thus,

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B}\lambda\,\delta C - k_{\rm B}\,C\,\delta\lambda \\ &= -k_{\rm B} \sum_n \Big[\ln P_n + 1 + \lambda \Big] \delta P_n - k_{\rm B}\,C\,\delta\lambda \equiv 0 \quad . \end{aligned} \tag{4.149}$$

We conclude that $C = 0$ and that

$$\ln P_n = -\big(1 + \lambda\big) \quad , \tag{4.150}$$

and we fix λ by the normalization condition $\sum_n P_n = 1$. This gives $P_n = 1/\Omega$, with

$$\Omega = \sum_n \Theta(E + \Delta E - E_n)\,\Theta(E_n - E) \quad , \tag{4.151}$$

i.e. the total number of energy states lying in the interval $[E, E + \Delta E]$.

4.7.2 OCE

We maximize S subject to the two constraints

$$C_1 = \sum_n P_n - 1 = 0 \quad , \quad C_2 = \sum_n E_n\,P_n - E = 0 \quad . \tag{4.152}$$

We now have two Lagrange multipliers. We write

$$S^* = S - k_{\rm B} \sum_{j=1}^{2} \lambda_j \, C_j \quad , \tag{4.153}$$

and we freely extremize over $\{P_n\}$ and $\{C_j\}$. We therefore have

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B} \sum_n \big(\lambda_1 + \lambda_2 \, E_n\big) \, \delta P_n - k_{\rm B} \sum_{j=1}^{2} C_j \, \delta\lambda_j \\ &= -k_{\rm B} \sum_n \Big[\ln P_n + 1 + \lambda_1 + \lambda_2 \, E_n \Big] \delta P_n - k_{\rm B} \sum_{j=1}^{2} C_j \, \delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.154}$$

Thus, $C_1 = C_2 = 0$ and

$$\ln P_n = -\big(1 + \lambda_1 + \lambda_2 \, E_n\big) \quad . \tag{4.155}$$

We define $\lambda_2 \equiv \beta$ and we fix λ_1 by normalization. This yields

$$P_n = \frac{1}{Z} \, e^{-\beta E_n} \qquad , \qquad Z = \sum_n e^{-\beta E_n} = \mathsf{Tr} \; e^{-\beta \hat{H}} \quad . \tag{4.156}$$

4.7.3 GCE

We maximize S subject to the three constraints

$$C_1 = \sum_n P_n - 1 = 0 \quad , \quad C_2 = \sum_n E_n \, P_n - E = 0 \quad , \quad C_3 = \sum_n N_n \, P_n - N = 0 \quad . \tag{4.157}$$

We now have three Lagrange multipliers. We write

$$S^* = S - k_{\rm B} \sum_{j=1}^{3} \lambda_j \, C_j \quad , \tag{4.158}$$

and hence

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B} \sum_n \big(\lambda_1 + \lambda_2 \, E_n + \lambda_3 \, N_n\big) \, \delta P_n - k_{\rm B} \sum_{j=1}^{3} C_j \, \delta\lambda_j \\ &= -k_{\rm B} \sum_n \Big[\ln P_n + 1 + \lambda_1 + \lambda_2 \, E_n + \lambda_3 \, N_n \Big] \delta P_n - k_{\rm B} \sum_{j=1}^{3} C_j \, \delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.159}$$

Thus, $C_1 = C_2 = C_3 = 0$ and

$$\ln P_n = -\big(1 + \lambda_1 + \lambda_2 \, E_n + \lambda_3 \, N_n\big) \quad . \tag{4.160}$$

We define $\lambda_2 \equiv \beta$ and $\lambda_3 \equiv -\beta\mu$, and we fix λ_1 by normalization. This yields

$$P_n = \frac{1}{\Xi} \, e^{-\beta(E_n - \mu N_n)} \qquad , \qquad \Xi = \sum_n e^{-\beta(E_n - \mu N_n)} = \mathsf{Tr} \; e^{-\beta(\hat{H} - \mu \hat{N})} \quad . \tag{4.161}$$

4.8 Ideal Gas Statistical Mechanics

The ordinary canonical partition function for the ideal gas was computed in eqn. 4.14. We found

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\prod_{i=1}^{N}\int\frac{d^dx_i\,d^dp_i}{(2\pi\hbar)^d}\,e^{-\beta\boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\,e^{-\beta p^2/2m}\right)^{Nd} = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^N \quad , \end{aligned} \tag{4.162}$$

where λ_T is the *thermal wavelength*:

$$\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T} \quad . \tag{4.163}$$

The physical interpretation of λ_T is that it is the de Broglie wavelength for a particle of mass m which has a kinetic energy of $k_{\rm B}T$.

In the GCE, we have

$$\begin{aligned} \Xi(T,V,\mu) &= \sum_{N=0}^{\infty} e^{\beta\mu N}\,Z(T,V,N) \\ &= \sum_{N=1}^{\infty}\frac{1}{N!}\left(\frac{Ve^{\mu/k_{\rm B}T}}{\lambda_T^d}\right)^N = \exp\left(\frac{Ve^{\mu/k_{\rm B}T}}{\lambda_T^d}\right) \quad . \end{aligned} \tag{4.164}$$

From $\Xi = e^{-\Omega/k_{\rm B}T}$, we have the grand potential is

$$\Omega(T,V,\mu) = -Vk_{\rm B}T\,e^{\mu/k_{\rm B}T}/\lambda_T^d \quad . \tag{4.165}$$

Since $\Omega = -pV$ (see §4.6.2), we have

$$p(T,\mu) = k_{\rm B}T\,\lambda_T^{-d}\,e^{\mu/k_{\rm B}T} \quad . \tag{4.166}$$

The number density can also be calculated:

$$n = \frac{N}{V} = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = \lambda_T^{-d}\,e^{\mu/k_{\rm B}T} \quad . \tag{4.167}$$

Combined, the last two equations recapitulate the ideal gas law, $pV = Nk_{\rm B}T$.

4.8.1 Maxwell velocity distribution

The distribution function for momenta is given by

$$g(\boldsymbol{p}) = \left\langle\frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{p}_i - \boldsymbol{p})\right\rangle \quad . \tag{4.168}$$

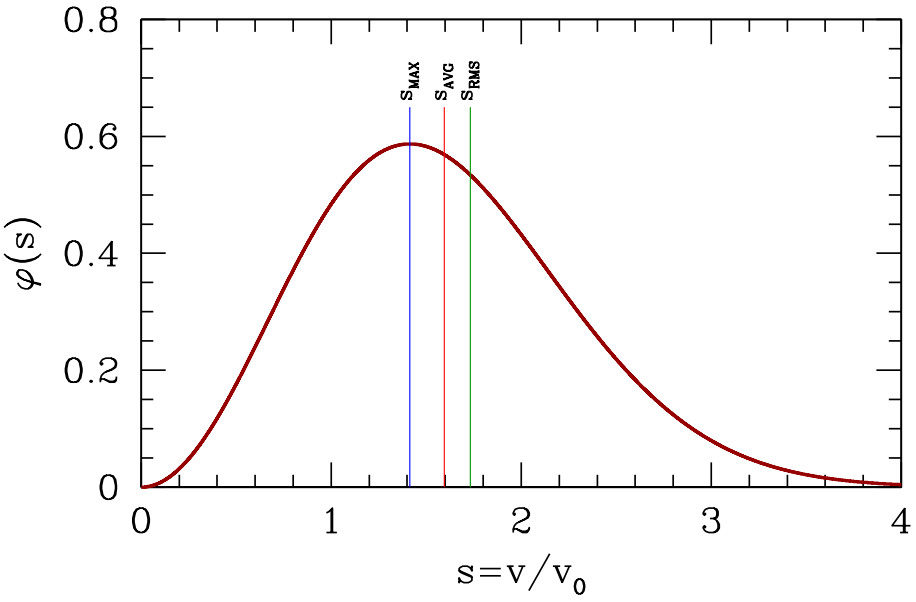

Figure 4.6: Maxwell distribution of speeds $\varphi(v/v_0)$. The most probable speed is $v_{\rm MAX} = \sqrt{2}\,v_0$. The average speed is $v_{\rm AVG} = \sqrt{\frac{8}{\pi}}\,v_0$. The RMS speed is $v_{\rm RMS} = \sqrt{3}\,v_0$.

Note that $g(\boldsymbol{p}) = \big\langle \delta(\boldsymbol{p}_i - \boldsymbol{p}) \big\rangle$ is the same for every particle, independent of its label i. We compute the average $\langle A \rangle = \mathsf{Tr}\,\left(A e^{-\beta \hat{H}}\right) / \mathsf{Tr}\, e^{-\beta \hat{H}}$. Setting $i = 1$, all the integrals other than that over $\boldsymbol{p}_1$ divide out between numerator and denominator. We then have

$$\begin{aligned} g(\boldsymbol{p}) &= \frac{\int d^3p_1\, \delta(\boldsymbol{p}_1 - \boldsymbol{p})\, e^{-\beta \boldsymbol{p}_1^2/2m}}{\int d^3p_1\, e^{-\beta \boldsymbol{p}_1^2/2m}} \\ &= (2\pi m k_{\rm B} T)^{-3/2}\, e^{-\beta \boldsymbol{p}^2/2m} \quad . \end{aligned} \tag{4.169}$$

Textbooks commonly refer to the *velocity distribution* $f(\boldsymbol{v})$, which is related to $g(\boldsymbol{p})$ by

$$f(\boldsymbol{v})\, d^3v = g(\boldsymbol{p})\, d^3p \quad . \tag{4.170}$$

Hence,

$$f(\boldsymbol{v}) = \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} e^{-m\boldsymbol{v}^2/2k_{\rm B}T} \quad . \tag{4.171}$$

This is known as the *Maxwell velocity distribution*. Note that the distributions are normalized, *viz.*

$$\int d^3p\, g(\boldsymbol{p}) = \int d^3v\, f(\boldsymbol{v}) = 1 \quad . \tag{4.172}$$

If we are only interested in averaging functions of $v = |\boldsymbol{v}|$ which are isotropic, then we can define the *Maxwell speed distribution*, $\tilde{f}(v)$, as

$$\tilde{f}(v) = 4\pi\, v^2 f(\boldsymbol{v}) = 4\pi \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} v^2\, e^{-mv^2/2k_{\rm B}T} \quad . \tag{4.173}$$

Note that $\tilde{f}(v)$ is normalized according to $\int_0^\infty dv\, \tilde{f}(v) = 1$. It is convenient to represent v in units of $v_0 = \sqrt{k_{\rm B}T/m}$, in which case

$$\tilde{f}(v) = \frac{1}{v_0}\,\varphi(v/v_0) \qquad , \qquad \varphi(s) = \sqrt{\tfrac{2}{\pi}}\, s^2\, e^{-s^2/2} \quad . \tag{4.174}$$

The distribution $\varphi(s)$ is shown in fig. 4.6. Computing averages, we have

$$C_k \equiv \langle s^k \rangle = \int_0^\infty ds\; s^k\, \varphi(s) = 2^{k/2} \cdot \frac{2}{\sqrt{\pi}}\, \Gamma\left(\tfrac{3}{2} + \tfrac{k}{2}\right) \quad . \tag{4.175}$$

Thus, $C_0 = 1$, $C_1 = \sqrt{\frac{8}{\pi}}$, $C_2 = 3$, *etc.* The speed averages are

$$\langle v^k \rangle = C_k \left(\frac{k_{\rm B}T}{m}\right)^{k/2} \quad . \tag{4.176}$$

Note that the average *velocity* is $\langle \boldsymbol{v} \rangle = 0$, but the average *speed* is $\langle v \rangle = \sqrt{8k_{\rm B}T/\pi m}$. The speed distribution is plotted in fig. 4.6.

4.8.2 Equipartition

The Hamiltonian for ballistic (*i.e.* massive nonrelativistic) particles is quadratic in the individual components of each momentum $\boldsymbol{p}_i$. There are other cases in which a classical degree of freedom appears quadratically in $\hat{H}$ as well. For example, an individual normal mode ξ of a system of coupled oscillators has the Lagrangian

$$L = \tfrac{1}{2}\dot{\xi}^2 - \tfrac{1}{2}\,\omega_0^2\,\xi^2 \quad , \tag{4.177}$$

where the dimensions of ξ are $[\xi] = M^{1/2}L$ by convention. The Hamiltonian for this normal mode is then

$$\hat{H} = \frac{p^2}{2} + \tfrac{1}{2}\,\omega_0^2\,\xi^2 \quad , \tag{4.178}$$

from which we see that both the kinetic as well as potential energy terms enter quadratically into the Hamiltonian. The classical rotational kinetic energy is also quadratic in the angular momentum components.

Let us compute the contribution of a single quadratic degree of freedom in $\hat{H}$ to the partition function. We'll call this degree of freedom ζ – it may be a position or momentum or angular momentum – and we'll write its contribution to $\hat{H}$ as $\hat{H}_\zeta = \frac{1}{2}K\zeta^2$, where K is some constant. Integrating over ζ yields the following factor in the partition function:

$$\int_{-\infty}^{\infty} d\zeta\; e^{-\beta K\zeta^2/2} = \left(\frac{2\pi}{K\beta}\right)^{1/2} \quad . \tag{4.179}$$

The contribution to the Helmholtz free energy is then

$$\Delta F_\zeta = \tfrac{1}{2} k_{\rm B} T \ln\left(\frac{K}{2\pi k_{\rm B} T}\right) \quad , \tag{4.180}$$

and therefore the contribution to the internal energy E is

$$\Delta E_\zeta = \frac{\partial}{\partial\beta}\big(\beta\,\Delta F_\zeta\big) = \frac{1}{2\beta} = \tfrac{1}{2} k_{\rm B} T \quad . \tag{4.181}$$

We have thus derived what is commonly called the *equipartition theorem* of classical statistical mechanics:

> *To each degree of freedom which enters the Hamiltonian quadratically is associated a contribution $\frac{1}{2}k_{\rm B}T$ to the internal energy of the system. This results in a concomitant contribution of $\frac{1}{2}k_{\rm B}$ to the heat capacity.*

We now see why the internal energy of a classical ideal gas with f degrees of freedom per molecule is $E = \frac{1}{2} f N k_{\rm B} T$, and $C_V = \frac{1}{2} N k_{\rm B}$. This result also has applications in the theory of solids. The atoms in a solid possess kinetic energy due to their motion, and potential energy due to the spring-like interatomic potentials which tend to keep the atoms in their preferred crystalline positions. Thus, for a three-dimensional crystal, there are *six* quadratic degrees of freedom (three positions and three momenta) per atom, and the classical energy should be $E = 3N k_{\rm B} T$, and the heat capacity $C_V = 3N k_{\rm B}$. As we shall see, quantum mechanics modifies this result considerably at temperatures below the highest normal mode (*i.e.* phonon) frequency, but the high temperature limit is given by the classical value $C_V = 3\nu R$ (where $\nu = N/N_{\rm A}$ is the number of moles) derived here, known as the *Dulong-Petit limit*.

4.9 Selected Examples

4.9.1 Spins in an external magnetic field

Consider a system of $N_{\rm s}$ spins , each of which can be either up ($\sigma = +1$) or down ($\sigma = -1$). The Hamiltonian for this system is

$$\hat{H} = -\mu_0 H \sum_{j=1}^{N_{\rm s}} \sigma_j \quad , \tag{4.182}$$

where now we write $\hat{H}$ for the Hamiltonian, to distinguish it from the external magnetic field H, and μ_0 is the magnetic moment per particle. We treat this system within the ordinary canonical ensemble. The partition function is

$$Z = \sum_{\sigma_1} \cdots \sum_{\sigma_{N_{\rm s}}} e^{-\beta\hat{H}} = \zeta^{N_{\rm s}} \quad , \tag{4.183}$$

where ζ is the single particle partition function:

$$\zeta = \sum_{\sigma=\pm 1} e^{\mu_0 H\sigma/k_{\rm B}T} = 2\cosh\left(\frac{\mu_0 H}{k_{\rm B} T}\right) \quad . \tag{4.184}$$

The Helmholtz free energy is then

$$F(T,H,N_{\rm s}) = -k_{\rm B}T\ln Z = -N_{\rm s}\,k_{\rm B}T\ln\left[2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.185}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,\,N_{\rm s}} = N_{\rm s}\,\mu_0\tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.186}$$

The energy is

$$E = \frac{\partial}{\partial\beta}\left(\beta F\right) = -N_{\rm s}\,\mu_0 H\tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.187}$$

Hence, $E = -HM$, which we already knew, from the form of $\hat{H}$ itself.

Each spin here is independent. The probability that a given spin has polarization σ is

$$P_\sigma = \frac{e^{\beta\mu_0 H\sigma}}{e^{\beta\mu_0 H} + e^{-\beta\mu_0 H}} \quad . \tag{4.188}$$

The total probability is unity, *i.e.* $P_\uparrow + P_\downarrow = 1$, and the average polarization is a weighted average of $\sigma = +1$ and $\sigma = -1$ contributions:

$$\langle\sigma\rangle = P_\uparrow - P_\downarrow = \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.189}$$

At low temperatures $T \ll \mu_0 H/k_{\rm B}$, we have $P_\uparrow \approx 1 - e^{-2\mu_0 H/k_{\rm B}T}$. At high temperatures $T > \mu_0 H/k_{\rm B}$, the two polarizations are equally likely, and $P_\sigma \approx \frac{1}{2}\left(1 + \frac{\sigma\mu_0 H}{k_{\rm B}T}\right)$.

The *isothermal magnetic susceptibility* is defined as

$$\chi_T = \frac{1}{N_{\rm s}}\left(\frac{\partial M}{\partial H}\right)_T = \frac{\mu_0^2}{k_{\rm B}T}\,\mathrm{sech}^2\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.190}$$

(Typically this is computed per unit volume rather than per particle.) At $H = 0$, we have $\chi_T = \mu_0^2/k_{\rm B}T$, which is known as the *Curie law*.

Aside

The energy $E = -HM$ here is not the same quantity we discussed in our study of thermodynamics. In fact, the thermodynamic energy for this problem vanishes! Here is why. To avoid confusion, we'll need to invoke a new symbol for the thermodynamic energy, $\mathcal{E}$. Recall that the thermodynamic energy $\mathcal{E}$ is a function of extensive quantities, meaning $\mathcal{E} = \mathcal{E}(S,M,N_{\rm s})$. It is obtained from the free energy $F(T,H,N_{\rm s})$ by a double Legendre transform:

$$\mathcal{E}(S,M,N_{\rm s}) = F(T,H,N_{\rm s}) + TS + HM \quad . \tag{4.191}$$

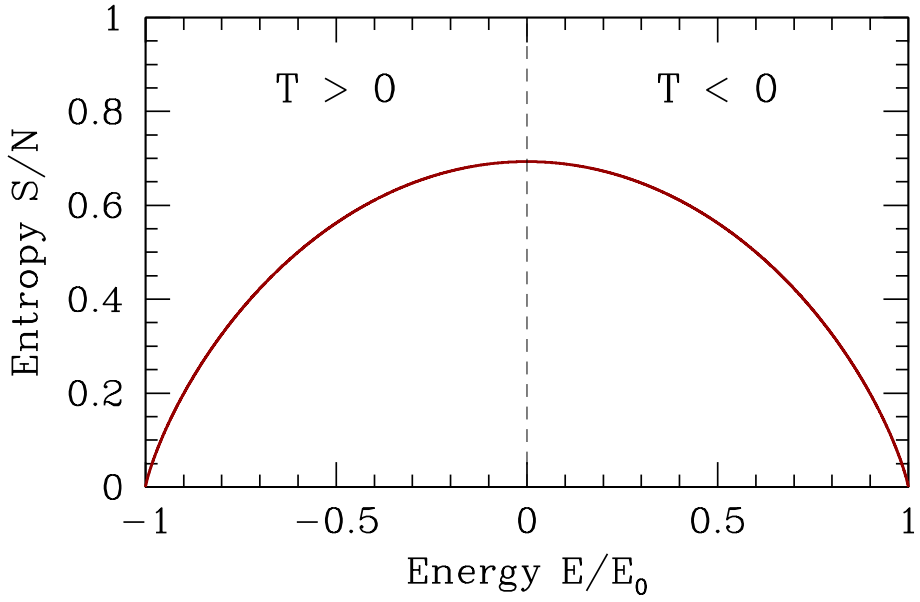

Figure 4.7: When entropy decreases with increasing energy, the temperature is negative. Typically, kinetic degrees of freedom prevent this peculiarity from manifesting in physical systems.

Now from eqn. 4.185 we derive the entropy

$$S = -\frac{\partial F}{\partial T} = N_{\rm s}\, k_{\rm B} \ln\left[2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right)\right] - N_{\rm s}\,\frac{\mu_0 H}{T}\,\tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.192}$$

Thus, using eqns. 4.185 and 4.186, we obtain $\mathcal{E}(S, M, N_{\rm s}) = 0$.

The potential confusion here arises from our use of the expression $F(T, H, N_{\rm s})$. In thermodynamics, it is the Gibbs free energy $G(T, p, N)$ which is a double Legendre transform of the energy: $G = \mathcal{E} - TS + pV$. By analogy, with magnetic systems we should perhaps write $G = \mathcal{E} - TS - HM$, but in keeping with many textbooks we shall use the symbol F and refer to it as the Helmholtz free energy. The quantity we've called E in eqn. 4.187 is in fact $E = \mathcal{E} - HM$, which means $\mathcal{E} = 0$. The energy $\mathcal{E}(S, M, N_{\rm s})$ vanishes here because the spins are noninteracting.

4.9.2 Negative temperature (!)

Consider again a system of $N_{\rm s}$ spins, each of which can be either up $(+)$ or down $(-)$. Let N_σ be the number of sites with spin σ, where $\sigma = \pm 1$. Clearly $N_+ + N_- = N_{\rm s}$. We now treat this system within the microcanonical ensemble.

The energy of the system is $E = -HM$, where H is an external magnetic field, and $M = (N_+ - N_-)\,\mu_0$ is the total magnetization. We now compute $S(E)$ using the ordinary canonical ensemble. The number of ways of arranging the system with N_+ up spins is

$$\Omega = \binom{N_{\rm s}}{N_+} \quad , \tag{4.193}$$

hence the entropy is

$$S = k_{\rm B} \ln \Omega = -N_{\rm s}\, k_{\rm B} \Big\{ x \ln x + (1-x) \ln(1-x) \Big\} \tag{4.194}$$

in the thermodynamic limit: $N_{\rm s} \to \infty$, $N_+ \to \infty$, $x = N_+/N_{\rm s}$ constant. Now the magnetization is $M = (N_+ - N_-)\mu_0 = (2N_+ - N_{\rm s})\mu_0$, hence if we define the maximum energy $E_0 \equiv N_{\rm s}\,\mu_0 H$, then

$$\frac{E}{E_0} = -\frac{M}{N_{\rm s}\,\mu_0} = 1 - 2x \qquad \Longrightarrow \qquad x = \frac{E_0 - E}{2E_0} \quad . \tag{4.195}$$

We therefore have

$$S(E, N_{\rm s}) = -N_{\rm s}\, k_{\rm B} \left[\left(\frac{E_0 - E}{2E_0} \right) \ln\left(\frac{E_0 - E}{2E_0} \right) + \left(\frac{E_0 + E}{2E_0} \right) \ln\left(\frac{E_0 + E}{2E_0} \right) \right] \quad . \tag{4.196}$$

We now have

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N_{\rm s}} = \frac{\partial S}{\partial x}\,\frac{\partial x}{\partial E} = \frac{N_{\rm s}\, k_{\rm B}}{2E_0} \ln\left(\frac{E_0 - E}{E_0 + E} \right) \quad . \tag{4.197}$$

We see that the temperature is positive for $-E_0 \le E < 0$ and is *negative* for $0 < E \le E_0$.

What has gone wrong? The answer is that *nothing* has gone wrong – all our calculations are perfectly correct. This system *does* exhibit the possibility of negative temperature. It is, however, unphysical in that we have neglected kinetic degrees of freedom, which result in an entropy function $S(E, N_{\rm s})$ which is an increasing function of energy. In this system, $S(E, N_{\rm s})$ achieves a maximum of $S_{\rm max} = N_{\rm s}\, k_{\rm B} \ln 2$ at $E = 0$ (*i.e.* $x = \frac{1}{2}$), and then turns over and starts decreasing. In fact, our results are completely consistent with eqn. 4.187 : the energy E is an odd function of temperature. Positive energy requires negative temperature! Another example of this peculiarity is provided in the appendix in §4.11.2.

4.9.3 Adsorption

PROBLEM: A surface containing $N_{\rm s}$ adsorption sites is in equilibrium with a monatomic ideal gas. Atoms adsorbed on the surface have an energy $-\Delta$ and no kinetic energy. Each adsorption site can accommodate at most one atom. Calculate the fraction f of occupied adsorption sites as a function of the gas density n, the temperature T, the binding energy Δ, and physical constants.

The grand partition function for the surface is

$$\begin{aligned} \Xi_{\rm surf} = e^{-\Omega_{\rm surf}/k_{\rm B}T} &= \sum_{j=0}^{N_{\rm s}} \binom{N_{\rm s}}{j} e^{j(\mu+\Delta)/k_{\rm B}T} \\ &= \left(1 + e^{\mu/k_{\rm B}T}\, e^{\Delta/k_{\rm B}T}\right)^{N_{\rm s}} \quad . \end{aligned} \tag{4.198}$$

The fraction of occupied sites is

$$f = \frac{\langle \hat{N}_{\rm surf} \rangle}{N_{\rm s}} = -\frac{1}{N_{\rm s}}\,\frac{\partial \Omega_{\rm surf}}{\partial \mu} = \frac{e^{\mu/k_{\rm B}T}}{e^{\mu/k_{\rm B}T} + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.199}$$

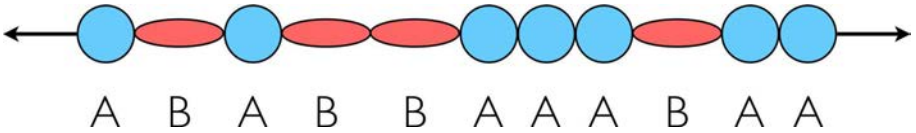

Figure 4.8: The monomers in wool are modeled as existing in one of two states. The low energy undeformed state is A, and the higher energy deformed state is B. Applying tension induces more monomers to enter the B state.

Since the surface is in equilibrium with the gas, its fugacity $z = \exp(\mu/k_{\rm B}T)$ and temperature T are the same as in the gas.

SOLUTION: For a monatomic ideal gas, the single particle partition function is $\zeta = V\lambda_T^{-3}$, where $\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T}$ is the thermal wavelength. Thus, the grand partition function, for indistinguishable particles, is

$$\Xi_{\rm gas} = \exp\left(V\lambda_T^{-3}\, e^{\mu/k_{\rm B}T}\right) \quad . \tag{4.200}$$

The gas density is

$$n = \frac{\langle \hat{N}_{\rm gas}\rangle}{V} = -\frac{1}{V}\,\frac{\partial\Omega_{\rm gas}}{\partial\mu} = \lambda_T^{-3}\, e^{\mu/k_{\rm B}T} \quad . \tag{4.201}$$

We can now solve for the fugacity: $z = e^{\mu/k_{\rm B}T} = n\lambda_T^3$. Thus, the fraction of occupied adsorption sites is

$$f = \frac{n\lambda_T^3}{n\lambda_T^3 + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.202}$$

Interestingly, the solution for f involves the constant $\hbar$.

It is always advisable to check that the solution makes sense in various limits. First of all, if the gas density tends to zero at fixed T and Δ, we have $f \to 0$. On the other hand, if $n \to \infty$ we have $f \to 1$, which also makes sense. At fixed n and T, if the adsorption energy is $(-\Delta) \to -\infty$, then once again $f = 1$ since every adsorption site wants to be occupied. Conversely, taking $(-\Delta) \to +\infty$ results in $n \to 0$, since the energetic cost of adsorption is infinitely high.

4.9.4 Elasticity of wool

Wool consists of interlocking protein molecules which can stretch into an elongated configuration, but reversibly so. This feature gives wool its very useful elasticity. Let us model a chain of these proteins by assuming they can exist in one of two states, which we will call A and B, with energies $\varepsilon_{\rm A}$ and $\varepsilon_{\rm B}$ and lengths $\ell_{\rm A}$ and $\ell_{\rm B}$. The situation is depicted in fig. 4.8. We model these conformational degrees of freedom by a spin variable $\sigma = \pm 1$ for each molecule, where $\sigma = +1$ in the A state and $\sigma = -1$ in the B state. Suppose a chain consisting of N monomers is placed under a tension τ. We then have

$$\hat{H} = \sum_{j=1}^{N}\left[\varepsilon_{\rm A}\,\delta_{\sigma_j,+1} + \varepsilon_{\rm B}\,\delta_{\sigma_j,-1}\right] \quad . \tag{4.203}$$

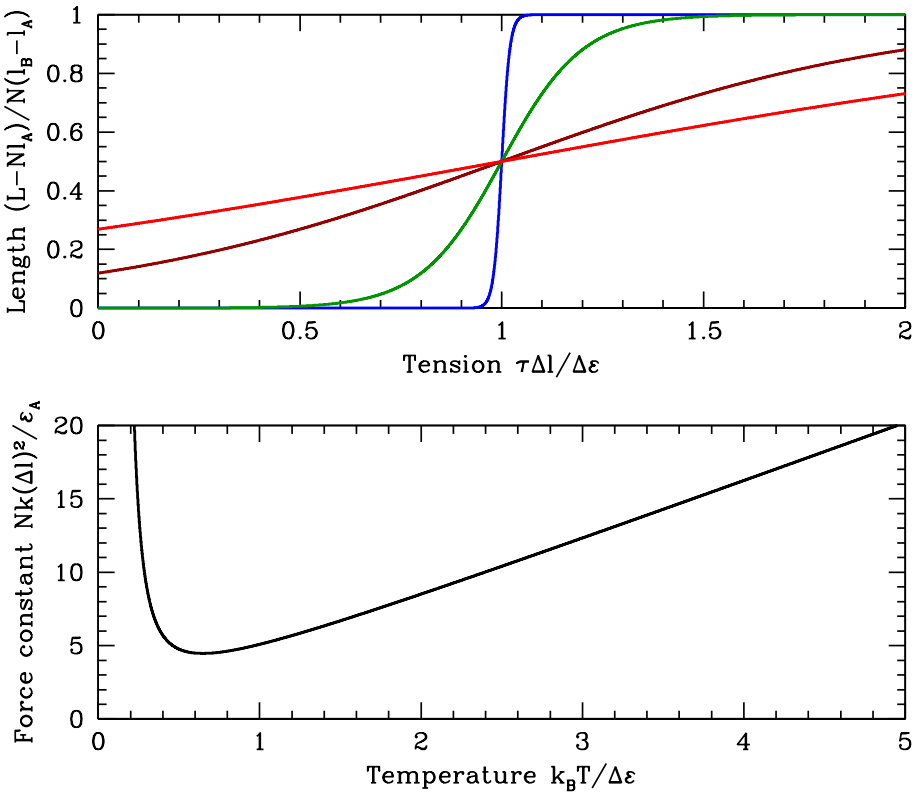

Figure 4.9: Upper panel: length $L(\tau, T)$ for $k_{\rm B}T/\tilde{\varepsilon} = 0.01$ (blue), 0.1 (green), 0.5 (dark red), and 1.0 (red). Bottom panel: dimensionless force constant $k/N(\Delta\ell)^2$ *versus* temperature.

Similarly, the length is

$$\hat{L} = \sum_{j=1}^{N} \Big[\ell_{\rm A}\, \delta_{\sigma_j,+1} + \ell_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.204}$$

The Gibbs partition function is $Y = {\rm Tr}\; e^{-\hat{K}/k_{\rm B}T}$, with $\hat{K} = \hat{H} - \tau\hat{L}$:

$$\hat{K} = \sum_{j=1}^{N} \Big[\tilde{\varepsilon}_{\rm A}\, \delta_{\sigma_j,+1} + \tilde{\varepsilon}_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad , \tag{4.205}$$

where $\tilde{\varepsilon}_{\rm A} \equiv \varepsilon_{\rm A} - \tau\ell_{\rm A}$ and $\tilde{\varepsilon}_{\rm B} \equiv \varepsilon_{\rm B} - \tau\ell_{\rm B}$. At $\tau = 0$ the A state is preferred for each monomer, but when τ exceeds τ^*, defined by the relation $\tilde{\varepsilon}_{\rm A} = \tilde{\varepsilon}_{\rm B}$, the B state is preferred. One finds

$$\tau^* = \frac{\varepsilon_{\rm B} - \varepsilon_{\rm A}}{\ell_{\rm B} - \ell_{\rm A}} \quad . \tag{4.206}$$

Once again, we have a set of N noninteracting spins. The partition function is $Y = \zeta^N$, where ζ is the single monomer partition function, $\zeta = {\rm Tr}\; e^{-\beta\hat{h}}$, where

$$\hat{h} = \tilde{\varepsilon}_{\rm A}\, \delta_{\sigma_j,1} + \tilde{\varepsilon}_{\rm B}\, \delta_{\sigma_j,-1} \tag{4.207}$$

is the single "spin" Hamiltonian. Thus,

$$\zeta = \mathrm{Tr}\; e^{-\beta\hat{h}} = e^{-\beta\tilde{\varepsilon}_{\rm A}} + e^{-\beta\tilde{\varepsilon}_{\rm B}} \quad , \tag{4.208}$$

It is convenient to define the differences

$$\Delta\varepsilon = \varepsilon_{\rm B} - \varepsilon_{\rm A} \quad , \quad \Delta\ell = \ell_{\rm B} - \ell_{\rm A} \quad , \quad \Delta\tilde{\varepsilon} = \tilde{\varepsilon}_{\rm B} - \tilde{\varepsilon}_{\rm A} \tag{4.209}$$

in which case the partition function Y is

$$Y(T,\tau,N) = e^{-N\beta\,\tilde{\varepsilon}_{\rm A}}\Big[1 + e^{-\beta\Delta\tilde{\varepsilon}}\Big]^N \tag{4.210}$$

$$G(T,\tau,N) = N\tilde{\varepsilon}_{\rm A} - Nk_{\rm B}T\ln\Big[1 + e^{-\Delta\tilde{\varepsilon}/k_{\rm B}T}\Big] \tag{4.211}$$

The average length is

$$\begin{aligned} L = \langle\hat{L}\rangle &= -\left(\frac{\partial G}{\partial\tau}\right)_{T,N} \\ &= N\ell_{\rm A} + \frac{N\Delta\ell}{e^{(\Delta\varepsilon-\tau\Delta\ell)/k_{\rm B}T}+1} \quad . \end{aligned} \tag{4.212}$$

The polymer behaves as a spring, and for small τ the spring constant is

$$k = \left.\frac{\partial\tau}{\partial L}\right|_{\tau=0} = \frac{4k_{\rm B}T}{N(\Delta\ell)^2}\,\cosh^2\!\left(\frac{\Delta\varepsilon}{2k_{\rm B}T}\right) \quad . \tag{4.213}$$

The results are shown in fig. 4.9. Note that length increases with temperature for $\tau < \tau^*$ and decreases with temperature for $\tau > \tau^*$. Note also that k diverges at both low and high temperatures. At low T, the energy gap $\Delta\varepsilon$ dominates and $L = N\ell_{\rm A}$, while at high temperatures $k_{\rm B}T$ dominates and $L = \frac{1}{2}N(\ell_{\rm A}+\ell_{\rm B})$.

4.9.5 Noninteracting spin dimers

Consider a system of noninteracting spin dimers as depicted in fig. 4.10. Each dimer contains two spins, and is described by the Hamiltonian

$$\hat{H}_{\rm dimer} = -J\,\sigma_1\sigma_2 - \mu_0 H\,(\sigma_1+\sigma_2) \quad . \tag{4.214}$$

Here, J is an *interaction energy* between the spins which comprise the dimer. If $J > 0$ the interaction is *ferromagnetic*, which prefers that the spins are aligned. That is, the lowest energy states are $|{\uparrow\uparrow}\rangle$ and $|{\downarrow\downarrow}\rangle$. If $J < 0$ the interaction is *antiferromagnetic*, which prefers that spins be anti-aligned: $|{\uparrow\downarrow}\rangle$ and $|{\downarrow\uparrow}\rangle$.[9]

[9] *Nota bene* we are concerned with classical spin configurations only – there is no superposition of states allowed in this model!

Suppose there are $N_{\rm d}$ dimers. Then the OCE partition function is $Z = \zeta^{N_{\rm d}}$, where $\zeta(T,H)$ is the single dimer partition function. To obtain $\zeta(T,H)$, we sum over the four possible states of the two spins, obtaining

$$\begin{aligned}\zeta &= \mathrm{Tr}\, e^{-\hat{H}_{\rm dimer}/k_{\rm B}T} \\ &= 2\, e^{-J/k_{\rm B}T} + 2\, e^{J/k_{\rm B}T} \cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right) \quad .\end{aligned}$$

Thus, the free energy is

$$F(T,H,N_{\rm d}) = -N_{\rm d}\, k_{\rm B}T \ln 2 - N_{\rm d}\, k_{\rm B}T \ln\left[e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T}\cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.215}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,N_{\rm d}} = 2N_{\rm d}\,\mu_0 \cdot \frac{e^{J/k_{\rm B}T}\sinh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)}{e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T}\cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)} \tag{4.216}$$

It is instructive to consider the zero field isothermal susceptibility per spin,

$$\chi_T = \frac{1}{2N_{\rm d}}\left.\frac{\partial M}{\partial H}\right|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T}\cdot\frac{2\, e^{J/k_{\rm B}T}}{e^{J/k_{\rm B}T} + e^{-J/k_{\rm B}T}} \quad . \tag{4.217}$$

The quantity $\mu_0^2/k_{\rm B}T$ is simply the Curie susceptibility for noninteracting classical spins. Note that we correctly recover the Curie result when $J = 0$, since then the individual spins comprising each dimer are in fact noninteracting. For the ferromagnetic case, if $J \gg k_{\rm B}T$, then we obtain

$$\chi_T(J \gg k_{\rm B}T) \approx \frac{2\mu_0^2}{k_{\rm B}T} \quad . \tag{4.218}$$

This has the following simple interpretation. When $J \gg k_{\rm B}T$, the spins of each dimer are effectively locked in parallel. Thus, each dimer has an effective magnetic moment $\mu_{\rm eff} = 2\mu_0$. On the other hand, there are only half as many dimers as there are spins, so the resulting Curie susceptibility per spin is $\frac{1}{2} \times (2\mu_0)^2/k_{\rm B}T$.

When $-J \gg k_{\rm B}T$, the spins of each dimer are effectively locked in one of the two antiparallel configurations. We then have

$$\chi_T(-J \gg k_{\rm B}T) \approx \frac{2\mu_0^2}{k_{\rm B}T}\, e^{-2|J|/k_{\rm B}T} \quad . \tag{4.219}$$

In this case, the individual dimers have essentially zero magnetic moment.

4.10 Statistical Mechanics of Molecular Gases

4.10.1 Separation of translational and internal degrees of freedom

The states of a noninteracting atom or molecule are labeled by its total momentum $\boldsymbol{p}$ and its internal quantum numbers, which we will simply write with a collective index α, specifying rotational, vibra-

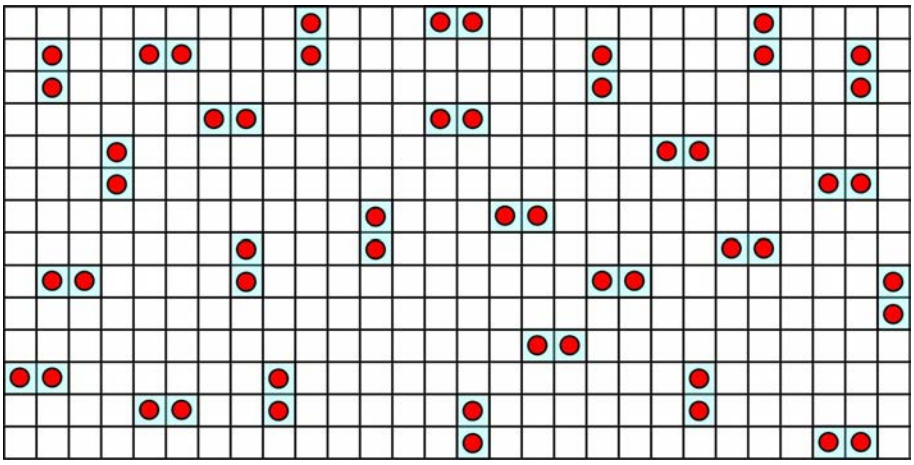

Figure 4.10: A model of noninteracting spin dimers on a lattice. Each red dot represents a classical spin for which $\sigma_j = \pm 1$.

tional, and electronic degrees of freedom. The single particle Hamiltonian is then

$$\hat{h} = \frac{\boldsymbol{p}^2}{2m} + \hat{h}_{\rm int} \quad , \tag{4.220}$$

with

$$\hat{h} \, \big| \, \boldsymbol{k} \, , \, \alpha \, \big\rangle = \left(\frac{\hbar^2 \boldsymbol{k}^2}{2m} + \varepsilon_\alpha \right) \big| \, \boldsymbol{k} \, , \, \alpha \, \big\rangle \quad . \tag{4.221}$$

The partition function is

$$\zeta = {\rm Tr} \; e^{-\beta \hat{h}} = \sum_{\boldsymbol{p}} e^{-\beta \boldsymbol{p}^2/2m} \sum_j g_j \, e^{-\beta \varepsilon_j} \quad . \tag{4.222}$$

Here we have replaced the internal label α with a label j of energy eigenvalues, with g_j being the degeneracy of the internal state with energy ε_j. To do the $\boldsymbol{p}$ sum, we quantize in a box of dimensions $L_1 \times L_2 \times \cdots \times L_d$, using periodic boundary conditions. Then

$$\boldsymbol{p} = \left(\frac{2\pi\hbar n_1}{L_1} \, , \, \frac{2\pi\hbar n_2}{L_2} \, , \, \ldots \, , \, \frac{2\pi\hbar n_d}{L_d} \right) \quad , \tag{4.223}$$

where each n_i is an integer. Since the differences between neighboring quantized $\boldsymbol{p}$ vectors are very tiny, we can replace the sum over $\boldsymbol{p}$ by an integral:

$$\sum_{\boldsymbol{p}} \longrightarrow \int \frac{d^d p}{\Delta p_1 \cdots \Delta p_d} \tag{4.224}$$

where the volume in momentum space of an elementary rectangle is

$$\Delta p_1 \cdots \Delta p_d = \frac{(2\pi\hbar)^d}{L_1 \cdots L_d} = \frac{(2\pi\hbar)^d}{V} \quad . \tag{4.225}$$

Thus,

$$\zeta = V\int\frac{d^dp}{(2\pi\hbar)^d}\,e^{-\boldsymbol{p}^2/2mk_{\rm B}T}\,\sum_j g_j\,e^{-\varepsilon_j/k_{\rm B}T} = V\lambda_T^{-d}\,\xi \tag{4.226}$$

$$\xi(T) = \sum_j g_j\,e^{-\varepsilon_j/k_{\rm B}T} \quad . \tag{4.227}$$

Here, $\xi(T)$ is the *internal coordinate partition function*. The full N-particle ordinary canonical partition function is then

$$Z_N = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^N \xi^N(T) \quad . \tag{4.228}$$

Using Stirling's approximation, we find the Helmholtz free energy $F = -k_{\rm B}T\ln Z$ is

$$\begin{aligned} F(T,V,N) &= -Nk_{\rm B}T\left[\ln\left(\frac{V}{N\lambda_T^d}\right) + 1 + \ln\xi(T)\right] \\ &= -Nk_{\rm B}T\left[\ln\left(\frac{V}{N\lambda_T^d}\right) + 1\right] + N\varphi(T) \quad , \end{aligned} \tag{4.229}$$

where

$$\varphi(T) = -k_{\rm B}T\ln\xi(T) \tag{4.230}$$

is the internal coordinate contribution to the single particle free energy. We could also compute the partition function in the Gibbs (T,p,N) ensemble:

$$\begin{aligned} Y(T,p,N) = e^{-\beta G(T,p,N)} &= \frac{1}{V_0}\int_0^\infty dV\;e^{-\beta pV}\;Z(T,V,N) \\ &= \left(\frac{k_{\rm B}T}{pV_0}\right)\left(\frac{k_{\rm B}T}{p\,\lambda_T^d}\right)^N\xi^N(T) \quad . \end{aligned} \tag{4.231}$$

Thus, in the thermodynamic limit,

$$\begin{aligned} \mu(T,p) = \frac{G(T,p,N)}{N} &= k_{\rm B}T\ln\left(\frac{p\,\lambda_T^d}{k_{\rm B}T}\right) - k_{\rm B}T\ln\xi(T) \\ &= k_{\rm B}T\ln\left(\frac{p\,\lambda_T^d}{k_{\rm B}T}\right) + \varphi(T) \quad . \end{aligned} \tag{4.232}$$

4.10.2 Ideal gas law

Since the internal coordinate contribution to the free energy is volume-independent, we have

$$V = \left(\frac{\partial G}{\partial p}\right)_{T,N} = \frac{Nk_{\rm B}T}{p} \quad , \tag{4.233}$$

and the ideal gas law applies. The entropy is

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N} = Nk_{\rm B}\left[\ln\left(\frac{k_{\rm B}T}{p\lambda_T^d}\right) + 1 + \tfrac{1}{2}d\right] - N\varphi'(T) \quad , \tag{4.234}$$

and therefore the heat capacity is

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_{p,N} = \left(\tfrac{1}{2}d+1\right)Nk_{\rm B} - NT\,\varphi''(T) \tag{4.235}$$

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_{V,N} = \tfrac{1}{2}dNk_{\rm B} - NT\,\varphi''(T) \quad . \tag{4.236}$$

Thus, any temperature variation in C_p must be due to the internal degrees of freedom.

4.10.3 The internal coordinate partition function

At energy scales of interest we can separate the internal degrees of freedom into distinct classes, writing

$$\hat{h}_{\rm int} = \hat{h}_{\rm rot} + \hat{h}_{\rm vib} + \hat{h}_{\rm elec} \tag{4.237}$$

as a sum over internal Hamiltonians governing rotational, vibrational, and electronic degrees of freedom. Then

$$\xi_{\rm int} = \xi_{\rm rot} \cdot \xi_{\rm vib} \cdot \xi_{\rm elec} \quad . \tag{4.238}$$

Associated with each class of excitation is a characteristic temperature Θ. Rotational and vibrational temperatures of a few common molecules are listed in table tab. 4.1.

4.10.4 Rotations

Consider a class of molecules which can be approximated as an axisymmetric top. The rotational Hamiltonian is then

$$\begin{aligned}\hat{h}_{\rm rot} &= \frac{\mathsf{L}_{\rm a}^2 + \mathsf{L}_{\rm b}^2}{2I_1} + \frac{\mathsf{L}_{\rm c}^2}{2I_3} \\ &= \frac{\hbar^2 L(L+1)}{2I_1} + \left(\frac{1}{2I_3} - \frac{1}{2I_1}\right)\mathsf{L}_c^2 \quad ,\end{aligned} \tag{4.239}$$

where $\hat{\boldsymbol{n}}_{\rm a,b,c}(t)$ are the principal axes, with $\hat{\boldsymbol{n}}_{\rm c}$ the symmetry axis, and $\mathsf{L}_{\rm a,b,c}$ are the components of the angular momentum vector $\boldsymbol{L}$ about these instantaneous body-fixed principal axes. The components of $\boldsymbol{L}$ along *space-fixed* axes $\{x, y, z\}$ are written as $L^{x,y,z}$. Note that

$$\left[L^\mu\,,\,\mathsf{L}_c\right] = n_{\rm c}^\nu\left[L^\mu\,,\,L^\nu\right] + \left[L^\mu\,,\,n_{\rm c}^\nu\right]L^\nu = i\epsilon_{\mu\nu\lambda}\,n_{\rm c}^\nu\,L^\lambda + i\epsilon_{\mu\nu\lambda}\,n_{\rm c}^\lambda\,L^\nu = 0 \quad , \tag{4.240}$$

molecule	$\Theta_{\rm rot}$(K)	$\Theta_{\rm vib}$(K)
H_2	85.4	6100
N_2	2.86	3340
H_2O	13.7 , 21.0 , 39.4	2290 , 5180 , 5400

Table 4.1: Some rotational and vibrational temperatures of common molecules.

which is equivalent to the statement that $\mathsf{L}_{\rm c} = \hat{\boldsymbol{n}}_{\rm c} \cdot \boldsymbol{L}$ is a rotational scalar. We can therefore simultaneously specify the eigenvalues of $\{\boldsymbol{L}^2, L^z, \mathsf{L}_{\rm c}\}$, which form a complete set of commuting observables (CSCO)[10]. The eigenvalues of L^z are $m\hbar$ with $m \in \{-L, \ldots, L\}$, while those of $\mathsf{L}_{\rm c}$ are $k\hbar$ with $k \in \{-L, \ldots, L\}$. There is a $(2L+1)$-fold degeneracy associated with the L^z quantum number.

We assume the molecule is prolate, so that $I_3 < I_1$. We can the define two temperature scales,

$$\Theta = \frac{\hbar^2}{2I_1 k_{\rm B}} \quad , \quad \widetilde{\Theta} = \frac{\hbar^2}{2I_3 k_{\rm B}} \quad . \tag{4.241}$$

Prolateness then means $\widetilde{\Theta} > \Theta$. We conclude that the rotational partition function for an axisymmetric molecule is given by

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty} (2L+1)\, e^{-L(L+1)\,\Theta/T} \sum_{k=-L}^{L} e^{-k^2\,(\widetilde{\Theta}-\Theta)/T} \tag{4.242}$$

In diatomic molecules, I_3 is extremely small, and $\widetilde{\Theta} \gg k_{\rm B}T$ at all relevant temperatures. Only the $k = 0$ term contributes to the partition sum, and we have

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty} (2L+1)\, e^{-L(L+1)\,\Theta/T} \quad . \tag{4.243}$$

When $T \ll \Theta$, only the first few terms contribute, and

$$\xi_{\rm rot}(T) = 1 + 3\, e^{-2\Theta/T} + 5\, e^{-6\Theta/T} + \ldots \tag{4.244}$$

In the high temperature limit, we have a slowly varying summand. The *Euler-MacLaurin summation formula* may be used to evaluate such a series:

$$\sum_{k=0}^{n} F_k = \int_0^n dk\, F(k) + \tfrac{1}{2}\big[F(0) + F(n)\big] + \sum_{j=1}^{\infty} \frac{B_{2j}}{(2j)!} \Big[F^{(2j-1)}(n) - F^{(2j-1)}(0)\Big] \tag{4.245}$$

where B_j is the $j^{\rm th}$ *Bernoulli number* where

$$B_0 = 1 \quad , \quad B_1 = -\tfrac{1}{2} \quad , \quad B_2 = \tfrac{1}{6} \quad , \quad B_4 = -\tfrac{1}{30} \quad , \quad B_6 = \tfrac{1}{42} \quad . \tag{4.246}$$

[10]Note that while we cannot simultaneously specify the eigenvalues of two components of $\boldsymbol{L}$ along axes fixed in space, we can simultaneously specify the components of $\boldsymbol{L}$ along one axis fixed in space and one axis rotating with a body. See Landau and Lifshitz, *Quantum Mechanics*, §103.

Thus,

$$\sum_{k=0}^{\infty} F_k = \int_0^\infty dx\, F(x) + \tfrac{1}{2}F(0) - \tfrac{1}{12}F'(0) - \frac{1}{720}\,F'''(0) + \dots \quad . \tag{4.247}$$

We have $F(x) = (2x+1)\,e^{-x(x+1)\Theta/T}$, for which $\int_0^\infty dx\, F(x) = \frac{T}{\Theta}$, hence

$$\xi_{\rm rot} = \frac{T}{\Theta} + \frac{1}{3} + \frac{1}{15}\,\frac{\Theta}{T} + \frac{4}{315}\left(\frac{\Theta}{T}\right)^2 + \dots \quad . \tag{4.248}$$

Recall that $\varphi(T) = -k_{\rm B}T\ln\xi(T)$. We conclude that $\varphi_{\rm rot}(T) \approx -3k_{\rm B}T\,e^{-2\Theta/T}$ for $T \ll \Theta$ and $\varphi_{\rm rot}(T) \approx -k_{\rm B}T\ln(T/\Theta)$ for $T \gg \Theta$. We have seen that the internal coordinate contribution to the heat capacity is $\Delta C_V = -NT\varphi''(T)$. For diatomic molecules, then, this contribution is exponentially suppressed for $T \ll \Theta$, while for high temperatures we have $\Delta C_V = Nk_{\rm B}$. One says that the rotational excitations are 'frozen out' at temperatures much below Θ. Including the first few terms, we have

$$\Delta C_V(T \ll \Theta) = 12\,Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2 e^{-2\Theta/T} + \dots \tag{4.249}$$

$$\Delta C_V(T \gg \Theta) = Nk_{\rm B}\left\{1 + \frac{1}{45}\left(\frac{\Theta}{T}\right)^2 + \frac{16}{945}\left(\frac{\Theta}{T}\right)^3 + \dots\right\} \quad . \tag{4.250}$$

Note that C_V overshoots its limiting value of $Nk_{\rm B}$ and asymptotically approaches it from above.

Special care must be taken in the case of homonuclear diatomic molecules, for then only even or odd L states are allowed, depending on the total nuclear spin. This is discussed below in §4.10.7.

For polyatomic molecules, the moments of inertia generally are large enough that the molecule's rotations can be considered classically. We then have

$$\varepsilon(\mathsf{L}_{\rm a}, \mathsf{L}_{\rm b}, \mathsf{L}_{\rm c}) = \frac{\mathsf{L}_{\rm a}^2}{2I_1} + \frac{\mathsf{L}_{\rm b}^2}{2I_2} + \frac{\mathsf{L}_{\rm c}^2}{2I_3} \quad . \tag{4.251}$$

We then have

$$\xi_{\rm rot}(T) = \frac{1}{g_{\rm rot}} \int \frac{d\mathsf{L}_{\rm a}\, d\mathsf{L}_{\rm b}\, d\mathsf{L}_{\rm c}\, d\phi\, d\theta\, d\psi}{(2\pi\hbar)^3}\, e^{-\varepsilon(\mathsf{L}_{\rm a}\,\mathsf{L}_{\rm b}\,\mathsf{L}_{\rm c})/k_{\rm B}T} \quad , \tag{4.252}$$

where $(\phi, \theta\,\psi)$ are the Euler angles. Recall $\phi \in [0, 2\pi]$, $\theta \in [0, \pi]$, and $\psi \in [0, 2\pi]$. The factor $g_{\rm rot}$ accounts for physically indistinguishable orientations of the molecule brought about by rotations, which can happen when more than one of the nuclei is the same. We then have

$$\xi_{\rm rot}(T) = \left(\frac{2k_{\rm B}T}{\hbar^2}\right)^{3/2} \sqrt{\pi I_1 I_2 I_3} \quad . \tag{4.253}$$

This leads to $\Delta C_V = \frac{3}{2}Nk_{\rm B}$.

4.10.5 Vibrations

Vibrational frequencies are often given in units of inverse wavelength, such as cm^{-1}, called a *wavenumber*. To convert to a temperature scale T^*, we write $k_{\rm B}T^* = h\nu = hc/\lambda$, hence $T^* = (hc/k_{\rm B})\,\lambda^{-1}$, and we multiply by

$$\frac{hc}{k_{\rm B}} = 1.436\,\mathrm{K}\cdot\mathrm{cm} \quad . \tag{4.254}$$

For example, infrared absorption (~ 50 cm^{-1} to 10^4 cm^{-1}) reveals that the 'asymmetric stretch' mode of the H_2O molecule has a vibrational frequency of $\nu = 3756\,\mathrm{cm}^{-1}$. The corresponding temperature scale is $T^* = 5394\,\mathrm{K}$.

Vibrations are normal modes of oscillations. A single normal mode Hamiltonian is of the form

$$\hat{h} = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 q^2 = \hbar\omega\big(a^\dagger a + \tfrac{1}{2}\big) \quad . \tag{4.255}$$

In general there are many vibrational modes, hence many normal mode frequencies ω_α. We then must sum over all of them, resulting in

$$\xi_{\rm vib} = \prod_\alpha \xi_{\rm vib}^{(\alpha)} \quad . \tag{4.256}$$

For each such normal mode, the contribution is

$$\begin{aligned}\xi &= \sum_{n=0}^\infty e^{-(n+\frac{1}{2})\hbar\omega/k_{\rm B}T} = e^{-\hbar\omega/2k_{\rm B}T}\sum_{n=0}^\infty \Big(e^{-\hbar\omega/k_{\rm B}T}\Big)^n \\ &= \frac{e^{-\hbar\omega/2k_{\rm B}T}}{1-e^{-\hbar\omega/k_{\rm B}T}} = \frac{1}{2\sinh(\Theta/2T)} \quad ,\end{aligned} \tag{4.257}$$

where $\Theta = \hbar\omega/k_{\rm B}$. Then

$$\begin{aligned}\varphi &= k_{\rm B}T\ln\Big(2\sinh(\Theta/2T)\Big) \\ &= \tfrac{1}{2}k_{\rm B}\Theta + k_{\rm B}T\ln\big(1-e^{-\Theta/T}\big) \quad .\end{aligned} \tag{4.258}$$

The contribution to the heat capacity is

$$\begin{aligned}\Delta C_V &= Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2\frac{e^{\Theta/T}}{(e^{\Theta/T}-1)^2} \\ &= \begin{cases} Nk_{\rm B}\,(\Theta/T)^2\,\exp(-\Theta/T) & (T\to 0) \\ Nk_{\rm B} & (T\to\infty)\end{cases}\end{aligned} \tag{4.259}$$

4.10.6 Two-level systems : Schottky anomaly

Consider now a two-level system, with energies ε_0 and ε_1. We define $\Delta \equiv \varepsilon_1 - \varepsilon_0$ and assume without loss of generality that $\Delta > 0$. The partition function is

$$\zeta = e^{-\beta\varepsilon_0} + e^{-\beta\varepsilon_1} = e^{-\beta\varepsilon_0}\big(1+e^{-\beta\Delta}\big) \quad . \tag{4.260}$$

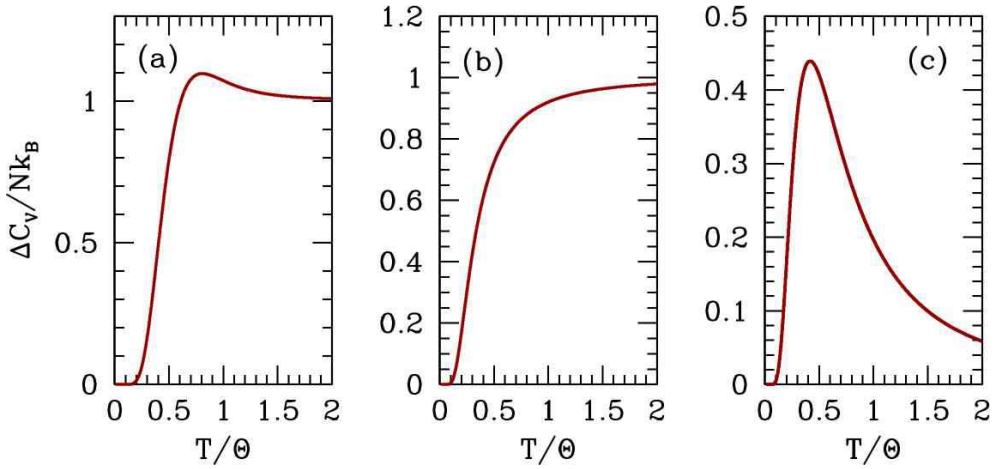

Figure 4.11: Heat capacity per molecule as a function of temperature for (a) heteronuclear diatomic gases, (b) a single vibrational mode, and (c) a single two-level system.

The free energy is

$$f = -k_{\rm B}T \ln \zeta = \varepsilon_0 - k_{\rm B}T \ln\left(1 + e^{-\Delta/k_{\rm B}T}\right) \quad . \tag{4.261}$$

The entropy for a given two level system is then

$$s = -\frac{\partial f}{\partial T} = k_{\rm B} \ln\left(1 + e^{-\Delta/k_{\rm B}T}\right) + \frac{\Delta}{T} \cdot \frac{1}{e^{\Delta/k_{\rm B}T} + 1} \tag{4.262}$$

and the heat capacity is $= T\,(\partial s/\partial T)$, *i.e.*

$$c(T) = \frac{\Delta^2}{k_{\rm B}T^2} \cdot \frac{e^{\Delta/k_{\rm B}T}}{\left(e^{\Delta/k_{\rm B}T} + 1\right)^2} \quad . \tag{4.263}$$

Thus,

$$c\,(T \ll \Delta) = \frac{\Delta^2}{k_{\rm B}T^2}\, e^{-\Delta/k_{\rm B}T} \tag{4.264}$$

$$c\,(T \gg \Delta) = \frac{\Delta^2}{4k_{\rm B}T^2} \quad . \tag{4.265}$$

We find that $c(T)$ has a characteristic peak at $T^* \approx 0.42\,\Delta/k_{\rm B}$. The heat capacity vanishes in both the low temperature and high temperature limits. At low temperatures, the gap to the excited state is much greater than $k_{\rm B}T$, and it is not possible to populate it and store energy. At high temperatures, both ground state and excited state are equally populated, and once again there is no way to store energy.

If we have a distribution of independent two-level systems, the heat capacity of such a system is a sum over the individual Schottky functions:

$$C(T) = \sum_i \widetilde{c}\,(\Delta_i/k_{\rm B}T) = N \int_0^\infty d\Delta\, P(\Delta)\, \widetilde{c}(\Delta/T) \quad , \tag{4.266}$$

where N is the number of two level systems, $\widetilde{c}(x) = k_{\rm B}\, x^2\, e^x/(e^x+1)^2$, and where $P(\Delta)$ is the normalized distribution function, which satisfies the normalization condition

$$\int_0^\infty d\Delta\, P(\Delta) = 1 \quad . \tag{4.267}$$

$N_{\rm s}$ is the total number of two level systems. If $P(\Delta) \propto \Delta^r$ for $\Delta \to 0$, then the low temperature heat capacity behaves as $C(T) \propto T^{1+r}$. Many amorphous or glassy systems contain such a distribution of two level systems, with $r \approx 0$ for glasses, leading to a linear low-temperature heat capacity. The origin of these two-level systems is not always so clear but is generally believed to be associated with local atomic configurations for which there are two low-lying states which are close in energy. The paradigmatic example is the mixed crystalline solid $({\rm KBr})_{1-x}({\rm KCN})_x$ which over the range $0.1 \lesssim x \lesssim 0.6$ forms an 'orientational glass' at low temperatures. The two level systems are associated with different orientation of the cyanide (CN) dipoles.

4.10.7 Electronic and nuclear excitations

For a monatomic gas, the internal coordinate partition function arises due to electronic and nuclear degrees of freedom. Let's first consider the electronic degrees of freedom. We assume that $k_{\rm B}T$ is small compared with energy differences between successive electronic shells. The atomic ground state is then computed by filling up the hydrogenic orbitals until all the electrons are used up. If the atomic number is a 'magic number' ($A = 2$ (He), 10 (Ne), 18 (Ar), 36 (Kr), 54 (Xe), *etc.*) then the atom has all shells filled and $L = 0$ and $S = 0$. Otherwise the last shell is partially filled and one or both of L and S will be nonzero. The atomic ground state configuration ${}^{2J+1}L_S$ is then determined by *Hund's rules*:

1. The LS multiplet with the *largest* S has the lowest energy.
2. If the largest value of S is associated with several multiplets, the multiplet with the *largest* L has the lowest energy.
3. If an incomplete shell is not more than half-filled, then the lowest energy state has $J = |L - S|$. If the shell is more than half-filled, then $J = L + S$.

The last of Hund's rules distinguishes between the $(2S+1)(2L+1)$ states which result upon fixing S and L as per rules #1 and #2. It arises due to the atomic spin-orbit coupling, whose effective Hamiltonian may be written $\hat{H} = \Lambda \boldsymbol{L} \cdot \boldsymbol{S}$, where Λ is the Russell-Saunders coupling. If the last shell is less than or equal to half-filled, then $\Lambda > 0$ and the ground state has $J = |L - S|$. If the last shell is more than half-filled, the coupling is *inverted*, *i.e.* $\Lambda < 0$, and the ground state has $J = L + S$.[11]

The electronic contribution to ξ is then

$$\xi_{\rm elec} = \sum_{J=|L-S|}^{L+S} (2J+1)\, e^{-\Delta\varepsilon(L,S,J)/k_{\rm B}T} \tag{4.268}$$

[11] See *e.g.* §72 of Landau and Lifshitz, *Quantum Mechanics*, which, in my humble estimation, is the greatest physics book ever written.

where

$$\Delta\varepsilon(L,S,J) = \tfrac{1}{2}\Lambda\Big[J(J+1) - L(L+1) - S(S+1)\Big] \quad . \tag{4.269}$$

At high temperatures, $k_{\rm B}T$ is larger than the energy difference between the different J multiplets, and we have $\xi_{\rm elec} \sim (2L+1)(2S+1)\,e^{-\beta\varepsilon_0}$, where ε_0 is the ground state energy. At low temperatures, a particular value of J is selected – that determined by Hund's third rule – and we have $\xi_{\rm elec} \sim (2J+1)\,e^{-\beta\varepsilon_0}$. If, in addition, there is a nonzero nuclear spin I, then we also must include a factor $\xi_{\rm nuc} = (2I+1)$, neglecting the small hyperfine splittings due to the coupling of nuclear and electronic angular momenta.

For *heteronuclear* diatomic molecules, *i.e.* molecules composed from two different atomic nuclei, the internal partition function simply receives a factor of $\xi_{\rm elec} \cdot \xi^{(1)}_{\rm nuc} \cdot \xi^{(2)}_{\rm nuc}$, where the first term is a sum over molecular electronic states, and the second two terms arise from the spin degeneracies of the two nuclei. For *homonuclear* diatomic molecules, the exchange of nuclear centers is a symmetry operation, and does not represent a distinct quantum state. To correctly count the electronic states, we first assume that the total electronic spin is $S = 0$. This is generally a very safe assumption. Exchange symmetry now puts restrictions on the possible values of the molecular angular momentum L, depending on the total nuclear angular momentum $I_{\rm tot}$. If $I_{\rm tot}$ is even, then the molecular angular momentum L must also be even. If the total nuclear angular momentum is odd, then L must be odd. This is so because the molecular ground state configuration is ${}^1\Sigma_g^+$.[12]

The total number of nuclear states for the molecule is $(2I+1)^2$, of which some are even under nuclear exchange, and some are odd. The number of even states, corresponding to even total nuclear angular momentum is written as g_g, where the subscript conventionally stands for the (mercifully short) German word *gerade*, meaning 'even'. The number of odd (Ger. *ungerade*) states is written g_u. Table 4.2 gives the values of $g_{g,u}$ corresponding to half-odd-integer I and integer I.

The final answer for the rotational component of the internal molecular partition function is then

$$\xi_{\rm rot}(T) = g_g\,\zeta_g + g_u\,\zeta_u \quad , \tag{4.270}$$

where

$$\begin{aligned} \zeta_g &= \sum_{L\ \text{even}} (2L+1)\,e^{-L(L+1)\,\Theta_{\rm rot}/T} \\ \zeta_u &= \sum_{L\ \text{odd}} (2L+1)\,e^{-L(L+1)\,\Theta_{\rm rot}/T} \quad . \end{aligned} \tag{4.271}$$

For hydrogen, the molecules with the larger nuclear statistical weight are called *orthohydrogen* and those with the smaller statistical weight are called *parahydrogen*. For H_2, we have $I = \frac{1}{2}$ hence the ortho state has $g_u = 3$ and the para state has $g_g = 1$. In D_2, we have $I = 1$ and the ortho state has $g_g = 6$ while the para state has $g_u = 3$. In equilibrium, the ratio of ortho to para states is then

$$\frac{N^{\rm ortho}_{\rm H_2}}{N^{\rm para}_{\rm H_2}} = \frac{g_u\,\zeta_u}{g_g\,\zeta_g} = \frac{3\,\zeta_u}{\zeta_g} \qquad , \qquad \frac{N^{\rm ortho}_{\rm D_2}}{N^{\rm para}_{\rm D_2}} = \frac{g_g\,\zeta_g}{g_u\,\zeta_u} = \frac{2\,\zeta_g}{\zeta_u} \quad . \tag{4.272}$$

[12] See Landau and Lifshitz, *Quantum Mechanics*, §86.

$2I$	g_g	g_u
odd	$I(2I+1)$	$(I+1)(2I+1)$
even	$(I+1)(2I+1)$	$I(2I+1)$

Table 4.2: Number of even (g_g) and odd (g_u) total nuclear angular momentum states for a homonuclear diatomic molecule. I is the ground state nuclear spin.

Incidentally, how do we derive the results in Tab. 4.2 ? The total nuclear angular momentum $I_{\rm tot}$ is the quantum mechanical sum of the two individual nuclear angular momenta, each of which are of magnitude I. From elementary addition of angular momenta, we have

$$I \otimes I = 0 \oplus 1 \oplus 2 \oplus \cdots \oplus 2I \quad . \tag{4.273}$$

The right hand side of the above equation lists all the possible multiplets. Thus, $I_{\rm tot} \in \{0, 1, \ldots, 2I\}$. Now let us count the total number of states with *even* $I_{\rm tot}$. If $2I$ is even, which is to say if I is an integer, we have

$$g_g^{(2I={\rm even})} = \sum_{n=0}^{I} \Big\{ 2 \cdot (2n) + 1 \Big\} = (I+1)(2I+1) \quad , \tag{4.274}$$

because the degeneracy of each multiplet is $2I_{\rm tot} + 1$. It follows that

$$g_u^{(2I={\rm even})} = (2I+1)^2 - g_g = I(2I+1) \quad . \tag{4.275}$$

On the other hand, if $2I$ is odd, which is to say I is a half odd integer, then

$$g_g^{(2I={\rm odd})} = \sum_{n=0}^{I-\frac{1}{2}} \Big\{ 2 \cdot (2n) + 1 \Big\} = I(2I+1) \quad . \tag{4.276}$$

It follows that

$$g_u^{(2I={\rm odd})} = (2I+1)^2 - g_g = (I+1)(2I+1) \quad . \tag{4.277}$$

4.11 Appendix I : Additional Examples

4.11.1 Three state system

Consider a spin-1 particle where $\sigma = -1, 0, +1$. We model this with the single particle Hamiltonian

$$\hat{h} = -\mu_0 H \, \sigma + \Delta(1 - \sigma^2) \quad . \tag{4.278}$$

We can also interpret this as describing a spin if $\sigma = \pm 1$ and a vacancy if $\sigma = 0$. The parameter Δ then represents the vacancy formation energy. The single particle partition function is

$$\zeta = {\rm Tr}\, e^{-\beta \hat{h}} = e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H) \quad . \tag{4.279}$$

With $N_{\rm s}$ distinguishable noninteracting spins (*e.g.* at different sites in a crystalline lattice), we have $Z = \zeta^{N_{\rm s}}$ and

$$F \equiv N_{\rm s} f = -k_{\rm B} T \ln Z = -N_{\rm s}\, k_{\rm B} T \ln \Big[e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H) \Big] \quad , \tag{4.280}$$

where $f = -k_{\rm B} T \ln \zeta$ is the free energy of a single particle. Note that

$$\hat{n}_{\rm V} = 1 - \sigma^2 = \frac{\partial \hat{h}}{\partial \Delta} \tag{4.281}$$

$$\hat{m} = \mu_0\, \sigma = -\frac{\partial \hat{h}}{\partial H} \tag{4.282}$$

are the vacancy number and magnetization, respectively. Thus,

$$n_{\rm V} = \big\langle\, \hat{n}_{\rm V} \,\big\rangle = \frac{\partial f}{\partial \Delta} = \frac{e^{-\Delta/k_{\rm B}T}}{e^{-\Delta/k_{\rm B}T} + 2\cosh(\mu_0 H/k_{\rm B}T)} \tag{4.283}$$

and

$$m = \big\langle\, \hat{m} \,\big\rangle = -\frac{\partial f}{\partial H} = \frac{2\mu_0\, \sinh(\mu_0 H/k_{\rm B}T)}{e^{-\Delta/k_{\rm B}T} + 2\cosh(\mu_0 H/k_{\rm B}T)} \quad . \tag{4.284}$$

At weak fields we can compute

$$\chi_T = \left.\frac{\partial m}{\partial H}\right|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T} \cdot \frac{2}{2 + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.285}$$

We thus obtain a modified Curie law. At temperatures $T \ll \Delta/k_{\rm B}$, the vacancies are frozen out and we recover the usual Curie behavior. At high temperatures, where $T \gg \Delta/k_{\rm B}$, the low temperature result is reduced by a factor of $\frac{2}{3}$, which accounts for the fact that one third of the time the particle is in a nonmagnetic state with $\sigma = 0$.

4.11.2 Spins and vacancies on a surface

PROBLEM: A collection of spin-$\frac{1}{2}$ particles is confined to a surface with N sites. For each site, let $\sigma = 0$ if there is a vacancy, $\sigma = +1$ if there is particle present with spin up, and $\sigma = -1$ if there is a particle present with spin down. The particles are non-interacting, and the energy for each site is given by $\varepsilon = -W\sigma^2$, where $-W < 0$ is the binding energy.

(a) Let $Q = N_\uparrow + N_\downarrow$ be the number of spins, and N_0 be the number of vacancies. The surface magnetization is $M = N_\uparrow - N_\downarrow$. Compute, in the microcanonical ensemble, the statistical entropy $S(Q, M)$.

(b) Let $q = Q/N$ and $m = M/N$ be the dimensionless particle density and magnetization density, respectively. Assuming that we are in the thermodynamic limit, where N, Q, and M all tend to infinity, but with q and m finite, Find the temperature $T(q, m)$. Recall Stirling's formula

$$\ln(N!) = N \ln N - N + \mathcal{O}(\ln N) \quad .$$

(c) Show explicitly that T can be negative for this system. What does negative T mean? What physical degrees of freedom have been left out that would avoid this strange property?

SOLUTION: There is a constraint on $N_\uparrow$, N_0, and $N_\downarrow$:

$$N_\uparrow + N_0 + N_\downarrow = Q + N_0 = N \quad . \tag{4.286}$$

The total energy of the system is $E = -WQ$.

(a) The number of states available to the system is

$$\Omega = \frac{N!}{N_\uparrow!\, N_0!\, N_\downarrow!} \quad . \tag{4.287}$$

Fixing Q and M, along with the above constraint, is enough to completely determine $\{N_\uparrow, N_0, N_\downarrow\}$:

$$N_\uparrow = \tfrac{1}{2}\,(Q+M) \qquad , \qquad N_0 = N - Q \qquad , \qquad N_\downarrow = \tfrac{1}{2}\,(Q-M) \quad , \tag{4.288}$$

whence

$$\Omega(Q,M) = \frac{N!}{\big[\frac{1}{2}(Q+M)\big]!\,\big[\frac{1}{2}(Q-M)\big]!\,(N-Q)!} \quad . \tag{4.289}$$

The statistical entropy is $S = k_{\rm B} \ln \Omega$:

$$S(Q,M) = k_{\rm B} \ln(N!) - k_{\rm B} \ln\big[\tfrac{1}{2}(Q+M)!\big] - k_{\rm B} \ln\big[\tfrac{1}{2}(Q-M)!\big] - k_{\rm B} \ln\big[(N-Q)!\big] \quad . \tag{4.290}$$

(b) Now we invoke Stirling's rule,

$$\ln(N!) = N \ln N - N + \mathcal{O}(\ln N) \quad , \tag{4.291}$$

to obtain

$$\begin{aligned} \ln \Omega(Q,M) = N \ln N - N &- \tfrac{1}{2}(Q+M) \ln\big[\tfrac{1}{2}(Q+M)\big] + \tfrac{1}{2}(Q+M) \\ &- \tfrac{1}{2}(Q-M) \ln\big[\tfrac{1}{2}(Q-M)\big] + \tfrac{1}{2}(Q-M) \\ &- (N-Q)\ln(N-Q) + (N-Q) \end{aligned} \tag{4.292}$$

$$= N \ln N - \tfrac{1}{2} Q \ln\Big[\tfrac{1}{4}(Q^2 - M^2)\Big] - \tfrac{1}{2} M \ln\left(\frac{Q+M}{Q-M}\right)$$

Combining terms,

$$\ln \Omega(Q,M) = -Nq \ln\Big[\tfrac{1}{2}\sqrt{q^2 - m^2}\Big] - \tfrac{1}{2} Nm \ln\left(\frac{q+m}{q-m}\right) - N(1-q)\ln(1-q) \quad , \tag{4.293}$$

where $Q = Nq$ and $M = Nm$. Note that the entropy $S = k_{\rm B} \ln \Omega$ is extensive. The statistical entropy per site is thus

$$s(q,m) = -k_{\rm B}\, q \ln\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] - \tfrac{1}{2} k_{\rm B}\, m \ln\left(\frac{q+m}{q-m}\right) - k_{\rm B}\,(1-q)\ln(1-q) \quad . \tag{4.294}$$

The temperature is obtained from the relation

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial E}\right)_M = \frac{1}{W}\left(\frac{\partial s}{\partial q}\right)_m \\ &= \frac{1}{W}\,\ln(1-q) - \frac{1}{W}\,\ln\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] \quad . \end{aligned} \tag{4.295}$$

Thus,

$$T = \frac{W/k_{\rm B}}{\ln\bigl[2(1-q)/\sqrt{q^2-m^2}\bigr]} \quad . \tag{4.296}$$

(c) We have $0 \le q \le 1$ and $-q \le m \le q$, so T is real (thank heavens!). But it is easy to choose $\{q, m\}$ such that $T < 0$. For example, when $m = 0$ we have $T = W/k_{\rm B} \ln(2q^{-1} - 2)$ and $T < 0$ for all $q \in \left(\frac{2}{3}, 1\right]$. The reason for this strange state of affairs is that the entropy S is bounded, and is not an monotonically increasing function of the energy E (or the dimensionless quantity Q). The entropy is maximized for $N_\uparrow = N_0 = N_\downarrow = \frac{1}{3}$, which says $m = 0$ and $q = \frac{2}{3}$. Increasing q beyond this point (with $m = 0$ fixed) starts to reduce the entropy, and hence $(\partial S/\partial E) < 0$ in this range, which immediately gives $T < 0$. What we've left out are kinetic degrees of freedom, such as vibrations and rotations, whose energies are unbounded, and which result in an increasing $S(E)$ function.

4.11.3 Fluctuating interface

Consider an interface between two dissimilar fluids. In equilibrium, in a uniform gravitational field, the denser fluid is on the bottom. Let $z = z(x, y)$ be the height the interface between the fluids, relative to equilibrium. The potential energy is a sum of gravitational and surface tension terms, with

$$U_{\rm grav} = \int d^2x \int_0^z dz'\, \Delta\rho\, g\, z' \tag{4.297}$$

$$U_{\rm surf} = \tfrac{1}{2}\sigma \int d^2x\, (\boldsymbol{\nabla} z)^2 \quad . \tag{4.298}$$

We won't need the kinetic energy in our calculations, but we can include it just for completeness. It isn't so clear how to model it *a priori* so we will assume a rather general form

$$T = \int d^2x \int d^2x'\, \tfrac{1}{2}\mu(\boldsymbol{x}, \boldsymbol{x}')\, \frac{\partial z(\boldsymbol{x}, t)}{\partial t}\, \frac{\partial z(\boldsymbol{x}', t)}{\partial t} \quad . \tag{4.299}$$

We assume that the (x, y) plane is a rectangle of dimensions $L_x \times L_y$. We also assume $\mu(\boldsymbol{x}, \boldsymbol{x}') = \mu\bigl(|\boldsymbol{x} - \boldsymbol{x}'|\bigr)$. We can then Fourier transform

$$z(\boldsymbol{x}) = \bigl(L_x\, L_y\bigr)^{-1/2} \sum_{\boldsymbol{k}} z_{\boldsymbol{k}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \quad , \tag{4.300}$$

where the wavevectors $\boldsymbol{k}$ are quantized according to

$$\boldsymbol{k} = \frac{2\pi n_x}{L_x}\,\hat{\boldsymbol{x}} + \frac{2\pi n_y}{L_y}\,\hat{\boldsymbol{y}} \quad , \tag{4.301}$$

with integer n_x and n_y, if we impose periodic boundary conditions (for calculational convenience). The Lagrangian is then

$$L = \frac{1}{2}\sum_{\boldsymbol{k}} \Big[\mu_{\boldsymbol{k}} \left| \dot{z}_{\boldsymbol{k}} \right|^2 - \left(g\,\Delta\rho + \sigma \boldsymbol{k}^2 \right) \left| z_{\boldsymbol{k}} \right|^2 \Big] \quad , \tag{4.302}$$

where

$$\mu_{\boldsymbol{k}} = \int d^2x\, \mu\big(|\boldsymbol{x}|\big)\, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \quad . \tag{4.303}$$

Since $z(\boldsymbol{x},t)$ is real, we have the relation $z_{-\boldsymbol{k}} = z_{\boldsymbol{k}}^*$, therefore the Fourier coefficients at $\boldsymbol{k}$ and $-\boldsymbol{k}$ are not independent. The canonical momenta are given by

$$p_{\boldsymbol{k}} = \frac{\partial L}{\partial \dot{z}_{\boldsymbol{k}}^*} = \mu_{\boldsymbol{k}}\,\dot{z}_{\boldsymbol{k}} \qquad , \qquad p_{\boldsymbol{k}}^* = \frac{\partial L}{\partial \dot{z}_{\boldsymbol{k}}} = \mu_{\boldsymbol{k}}\,\dot{z}_{\boldsymbol{k}}^* \tag{4.304}$$

The Hamiltonian is then

$$\hat{H} = {\sum_{\boldsymbol{k}}}' \Big[p_{\boldsymbol{k}}\, z_{\boldsymbol{k}}^* + p_{\boldsymbol{k}}^*\, z_{\boldsymbol{k}} \Big] - L \tag{4.305}$$

$$= {\sum_{\boldsymbol{k}}}' \left[\frac{|p_{\boldsymbol{k}}|^2}{\mu_{\boldsymbol{k}}} + \left(g\,\Delta\rho + \sigma\boldsymbol{k}^2 \right) |z_{\boldsymbol{k}}|^2 \right] \quad , \tag{4.306}$$

where the prime on the $\boldsymbol{k}$ sum indicates that only one of the pair $\{\boldsymbol{k}, -\boldsymbol{k}\}$ is to be included, for each $\boldsymbol{k}$.

We may now compute the ordinary canonical partition function:

$$\begin{aligned} Z &= {\prod_{\boldsymbol{k}}}' \int \frac{d^2p_{\boldsymbol{k}}\, d^2z_{\boldsymbol{k}}}{(2\pi\hbar)^2}\, e^{-|p_{\boldsymbol{k}}|^2/\mu_{\boldsymbol{k}} k_{\rm B}T}\, e^{-(g\,\Delta\rho + \sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= {\prod_{\boldsymbol{k}}}' \left(\frac{k_{\rm B}T}{2\hbar} \right)^2 \left(\frac{\mu_{\boldsymbol{k}}}{g\,\Delta\rho + \sigma\boldsymbol{k}^2} \right) \quad . \end{aligned} \tag{4.307}$$

Thus,

$$F = -k_{\rm B}T \sum_{\boldsymbol{k}} \ln\left(\frac{k_{\rm B}T}{2\hbar\Omega_{\boldsymbol{k}}} \right) \quad , \tag{4.308}$$

where[13]

$$\Omega_{\boldsymbol{k}} = \left(\frac{g\,\Delta\rho + \sigma\boldsymbol{k}^2}{\mu_{\boldsymbol{k}}} \right)^{1/2} \quad . \tag{4.309}$$

is the normal mode frequency for surface oscillations at wavevector $\boldsymbol{k}$. For deep water waves, it is appropriate to take $\mu_{\boldsymbol{k}} = \Delta\rho/|\boldsymbol{k}|$, where $\Delta\rho = \rho_{\rm L} - \rho_{\rm G} \approx \rho_{\rm L}$ is the difference between the densities of water and air.

[13]Note that there is no prime on the $\boldsymbol{k}$ sum for F, as we have divided the logarithm of Z by two and replaced the half sum by the whole sum.

It is now easy to compute the thermal average

$$\begin{aligned}\left\langle\, |z_{\boldsymbol{k}}|^2 \,\right\rangle &= \int d^2z_{\boldsymbol{k}}\, |z_{\boldsymbol{k}}|^2\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \Big/ \int d^2z_{\boldsymbol{k}}\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= \frac{k_{\rm B}T}{g\,\Delta\rho+\sigma\boldsymbol{k}^2} \quad .\end{aligned} \tag{4.310}$$

Note that this result does not depend on $\mu_{\boldsymbol{k}}$, *i.e.* on our choice of kinetic energy. One defines the *correlation function*

$$\begin{aligned}C(\boldsymbol{x}) \equiv \left\langle\, z(\boldsymbol{x})\, z(0) \,\right\rangle &= \frac{1}{L_xL_y}\sum_{\boldsymbol{k}} \left\langle\, |z_{\boldsymbol{k}}|^2 \,\right\rangle e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \int\frac{d^2k}{(2\pi)^2}\left(\frac{k_{\rm B}T}{g\,\Delta\rho+\sigma\boldsymbol{k}^2}\right)e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \\ &= \frac{k_{\rm B}T}{4\pi\sigma}\int\limits_0^\infty dq\,\frac{e^{ik|\boldsymbol{x}|}}{\sqrt{q^2+\xi^2}} = \frac{k_{\rm B}T}{4\pi\sigma}\,K_0\big(|\boldsymbol{x}|/\xi\big) \quad ,\end{aligned} \tag{4.311}$$

where $\xi = \sqrt{g\,\Delta\rho/\sigma}$ is the correlation length, and where $K_0(z)$ is the Bessel function of imaginary argument. The asymptotic behavior of $K_0(z)$ for small z is $K_0(z) \sim \ln(2/z)$, whereas for large z one has $K_0(z) \sim (\pi/2z)^{1/2}\,e^{-z}$. We see that on large length scales the correlations decay exponentially, but on small length scales they diverge. This divergence is due to the improper energetics we have assigned to short wavelength fluctuations of the interface. Roughly, it can cured by imposing a cutoff on the integral, or by insisting that the shortest distance scale is a molecular diameter.

4.11.4 Dissociation of molecular hydrogen

Consider the reaction

$$\mathrm{H} \rightleftharpoons \mathrm{p}^+ + \mathrm{e}^- \quad . \tag{4.312}$$

In equilibrium, we have

$$\mu_{\rm H} = \mu_{\rm p} + \mu_{\rm e} \quad . \tag{4.313}$$

What is the relationship between the temperature T and the fraction x of hydrogen which is dissociated?

Let us assume a fraction x of the hydrogen is dissociated. Then the densities of H, p, and e are then

$$n_{\rm H} = (1-x)\,n \qquad , \qquad n_{\rm p} = xn \qquad , \qquad n_{\rm e} = xn \quad . \tag{4.314}$$

The single particle partition function for each species is

$$\zeta = \frac{g^N}{N!}\left(\frac{V}{\lambda_T^3}\right)^N e^{-N\varepsilon_{\rm int}/k_{\rm B}T} \quad , \tag{4.315}$$

where g is the degeneracy and $\varepsilon_{\rm int}$ the internal energy for a given species. We have $\varepsilon_{\rm int} = 0$ for p and e, and $\varepsilon_{\rm int} = -\Delta$ for H, where $\Delta = e^2/2a_{\rm B} = 13.6\,\text{eV}$, the binding energy of hydrogen. Neglecting hyperfine splittings[14], we have $g_{\rm H} = 4$, while $g_{\rm e} = g_{\rm p} = 2$ because each has spin $S = \frac{1}{2}$. Thus, the

[14]The hyperfine splitting in hydrogen is on the order of $(m_{\rm e}/m_{\rm p})\,\alpha^4\,m_{\rm e}c^2 \sim 10^{-6}$ eV, which is on the order of $0.01\,\text{K}$. Here $\alpha = e^2/\hbar c$ is the fine structure constant.

associated grand potentials are

$$\Omega_{\rm H} = -g_{\rm H}\, V k_{\rm B} T\, \lambda_{T,{\rm H}}^{-3}\, e^{(\mu_{\rm H}+\Delta)/k_{\rm B}T} \tag{4.316}$$

$$\Omega_{\rm p} = -g_{\rm p}\, V k_{\rm B} T\, \lambda_{T,{\rm p}}^{-3}\, e^{\mu_{\rm p}/k_{\rm B}T} \tag{4.317}$$

$$\Omega_{\rm e} = -g_{\rm e}\, V k_{\rm B} T\, \lambda_{T,{\rm e}}^{-3}\, e^{\mu_{\rm e}/k_{\rm B}T} \quad , \tag{4.318}$$

where

$$\lambda_{T,a} = \sqrt{\frac{2\pi\hbar^2}{m_a k_{\rm B} T}} \tag{4.319}$$

for species a. The corresponding number densities are

$$n = \frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = g\,\lambda_T^{-3}\, e^{(\mu-\varepsilon_{\rm int})/k_{\rm B}T} \quad , \tag{4.320}$$

and the fugacity $z = e^{\mu/k_{\rm B}T}$ of a given species is given by

$$z = g^{-1} n\, \lambda_T^3\, e^{\varepsilon_{\rm int}/k_{\rm B}T} \quad . \tag{4.321}$$

We now invoke $\mu_{\rm H} = \mu_{\rm p} + \mu_{\rm e}$, which says $z_{\rm H} = z_{\rm p}\, z_{\rm e}$, or

$$g_{\rm H}^{-1} n_{\rm H}\, \lambda_{T,{\rm H}}^3\, e^{-\Delta/k_{\rm B}T} = \left(g_{\rm p}^{-1} n_{\rm p}\, \lambda_{T,{\rm p}}^3\right)\left(g_{\rm e}^{-1} n_{\rm e}\, \lambda_{T,{\rm e}}^3\right) \quad , \tag{4.322}$$

which yields

$$\left(\frac{x^2}{1-x}\right) n\tilde{\lambda}_T^3 = e^{-\Delta/k_{\rm B}T} \quad , \tag{4.323}$$

where $\tilde{\lambda}_T = \sqrt{2\pi\hbar^2/m^* k_{\rm B} T}$, with $m^* = m_{\rm p} m_{\rm e}/m_{\rm H} \approx m_{\rm e}$. Note that

$$\tilde{\lambda}_T = a_{\rm B}\sqrt{\frac{4\pi m_{\rm H}}{m_{\rm p}}}\sqrt{\frac{\Delta}{k_{\rm B}T}} \quad , \tag{4.324}$$

where $a_{\rm B} = 0.529\,\text{Å}$ is the Bohr radius. Thus, we have

$$\left(\frac{x^2}{1-x}\right)\cdot(4\pi)^{3/2}\,\nu = \left(\frac{T}{T_0}\right)^{3/2} e^{-T_0/T} \quad , \tag{4.325}$$

where $T_0 = \Delta/k_{\rm B} = 1.578\times10^5\,{\rm K}$ and $\nu = n a_{\rm B}^3$. Consider for example a temperature $T = 3000\,{\rm K}$, for which $T_0/T = 52.6$, and assume that $x = \frac{1}{2}$. We then find $\nu = 1.69\times10^{-27}$, corresponding to a density of $n = 1.14\times10^{-2}\,{\rm cm}^{-3}$. At this temperature, the fraction of hydrogen molecules in their first excited (2s) state is $x' \sim e^{-T_0/2T} = 3.8\times10^{-12}$. This is quite striking: half the hydrogen atoms are completely dissociated, which requires an energy of Δ, yet the number in their first excited state, requiring energy $\frac{1}{2}\Delta$, is twelve orders of magnitude smaller. The student should reflect on why this can be the case.