Contents

Contents i

List of Tables iii

List of Figures iv

4 Statistical Ensembles 1

4.1 References . . . 1

4.2 Microcanonical Ensemble (μCE) . . . 2

4.2.1 The microcanonical distribution function . . . 2

4.2.2 Density of states . . . 3

4.2.3 Arbitrariness in the definition of $S(E)$ . . . 6

4.2.4 Ultra-relativistic ideal gas . . . 6

4.2.5 Discrete systems . . . 7

4.2.6 Two systems in thermal contact . . . 7

4.3 The Quantum Mechanical Trace . . . 9

4.3.1 The density matrix . . . 9

4.3.2 Averaging the DOS . . . 12

4.3.3 Coherent states . . . 13

4.4 Ordinary Canonical Ensemble (OCE) . . . 15

4.4.1 Canonical distribution and partition function . . . 15

4.4.2 The difference between $P(E_n)$ and P_n . . . 16

4.4.3 Additional remarks . . . 16
4.4.4 Averages within the OCE . . . 17
4.4.5 Entropy and free energy . . . 17
4.4.6 Fluctuations in the OCE . . . 18
4.4.7 Energy distribution within the OCE . . . 18
4.4.8 Thermodynamics revisited . . . 19
4.4.9 Generalized susceptibilities . . . 20
4.5 Grand Canonical Ensemble (GCE) . . . 21
4.5.1 Grand canonical distribution and partition function . . . 21
4.5.2 Entropy and Gibbs-Duhem relation . . . 22
4.5.3 Fluctuations in the GCE . . . 23
4.5.4 Generalized susceptibilities in the GCE . . . 24
4.5.5 Gibbs ensemble . . . 25
4.6 Statistical Ensembles from Maximum Entropy . . . 25
4.6.1 Microcanonical ensemble . . . 26
4.6.2 Ordinary canonical ensemble . . . 26
4.6.3 Grand canonical ensemble . . . 27
4.6.4 Generalized Gibbs ensembles . . . 27
4.7 Ideal Gas Statistical Mechanics . . . 29
4.7.1 Maxwell velocity distribution . . . 29
4.7.2 Equipartition . . . 31
4.8 Selected Examples . . . 32
4.8.1 Spins in an external magnetic field . . . 32
4.8.2 Negative temperature (!) . . . 34
4.8.3 Adsorption . . . 35
4.8.4 Dipole gas . . . 36
4.8.5 Elasticity of wool . . . 38
4.9 Statistical Mechanics of Molecular Gases . . . 40

4.9.1 Separation of translational and internal degrees of freedom 40
4.9.2 The internal coordinate partition function 41
4.9.3 Quantum theory of rotations 42
4.9.4 Vibrations 44
4.9.5 Two-level systems : Schottky anomaly 45
4.9.6 Electronic and nuclear excitations 46
4.10 Appendix I: Constrained Extremization of Functions 49
4.10.1 Example : volume of a cylinder 50
4.11 Appendix II: Additional Examples of Statistical Ensembles 50
4.11.1 Noninteracting spin dimers 50
4.11.2 Three-state system 52
4.11.3 Spins and vacancies on a surface 53
4.11.4 Fluctuating interface 55
4.11.5 Dissociation of atomic hydrogen 57

List of Tables

4.1 Rotational and vibrational temperatures of common molecules 42

4.2 Nuclear angular momentum states for homonuclear diatomic molecules 48

List of Figures

4.1 Complex integration contours $\mathcal{C}$ for inverse Laplace transform 4
4.2 Two systems in thermal contact 8
4.3 A system S in contact with a world W 10
4.4 Averaging the quantum mechanical discrete density of states 13
4.5 Microscopic interpretation of the First Law of Thermodynamics 20
4.6 Maxwell distribution of speeds $\varphi(v/v_0)$ 31
4.7 Entropy *versus* energy for a spin system 35
4.8 A crude model for the elasticity of wool 38
4.9 Length *versus* temperature and force 39
4.10 Heat capacity per molecule as a function of temperature 46
4.11 Noninteracting spin dimers on a lattice 51

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.

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[1]:

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

where

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

and

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

Here $\hat{H}(\boldsymbol{\varphi}) = \hat{H}(\boldsymbol{q},\boldsymbol{p})$ is the Hamiltonian, and $\delta(x)$ is the Dirac δ-function[2]. The dynamics are presumed to conserve energy, with $\hat{H}(\boldsymbol{\varphi}(t=0)) = E$. 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 $\varrho(\Lambda_1,\ldots,\Lambda_k)$ which is a function of conserved quantities $\Lambda_a(\boldsymbol{\varphi})$ 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 *indistinguishable particles*[3]. The normalization factor $(2\pi\hbar)^{-Nd}$ renders the trace dimensionless. Again,

[1] Technically this is the case only for suitable observables and almost every initial condition on the invariant energy surface.

[2] We write the Hamiltonian as $\hat{H}$ (classical or quantum) in order to distinguish it from magnetic field, H.

[3] More on this in chapter 5.

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

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

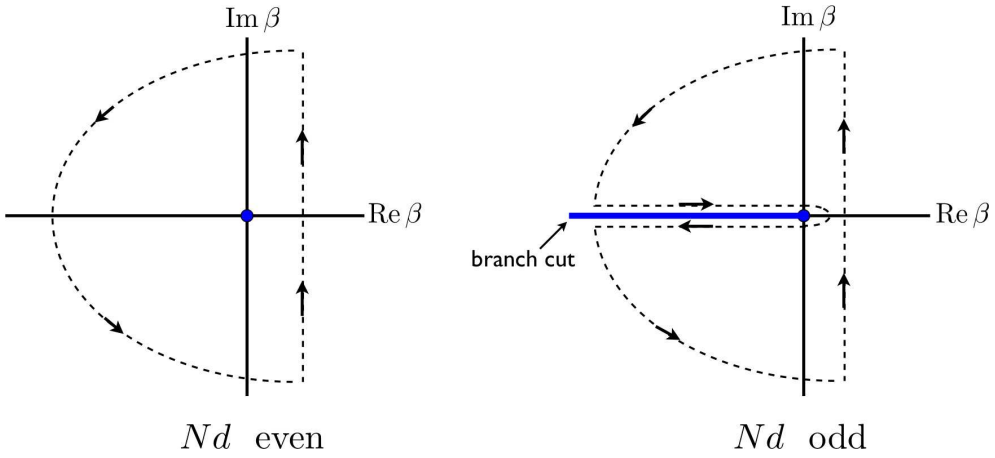


Figure 4.1: Complex integration contours $\mathcal{C}$ for inverse Laplace transform $\mathcal{L}^{-1}\big[Z(\beta)\big] = D(E)$. When the product dN is odd, there is a branch cut along the negative $\mathrm{Re}\,\beta$ axis.

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 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,

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

we may define the *statistical entropy*[4],

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

where

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

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:[5]

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

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

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

[4]Not that $D(E,V,N)$ has dimensions of E^{-1} so we should first multiply by an energy scale before taking the logarithm. This is discussed in §4.2.3 below.

[5]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 $u = |\boldsymbol{u}| \geq 0$, the second term can be dropped.

where $s = u^2$, and where

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

is the Gamma function, which satisfies $z\,\Gamma(z) = \Gamma(z+1)$.[6] 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 $S(E;\Delta E) \equiv k_{\rm B} \log \tilde{D}(E;\Delta E)$. The definition of statistical entropy then involves the arbitrary parameter ΔE, however this only affects $S(E)$ in an additive way. That is,

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

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 $S(E,V,N) = k_{\rm B} \log D(E,V,N) = N k_{\rm B}\, \phi\big(\frac{E}{N}, \frac{V}{N}\big)$, with

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

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

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 of configuring the system at fixed energy E is then

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

where the sum is over all possible configurations, and where $\boldsymbol{\lambda}$ is a vector of parameters which enter into $\hat{H}(\boldsymbol{\sigma})$. 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 B,[7] so the individual particle energy is $\varepsilon_i = -\mu_0 B \sigma_i$, where $\sigma = \pm 1$, then in a configuration in which N_+ particles have $\sigma_i = +1$ and $N_- = N - N_+$ particles have $\sigma_i = -1$, the energy is $E = (N_- - N_+)\,\mu_0 B$. The number of configurations at fixed energy E is

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

We may write $N_\pm = \frac{1}{2}N(1 \pm m)$ where $m = (N_+ - N_-)/N \in [-1, 1]$ is the "magnetization." Thus $m = -E/N\mu_0 B$, and it is left as an exercise to the reader to show, using Stirling's formula, that the statistical entropy is

$$S(E, N, B) = k_{\rm B} \log \Omega(E, N, B) = -N k_{\rm B} \left[\left(\frac{1+m}{2} \right) \log \left(\frac{1+m}{2} \right) + \left(\frac{1-m}{2} \right) \log \left(\frac{1-m}{2} \right) \right] \quad . \tag{4.32}$$

4.2.6 Two systems in thermal contact

Consider two systems in thermal contact, as depicted in fig. 4.2. 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

$$D(E) = \int dE_1\, D_1(E_1)\, D_2(E - E_1) \quad . \tag{4.33}$$

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

$$P_1(E_1) = \frac{D_1(E_1)\, D_2(E - E_1)}{D(E)} \quad . \tag{4.34}$$

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

$$\begin{aligned} 0 &= \frac{1}{P_1(E_1)} \frac{dP_1(E_1)}{dE_1} = \frac{\partial}{\partial E_1} \log P_1(E_1) \\ &= \frac{\partial}{\partial E_1} \log D_1(E_1) + \frac{\partial}{\partial E_1} \log D_2(E - E_1) \quad . \end{aligned} \tag{4.35}$$

[7] Properly, we should use H here rather than B, but to obviate any confusion between H and the Hamiltonian $\hat{H}$, we use B instead.

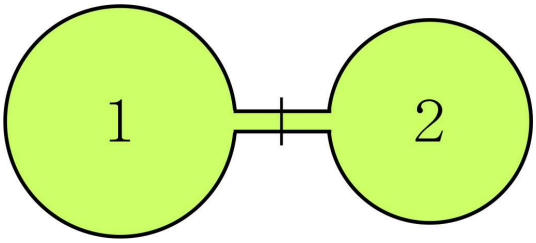


Figure 4.2: Two systems in thermal contact.

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

$$\frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2} \quad . \tag{4.36}$$

This guarantees that

$$S(E, E_1) = S_1(E_1) + S_2(E - E_1) \tag{4.37}$$

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

The *temperature* T is defined as

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N} \quad , \tag{4.38}$$

a result familiar from thermodynamics. The difference is now we have a more rigorous definition of the entropy. When the total entropy S is extremized we have that $T_1 = T_2$. Together with the usual stability/concavity condition, this entails a maximum for S. Once again, two systems in thermal contact can exchange energy and in equilibrium will 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

$$\begin{aligned} S_{\rm NR} &= \tfrac{1}{2} N d\, k_{\rm B} \log\left(\frac{E}{N}\right) + N k_{\rm B} \log\left(\frac{V}{N}\right) + \text{const.} \\ S_{\rm UR} &= N d\, k_{\rm B} \log\left(\frac{E}{N}\right) + N k_{\rm B} \log\left(\frac{V}{N}\right) + \text{const.} \quad . \end{aligned} \tag{4.39}$$

Invoking eqn. 4.38, 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. 4.35. We then have

$$\log P_1(E_1^* + \Delta E_1) = \log P_1(E_1^*) + \frac{1}{2k_{\rm B}} \frac{\partial^2 S_1}{\partial E_1^2}\bigg|_{E_1^*} (\Delta E_1)^2 + \frac{1}{2k_{\rm B}} \frac{\partial^2 S_2}{\partial E_2^2}\bigg|_{E_2^*} (\Delta E_1)^2 + \ldots \quad , \tag{4.40}$$

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

$$\frac{\partial^2 S}{\partial E^2} = \frac{\partial}{\partial E}\left(\frac{1}{T}\right) = -\frac{1}{T^2}\left(\frac{\partial T}{\partial E}\right)_{V,N} = -\frac{1}{T^2\, C_V} \quad , \tag{4.41}$$

where $C_V = \big(\partial E/\partial T\big)_{V,N}$ is the heat capacity. Thus, in the vicinity of a stable entropy maximum,

$$P_1 = P_1^* \, e^{-(\Delta E_1)^2/2k_{\rm B}T^2\bar{C}_V} \quad , \tag{4.42}$$

where

$$\bar{C}_V = \frac{C_{V,1}\,C_{V,2}}{C_{V,1}+C_{V,2}} \quad . \tag{4.43}$$

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

$$\big\langle (\Delta E_1)^2 \big\rangle = k_{\rm B}T^2\,\bar{C}_V \qquad \Longrightarrow \qquad (\Delta E_1)_{\rm RMS} = T\sqrt{k_{\rm B}\bar{C}_V} \quad . \tag{4.44}$$

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*.

4.3 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

Consider a system S in contact with a world W as depicted in fig. 4.3. Their union $\mathsf{U} = \mathsf{W} \cup \mathsf{S}$ forms the universe. Let us choose orthonormal bases $|\sigma\rangle$ for the system and $|i\rangle$ for the world, with $\sigma \in \{1, \ldots, D_{\mathsf{S}}\}$ and $i \in \{1, \ldots, D_{\mathsf{W}}\}$. We may use the states $|\sigma\rangle \otimes |i\rangle$ as a basis for the universe, for which a general wavefunction may be written

$$|\,\Psi\,\rangle = \sum_{\sigma=1}^{D_{\mathsf{S}}} \sum_{i=1}^{D_{\mathsf{W}}} \Psi_{\sigma i}\,|\,\sigma\,\rangle \otimes |\,i\,\rangle \quad . \tag{4.45}$$

Now consider an operator $\hat{A}$ which only acts within the system's Hilbert space $\mathcal{H}_{\mathsf{S}}$. The expectation of $\hat{A}$ is given by

$$\langle\,\Psi\,|\,\hat{A}\,|\,\Psi\,\rangle = \sum_{\sigma,i}\sum_{\tau,j} \Psi_{\sigma i}\Psi^*_{\tau j}\,\langle\,\tau\,|\,\hat{A}\,|\,\sigma\,\rangle\langle\,j\,|\,i\,\rangle = \sum_{\sigma,\tau} \varrho_{\sigma\tau}\,A_{\tau\sigma} \tag{4.46}$$

where $\varrho_{\sigma\tau} = \sum_j \Psi_{\sigma j}\Psi^*_{\tau j}$ is the *reduced density matrix* of the system. In operator form,

$$\hat{\varrho} = \sum_{\sigma,\tau} \varrho_{\sigma\tau}\,|\,\sigma\,\rangle\langle\,\tau\,| \qquad \Rightarrow \qquad \langle\hat{A}\rangle = \mathsf{Tr}\,(\hat{\varrho}\,\hat{A}) \quad . \tag{4.47}$$

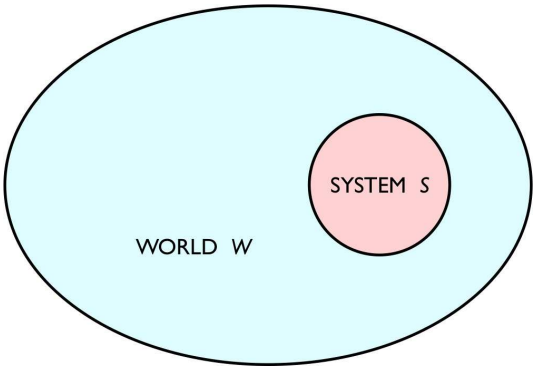


Figure 4.3: A system S in contact with a world W. The union of the two is the universe U = W ∪ S.

Note that $\varrho_{\sigma\tau} = \varrho^*_{\tau\sigma}$ hence $\hat{\varrho} = \hat{\varrho}^\dagger$ is Hermitian. Note also that $\hat{\varrho}$ is normalized:

$$\mathsf{Tr}\ \hat{\varrho} = \sum_\sigma \varrho_{\sigma\sigma} = \sum_{\sigma,j} \Psi_{\sigma j}\Psi^*_{\sigma j} = \langle\Psi|\Psi\rangle = 1 \quad . \tag{4.48}$$

The Hermitian operator $\hat{\varrho}$ has D_{S} eigenvalues and D_{S} eigenvectors. In the eigenbasis of $\hat{\varrho}$, we have

$$\hat{\varrho} = \sum_{\alpha=1}^{D_{\mathsf{S}}} \lambda_\alpha\, |\,\alpha\,\rangle\langle\,\alpha\,| \qquad \text{with} \qquad \mathsf{Tr}\ \hat{\varrho} = \sum_{\alpha=1}^{D_{\mathsf{S}}} \lambda_\alpha = 1 \tag{4.49}$$

and the matrix $\varrho_{\sigma\tau} = \lambda_\sigma\,\delta_{\sigma\tau}$ (no sum on σ) is diagonal. Another constraint is that each $\lambda_\alpha \in [0,1]$. This is because $\hat{\varrho}$ is positive semidefinite:

$$\langle\,\phi\,|\,\hat{\varrho}\,|\,\phi\,\rangle = \sum_j \left|\sum_\sigma \phi^*_\sigma \Psi_{\sigma j}\right|^2 \geq 0 \quad . \tag{4.50}$$

This entails each $\lambda_\alpha \geq 0$, and from $\mathsf{Tr}\ \hat{\varrho} = \sum_\alpha \lambda_\alpha = 1$ we conclude $0 \leq \lambda_\alpha \leq 1$. This further entails the (weak) inequality $\mathsf{Tr}\ \hat{\varrho}^2 \leq 1$.

One example of a density matrix is a simple projection operator $\hat{P}_\sigma = |\sigma\rangle\langle\sigma|$. All projectors $\hat{P}$ are *idempotent*, meaning $\hat{P}^2 = \hat{P}$. If $\hat{\varrho} = \hat{P}_\sigma$, then $|\sigma\rangle$ is an eigenvector of $\hat{\varrho}$ with eigenvalue $\lambda = 1$, and all other vectors in the $(D_{\mathsf{S}} - 1)$-dimensional subspace of $\mathcal{H}_{\mathsf{S}}$ orthogonal to $|\sigma\rangle$ are annihilated by $\hat{\varrho}$, *i.e.* they are null eigenvectors. The eigenspectrum of $\hat{\varrho}$ thus consists of $\lambda_1 = 1$ and $\lambda_2 = \cdots = \lambda_{D_{\mathsf{S}}} = 0$. The system is then said to be in a *pure state*. This is a very special state of affairs. In the space of allowed density matrices, the vast majority describe *mixed states*. As an example of a mixed state, let $\hat{\varrho} = x\,|1\rangle\langle 1| + (1-x)\,|2\rangle\langle 2|$. Here we have $\lambda_1 = x$, $\lambda_2 = 1 - x$, and $\lambda_3 = \cdots = \lambda_{D_S} = 0$, with $x \in (0,1)$. A crisp distinction between pure and mixed states is that while $\mathsf{Tr}\,\hat{\varrho} = 1$ in all cases, for pure states $\mathsf{Tr}\,\hat{\varrho}^2 = 1$ while for mixed states $\mathsf{Tr}\,\hat{\varrho}^2 < 1$.

How does $\hat{\varrho}(t)$ evolve in time? If the Hamiltonian of the universe is given by $\hat{H}_{\mathsf{U}} = \hat{H}_{\mathsf{S}} + \hat{H}_{\mathsf{W}}$ and involves no coupling between the system and the world, then the analysis is simple. Let's work in eigenbases where $\hat{H}_{\mathsf{S}}\,|\alpha\rangle = \omega_\alpha\,|\alpha\rangle$ and $\hat{H}_{\mathsf{W}}\,|k\rangle = \varepsilon_k\,|k\rangle$. Then

$$|\,\Psi(t)\,\rangle = \sum_{\alpha,k} \Psi_{\alpha k}(t)\,|\,\alpha\,\rangle \otimes |\,k\,\rangle \tag{4.51}$$

where $\Psi_{\alpha k}(t) = e^{-i(\omega_\alpha+\varepsilon_k)t/\hbar}\,\Psi_{\alpha k}(0)$. The reduced density matrix for the system then obeys *Liouville's equation*,

$$i\hbar\,\frac{\partial\hat\varrho}{\partial t} = \big[\hat H_{\mathsf S},\hat\varrho\big] \quad . \tag{4.52}$$

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

$$\hat A(t) = e^{+iHt/\hbar}\,A\,e^{-i\hat Ht/\hbar} \qquad\Longrightarrow\qquad i\hbar\,\frac{\partial\hat A}{\partial t} = \big[\hat A,\hat H\big] = -\big[\hat H,\hat A\big] \quad . \tag{4.53}$$

In the eigenbasis of $\hat H_{\mathsf S}$, Liouville's equation says $\varrho_{\alpha\beta}(t) = e^{i(\omega_\beta-\omega_\alpha)t/\hbar}\varrho_{\alpha\beta}(0)$, hence both $\mathsf{Tr}\;\hat\varrho(t) = \mathsf{Tr}\;\hat\varrho(0)$ and $\mathsf{Tr}\;\hat\varrho^2(t) = \mathsf{Tr}\;\hat\varrho^2(0)$. In other words, pure states evolve into pure states and mixed states evolve into mixed states. This is true for any isolated quantum system evolving purely internally, with no external couplings. What happens if there are such couplings between S and W? Consider an initial density matrix

$$\hat\varrho_{\mathsf U}(0) = \hat\varrho_{\mathsf S}(0)\otimes\hat\varrho_{\mathsf W}(0) \tag{4.54}$$

and let the (time-independent) Hamiltonian be $\hat H_{\mathsf U} = \hat H_{\mathsf S} + \hat H_{\mathsf W} + \hat H_{\mathsf{SW}}$ where the last term couples S and W. The initial density matrices $\hat\varrho_{\mathsf S}(0)$ and $\hat\varrho_{\mathsf W}(0)$ are arbitrary – they could be pure or mixed. The evolution operator $\hat U_{\mathsf U}(t) = \exp(-i\hat H_{\mathsf U}t/\hbar)$ acts on the full Hilbert space $\mathcal{H}_{\mathsf U} = \mathcal{H}_{\mathsf S}\otimes\mathcal{H}_{\mathsf W}$. The reduced density matrix for S at time t is given by

$$\hat\varrho_{\mathsf S}(t) = \operatorname*{\mathsf{Tr}}_{\mathsf W}\;\hat U(t)\,\hat\varrho_{\mathsf U}(0)\,\hat U^\dagger(t) \tag{4.55}$$

Let $|k\rangle$ denote an eigenstate of $\hat\varrho_{\mathsf W}(0)$ with eigenvalue $\xi_k\in[0,1]$ and $\sum_k\xi_k = 1$. We may then spectrally decompose $\hat\varrho_{\mathsf W}(0)$ as

$$\hat\varrho_{\mathsf W}(0) = \sum_k \xi_k\,|\,k\,\rangle\langle\,k\,| \quad . \tag{4.56}$$

Now define the operators

$$\hat V_{kl}(t) \equiv \langle\,k\,|\,\hat U_{\mathsf U}(t)\,|\,l\,\rangle \quad , \tag{4.57}$$

each of which acts as an *operator* within $\mathcal{H}_{\mathsf S}$. Note $\hat V_{kl}^\dagger(t) = \langle\,l\,|\,\hat U_{\mathsf U}^\dagger(t)\,|\,k\,\rangle$. We then have

$$\hat\varrho_{\mathsf S}(t) = \sum_{k,l}\xi_l\,\hat V_{kl}(t)\,\hat\varrho_{\mathsf S}(0)\,\hat V_{kl}^\dagger(t) = \sum_r \hat K_r(t)\,\hat\varrho_{\mathsf S}(0)\,\hat K_r^\dagger(t) \tag{4.58}$$

where $r\equiv(k,l)$ is a composite index and

$$\hat K_r(t) = \xi_l^{1/2}\,\hat V_{kl}(t) \qquad , \qquad \hat K_r^\dagger(t) = \xi_l^{1/2}\,\hat V_{kl}^\dagger(t) \quad , \tag{4.59}$$

which satisfy

$$\sum_r \hat K_r^\dagger(t)\,\hat K_r(t) = \mathbb{1}_{\mathsf S} \quad . \tag{4.60}$$

Fixing and suppressing t, eqns. 4.58 and 4.60 define the *Kraus map* on the space of density matrices, *viz.*

$$\hat\varrho' = \sum_{r=1}^M \hat K_r\,\hat\varrho\,\hat K_r^\dagger \qquad , \qquad \sum_{r=1}^M \hat K_r^\dagger\hat K_r = \mathbb{1} \quad . \tag{4.61}$$

Here M is the total number of Kraus operators; in the above discussion we have $M = D_{\mathsf{W}}^2$ as there is one per (k,l) pair. The Kraus map is linear, trace-preserving, Hermiticity preserving, and satisfies an additional desideratum, called "complete positivity," which we will not discuss here. When there is only one Kraus operator $\hat{K}$, we have $\hat{K}^\dagger \hat{K} = \mathbb{1}$ and unitary evolution. When there is more than one Kraus operator $\hat{K}_r$, the image $\hat{\varrho}'$ of $\hat{\varrho}$ under the Kraus map is Hermitian and satisfies $\mathsf{Tr}\ \hat{\varrho}' = 1$, but unlike unitary evolution the Kraus map does not separately preserve pure and mixed states. Owing to the environmental coupling, a pure state of S can evolve into a mixed state and *vice versa*. If $\varrho_{\mathsf{S}}(0)$ is pure, its evolution into a mixed state is a signature of *quantum entanglement* between S and W. In general, environmental coupling allows a pure reduced state to evolve into a mixed one, although subsequent recoherence or disentanglement is also possible.

In order to iterate the Kraus map using the same Kraus operators, we adopt the following protocol. We start with $\hat{\varrho}_{\mathsf{U}} \equiv \hat{\varrho}_n \otimes \hat{\varrho}^0_{\mathsf{W}}$. We evolve $\hat{\varrho}_{\mathsf{U}}$ with $\hat{U}_{\mathsf{U}}(\Delta t)$ over a fixed time interval Δt, then trace out over the W degrees of freedom, leaving us with the next iterate of the reduced density matrix of S, *i.e.* $\hat{\varrho}_{n+1}$. Then we disentangle S and W, reset $\hat{\varrho}_{\mathsf{U}} \equiv \hat{\varrho}_{n+1} \otimes \hat{\varrho}^0_{\mathsf{W}}$, and repeat the process. Physically, we can imagine that the degrees of freedom (*i.e.* qubits) in W that were entangled with S during the previous iteration have been advected away from the interface of S and W, and new unentangled qubits described by the same W density matrix take their place. We then have $\hat{\varrho}_{n+1} = \sum_r \hat{K}_r\, \hat{\varrho}_n\, \hat{K}_r^\dagger$, with the same set of Kraus operators. Lather, rinse, repeat.

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

$$\hat{L} = -i \sum_{j=1}^{Nd} \left\{ \frac{\partial \hat{H}}{\partial p_j} \frac{\partial}{\partial q_j} - \frac{\partial \hat{H}}{\partial q_j} \frac{\partial}{\partial p_j} \right\} \quad . \tag{4.62}$$

Accordingly, any distribution $\varrho(\mathit{\Lambda}_1, \ldots, \mathit{\Lambda}_k)$ which is a function of constants of the motion $\mathit{\Lambda}_a(q,p)$ is a stationary solution to the Liouville equation: $\partial_t\, \varrho(\mathit{\Lambda}_1, \ldots, \mathit{\Lambda}_k) = 0$. Similarly, any quantum mechanical density matrix which commutes with the Hamiltonian is a stationary solution to eqn. 4.52. The corresponding microcanonical distribution is $\hat{\varrho}_E = \delta\big(E - \hat{H}\big)/D(E)$ where $D(E)$ is the density of states.

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

$$D(E) = \mathsf{Tr}\ \delta\big(E - \hat{H}\big) = \sum_l \delta(E - E_l) \quad , \tag{4.63}$$

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:

$$\overline{D(E)} = \frac{1}{\Delta E} \int_E^{E+\Delta E} dE'\, D(E') \quad . \tag{4.64}$$

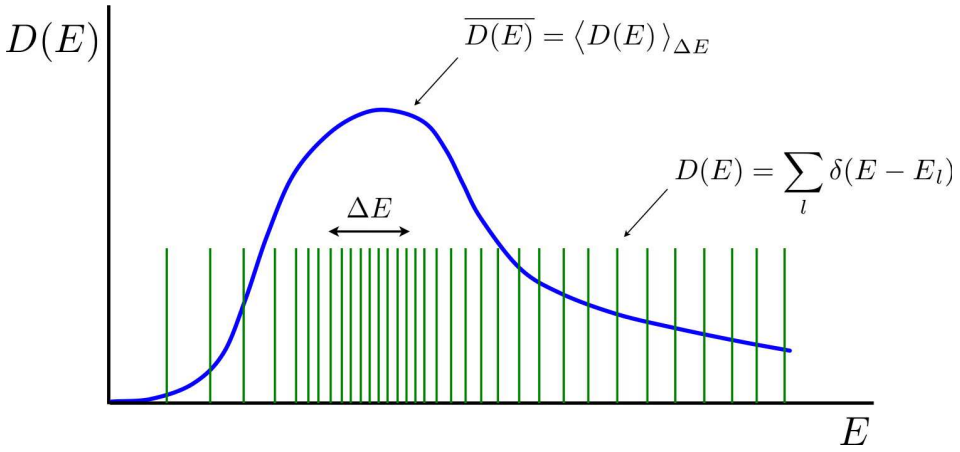


Figure 4.4: Averaging the quantum mechanical discrete density of states yields a continuous curve.

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. 4.4. We will in general drop the bar and refer to this function as $D(E)$. Note that $\delta E \sim 1/D(E) = \exp\bigl[-N\phi(\varepsilon, v)\bigr]$ 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

$$\hat{H}_0 = \frac{p^2}{2m} + \tfrac{1}{2} m\,\omega_0^2\,q^2 = \hbar\omega_0\left(a^\dagger a + \tfrac{1}{2}\right) \quad , \tag{4.65}$$

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

$$a = \ell\,\frac{\partial}{\partial q} + \frac{q}{2\ell} \qquad , \qquad a^\dagger = -\ell\,\frac{\partial}{\partial q} + \frac{q}{2\ell} \quad , \tag{4.66}$$

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

$$\psi_0(q) = (2\pi\ell^2)^{-1/4}\,e^{-q^2/4\ell^2} \quad . \tag{4.67}$$

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

$$|\,z\,\rangle = e^{-\frac{1}{2}|z|^2}\,e^{za^\dagger}\,|\,0\,\rangle = e^{-\frac{1}{2}|z|^2}\sum_{n=0}^{\infty}\frac{z^n}{\sqrt{n!}}\,|\,n\,\rangle \quad . \tag{4.68}$$

The overlap of coherent states is given by

$$\langle\,z_1\,|\,z_2\,\rangle = e^{-\frac{1}{2}|z_1|^2}\,e^{-\frac{1}{2}|z_2|^2}\,e^{\bar{z}_1 z_2} \quad , \tag{4.69}$$

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

$$1 = \int \frac{d^2z}{2\pi i} \, |\, z \,\rangle\langle\, z \,| \quad ; \quad \frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re}z\; d\,\mathrm{Im}z}{\pi} \tag{4.70}$$

which is straightforward to establish.

To gain some physical intuition about the coherent states, define

$$z \equiv \frac{Q}{2\ell} + \frac{i\ell P}{\hbar} \tag{4.71}$$

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

$$\psi_{Q,P}(q) = \langle\, q \,|\, z \,\rangle = (2\pi\ell^2)^{-1/4}\, e^{-iPQ/2\hbar}\, e^{iPq/\hbar}\, e^{-(q-Q)^2/4\ell^2} \quad , \tag{4.72}$$

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

$$\langle\, Q, P \,|\, q \,|\, Q, P \,\rangle = \langle\, z \,|\, \ell\,(a + a^\dagger) \,|\, z \,\rangle = 2\ell\, \mathrm{Re}\, z = Q \tag{4.73}$$

$$\langle\, Q, P \,|\, p \,|\, Q, P \,\rangle = \langle\, z \,|\, \frac{\hbar}{2i\ell}\,(a - a^\dagger) \,|\, z \,\rangle = \frac{\hbar}{\ell}\, \mathrm{Im}\, z = P \tag{4.74}$$

as well as

$$\langle\, Q, P \,|\, q^2 \,|\, Q, P \,\rangle = \langle\, z \,|\, \ell^2\,(a + a^\dagger)^2 \,|\, z \,\rangle = Q^2 + \ell^2 \tag{4.75}$$

$$\langle\, Q, P \,|\, p^2 \,|\, Q, P \,\rangle = -\langle\, z \,|\, \frac{\hbar^2}{4\ell^2}(a - a^\dagger)^2 \,|\, z \,\rangle = P^2 + \frac{\hbar^2}{4\ell^2} \quad . \tag{4.76}$$

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

$$\Delta q = \ell = \sqrt{\frac{\hbar}{2m\omega_0}} \qquad , \qquad \Delta p = \frac{\hbar}{2\ell} = \sqrt{\frac{m\hbar\omega_0}{2}} \quad , \tag{4.77}$$

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

$$\langle\, Q, P \,|\, \hat{A}(q, p) \,|\, Q, P \,\rangle = A(Q, P) + \mathcal{O}(\hbar) \quad , \tag{4.78}$$

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

Since

$$\frac{d^2z}{2\pi i} \equiv \frac{d\,\mathrm{Re}z\; d\,\mathrm{Im}z}{\pi} = \frac{dQ\, dP}{2\pi\hbar} \quad , \tag{4.79}$$

we can write the trace using coherent states as

$$\mathsf{Tr}\ \hat{A} = \frac{1}{2\pi\hbar} \int\limits_{-\infty}^{\infty} dQ \int\limits_{-\infty}^{\infty} dP\, \langle\, Q, P \,|\, \hat{A} \,|\, Q, P \,\rangle \quad . \tag{4.80}$$

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 Ordinary Canonical Ensemble (OCE)

4.4.1 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. 4.3. 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

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0} \frac{D_{\rm W}(E_{\rm U} - E_n)\,\Delta E}{D_{\rm U}(E_{\rm U})\,\Delta E} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.81}$$

Then

$$\begin{aligned} \log P_n &= \log D_{\rm W}(E_{\rm U} - E_n) - \log D_{\rm U}(E_{\rm U}) \\ &= \log D_{\rm W}(E_{\rm U}) - \log D_{\rm U}(E_{\rm U}) - E_n \left.\frac{\partial \log D_{\rm W}(E)}{\partial E}\right|_{E=E_{\rm U}} + \ldots \quad . \end{aligned} \tag{4.82}$$

The higher order terms are negligible if $\mathsf{vol}(S) \ll \mathsf{vol}(W)$. In this case we have $\log P_n = -\alpha - \beta E_n$, with

$$\beta = \left.\frac{\partial \log D_{\rm W}(E)}{\partial E}\right|_{E=E_{\rm U}} \equiv \frac{1}{k_{\rm B}T} \quad . \tag{4.83}$$

The constant α is fixed by the requirement that $\sum_n P_n = 1$, and thus we obtain

$$P_n = \frac{1}{Z}\, e^{-E_n/k_{\rm B}T} \qquad , \qquad Z(T,V,N) = \mathsf{Tr}\; e^{-\beta\hat{H}} = \sum_n e^{-E_n/k_{\rm B}T} \quad . \tag{4.84}$$

We define the Helmholtz free energy $F(T,V,N)$ as

$$F(T,V,N) = -k_{\rm B}T \log Z(T,V,N) \quad . \tag{4.85}$$

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

$$\hat{\varrho} = \frac{e^{-\beta\hat{H}}}{\mathsf{Tr}\; e^{-\beta\hat{H}}} \quad , \tag{4.86}$$

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.

4.4.2 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

$$P(E_{\rm S}, E_{\rm W}) = D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm W})\, \delta(E_{\rm U} - E_{\rm S} - E_{\rm W}) / D_{\rm U}(E_{\rm U}) \quad , \tag{4.87}$$

where

$$D_{\rm U}(E_{\rm U}) = \int\limits_{-\infty}^{\infty} dE_{\rm S}\; D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S}) \quad , \tag{4.88}$$

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

$$P(E_{\rm S}) = \frac{D_{\rm S}(E_{\rm S})\, D_{\rm W}(E_{\rm U} - E_{\rm S})}{D_{\rm U}(E_{\rm U})} \quad , \tag{4.89}$$

as we have in eqn. 4.81. Suppose we wish to know the probability P_n that the system is in a *particular* state $|\, n \,\rangle$ with energy E_n. Clearly

$$P_n = \lim_{\Delta E \to 0} \frac{\text{probability that } E_{\rm S} \in [E_n, E_n + \Delta E]}{\text{\# of S states with } E_{\rm S} \in [E_n, E_n + \Delta E]} = \frac{P(E_n)\, \Delta E}{D_{\rm S}(E_n)\, \Delta E} = \frac{D_{\rm W}(E_{\rm U} - E_n)}{D_{\rm U}(E_{\rm U})} \quad . \tag{4.90}$$

4.4.3 Additional remarks

The formula of eqn. 4.81 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[8], 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

$$\begin{aligned} D_{\rm W}(E_{\rm U} - E_n)\, \Delta E \to e^{S_{\rm W}(E_{\rm U} - E_n\,,\, \Delta E)} &\equiv \mathop{\rm Tr}_{\rm W} \int\limits_{E_{\rm U} - E_n}^{E_{\rm U} - E_n + \Delta E} dE\; \delta(E - \hat{H}_{\rm W}) \\ D_{\rm U}(E_{\rm U})\, \Delta E \to e^{S_{\rm U}(E_{\rm U}\,,\, \Delta E)} &\equiv \mathop{\rm Tr}_{\rm U} \int\limits_{E_{\rm U}}^{E_{\rm U} + \Delta E} dE\; \delta(E - \hat{H}_{\rm U}) \quad . \end{aligned} \tag{4.91}$$

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

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

4.4.4 Averages within the OCE

To compute averages within the OCE,

$$\big\langle \hat{A} \big\rangle = \mathsf{Tr}\,\big(\hat{\varrho}\,\hat{A}\big) = \frac{\sum_n \langle n|\hat{A}|n\rangle\, e^{-\beta E_n}}{\sum_j e^{-\beta E_j}} \quad , \tag{4.92}$$

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

$$\varrho(\boldsymbol{\varphi}) = \frac{1}{Z}\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \quad Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \int d\mu\, e^{-\beta \hat{H}(\boldsymbol{\varphi})} \quad , \tag{4.93}$$

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,

$$\langle A \rangle = \mathsf{Tr}\,(\varrho A) = \frac{\int d\mu\; A(\boldsymbol{\varphi})\, e^{-\beta \hat{H}(\boldsymbol{\varphi})}}{\int d\mu\, e^{-\beta \hat{H}(\boldsymbol{\varphi})}} \quad . \tag{4.94}$$

4.4.5 Entropy and free energy

The *Boltzmann entropy* is defined by

$$S = -k_{\rm B}\mathsf{Tr}\,(\hat{\varrho}\log\hat{\varrho}) = -k_{\rm B}\sum_n P_n\, \log P_n \quad . \tag{4.95}$$

The Boltzmann entropy and the statistical entropy $S = k_{\rm B}\log D(E)$ are identical in the thermodynamic limit. Since $\log P_n = \beta(F - E_n)$, we have

$$S = -k_{\rm B}\sum_n P_n\,\big(\beta F - \beta E_n\big) = -\frac{F}{T} + \frac{\langle\, \hat{H}\,\rangle}{T} \quad , \tag{4.96}$$

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

$$E = \sum_n P_n\, E_n = \frac{\mathsf{Tr}\;\hat{H}\, e^{-\beta \hat{H}}}{\mathsf{Tr}\; e^{-\beta \hat{H}}} \tag{4.97}$$

is the average energy. We also see that

$$Z = \mathsf{Tr}\; e^{-\beta \hat{H}} = \sum_n e^{-\beta E_n} \quad \Longrightarrow \quad E = \frac{\sum_n E_n\, e^{-\beta E_n}}{\sum_j e^{-\beta E_j}} = -\frac{\partial}{\partial\beta}\,\log Z = \frac{\partial}{\partial\beta}\big(\beta F\big) \quad . \tag{4.98}$$

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

$$dF = -S\,dT - p\,dV + \mu\,dN \quad , \tag{4.99}$$

which means

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \quad , \quad p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \quad , \quad \mu = +\left(\frac{\partial F}{\partial N}\right)_{T,V} \quad . \tag{4.100}$$

4.4.6 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

$$\begin{aligned} -\frac{\partial E}{\partial \beta} &= k_{\rm B} T^2 \, \frac{\partial E}{\partial T} = \frac{\partial^2 \log Z}{\partial \beta^2} \\ &= \frac{\mathsf{Tr}\ \hat{H}^2 \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} - \left(\frac{\mathsf{Tr}\ \hat{H} \, e^{-\beta \hat{H}}}{\mathsf{Tr}\ e^{-\beta \hat{H}}} \right)^2 \\ &= \big\langle \hat{H}^2 \big\rangle - \big\langle \hat{H} \big\rangle^2 \quad . \end{aligned} \tag{4.101}$$

Thus, the heat capacity is related to the fluctuations in the energy, just as we saw in §4.2.6:

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{V,N} = \frac{1}{k_{\rm B} T^2} \left(\big\langle \hat{H}^2 \big\rangle - \big\langle \hat{H} \big\rangle^2 \right) \qquad \Rightarrow \qquad \frac{(\Delta E)_{\rm RMS}}{E} = \frac{\sqrt{k_{\rm B} T^2 \, C_V}}{E} \quad . \tag{4.102}$$

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

$$\left(\frac{(\Delta E)_{\rm RMS}}{E} \right)_{\rm NRIG} = \sqrt{\frac{2}{Nd}} \quad , \tag{4.103}$$

which scales as $N^{-1/2}$ and thus vanishes in the thermodynamic limit.

4.4.7 Energy distribution within the OCE

The full distribution function for the energy within the ordinary canonical ensemble is

$$P(\mathcal{E}) = \big\langle \delta(\mathcal{E} - \hat{H}) \big\rangle = \frac{\mathsf{Tr} \big[\delta(\mathcal{E} - \hat{H}) \, e^{-\beta \hat{H}} \big]}{\mathsf{Tr}\ e^{-\beta \hat{H}}} = \frac{1}{Z} \, D(\mathcal{E}) \, e^{-\beta \mathcal{E}} \quad . \tag{4.104}$$

Thus,

$$P(\mathcal{E}) = \frac{e^{-\beta [\mathcal{E} - TS(\mathcal{E})]}}{\int d\mathcal{E}' \, e^{-\beta [\mathcal{E}' - TS(\mathcal{E}')]}} \quad , \tag{4.105}$$

where $S(\mathcal{E}) = k_{\rm B} \log D(\mathcal{E})$ 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

$$S(E + \delta\mathcal{E}) = S(E) + \frac{\delta\mathcal{E}}{T} - \frac{(\delta\mathcal{E})^2}{2T^2 \, C_V} + \ldots \tag{4.106}$$

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

$$\mathcal{E} - T\,S(\mathcal{E}) = E - T\,S(E) + \frac{(\delta\mathcal{E})^2}{2T\,C_V} + \mathcal{O}\big((\delta\mathcal{E})^3\big) \quad . \tag{4.107}$$

Applying this to both numerator and denominator of eqn. 4.105, we obtain[9]

$$P(\mathcal{E}) = \mathcal{N} \exp\left[- \frac{(\delta\mathcal{E})^2}{2k_{\rm B}T^2\,C_V} \right] \quad , \tag{4.108}$$

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.

4.4.8 Thermodynamics revisited

The average energy within the OCE is

$$E = \sum_n E_n P_n \quad , \tag{4.109}$$

and therefore

$$dE = \sum_n E_n\, dP_n + \sum_n P_n\, dE_n = đQ - đW \quad , \tag{4.110}$$

where

$$đQ = \sum_n E_n\, dP_n \qquad , \qquad đW = -\sum_n P_n\, dE_n \quad . \tag{4.111}$$

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

$$E_n = -k_{\rm B}T \log Z - k_{\rm B}T \log P_n \quad , \tag{4.112}$$

with which we obtain

$$\begin{aligned} đQ = \sum_n E_n\, dP_n &= -k_{\rm B}T \log Z \sum_n dP_n - k_{\rm B}T \sum_n \log P_n\, dP_n \\ &= T\, d\Big(- k_{\rm B} \sum_n P_n \log P_n \Big) = T\, dS \quad . \end{aligned} \tag{4.113}$$

Note also that

$$\begin{aligned} đW = -\sum_n P_n\, dE_n &= -\sum_n P_n \left(\sum_i \frac{\partial E_n}{\partial X_i}\, dX_i \right) \\ &= -\sum_{n,i} P_n \, \langle\, n \,|\, \frac{\partial \hat{H}}{\partial X_i} \,|\, n \,\rangle\, dX_i \equiv \sum_i F_i\, dX_i \quad , \end{aligned} \tag{4.114}$$

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

$$F_i = -\sum_n P_n \frac{\partial E_n}{\partial X_i} = -\left\langle \frac{\partial \hat{H}}{\partial X_i} \right\rangle \quad . \tag{4.115}$$

[9] In applying eqn. 4.107 to the denominator of eqn. 4.105, 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.

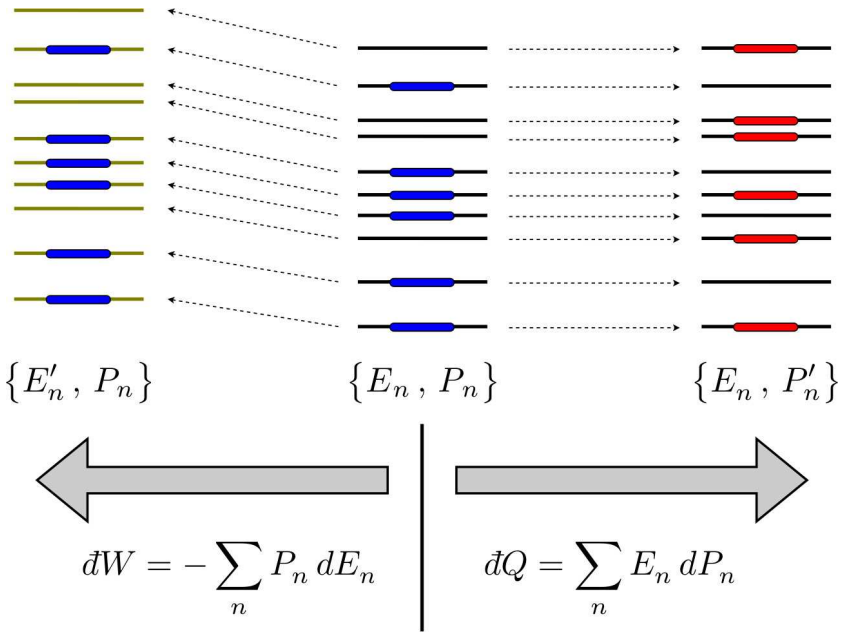


Figure 4.5: Microscopic, statistical interpretation of the First Law of Thermodynamics.

This is the force acting *on* the system[10]. In the chapter on thermodynamics, we defined the generalized force conjugate to X_i as $y_i \equiv -F_i$.

Thus we see from eqn. 4.110 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.

4.4.9 Generalized susceptibilities

Suppose our Hamiltonian is of the form

$$\hat{H} = \hat{H}(\lambda) = \hat{H}_0 - \lambda\,\hat{Q} \quad , \tag{4.116}$$

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

$$\frac{1}{Z}\frac{\partial Z}{\partial \lambda} = \beta \cdot \frac{1}{Z}\,\mathrm{Tr}\left(\hat{Q}\,e^{-\beta\hat{H}(\lambda)}\right) = \beta\,\langle\hat{Q}\rangle \quad . \tag{4.117}$$

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

$$Q(\lambda, T) = \langle\,\hat{Q}\,\rangle = -\left(\frac{\partial F}{\partial \lambda}\right)_T \quad . \tag{4.118}$$

[10] In deriving eqn. 4.115, 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.

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

$$\chi = \frac{1}{V}\frac{\partial Q}{\partial \lambda} = -\frac{1}{V}\frac{\partial^2 F}{\partial \lambda^2} \quad . \tag{4.119}$$

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

$$\hat{H}(\boldsymbol{\lambda}) = \hat{H}_0 - \sum_k \lambda_k \hat{Q}_k \quad , \tag{4.120}$$

we have that

$$Q_k(\boldsymbol{\lambda}, T) = \langle\, \hat{Q}_k \,\rangle = -\left(\frac{\partial F}{\partial \lambda_k}\right)_{T,\,\{N_a\},\,\{\lambda_j (\neq k)\}} \tag{4.121}$$

and we may define an entire matrix of susceptibilities,

$$\chi_{kl} \equiv \frac{1}{V}\frac{\partial Q_k}{\partial \lambda_l} = -\frac{1}{V}\frac{\partial^2 F}{\partial \lambda_k\,\partial \lambda_l} \quad . \tag{4.122}$$

4.5 Grand Canonical Ensemble (GCE)

4.5.1 Grand canonical distribution and partition function

Consider once again the situation depicted in fig. 4.3, 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

$$\begin{aligned} P_n &= \lim_{\Delta E \to 0}\,\lim_{\Delta N \to 0}\,\frac{D_{\rm W}(E_{\rm U} - E_n\,,\,N_{\rm U} - N_n)\,\Delta E\,\Delta N}{D_{\rm U}(E_{\rm U}, N_{\rm U})\,\Delta E\,\Delta N} \\ &= \frac{\text{\# of states accessible to } W \text{ given that } E_{\rm S} = E_n \text{ and } N_{\rm S} = N_n}{\text{total \# of states in } U} \quad . \end{aligned} \tag{4.123}$$

Then

$$\begin{aligned} \log P_n &= \log D_{\rm W}(E_{\rm U} - E_n\,,\,N_{\rm U} - N_n) - \log D_{\rm U}(E_{\rm U}, N_{\rm U}) \\ &= \log D_{\rm W}(E_{\rm U}, N_{\rm U}) - \log D_{\rm U}(E_{\rm U}, N_{\rm U}) \\ &\qquad - E_n \left.\frac{\partial \log D_{\rm W}(E,N)}{\partial E}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} - N_n \left.\frac{\partial \log D_{\rm W}(E,N)}{\partial N}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} + \ldots \end{aligned} \tag{4.124}$$

and thus $\log P_n = -\alpha - \beta E_n + \beta\mu N_n$, with

$$\beta = \left.\frac{\partial \log D_{\rm W}(E,N)}{\partial E}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} = \frac{1}{k_{\rm B}T} \qquad , \qquad \mu = -k_{\rm B}T\left.\frac{\partial \log D_{\rm W}(E,N)}{\partial N}\right|_{\substack{E=E_{\rm U}\\N=N_{\rm U}}} \quad . \tag{4.125}$$

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 $\sum_n P_n = 1$:

$$\begin{aligned} P_n &= \frac{1}{\Xi}\, e^{-\beta(E_n-\mu N_n)} \\ \Xi(\beta,V,\mu) &= \sum_n e^{-\beta(E_n-\mu N_n)} = \mathsf{Tr}\; e^{-\beta\hat{K}} \quad . \end{aligned} \tag{4.126}$$

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

$$\hat{\varrho} = \frac{e^{-\beta\hat{K}}}{\mathsf{Tr}\; e^{-\beta\hat{K}}} \quad , \tag{4.127}$$

with $\hat{K} = \hat{H} - \mu\hat{N}$. 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:

$$\Omega(T,V,\mu) = -k_{\rm B}T\log\Xi(T,V,\mu) \qquad \Longleftrightarrow \qquad \Xi = \exp(-\Omega/k_{\rm B}T) \quad , \tag{4.128}$$

where $\Omega(T,V,\mu)$ is the *grand potential*, also known as the *Landau free energy*. The dimensionless quantity $z \equiv \exp(\mu/k_{\rm B}T)$ is called the *fugacity*. Thus, $\log P_n = (\Omega - E_n + \mu N_n)/k_{\rm B}T$.

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

$$\Xi(T,V,\mu) = \sum_N e^{\beta\mu N}\, Z(T,V,N) \quad . \tag{4.129}$$

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

$$\hat{K} = \hat{H} - \sum_a \mu_a \hat{N}_a \quad , \tag{4.130}$$

with $\Xi = \mathsf{Tr}\; e^{-\beta\hat{K}}$ as before. To compute averages within the GCE, we use the grand canonical density matrix $\hat{\varrho}$:

$$\left\langle \hat{A} \right\rangle = \mathsf{Tr}\left(\hat{\varrho}\,\hat{A}\right) = \frac{\sum_n \langle n|\hat{A}|n\rangle\, e^{-\beta(E_n-\mu N_n)}}{\sum_j e^{-\beta(E_j-\mu N_j)}} \quad , \tag{4.131}$$

4.5.2 Entropy and Gibbs-Duhem relation

In the GCE, the Boltzmann entropy is

$$\begin{aligned} S = -k_{\rm B}\sum_n P_n \log P_n &= -k_{\rm B}\sum_n P_n\left(\beta\Omega - \beta E_n + \beta\mu N_n\right) \\ &= -\frac{\Omega}{T} + \frac{\langle\hat{H}\rangle}{T} - \frac{\mu\,\langle\hat{N}\rangle}{T} \quad , \end{aligned} \tag{4.132}$$

which says

$$\Omega = E - TS - \mu N \quad , \tag{4.133}$$

where

$$E = \sum_n E_n \, P_n = \mathsf{Tr}\big(\hat{\varrho}\, \hat{H}\big) \qquad , \qquad N = \sum_n N_n \, P_n = \mathsf{Tr}\big(\hat{\varrho}\, \hat{N}\big) \quad . \tag{4.134}$$

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

$$d\Omega = -S\, dT - p\, dV - N d\mu \quad , \tag{4.135}$$

which entails

$$S = -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} \qquad , \qquad p = -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \qquad , \qquad N = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} \quad . \tag{4.136}$$

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:

$$\frac{\partial \Omega}{\partial V} = -\frac{k_{\rm B} T}{\Xi}\left(\frac{\partial \Xi}{\partial V}\right)_{T,\mu} = \frac{1}{\Xi}\sum_n \frac{\partial E_n}{\partial V}\, e^{-\beta(E_n - \mu N_n)} = \left(\frac{\partial E}{\partial V}\right)_{T,\mu} = -p(T, \mu) \quad . \tag{4.137}$$

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,

$$d\Omega = -S\, dT - p\, dV - N d\mu = -p\, dV - V dp \quad \Rightarrow \quad S\, dT - V dp + N d\mu = 0 \quad . \tag{4.138}$$

4.5.3 Fluctuations in the GCE

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

$$N = \langle\, \hat{N}\, \rangle = \frac{\mathsf{Tr}\ \hat{N}\, e^{-\beta(\hat{H} - \mu \hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H} - \mu \hat{N})}} = \frac{1}{\beta}\frac{\partial}{\partial \mu}\log \Xi = -\frac{\partial \Omega}{\partial \mu} \quad . \tag{4.139}$$

Therefore,

$$\frac{1}{\beta}\frac{\partial N}{\partial \mu} = -\frac{1}{\beta}\frac{\partial^2 \Omega}{\partial \mu^2} = \frac{\mathsf{Tr}\ \hat{N}^2\, e^{-\beta(\hat{H} - \mu \hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H} - \mu \hat{N})}} - \left(\frac{\mathsf{Tr}\ \hat{N}\, e^{-\beta(\hat{H} - \mu \hat{N})}}{\mathsf{Tr}\ e^{-\beta(\hat{H} - \mu \hat{N})}}\right)^2 = \big\langle \hat{N}^2 \big\rangle - \big\langle \hat{N} \big\rangle^2 \quad , \tag{4.140}$$

and thus

$$\frac{\big\langle \hat{N}^2 \big\rangle - \big\langle \hat{N} \big\rangle^2}{\big\langle \hat{N} \big\rangle^2} = \frac{k_{\rm B} T}{N^2}\left(\frac{\partial N}{\partial \mu}\right)_{T,V} = \frac{k_{\rm B} T}{V}\, \kappa_T \quad , \tag{4.141}$$

where κ_T is the isothermal compressibility. Note:

$$\begin{aligned}\left(\frac{\partial N}{\partial \mu}\right)_{T,V} &= \frac{\partial(N,T,V)}{\partial(\mu,T,V)} = -\frac{\partial(N,T,V)}{\partial(V,T,\mu)} \\ &= -\frac{\partial(N,T,V)}{\partial(N,T,p)} \cdot \overbrace{\frac{\partial(N,T,p)}{\partial(V,T,p)}}^{N/V} \cdot \overbrace{\frac{\partial(V,T,p)}{\partial(N,T,\mu)}}^{1} \cdot \overbrace{\frac{\partial(N,T,\mu)}{\partial(V,T,\mu)}}^{N/V} \\ &= -\frac{N^2}{V^2}\left(\frac{\partial V}{\partial p}\right)_{T,N} = \frac{N^2}{V}\,\kappa_T \quad .\end{aligned} \tag{4.142}$$

We thus arrive at the result

$$\frac{(\Delta N)_{\rm RMS}}{N} = \sqrt{\frac{k_{\rm B}T\,\kappa_T}{V}} \quad , \tag{4.143}$$

which again scales as $V^{-1/2}$. For the nonrelativistic ideal gas $\kappa_T = 1/p$ and the ratio is $N^{-1/2}$. Compare with the OCE result in eqn. 4.102.

4.5.4 Generalized susceptibilities in the GCE

We can appropriate the results from §4.4.9 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. We define the grand canonical Hamiltonian,

$$\hat{K} = \hat{H}_0 - \sum_a \mu_a\,\hat{N}_a - \sum_k \lambda_k\,\hat{Q}_k \quad . \tag{4.144}$$

We then have

$$\Omega\big(T,V,\{\mu_a\},\{\lambda_k\}\big) = -k_{\rm B}T\log\mathsf{Tr}\;e^{-\hat{K}/k_{\rm B}T} \quad , \tag{4.145}$$

whence

$$d\Omega = -S\,dT - p\,dV - \sum_a N_a\,d\mu_a - \sum_k Q_k\,d\lambda_k \quad , \tag{4.146}$$

where the un-hatted quantities $\{N_a, Q_k\}$ are statistical averages within the GCE, *viz.*

$$N_a = \langle\hat{N}_a\rangle = -\frac{\partial\Omega}{\partial\mu_a} \qquad , \qquad Q_k = \langle\hat{Q}_k\rangle = -\frac{\partial\Omega}{\partial\lambda_k} \quad . \tag{4.147}$$

This leads to various generalized susceptibilities and cross-susceptibilities,

$$\frac{\partial N_a}{\partial\mu_b} = -\frac{\partial^2\Omega}{\partial\mu_a\partial\mu_b} \qquad , \qquad \frac{\partial Q_k}{\partial\lambda_l} = -\frac{\partial^2\Omega}{\partial\lambda_k\partial\lambda_l} \qquad , \qquad \frac{\partial N_a}{\partial\lambda_k} = \frac{\partial Q_k}{\partial\mu_a} = -\frac{\partial^2\Omega}{\partial\mu_a\partial\lambda_k} \quad . \tag{4.148}$$

Note that the mixed second derivatives are independent of order and yield various Maxwell relations.

4.5.5 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

$$P_n = \lim_{\Delta E\to 0}\lim_{\Delta V\to 0}\frac{D_{\rm W}(E_{\rm U}-E_n\,,\,V_{\rm U}-V_n)\,\Delta E\,\Delta V}{D_{\rm U}(E_{\rm U},V_{\rm U})\,\Delta E\,\Delta V}. \tag{4.149}$$

Then

$$\begin{aligned}\log P_n &= \log D_{\rm W}(E_{\rm U}-E_n\,,\,V_{\rm U}-V_n) - \log D_{\rm U}(E_{\rm U},V_{\rm U}) \qquad (4.150)\\ &= \log D_{\rm W}(E_{\rm U},V_{\rm U}) - \log D_{\rm U}(E_{\rm U},V_{\rm U}) - E_n\,\frac{\partial\log D_{\rm W}(E,V)}{\partial E}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} - V_n\,\frac{\partial\log D_{\rm W}(E,V)}{\partial V}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} + \ldots,\end{aligned}$$

and thus $\log P_n = -\alpha-\beta E_n-\beta p\,V_n$, where the constants β and p are given by

$$\beta = \frac{\partial\log D_{\rm W}(E,V)}{\partial E}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}} = \frac{1}{k_{\rm B}T}\quad,\qquad p = k_{\rm B}T\,\frac{\partial\log D_{\rm W}(E,V)}{\partial V}\Big|_{\substack{E=E_{\rm U}\\V=V_{\rm U}}}\quad. \tag{4.151}$$

The corresponding partition function is

$$Y(T,p,N) = \mathsf{Tr}\;e^{-\beta(\hat H+pV)} = \frac{1}{V_0}\int_0^\infty dV\,e^{-\beta pV}\,Z(T,V,N) \equiv e^{-\beta G(T,p,N)}\quad, \tag{4.152}$$

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 $G(V_0') = G(V_0) + k_{\rm B}T\log(V_0'/V_0)$, 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.*

$$\begin{aligned}G &= E - TS + pV\\ dG &= -S\,dT + V\,dp + \mu\,dN\quad,\end{aligned} \tag{4.153}$$

which entails

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N}\quad,\qquad V = +\left(\frac{\partial G}{\partial p}\right)_{T,N}\quad,\qquad \mu = +\left(\frac{\partial G}{\partial N}\right)_{T,p}\quad. \tag{4.154}$$

4.6 Statistical Ensembles from Maximum Entropy

The basic principle: maximize the Boltzmann entropy,

$$S = -k_{\rm B}\sum_n P_n\log P_n\quad, \tag{4.155}$$

subject to a set of *constraints*. Constrained extremization using Lagrange's method of undetermined multipliers is reviewed in the appendix, §4.10.

4.6.1 Microcanonical ensemble

We maximize S subject to the single constraint

$$C = \sum_n P_n - 1 = 0 \quad . \tag{4.156}$$

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

$$S^* = S - k_{\rm B}\,\lambda\,C \quad , \tag{4.157}$$

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

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B}\lambda\,\delta C - k_{\rm B}\,C\,\delta\lambda \\ &= -k_{\rm B}\sum_n \Big[\log P_n + 1 + \lambda\Big]\delta P_n - k_{\rm B}\,C\,\delta\lambda \equiv 0 \quad . \end{aligned} \tag{4.158}$$

We conclude that $C = 0$ and that

$$\log P_n = -\big(1 + \lambda\big) \quad , \tag{4.159}$$

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

$$\Omega = \sum_n \Theta(E + \Delta E - E_n)\,\Theta(E_n - E) \quad , \tag{4.160}$$

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

4.6.2 Ordinary canonical ensemble

We maximize S subject to the two constraints

$$C_1 = \sum_n P_n - 1 = 0 \qquad , \qquad C_2 = \sum_n E_n\,P_n - E = 0 \quad . \tag{4.161}$$

We now have two Lagrange multipliers. We write

$$S^* = S - k_{\rm B}\sum_{j=1}^{2}\lambda_j\,C_j \quad , \tag{4.162}$$

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

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B}\sum_n \big(\lambda_1 + \lambda_2\,E_n\big)\,\delta P_n - k_{\rm B}\sum_{j=1}^{2} C_j\,\delta\lambda_j \\ &= -k_{\rm B}\sum_n \Big[\log P_n + 1 + \lambda_1 + \lambda_2\,E_n\Big]\delta P_n - k_{\rm B}\sum_{j=1}^{2} C_j\,\delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.163}$$

Thus, $C_1 = C_2 = 0$ and

$$\log P_n = -\big(1 + \lambda_1 + \lambda_2\,E_n\big) \quad . \tag{4.164}$$

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

$$P_n = \frac{1}{Z}\,e^{-\beta E_n} \qquad , \qquad Z = \sum_n e^{-\beta E_n} = \mathsf{Tr}\;e^{-\beta\hat{H}} \quad . \tag{4.165}$$

4.6.3 Grand canonical ensemble

We maximize S subject to the three constraints

$$C_1 = \sum_n P_n - 1 = 0 \quad , \quad C_2 = \sum_n E_n\,P_n - E = 0 \quad , \quad C_3 = \sum_n N_n\,P_n - N = 0 \quad . \tag{4.166}$$

We now have three Lagrange multipliers. We write

$$S^* = S - k_{\rm B} \sum_{j=1}^{3} \lambda_j\,C_j \quad , \tag{4.167}$$

and hence

$$\begin{aligned} \delta S^* &= \delta S - k_{\rm B} \sum_n \big(\lambda_1 + \lambda_2\,E_n + \lambda_3\,N_n\big)\,\delta P_n - k_{\rm B} \sum_{j=1}^{3} C_j\,\delta\lambda_j \\ &= -k_{\rm B} \sum_n \Big[\log P_n + 1 + \lambda_1 + \lambda_2\,E_n + \lambda_3\,N_n\Big]\delta P_n - k_{\rm B} \sum_{j=1}^{3} C_j\,\delta\lambda_j \equiv 0 \quad . \end{aligned} \tag{4.168}$$

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

$$\log P_n = -\big(1 + \lambda_1 + \lambda_2\,E_n + \lambda_3\,N_n\big) \quad . \tag{4.169}$$

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

$$P_n = \frac{1}{\Xi}\,e^{-\beta(E_n - \mu N_n)} \qquad , \qquad \Xi = \sum_n e^{-\beta(E_n - \mu N_n)} = \mathsf{Tr}\; e^{-\beta(\hat{H} - \mu\hat{N})} \quad . \tag{4.170}$$

4.6.4 Generalized Gibbs ensembles

Suppose we have p constraints of the form $\langle \hat{Q}_k \rangle = Q_k$ with $k \in \{1, \ldots, \mathsf{p}\}$ and c constraints of the form $\langle \hat{N}_a \rangle = N_a$ with $a \in \{1, \ldots, \mathsf{c}\}$, in addition to the constraint of overall normalization of the probability distribution P_n. Then construct the extended function S^*, with

$$\begin{aligned} \frac{1}{k_{\rm B}}\,S^*(P_1, \ldots, P_\Omega\,,\,\lambda\,,\,\eta_1\,,\,\ldots,\eta_{\mathsf{c}}\,,\,\theta_1, \ldots, \theta_{\mathsf{p}}) = &- \sum_n P_n \log P_n - \lambda\Big(\sum_n P_n - 1\Big) \\ &- \sum_{k=1}^{\mathsf{p}} \theta_k \Big(\sum_n P_n\,Q_{k,n} - Q_k\Big) - \sum_{a=1}^{\mathsf{c}} \eta_a \Big(\sum_n P_n\,N_{a,n} - N_a\Big) \quad , \end{aligned} \tag{4.171}$$

where λ, $\{\theta_k\}$ and $\{\eta_a\}$ comprise $1 + \mathsf{p} + \mathsf{c}$ undetermined Lagrange multipliers. The total number of states is taken to be Ω, *i.e.* $n \in \{1, \ldots, \Omega\}$. Setting the variation $\delta S^* = 0$, we obtain the following $\Omega + 1 + \mathsf{p} + \mathsf{c}$

equations:

$$\begin{aligned}
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial P_n} = -\log P_n - (1+\lambda) - \sum_{k=1}^{\mathsf{p}} \theta_k\, Q_{k,n} - \sum_{a=1}^{\mathsf{c}} \eta_a\, N_{a,n} \\
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial \lambda} = 1 - \sum_n P_n \\
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial \theta_k} = Q_k - \sum_n P_n\, Q_{k,n} \\
0 &= \frac{1}{k_{\rm B}}\frac{\partial S^*}{\partial \eta_a} = N_a - \sum_n P_n\, N_{a,n} \quad .
\end{aligned} \tag{4.172}$$

Thus, the probability distribution is given by

$$P_n(\boldsymbol{\theta},\boldsymbol{\eta}) = \frac{1}{\mathcal{Z}(\boldsymbol{\theta},\boldsymbol{\eta})}\,\exp\left\{-\sum_{k=1}^{\mathsf{p}}\theta_k\,Q_{k,n} - \sum_{a=1}^{\mathsf{c}}\eta_a\,N_{a,n}\right\} \tag{4.173}$$

with

$$\mathcal{Z}(\boldsymbol{\theta},\boldsymbol{\eta}) = \sum_n \exp\left\{-\sum_{k=1}^{\mathsf{p}}\theta_k\,Q_{k,n} - \sum_{a=1}^{\mathsf{c}}\eta_a\,N_{a,n}\right\} \tag{4.174}$$

as well as

$$Q_k = \sum_n P_n(\boldsymbol{\theta},\boldsymbol{\eta})\,Q_{k,n} \qquad , \qquad N_a = \sum_n P_n(\boldsymbol{\theta},\boldsymbol{\eta})\,N_{a,n} \quad , \tag{4.175}$$

where $Q_{k,n} = \langle\, n \,|\, \hat{Q}_k \,|\, n \,\rangle$ and $N_{a,n} = \langle\, n \,|\, \hat{N}_a \,|\, n \,\rangle$.

If for $k = 1$ we have $\hat{Q}_{k=1} = \hat{H}$, then $Q_{1,n} = E_n$ and we may define $\theta_1 \equiv \beta = 1/k_{\rm B}T$ and $\theta_k \equiv -\beta\lambda_k$ for $k \geq 2$. Similarly we may define $\eta_a \equiv -\beta\mu_a$. Further defining the generalized grand canonical Hamiltonian as

$$\hat{K}(\boldsymbol{\lambda},\boldsymbol{\mu}) \equiv \hat{H} - \sum_{k=2}^{\mathsf{p}}\lambda_k\,\hat{Q}_k - \sum_{a=1}^{\mathsf{c}}\mu_a\,\hat{N}_a \quad , \tag{4.176}$$

we have

$$\mathcal{Z}(T,\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathsf{Tr}\; e^{-\beta\hat{K}} = \sum_n e^{-K_n/k_{\rm B}T} \tag{4.177}$$

and

$$P_n(T,\boldsymbol{\lambda},\boldsymbol{\mu}) = \mathcal{Z}^{-1}\,e^{-K_n/k_{\rm B}T} \qquad , \qquad K_n = \langle\, n \,|\, \hat{K} \,|\, n \,\rangle \quad . \tag{4.178}$$

Note the correspondence of these results with those of §4.5.4.

4.7 Ideal Gas Statistical Mechanics

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

$$\begin{aligned} Z(T,V,N) &= \frac{1}{N!}\prod_{i=1}^{N}\int\frac{d^dx_i\,d^dp_i}{(2\pi\hbar)^d}\,e^{-\beta\boldsymbol{p}_i^2/2m} \\ &= \frac{V^N}{N!}\left(\int_{-\infty}^{\infty}\frac{dp}{2\pi\hbar}\,e^{-\beta p^2/2m}\right)^{Nd} = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^N \quad , \end{aligned} \tag{4.179}$$

where λ_T is the *thermal wavelength*,

$$\lambda_T = \sqrt{2\pi\hbar^2/mk_{\rm B}T} \quad . \tag{4.180}$$

The physical interpretation of λ_T is that it is proportional to 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

$$\begin{aligned} \Xi(T,V,\mu) &= \sum_{N=0}^{\infty} e^{\beta\mu N}\,Z(T,V,N) \\ &= \sum_{N=0}^{\infty}\frac{1}{N!}\left(\frac{Ve^{\mu/k_{\rm B}T}}{\lambda_T^d}\right)^N = \exp\left(\frac{Ve^{\mu/k_{\rm B}T}}{\lambda_T^d}\right) \quad . \end{aligned} \tag{4.181}$$

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

$$\Omega(T,V,\mu) = -Vk_{\rm B}T\,e^{\mu/k_{\rm B}T}/\lambda_T^d \quad . \tag{4.182}$$

Since $\Omega = -pV$ (see §4.5.2), we have

$$p(T,\mu) = k_{\rm B}T\,\lambda_T^{-d}\,e^{\mu/k_{\rm B}T} \quad . \tag{4.183}$$

The number density can also be calculated:

$$n = \frac{N}{V} = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = \lambda_T^{-d}\,e^{\mu/k_{\rm B}T} \quad . \tag{4.184}$$

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

4.7.1 Maxwell velocity distribution

The distribution function for momenta is given by

$$g(\boldsymbol{p}) = \left\langle\frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{p}_i - \boldsymbol{p})\right\rangle \quad . \tag{4.185}$$

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 $\langle A \rangle = \mathrm{Tr}\,\big(A e^{-\beta\hat{H}}\big) / \mathrm{Tr}\; e^{-\beta\hat{H}}$. Setting $i = 1$, all the integrals other than that over $\boldsymbol{p}_1$ divide out between numerator and denominator. We then have

$$g(\boldsymbol{p}) = \frac{\int d^3p_1\; \delta(\boldsymbol{p}_1 - \boldsymbol{p})\, e^{-\beta \boldsymbol{p}_1^2/2m}}{\int d^3p_1\; e^{-\beta \boldsymbol{p}_1^2/2m}} = (2\pi m k_{\rm B} T)^{-3/2}\, e^{-\beta \boldsymbol{p}^2/2m} \quad . \tag{4.186}$$

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

$$f(\boldsymbol{v})\, d^3v = g(\boldsymbol{p})\, d^3p \quad . \tag{4.187}$$

Hence,

$$f(\boldsymbol{v}) = \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} e^{-m\boldsymbol{v}^2/2k_{\rm B}T} \quad . \tag{4.188}$$

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

$$\int d^3p\; g(\boldsymbol{p}) = \int d^3v\; f(\boldsymbol{v}) = 1 \quad . \tag{4.189}$$

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

$$\tilde{f}(v) = 4\pi\, v^2 f(\boldsymbol{v}) = 4\pi \left(\frac{m}{2\pi k_{\rm B} T}\right)^{3/2} v^2\, e^{-mv^2/2k_{\rm B}T} \quad . \tag{4.190}$$

Note that $\tilde{f}(v)$ is normalized according to $\int\limits_0^\infty dv\, \tilde{f}(v) = 1$. It is convenient to represent v in units of $v_0 = \sqrt{k_{\rm B}T/m}$, in which case

$$\tilde{f}(v) = \frac{1}{v_0}\, \varphi(v/v_0) \qquad , \qquad \varphi(s) = \sqrt{\tfrac{2}{\pi}}\, s^2\, e^{-s^2/2} \quad . \tag{4.191}$$

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

$$C_k \equiv \langle s^k \rangle = \int\limits_0^\infty ds\; s^k\, \varphi(s) = \frac{2}{\sqrt{\pi}} \cdot 2^{k/2}\, \Gamma\left(\tfrac{3}{2} + \tfrac{k}{2}\right) \quad . \tag{4.192}$$

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

$$\big\langle v^k \big\rangle = C_k \left(\frac{k_{\rm B}T}{m}\right)^{k/2} \quad . \tag{4.193}$$

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.

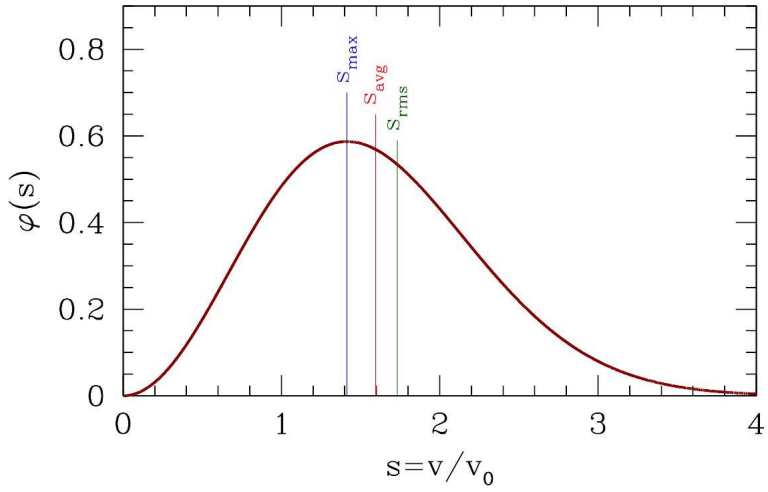


Figure 4.6: Maxwell distribution of speeds $\varphi(v/v_0)$. The most probable speed is $v_{\rm max} = \sqrt{2}\,v_0$. The average speed is $v_{\rm avg} = \sqrt{8/\pi}\,v_0$. The RMS speed is $v_{\rm rms} = \sqrt{3}\,v_0$.

4.7.2 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 $L = \frac{1}{2}\dot{\xi}^2 - \frac{1}{2}\,\omega_0^2\,\xi^2$, where the dimensions of ξ are $[\xi] = M^{1/2}L$ by convention. The Hamiltonian for this normal mode is then $\hat{H} = \frac{1}{2}p^2 + \frac{1}{2}\,\omega_0^2\,\xi^2$, 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 u – it may be a position or momentum or an angle or an angular momentum or some other generalized coordinate or conjugate momentum – and we'll write its contribution to $\hat{H}$ as $\hat{H}_u = \frac{1}{2}Ku^2$, where K is some constant. Integrating over u yields the following factor in the partition function:

$$\zeta(\beta) \equiv \int\limits_{-\infty}^{\infty} du\, e^{-\beta K u^2/2} = \left(\frac{2\pi}{K\beta}\right)^{1/2} \quad , \tag{4.194}$$

where $\beta = 1/k_{\rm B}T$. The contribution to the Helmholtz free energy is then

$$\Delta F_u = -k_{\rm B}T\log\zeta = \tfrac{1}{2}k_{\rm B}T\log\left(\frac{K}{2\pi k_{\rm B}T}\right) \quad , \tag{4.195}$$

and therefore the contribution to the internal energy E is

$$\Delta E_u = -\frac{\partial\log\zeta}{\partial\beta} = \tfrac{1}{2}k_{\rm B}T \quad . \tag{4.196}$$

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 $E = \frac{1}{2}fNk_{\rm B}T$, and $C_V = \frac{1}{2}fNk_{\rm B}$. 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 $E = 3Nk_{\rm B}T$, and the heat capacity $C_V = 3Nk_{\rm B}$. 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*.

For a degree of freedom which enters the Hamiltonian as a power, such as $\hat{H}_u = \frac{1}{2}K|u|^\sigma$ we have

$$\zeta(\beta) = 2\int_0^\infty du\, e^{-\beta K u^\sigma/2} = \frac{2^{1+\sigma^{-1}}\,\Gamma(1/\sigma)}{\sigma K^{1/\sigma}}\,\beta^{-1/\sigma} \tag{4.197}$$

after substituting $u(t) = (2t/\beta K)^{1/\sigma}$ and integrating over t. We then obtain $\Delta E_u = k_{\rm B}T/\sigma$.

4.8 Selected Examples

4.8.1 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

$$\hat{H} = -\mu_0 H \sum_{j=1}^{N_{\rm s}} \sigma_j \quad , \tag{4.198}$$

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

$$Z = \sum_{\sigma_1} \cdots \sum_{\sigma_{N_{\rm s}}} e^{-\beta\hat{H}} = \zeta^{N_{\rm s}} \quad , \tag{4.199}$$

where ζ is the single particle partition function:

$$\zeta = \sum_{\sigma=\pm1} e^{\mu_0 H\sigma/k_{\rm B}T} = 2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.200}$$

The Helmholtz free energy is then

$$F(T,H,N_{\rm s}) = -k_{\rm B}T\log Z = -N_{\rm s}k_{\rm B}T\log\left[2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.201}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T,\,N_{\rm s}} = N_{\rm s}\mu_0 \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.202}$$

The energy is

$$E = \frac{\partial}{\partial\beta}\left(\beta F\right) = -N_{\rm s}\mu_0 H \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.203}$$

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

$$P_\sigma = \frac{e^{\beta\mu_0 H\sigma}}{e^{\beta\mu_0 H} + e^{-\beta\mu_0 H}} \quad . \tag{4.204}$$

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

$$\langle\sigma\rangle = P_+ - P_- = \tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.205}$$

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 $T \gg \mu_0 H/k_{\rm B}$, 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

$$\chi_T = \frac{1}{N}\left(\frac{\partial M}{\partial H}\right)_T = \frac{\mu_0^2}{k_{\rm B}T}\,\mathrm{sech}^2\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.206}$$

(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:

$$\mathcal{E}(S, M, N_{\rm s}) = F(T, H, N_{\rm s}) + TS + HM \quad . \tag{4.207}$$

Now from eqn. 4.201 we derive the entropy

$$S = -\frac{\partial F}{\partial T} = N_{\rm s}k_{\rm B}\log\left[2\cosh\left(\frac{\mu_0 H}{k_{\rm B}T}\right)\right] - \frac{N_{\rm s}\mu_0 H}{T}\,\tanh\left(\frac{\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.208}$$

Thus, using eqns. 4.201 and 4.202, 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 $G(T, p, N_{\rm s})$ 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. 4.203 is in fact $E = \mathcal{E} - HM$, which means $\mathcal{E} = 0$. The energy $\mathcal{E}(S, M, N)$ vanishes here because the spins are noninteracting.

4.8.2 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 microcanonical ensemble, as described in §4.2.5. The number of ways of arranging the system with N_+ up spins is

$$\Omega = \binom{N_{\rm s}}{N_+} = \frac{N_{\rm s}!}{N_+!\,N_-!} \quad . \tag{4.209}$$

Using Stirling's expression $\log K! = K \log K - K + \mathcal{O}(\log K)$ for large K, we have

$$\begin{aligned} S = k_{\rm B} \log \Omega &= k_{\rm B}(N_{\rm s} \log N_{\rm s} - N_{\rm s}) - k_{\rm B}(N_+ \log N_+ - N_+) - k_{\rm B}(N_- \log N_- - N_-) \\ &= -N_{\rm s} k_{\rm B} \Big\{ x \log x + (1-x) \log(1-x) \Big\} \end{aligned} \tag{4.210}$$

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

$$\frac{E}{E_0} = -\frac{M}{N_{\rm s}\mu_0} = 1 - 2x \qquad \Longrightarrow \qquad x = \frac{E_0 - E}{2E_0} \quad . \tag{4.211}$$

We therefore have

$$S(E, N_{\rm s}) = -N_{\rm s} k_{\rm B} \left[\left(\frac{E_0 - E}{2E_0} \right) \log\left(\frac{E_0 - E}{2E_0} \right) + \left(\frac{E_0 + E}{2E_0} \right) \log\left(\frac{E_0 + E}{2E_0} \right) \right] \quad . \tag{4.212}$$

We now have

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{N_{\rm s}} = \frac{\partial S}{\partial x}\,\frac{\partial x}{\partial E} = \frac{N_{\rm s} k_{\rm B}}{2E_0} \log\left(\frac{E_0 - E}{E_0 + E} \right) \quad . \tag{4.213}$$

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 $S_{\rm max} = N_{\rm s} k_{\rm B} \log 2$ 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. 4.203 : 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 §4.11.3.

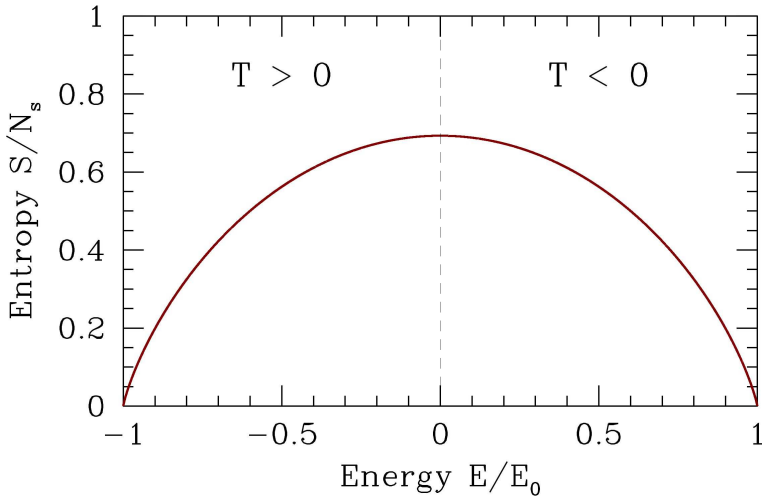


Figure 4.7: When entropy decreases with increasing energy, the temperature is negative. Typically, kinetic degrees of freedom prevent this peculiarity from manifesting in physical systems.

4.8.3 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.

SOLUTION: The grand partition function for the surface is

$$\Xi_{\rm surface} = e^{-\Omega_{\rm surface}/k_{\rm B}T} = \sum_{j=0}^{N_{\rm s}} \binom{N_{\rm s}}{j} e^{j(\mu+\Delta)/k_{\rm B}T} = \left(1 + e^{\mu/k_{\rm B}T}\, e^{\Delta/k_{\rm B}T}\right)^{N_{\rm s}} \quad . \tag{4.214}$$

The fraction of occupied sites is

$$f = \frac{\langle \hat{N}_{\rm surface} \rangle}{N_{\rm s}} = -\frac{1}{N_{\rm s}} \frac{\partial \Omega_{\rm surface}}{\partial \mu} = \frac{e^{\mu/k_{\rm B}T}}{e^{\mu/k_{\rm B}T} + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.215}$$

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.

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

$$\Xi_{\rm gas} = \exp\left(V\lambda_T^{-3}\, e^{\mu/k_{\rm B}T}\right) \quad . \tag{4.216}$$

The gas density is

$$n = \frac{\langle \hat{N}_{\rm gas} \rangle}{V} = -\frac{1}{V} \frac{\partial \Omega_{\rm gas}}{\partial \mu} = \lambda_T^{-3}\, e^{\mu/k_{\rm B}T} \quad . \tag{4.217}$$

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

$$f = \frac{n\lambda_T^3}{n\lambda_T^3 + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.218}$$

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 $f \to 0$, since the energetic cost of adsorption is infinitely high.

4.8.4 Dipole gas

Consider a gas of polar diatomic molecules, each endowed with a dipole moment $\boldsymbol{d}_j = \mu_0 \, \hat{\boldsymbol{n}}_j$. Here $\hat{\boldsymbol{n}}_j = (\sin\theta_j \cos\phi_j, \sin\theta_j \sin\phi_j, \cos\theta_j)$ is the unit orientation vector for the symmetry axis of molecule j, which points in the direction of the molecule's dipole moment. The single molecule Hamiltonian is then

$$\hat{h} = \frac{\boldsymbol{p}^2}{2m} + \frac{p_\theta^2}{2I} + \frac{p_\phi^2}{2I\sin^2\theta} - \mu_0 \boldsymbol{E} \cdot \hat{\boldsymbol{n}} \quad . \tag{4.219}$$

Without loss of generality we may choose $\boldsymbol{E} = E\hat{\boldsymbol{z}}$, hence the dipole energy function is $u(\theta) = -\mu_0 E \cos\theta$. Here $\boldsymbol{p} = (p_x, p_y, p_z)$ is the vector of linear momenta and $(p_\phi, p_\theta, p_\psi)$ are the *angular* momenta conjugate to the Euler angles ϕ, θ, and ψ. For a symmetric top, the Euler angles (θ, ϕ) are, respectively, the polar and azimuthal angles of the symmetry axis, and ψ is the angle of rotation about the symmetry axis. The inertia tensor in the center-of-mass frame is then $\mathsf{diag}(I_1, I_1, I_3)$, and for diatomic molecules I_3 is on the order of $ma_{\rm nuc}^2$, where $a_{\rm nuc}$ is the nuclear diameter. Thus I_3 is so small that the lowest nonzero quantized rotation energies $\propto \hbar^2/2I_3$ are much larger than $k_{\rm B}T$ in all conceivable applications, and we may forget about the $\{\psi, p_\psi\}$ degrees of freedom and set $I_1 = I_2 \equiv I$. This results in the rotational kinetic energy term in the above molecular Hamiltonian and a phase space metric

$$d\mu = \frac{dx\,dp_x}{h} \cdot \frac{dy\,dp_y}{h} \cdot \frac{dz\,dp_z}{h} \cdot \frac{d\theta\,dp_\theta}{h} \cdot \frac{d\phi\,dp_\phi}{h} \tag{4.220}$$

and the single molecule partition function is

$$\begin{aligned} \zeta(T,V,E) &= \int d\mu \, e^{-\hat{h}/k_{\rm B}T} = \frac{V}{\lambda_T^3} \int_0^{2\pi} d\phi \int_0^{\pi} d\theta \int_{-\infty}^{\infty} \frac{dp_\theta}{h} \int_{-\infty}^{\infty} \frac{dp_\phi}{h} \, e^{-\beta p_\theta^2/2I} \, e^{-\beta p_\phi^2/2I\sin^2\theta} \, e^{-\beta u(\theta)} \\ &= \frac{V}{\lambda_T^3} \cdot \frac{Ik_{\rm B}T}{\hbar^2} \int_0^{\pi} d\theta \, \sin\theta \, e^{-\beta u(\theta)} = \underbrace{\frac{V}{\lambda_T^3}}_{\zeta_{\rm trans}} \cdot \underbrace{\frac{2Ik_{\rm B}T}{\hbar^2}}_{\zeta_{\rm rot}} \cdot \underbrace{\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}}_{\zeta_{\rm pol}} \quad , \end{aligned} \tag{4.221}$$

where

$$\zeta_{\rm trans}(T,V) = \frac{V}{\lambda_T^3} \qquad , \qquad \zeta_{\rm rot}(T) = \frac{2Ik_{\rm B}T}{\hbar^2} \qquad , \qquad \zeta_{\rm pol}(T,E) = \frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T} \tag{4.222}$$

are the contributions to the single molecule partition function from translational, rotational, and polarization energies, respectively. The N-molecule partition function is $Z(T,V,N,E) = \zeta(T,V,E)^N/N!$ and thus in the thermodynamic limit, using Stirling's rule,

$$\begin{aligned} F(T,V,N,E) &= -Nk_{\rm B}T\log\left[\zeta(T,V,E)/N\right] - Nk_{\rm B}T \\ &= -Nk_{\rm B}T\left\{\log\left(\frac{V}{N\lambda_T^3}\right) + \log\left(\frac{2Ik_{\rm B}T}{\hbar^2}\right) + \log\left(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\right) + 1\right\} \quad . \end{aligned} \tag{4.223}$$

The average dipole moment per molecule is given by

$$\langle \boldsymbol{d}\rangle = \mu_0\,\langle\hat{\boldsymbol{n}}\rangle = -\frac{1}{N}\left(\frac{\partial F}{\partial \boldsymbol{E}}\right)_{T,V,N} = \mu_0\left\{\mathrm{ctnh}\left(\frac{\mu_0 E}{k_{\rm B}T}\right) - \frac{k_{\rm B}T}{\mu_0 E}\right\}\widehat{\boldsymbol{E}} \quad , \tag{4.224}$$

where $\widehat{\boldsymbol{E}} = \hat{\boldsymbol{z}}$.

Let's work out the physics in the GCE. The grand partition function is given by

$$\begin{aligned} \Xi(T,V,\mu,E) &= \sum_{N=0}^{\infty} z^N Z(T,V,N,E) = \exp\bigl(z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\,\zeta_{\rm pol}\bigr) \\ &= \exp\left\{\left(\frac{z\,V}{\lambda_T^3}\right)\left(\frac{2Ik_{\rm B}T}{\hbar^2}\right)\left(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\right)\right\} \quad , \end{aligned} \tag{4.225}$$

where $z = \exp(\mu/k_{\rm B}T)$ is the fugacity. Thus, the grand potential is

$$\begin{aligned} \Omega(T,V,\mu,E) &= -k_{\rm B}T\log\Xi(T,V,\mu,E) = -k_{\rm B}T\,z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\,\zeta_{\rm pol} \\ &= -k_{\rm B}T\left(\frac{z\,V}{\lambda_T^3}\right)\left(\frac{2Ik_{\rm B}T}{\hbar^2}\right)\left(\frac{\sinh(\mu_0 E/k_{\rm B}T)}{\mu_0 E/k_{\rm B}T}\right) \quad . \end{aligned} \tag{4.226}$$

As a sanity check, we compute $\langle\boldsymbol{d}\rangle$ within the GCE and make sure we get the same result in the thermodynamic limit. We have the average *total* dipole moment is

$$\boldsymbol{D} = \Big\langle\sum_{j=1}^{N}\boldsymbol{d}_j\Big\rangle = -\left(\frac{\partial\Omega}{\partial\boldsymbol{E}}\right)_{T,V,\mu} = k_{\rm B}T\,z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\left(\frac{\partial\zeta_{\rm pol}}{\partial\boldsymbol{E}}\right)_{T,V,\mu} \quad . \tag{4.227}$$

The average number of particles is

$$N = -\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V,\boldsymbol{E}} = z\,\zeta_{\rm trans}\,\zeta_{\rm rot}\,\zeta_{\rm pol} \tag{4.228}$$

and therefore

$$\boldsymbol{d} = \frac{\boldsymbol{D}}{N} = k_{\rm B}T\left(\frac{\partial\log\zeta_{\rm pol}}{\partial\boldsymbol{E}}\right)_{T,V,\mu} = \mu_0\left\{\mathrm{ctnh}\left(\frac{\mu_0 E}{k_{\rm B}T}\right) - \frac{k_{\rm B}T}{\mu_0 E}\right\}\widehat{\boldsymbol{E}} \quad . \tag{4.229}$$

4.8.5 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 $N_{\rm m}$ monomers is placed under a tension τ. We then have

$$\hat{H} = \sum_{j=1}^{N_{\rm m}} \Big[\varepsilon_{\rm A}\, \delta_{\sigma_j,+1} + \varepsilon_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.230}$$

Similarly, the length is

$$\hat{L} = \sum_{j=1}^{N_{\rm m}} \Big[\ell_{\rm A}\, \delta_{\sigma_j,+1} + \ell_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad . \tag{4.231}$$

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

$$\hat{K} = \sum_{j=1}^{N_{\rm m}} \Big[\tilde{\varepsilon}_{\rm A}\, \delta_{\sigma_j,+1} + \tilde{\varepsilon}_{\rm B}\, \delta_{\sigma_j,-1} \Big] \quad , \tag{4.232}$$

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

$$\tau^* = \frac{\varepsilon_{\rm B} - \varepsilon_{\rm A}}{\ell_{\rm B} - \ell_{\rm A}} \quad . \tag{4.233}$$

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

$$\hat{h} = \tilde{\varepsilon}_{\rm A}\, \delta_{\sigma,1} + \tilde{\varepsilon}_{\rm B}\, \delta_{\sigma,-1} \tag{4.234}$$

is the single "spin" Hamiltonian. Thus,

$$\zeta = {\rm Tr}\; e^{-\beta \hat{h}} = e^{-\beta \tilde{\varepsilon}_{\rm A}} + e^{-\beta \tilde{\varepsilon}_{\rm B}} \quad , \tag{4.235}$$

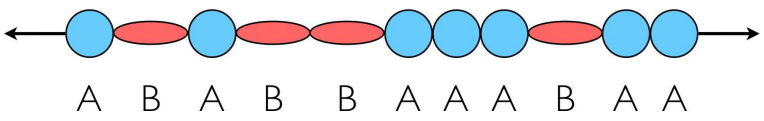


Figure 4.8: The monomers in wool are modeled as existing in one of two states. The low energy undeformed state is A, and the higher energy deformed state is B. Applying tension induces more monomers to enter the B state.

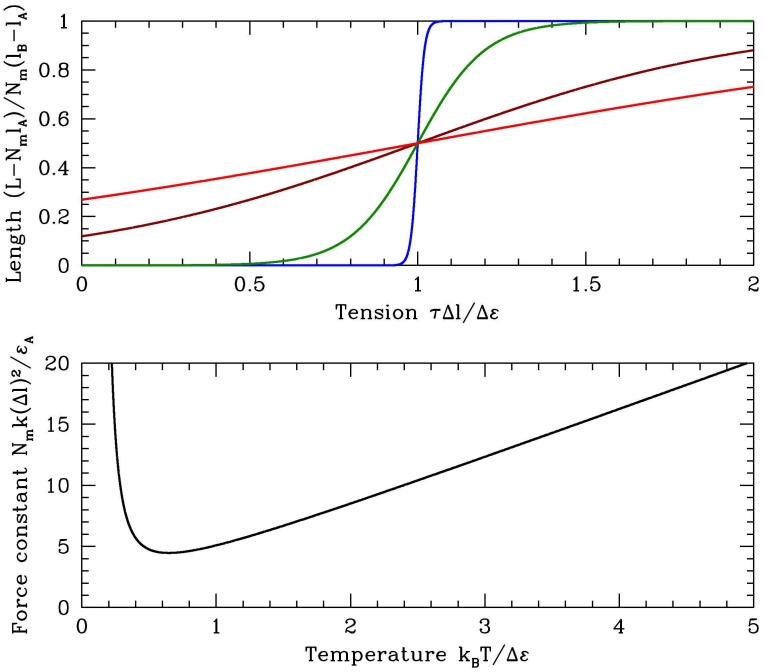


Figure 4.9: Upper panel: length $L(\tau, T)$ for $k_{\rm B}T/\tilde{\varepsilon} = 0.01$ (blue), 0.1 (green), 0.5 (dark red), and 1.0 (red). Bottom panel: dimensionless force constant $N_{\rm m}(\Delta\ell)^2 k/\varepsilon_{\rm A}$ *versus* dimensionless temperature $k_{\rm B}T/\Delta\varepsilon$.

It is convenient to define the differences

$$\Delta\varepsilon = \varepsilon_{\rm B} - \varepsilon_{\rm A} \quad , \quad \Delta\ell = \ell_{\rm B} - \ell_{\rm A} \quad , \quad \Delta\tilde{\varepsilon} = \tilde{\varepsilon}_{\rm B} - \tilde{\varepsilon}_{\rm A} \tag{4.236}$$

in which case the partition function Y and Gibbs free energy G are

$$\begin{aligned} Y(T,\tau,N_{\rm m}) &= e^{-N_{\rm m}\beta\,\tilde{\varepsilon}_{\rm A}}\Big[1 + e^{-\beta\Delta\tilde{\varepsilon}}\Big]^{N_{\rm m}} \\ G(T,\tau,N_{\rm m}) &= N_{\rm m}\,\tilde{\varepsilon}_{\rm A} - N_{\rm m}\,k_{\rm B}T\log\Big[1 + e^{-\Delta\tilde{\varepsilon}/k_{\rm B}T}\Big] \end{aligned} \tag{4.237}$$

The average length is

$$L = \langle\hat{L}\rangle = -\left(\frac{\partial G}{\partial\tau}\right)_{T,N_{\rm m}} = N_{\rm m}\,\ell_{\rm A} + \frac{N_{\rm m}\,\Delta\ell}{e^{(\Delta\varepsilon-\tau\Delta\ell)/k_{\rm B}T}+1} \quad . \tag{4.238}$$

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

$$k = \left.\frac{\partial\tau}{\partial L}\right|_{\tau=0} = \frac{4k_{\rm B}T}{N_{\rm m}(\Delta\ell)^2}\,\cosh^2\left(\frac{\Delta\varepsilon}{2k_{\rm B}T}\right) \quad . \tag{4.239}$$

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 all monomers are in the A state, hence $L = N_{\rm m}\,\ell_{\rm A}$. At high temperatures $k_{\rm B}T$ dominates and both A and B configurations are equally likely. The mean length is then $L = \frac{1}{2}N_{\rm m}(\ell_{\rm A} + \ell_{\rm B})$.

4.9 Statistical Mechanics of Molecular Gases

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

$$\hat{h} = \frac{\boldsymbol{p}^2}{2m} + \hat{h}_{\rm int} \quad , \tag{4.240}$$

with

$$\hat{h}\,|\,\boldsymbol{k}\,,\,\alpha\,\rangle = \left(\frac{\hbar^2\boldsymbol{k}^2}{2m} + \varepsilon_\alpha\right)|\,\boldsymbol{k}\,,\,\alpha\,\rangle \quad . \tag{4.241}$$

The partition function is

$$\zeta = {\rm Tr}\; e^{-\beta\hat{h}} = \sum_{\boldsymbol{p}} e^{-\beta\boldsymbol{p}^2/2m} \sum_j g_j\, e^{-\beta\varepsilon_j} \quad . \tag{4.242}$$

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

$$\boldsymbol{p} = \left(\frac{2\pi\hbar n_1}{L_1}\,,\,\frac{2\pi\hbar n_2}{L_2}\,,\,\ldots\,,\,\frac{2\pi\hbar n_d}{L_d}\right) \quad , \tag{4.243}$$

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:

$$\sum_{\boldsymbol{p}} \longrightarrow \int \frac{d^dp}{\Delta p_1 \cdots \Delta p_d} \tag{4.244}$$

where the volume in momentum space of an elementary rectangle is

$$\Delta p_1 \cdots \Delta p_d = \frac{(2\pi\hbar)^d}{L_1 \cdots L_d} = \frac{(2\pi\hbar)^d}{V} \quad . \tag{4.245}$$

Thus,

$$\zeta = V\!\int\!\frac{d^dp}{(2\pi\hbar)^d}\, e^{-\boldsymbol{p}^2/2mk_{\rm B}T} \sum_j g_j\, e^{-\varepsilon_j/k_{\rm B}T} = V\lambda_T^{-d}\,\xi(T) \tag{4.246}$$

where

$$\xi(T) = \sum_j g_j\, e^{-\varepsilon_j/k_{\rm B}T} \tag{4.247}$$

is the *internal coordinate partition function*. The N-particle ordinary canonical partition function is then

$$Z_N = \frac{1}{N!}\left(\frac{V}{\lambda_T^d}\right)^N \xi^N(T) \quad . \tag{4.248}$$

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

$$F(T,V,N) = -Nk_{\rm B}T\left[\log\left(\frac{V}{N\lambda_T^d}\right) + 1 + \log\xi(T)\right] \quad , \tag{4.249}$$

which entails the ideal gas law $pV = Nk_{\rm B}T$ since $dF = -S\,dT - p\,dV + \mu\,dN$. Making the Legendre transform to $G(T,p,N) = F(T,V,N) + pV$, we have

$$G(T,p,N) = Nk_{\rm B}T\log\left(\frac{p\,\lambda_T^d}{k_{\rm B}T}\right) + N\varphi(T) \quad , \tag{4.250}$$

where

$$\varphi(T) = -k_{\rm B}T\log\xi(T) \tag{4.251}$$

is the internal coordinate contribution to the single particle free energy. Recall that $G = \mu N$, which is to say $\mu(T,p) = G(T,p,N)/N$.

The entropy is

$$S = -\left(\frac{\partial G}{\partial T}\right)_{p,N} = Nk_{\rm B}\left[\log\left(\frac{k_{\rm B}T}{p\lambda_T^d}\right) + 1 + \tfrac{1}{2}d\right] - N\varphi'(T) \quad , \tag{4.252}$$

and therefore the heat capacities are

$$\begin{aligned} C_{p,N} &= T\left(\frac{\partial S}{\partial T}\right)_{p,N} = \left(\tfrac{1}{2}d+1\right)Nk_{\rm B} - NT\varphi''(T) \\ C_{V,N} &= T\left(\frac{\partial S}{\partial T}\right)_{V,N} = \tfrac{1}{2}dNk_{\rm B} - NT\varphi''(T) \quad , \end{aligned} \tag{4.253}$$

and we have $C_{p,N} - C_{V,N} = Nk_{\rm B}$. Furthermore, any temperature variation in the specific heats must be due to the internal degrees of freedom.

4.9.2 The internal coordinate partition function

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

$$\hat{h}_{\rm int} = \hat{h}_{\rm rot} + \hat{h}_{\rm vib} + \hat{h}_{\rm elec} \tag{4.254}$$

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

$$\xi_{\rm int}(T) = \xi_{\rm rot}(T)\cdot\xi_{\rm vib}(T)\cdot\xi_{\rm elec}(T) \quad . \tag{4.255}$$

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

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.

4.9.3 Quantum theory of rotations

Consider a class of molecules which can be approximated as axisymmetric tops. Let $\boldsymbol{L}$ be the molecular angular momentum vector with components $L^{x,y,z}$ along *space-fixed* axes. We write $\vec{\mathfrak{L}} = (\mathfrak{L}_{\rm a}, \mathfrak{L}_{\rm b}, \mathfrak{L}_{\rm c})$ as the angular momentum vector about instantaneous *body-fixed* principal axes. Thus $\mathfrak{L}_{\rm a} = \hat{\boldsymbol{n}}_{\rm a} \cdot \boldsymbol{L}$, $\mathfrak{L}_{\rm b} = \hat{\boldsymbol{n}}_{\rm b} \cdot \boldsymbol{L}$, and $\mathfrak{L}_{\rm c} = \hat{\boldsymbol{n}}_{\rm c} \cdot \boldsymbol{L}$, where $\hat{\boldsymbol{n}}_{\rm a,b,c}(t)$ are the instantaneous principal axes. We take $\hat{\boldsymbol{n}}_{\rm c}$ to be the symmetry axis of the molecule. The rotational Hamiltonian is then

$$\begin{aligned}\hat{h}_{\rm rot} &= \frac{\mathfrak{L}_{\rm a}^2 + \mathfrak{L}_{\rm b}^2}{2I_1} + \frac{\mathfrak{L}_{\rm c}^2}{2I_3} \\ &= \frac{\hbar^2 L(L+1)}{2I_1} + \left(\frac{1}{2I_3} - \frac{1}{2I_1}\right)\mathfrak{L}_{\rm c}^2 \quad .\end{aligned} \tag{4.256}$$

Note that

$$\left[L^\mu\,,\,\mathfrak{L}_c\right] = n_{\rm c}^\nu\left[L^\mu\,,\,L^\nu\right] + \left[L^\mu\,,\,n_{\rm c}^\nu\right]L^\nu = i\epsilon_{\mu\nu\lambda}\,\hbar n_{\rm c}^\nu\,L^\lambda + i\epsilon_{\mu\nu\lambda}\,\hbar n_{\rm c}^\lambda\,L^\nu = 0 \quad , \tag{4.257}$$

which is equivalent to the statement that $\mathfrak{L}_{\rm c} = \hat{\boldsymbol{n}}_{\rm c} \cdot \boldsymbol{L}$ is a rotational scalar. We can therefore simultaneously specify the eigenvalues of $\{\boldsymbol{L}^2, L^z, \mathfrak{L}_{\rm c}\}$, which form a complete set of commuting observables (CSCO)[11]. The eigenvalues of L^z are $m\hbar$ with $m \in \{-L, \ldots, L\}$, while those of $\mathfrak{L}_{\rm c}$ 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 then define two temperature scales,

$$\Theta = \frac{\hbar^2}{2I_1 k_{\rm B}} \qquad , \qquad \Theta_3 = \frac{\hbar^2}{2I_3 k_{\rm B}} \quad . \tag{4.258}$$

Prolateness then means $\Theta > \Theta_3$. We conclude that the rotational partition function for an axisymmetric molecule is given by

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty}(2L+1)\,e^{-L(L+1)\,\Theta/T}\sum_{k=-L}^{L} e^{-k^2\,(\Theta-\Theta_3)/T} \tag{4.259}$$

[11] 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.

In diatomic molecules, I_3 is extremely small, and $\Theta_3 \gg k_{\rm B}T$ at all relevant temperatures. Only the $k = 0$ term contributes to the partition sum, and we have

$$\xi_{\rm rot}(T) = \sum_{L=0}^{\infty} (2L+1)\, e^{-L(L+1)\,\Theta/T} \quad . \tag{4.260}$$

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

$$\xi_{\rm rot}(T) = 1 + 3\, e^{-2\Theta/T} + 5\, e^{-6\Theta/T} + \ldots \tag{4.261}$$

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

$$\sum_{k=0}^{n} F_k = \int_0^n dk\, F(k) + \tfrac{1}{2}\big[F(0) + F(n)\big] + \sum_{j=1}^{\infty} \frac{\mathsf{B}_{2j}}{(2j)!}\Big[F^{(2j-1)}(n) - F^{(2j-1)}(0)\Big] \tag{4.262}$$

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

$$\mathsf{B}_0 = 1 \quad , \quad \mathsf{B}_1 = -\tfrac{1}{2} \quad , \quad \mathsf{B}_2 = \tfrac{1}{6} \quad , \quad \mathsf{B}_4 = -\tfrac{1}{30} \quad , \quad \mathsf{B}_6 = \tfrac{1}{42} \quad . \tag{4.263}$$

Thus,

$$\sum_{k=0}^{\infty} F_k = \int_0^{\infty} dx\, F(x) + \tfrac{1}{2}F(0) - \tfrac{1}{12}F'(0) - \frac{1}{720}\,F'''(0) + \ldots \quad . \tag{4.264}$$

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

$$\xi_{\rm rot} = \frac{T}{\Theta} + \frac{1}{3} + \frac{1}{15}\,\frac{\Theta}{T} + \frac{4}{315}\left(\frac{\Theta}{T}\right)^2 + \ldots \quad . \tag{4.265}$$

Recall that $\varphi(T) = -k_{\rm B}T \log \xi(T)$. We conclude that $\varphi_{\rm rot}(T) \approx -3k_{\rm B}T\, e^{-2\Theta/T}$ for $T \ll \Theta$ and $\varphi_{\rm rot}(T) \approx -k_{\rm B}T \log(T/\Theta)$ 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

$$\begin{aligned} \Delta C_V(T \ll \Theta) &= 12\, Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2 e^{-2\Theta/T} + \ldots \\ \Delta C_V(T \gg \Theta) &= Nk_{\rm B}\left\{1 + \frac{1}{45}\left(\frac{\Theta}{T}\right)^2 + \frac{16}{945}\left(\frac{\Theta}{T}\right)^3 + \ldots\right\} \quad . \end{aligned} \tag{4.266}$$

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 §4.9.6.

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

$$K(\mathfrak{L}_{\rm a},\mathfrak{L}_{\rm b},\mathfrak{L}_{\rm c}) = \frac{\mathfrak{L}_{\rm a}^2}{2I_1} + \frac{\mathfrak{L}_{\rm b}^2}{2I_2} + \frac{\mathfrak{L}_{\rm c}^2}{2I_3} \quad . \tag{4.267}$$

We then have

$$\xi_{\rm rot}(T) = \frac{1}{g_{\rm rot}} \int \frac{d\mathfrak{L}_{\rm a}\, d\mathfrak{L}_{\rm b}\, d\mathfrak{L}_{\rm c}\, d\phi\, d\theta\, d\psi}{(2\pi\hbar)^3}\, e^{-K(\mathfrak{L}_{\rm a}\,\mathfrak{L}_{\rm b}\,\mathfrak{L}_{\rm c})/k_{\rm B}T} \quad , \tag{4.268}$$

where (ϕ,θ,ψ) 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 occur when more than one of the nuclei is the same. We then have

$$\xi_{\rm rot}(T) = \frac{1}{g_{\rm rot}} \left(\frac{2k_{\rm B}T}{\hbar^2}\right)^{3/2} \sqrt{\pi I_1 I_2 I_3} \quad . \tag{4.269}$$

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

Quantum mechanically, the angular momentum operators with respect to principal axes may be expressed as differential operators in the Euler angles, *viz.*

$$\begin{aligned}
\mathfrak{L}_1 &= -i\hbar \left\{ \cos\psi \,\mathrm{ctn}\,\theta\, \frac{\partial}{\partial\psi} + \sin\psi\, \frac{\partial}{\partial\theta} - \cos\psi \,\mathrm{csc}\,\theta\, \frac{\partial}{\partial\phi} \right\} \\
\mathfrak{L}_2 &= -i\hbar \left\{ -\sin\psi \,\mathrm{ctn}\,\theta\, \frac{\partial}{\partial\psi} + \cos\psi\, \frac{\partial}{\partial\theta} + \sin\psi \,\mathrm{csc}\,\theta\, \frac{\partial}{\partial\phi} \right\} \\
\mathfrak{L}_3 &= -i\hbar\, \frac{\partial}{\partial\psi} \quad ,
\end{aligned} \tag{4.270}$$

for which $[\mathfrak{L}_i,\mathfrak{L}_j] = i\hbar\,\epsilon_{ijk}\,\mathfrak{L}_k$. The rigid body Hamiltonian is then

$$\hat{h}_{\rm rot} = \sum_{i=1}^{3} \frac{\mathfrak{L}_i^2}{2I_i} + u(\phi,\theta,\psi) \quad , \tag{4.271}$$

where, for the sake of generality, we include a potential energy $u(\phi,\theta,\psi)$.

4.9.4 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

$$\frac{hc}{k_{\rm B}} = 1.436\,\mathrm{K}\cdot\mathrm{cm} \quad . \tag{4.272}$$

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

$$\hat{h} = \frac{p^2}{2m} + \tfrac{1}{2}m\omega^2 q^2 = \hbar\omega\big(a^\dagger a + \tfrac{1}{2}\big) \quad . \tag{4.273}$$

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

$$\xi_{\rm vib} = \prod_\alpha \xi_{\rm vib}^{(\alpha)} \quad . \tag{4.274}$$

For each such normal mode, the contribution is

$$\begin{aligned}\xi &= \sum_{n=0}^\infty e^{-(n+\frac{1}{2})\hbar\omega/k_{\rm B}T} = e^{-\hbar\omega/2k_{\rm B}T}\sum_{n=0}^\infty \Big(e^{-\hbar\omega/k_{\rm B}T}\Big)^n \\ &= \frac{e^{-\hbar\omega/2k_{\rm B}T}}{1-e^{-\hbar\omega/k_{\rm B}T}} = \frac{1}{2\sinh(\Theta/2T)} \quad ,\end{aligned} \tag{4.275}$$

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

$$\varphi = k_{\rm B}T\log\Big(2\sinh(\Theta/2T)\Big) = \tfrac{1}{2}k_{\rm B}\Theta + k_{\rm B}T\log\big(1-e^{-\Theta/T}\big) \quad . \tag{4.276}$$

The contribution to the heat capacity is

$$\Delta C_V = Nk_{\rm B}\left(\frac{\Theta}{T}\right)^2\frac{e^{\Theta/T}}{(e^{\Theta/T}-1)^2} = \begin{cases} Nk_{\rm B}\,(\Theta/T)^2\,\exp(-\Theta/T) & (T\to 0) \\ Nk_{\rm B} & (T\to\infty) \end{cases} \quad . \tag{4.277}$$

4.9.5 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

$$\zeta = e^{-\beta\varepsilon_0} + e^{-\beta\varepsilon_1} = e^{-\beta\varepsilon_0}\big(1+e^{-\beta\Delta}\big) \quad . \tag{4.278}$$

The free energy is

$$f = -k_{\rm B}T\log\zeta = \varepsilon_0 - k_{\rm B}T\log\big(1+e^{-\Delta/k_{\rm B}T}\big) \quad . \tag{4.279}$$

The entropy for a given two level system is then

$$s = -\frac{\partial f}{\partial T} = k_{\rm B}\log\big(1+e^{-\Delta/k_{\rm B}T}\big) + \frac{\Delta}{T}\cdot\frac{1}{e^{\Delta/k_{\rm B}T}+1} \tag{4.280}$$

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

$$c(T) = \frac{\Delta^2}{k_{\rm B}T^2}\cdot\frac{e^{\Delta/k_{\rm B}T}}{\big(e^{\Delta/k_{\rm B}T}+1\big)^2} \quad . \tag{4.281}$$

Thus,

$$c\,(k_{\rm B}T \ll \Delta) = \frac{\Delta^2}{k_{\rm B}T^2}\,e^{-\Delta/k_{\rm B}T} \qquad , \qquad c\,(k_{\rm B}T \gg \Delta) = \frac{\Delta^2}{4k_{\rm B}T^2} \quad . \tag{4.282}$$

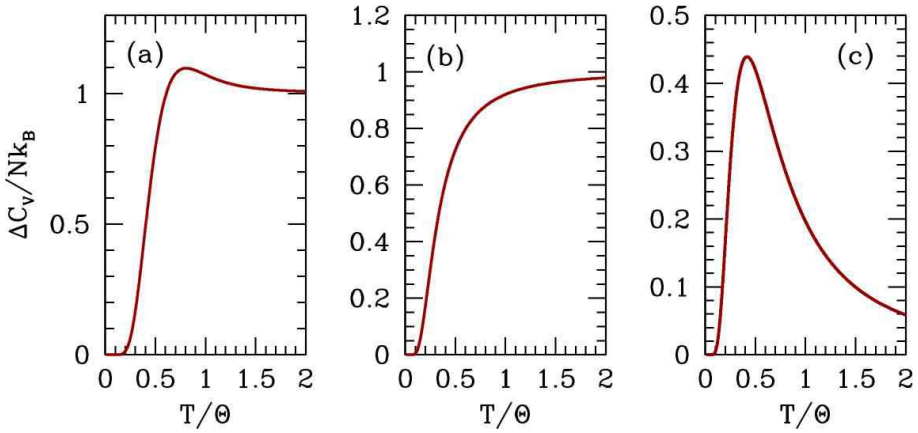


Figure 4.10: Heat capacity per molecule as a function of temperature for (a) heteronuclear diatomic gases, (b) a single vibrational mode, and (c) a single two-level system.

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:

$$C(T) = \sum_i \tilde{c}\,(\Delta_i/k_{\rm B}T) = N\int\limits_0^\infty d\Delta\,P(\Delta)\,\tilde{c}(\Delta/k_{\rm B}T) \quad , \tag{4.283}$$

where N is the number of two level systems, $\tilde{c}(x) = k_{\rm B}\,x^2\,e^x/(e^x+1)^2$, and where $P(\Delta)$ is the normalized distribution function, which satisfies the normalization condition $\int_0^\infty d\Delta\,P(\Delta) = 1$. 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 orientations of the cyanide (CN) dipoles.

4.9.6 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 atom has a closed-shell configuration, as for $Z = 2$ (He), 10 (Ne), 18 (Ar), 36 (Kr), 54 (Xe), *etc.*, then $L = 0$ 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 ${}^{2S+1}L_J$ 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 $J = L + S$.[12]

The electronic contribution to ξ is then

$$\xi_{\rm elec} = \sum_{J=|L-S|}^{L+S} (2J+1)\, e^{-\Delta\varepsilon(L,S,J)/k_{\rm B}T} \tag{4.284}$$

where

$$\Delta\varepsilon(L,S,J) = \tfrac{1}{2}\Lambda\Big[J(J+1) - L(L+1) - S(S+1)\Big] \quad . \tag{4.285}$$

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 ${}^1\Sigma_g^+$.[13]

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

[12] See *e.g.* §72 of Landau and Lifshitz, *Quantum Mechanics*.
[13] *Ibid.* §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.

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

$$\xi_{\rm rot}(T) = g_g\,\zeta_g(T) + g_u\,\zeta_u(T) \quad , \tag{4.286}$$

where

$$\zeta_g(T) = \sum_{L\ {\rm even}} (2L+1)\,e^{-L(L+1)\,\Theta_{\rm rot}/T} \quad , \qquad \zeta_u(T) = \sum_{L\ {\rm odd}} (2L+1)\,e^{-L(L+1)\,\Theta_{\rm rot}/T} \quad . \tag{4.287}$$

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

$$\frac{N^{\rm ortho}_{H_2}}{N^{\rm para}_{H_2}} = \frac{g_u\,\zeta_u(T)}{g_g\,\zeta_g(T)} = \frac{3\,\zeta_u(T)}{\zeta_g(T)} \quad , \qquad \frac{N^{\rm ortho}_{D_2}}{N^{\rm para}_{D_2}} = \frac{g_g\,\zeta_g(T)}{g_u\,\zeta_u(T)} = \frac{2\,\zeta_g(T)}{\zeta_u(T)} \quad . \tag{4.288}$$

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

$$I \otimes I = 0 \oplus 1 \oplus 2 \oplus \cdots \oplus 2I \quad . \tag{4.289}$$

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

$$g_g^{\rm even} = \sum_{n=0}^{I} \Big\{ 2\cdot(2n) + 1 \Big\} = (I+1)(2I+1) \quad , \tag{4.290}$$

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

$$g_u^{\rm even} = (2I+1)^2 - g_g^{\rm even} = I(2I+1) \quad . \tag{4.291}$$

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

$$g_g^{\text{odd}} = \sum_{n=0}^{I-\frac{1}{2}} \Big\{ 2 \cdot (2n) + 1 \Big\} = I(2I+1) \quad . \tag{4.292}$$

It follows that

$$g_u^{\text{odd}} = (2I+1)^2 - g_g^{\text{odd}} = (I+1)(2I+1) \quad . \tag{4.293}$$

4.10 Appendix I: Constrained Extremization of Functions

Given $F(x_1, \ldots, x_n)$ to be extremized subject to k constraints of the form $G_j(x_1, \ldots, x_n) = 0$ where $j = 1, \ldots, k$, construct

$$F^*\big(x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_k\big) \equiv F(x_1, \ldots, x_n) + \sum_{j=1}^{k} \lambda_j \, G_j(x_1, \ldots, x_n) \tag{4.294}$$

which is a function of the $(n+k)$ variables $\big\{x_1, \ldots, x_n; \lambda_1, \ldots, \lambda_k\big\}$, where the quantities $\{\lambda_1, \ldots, \lambda_k\}$ are *Lagrange undetermined multipliers*. We now *freely* extremize the extended function F^*:

$$\begin{aligned} dF^* &= \sum_{\sigma=1}^{n} \frac{\partial F^*}{\partial x_\sigma} \, dx_\sigma + \sum_{j=1}^{k} \frac{\partial F^*}{\partial \lambda_j} \, d\lambda_j \\ &= \sum_{\sigma=1}^{n} \left(\frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^{k} \lambda_j \, \frac{\partial G_j}{\partial x_\sigma} \right) dx_\sigma + \sum_{j=1}^{k} G_j \, d\lambda_j = 0 \end{aligned} \tag{4.295}$$

This results in the $(n+k)$ equations

$$\begin{aligned} \frac{\partial F}{\partial x_\sigma} + \sum_{j=1}^{k} \lambda_j \, \frac{\partial G_j}{\partial x_\sigma} = 0 \qquad & (\sigma = 1, \ldots, n) \\ G_j = 0 \qquad & (j = 1, \ldots, k) \quad . \end{aligned} \tag{4.296}$$

The interpretation of all this is as follows. The first n equations in 4.296 can be written in vector form as

$$\boldsymbol{\nabla} F + \sum_{j=1}^{k} \lambda_j \, \boldsymbol{\nabla} G_j = 0 \quad . \tag{4.297}$$

This says that the (n-component) vector $\boldsymbol{\nabla} F$ is linearly dependent upon the k vectors $\boldsymbol{\nabla} G_j$. Thus, any movement in the direction of $\boldsymbol{\nabla} F$ must necessarily entail movement along one or more of the directions $\boldsymbol{\nabla} G_j$. This would require violating the constraints, since movement along $\boldsymbol{\nabla} G_j$ takes us off the level set $G_j = 0$. Were $\boldsymbol{\nabla} F$ linearly *independent* of the set $\{\boldsymbol{\nabla} G_j\}$, this would mean that we could find a

differential displacement $d\boldsymbol{x}$ which has finite overlap with $\boldsymbol{\nabla} F$ but zero overlap with each $\boldsymbol{\nabla} G_j$. Thus $\boldsymbol{x} + d\boldsymbol{x}$ would still satisfy $G_j(\boldsymbol{x} + d\boldsymbol{x}) = 0$, but F would change by the finite amount $dF = \boldsymbol{\nabla} F(\boldsymbol{x}) \cdot d\boldsymbol{x}$.

Put another way, when we extremize $F(\boldsymbol{x})$ without constraints, we identify points $\boldsymbol{x} \in \mathbb{R}^n$ where the gradient $\boldsymbol{\nabla} F$ vanishes. However, when we have k constraints of the form $G_j(\boldsymbol{x}) = 0$, the subset

$$\Sigma = \big\{ \boldsymbol{x} \in \mathbb{R}^n \,|\, G_j(\boldsymbol{x}) = 0 \;\forall\; j \in \{1, \ldots, k\} \big\} \tag{4.298}$$

is a hypersurface of dimension $n - k$. Generically we should not expect any of the solutions to $\boldsymbol{\nabla} F = 0$ to lie within the subspace Σ. Extremizing $F(\boldsymbol{x})$ subject to the k constraints $G_j(\boldsymbol{x}) = 0$ means that we must find the extrema of $F(\boldsymbol{x})$ for $\boldsymbol{x} \in \Sigma \subset \mathbb{R}^n$. All such extrema satisfy that $\boldsymbol{\nabla} F(\boldsymbol{x})$ is *perpendicular* to the hypersurface Σ, *i.e.* $\boldsymbol{\nabla} F(\boldsymbol{x})$ must lie in the k-dimensional subspace spanned by the vectors $\boldsymbol{\nabla} G_j(\boldsymbol{x})$.

4.10.1 Example : volume of a cylinder

To see how this formalism works in practice, let's extremize the volume $V = \pi a^2 h$ of a cylinder of radius a and height h, subject to the constraint

$$G(a, h) = 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \tag{4.299}$$

Here, b and ℓ are constant parameters, each of which has dimensions of length.

Following Lagrange's method, we define the extended function

$$V^*(a, h, \lambda) \equiv V(a, h) + \lambda\, G(a, h) \quad , \tag{4.300}$$

and set

$$\begin{aligned} \frac{\partial V^*}{\partial a} &= 2\pi a h + 2\pi\lambda = 0 \\ \frac{\partial V^*}{\partial h} &= \pi a^2 + 2\lambda\,\frac{h}{b} = 0 \\ \frac{\partial V^*}{\partial \lambda} &= 2\pi a + \frac{h^2}{b} - \ell = 0 \quad . \end{aligned} \tag{4.301}$$

Solving these three equations simultaneously gives

$$a = \frac{2\ell}{5\pi} \quad , \quad h = \sqrt{\frac{b\ell}{5}} \quad , \quad \lambda = -\frac{2}{5^{3/2}\pi}\, b^{1/2}\, \ell^{3/2} \quad , \quad V^* = \frac{4}{5^{5/2}\,\pi}\, \ell^{5/2}\, b^{1/2} \quad . \tag{4.302}$$

4.11 Appendix II: Additional Examples of Statistical Ensembles

4.11.1 Noninteracting spin dimers

Consider a system of noninteracting spin dimers as depicted in fig. 4.11. Each dimer contains two spins, and is described by the Hamiltonian

$$\hat{h} = -J\,\sigma_1\sigma_2 - \mu_0 H\,(\sigma_1 + \sigma_2) \quad . \tag{4.303}$$

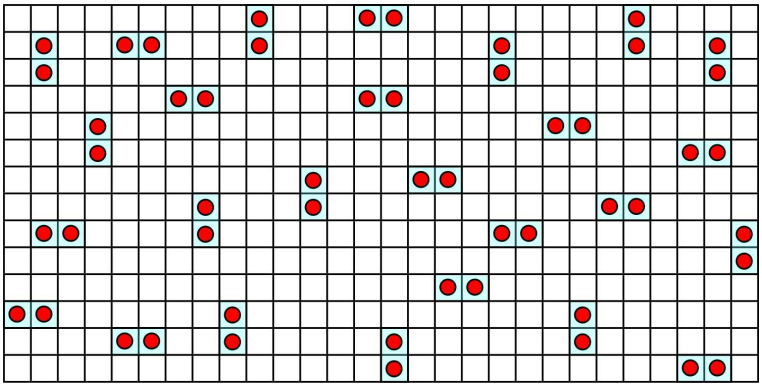


Figure 4.11: A model of noninteracting spin dimers on a lattice. Each red dot represents a classical spin for which $\sigma_j = \pm 1$.

Here, J is an *interaction energy* between the spins which comprise the dimer. If $J > 0$ the interaction is *ferromagnetic*, which prefers that the dimer 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 the dimer spins are anti-aligned, *i.e.* in configurations $|\uparrow\downarrow\rangle$ or $|\downarrow\uparrow\rangle$[14].

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

$$\zeta = {\rm Tr}\, e^{-\hat{h}/k_{\rm B}T} = 2\, e^{-J/k_{\rm B}T} + 2\, e^{J/k_{\rm B}T} \cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right) \quad . \tag{4.304}$$

Thus, the free energy is

$$F(T, H, N_{\rm d}) = -N_{\rm d}\, k_{\rm B}T \log 2 - N_{\rm d}\, k_{\rm B}T \log\left[e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T} \cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)\right] \quad . \tag{4.305}$$

The magnetization is

$$M = -\left(\frac{\partial F}{\partial H}\right)_{T, N_{\rm d}} = 2N_{\rm d}\, \mu_0 \cdot \frac{e^{J/k_{\rm B}T} \sinh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)}{e^{-J/k_{\rm B}T} + e^{J/k_{\rm B}T} \cosh\left(\frac{2\mu_0 H}{k_{\rm B}T}\right)} \tag{4.306}$$

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

$$\chi_T = \frac{1}{2N_{\rm d}} \frac{\partial M}{\partial H}\bigg|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T} \cdot \frac{2\, e^{J/k_{\rm B}T}}{e^{J/k_{\rm B}T} + e^{-J/k_{\rm B}T}} \quad . \tag{4.307}$$

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

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

are in fact noninteracting. For the ferromagnetic case, if $J \gg k_{\rm B}T$, then we obtain

$$\chi_T(J \gg k_{\rm B}T) \approx \frac{2\mu_0^2}{k_{\rm B}T} \quad . \tag{4.308}$$

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

$$\chi_T(-J \gg k_{\rm B}T) \approx \frac{2\mu_0^2}{k_{\rm B}T}\, e^{-2|J|/k_{\rm B}T} \quad . \tag{4.309}$$

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

4.11.2 Three-state system

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

$$\hat{h} = -\mu_0 H\, \sigma + \Delta(1 - \sigma^2) \quad . \tag{4.310}$$

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

$$\zeta = \mathsf{Tr}\; e^{-\beta\hat{h}} = e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H) \quad . \tag{4.311}$$

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

$$F \equiv Nf = -k_{\rm B}T \log Z = -Nk_{\rm B}T \log\Big[e^{-\beta\Delta} + 2\cosh(\beta\mu_0 H)\Big] \quad , \tag{4.312}$$

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

$$\hat{n}_{\rm V} = 1 - \sigma^2 = \frac{\partial\hat{h}}{\partial\Delta} \qquad , \qquad \hat{m} = \mu_0\,\sigma = -\frac{\partial\hat{h}}{\partial H} \quad , \tag{4.313}$$

are the vacancy number and magnetization, respectively. Thus,

$$n_{\rm V} = \big\langle\, \hat{n}_{\rm V} \,\big\rangle = \frac{\partial f}{\partial\Delta} = \frac{e^{-\Delta/k_{\rm B}T}}{e^{-\Delta/k_{\rm B}T} + 2\cosh(\mu_0 H/k_{\rm B}T)} \tag{4.314}$$

and

$$m = \big\langle\, \hat{m} \,\big\rangle = -\frac{\partial f}{\partial H} = \frac{2\mu_0\,\sinh(\mu_0 H/k_{\rm B}T)}{e^{-\Delta/k_{\rm B}T} + 2\cosh(\mu_0 H/k_{\rm B}T)} \quad . \tag{4.315}$$

At weak fields we can compute

$$\chi_T = \left.\frac{\partial m}{\partial H}\right|_{H=0} = \frac{\mu_0^2}{k_{\rm B}T} \cdot \frac{2}{2 + e^{-\Delta/k_{\rm B}T}} \quad . \tag{4.316}$$

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 $\sigma = 0$.

4.11.3 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 a 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

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

(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 $N_\downarrow$:

$$N_\uparrow + N_0 + N_\downarrow = Q + N_0 = N \quad . \tag{4.317}$$

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

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

$$\Omega = \frac{N!}{N_\uparrow!\, N_0!\, N_\downarrow!} \quad . \tag{4.318}$$

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

$$N_\uparrow = \tfrac{1}{2}\,(Q + M) \qquad , \qquad N_0 = N - Q \qquad , \qquad N_\downarrow = \tfrac{1}{2}\,(Q - M) \quad , \tag{4.319}$$

whence

$$\Omega(Q, M) = \frac{N!}{\left[\frac{1}{2}(Q+M)\right]!\,\left[\frac{1}{2}(Q-M)\right]!\,(N-Q)!} \quad . \tag{4.320}$$

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

$$S(Q, M) = k_{\rm B} \log(N!) - k_{\rm B} \log\left[\tfrac{1}{2}(Q+M)!\right] - k_{\rm B} \log\left[\tfrac{1}{2}(Q-M)!\right] - k_{\rm B} \log\left[(N-Q)!\right] \quad . \tag{4.321}$$

(b) Now we invoke Stirling's rule,

$$\log(N!) = N\log N - N + \mathcal{O}(\log N) \quad , \tag{4.322}$$

to obtain

$$\log\Omega(Q,M) = N\log N - N - \tfrac{1}{2}(Q+M)\log\left[\tfrac{1}{2}(Q+M)\right] + \tfrac{1}{2}(Q+M) \tag{4.323}$$

$$- \tfrac{1}{2}(Q-M)\log\left[\tfrac{1}{2}(Q-M)\right] + \tfrac{1}{2}(Q-M) - (N-Q)\log(N-Q) + (N-Q)$$

$$= N\log N - \tfrac{1}{2}Q\log\left[\tfrac{1}{4}(Q^2-M^2)\right] - \tfrac{1}{2}M\log\left(\frac{Q+M}{Q-M}\right) \tag{4.324}$$

Combining terms,

$$\log\Omega(Q,M) = -Nq\log\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] - \tfrac{1}{2}Nm\log\left(\frac{q+m}{q-m}\right) - N(1-q)\log(1-q) \quad , \tag{4.325}$$

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

$$s(q,m) = -k_{\rm B}\,q\log\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] - \tfrac{1}{2}k_{\rm B}\,m\log\left(\frac{q+m}{q-m}\right) - k_{\rm B}\,(1-q)\log(1-q) \quad . \tag{4.326}$$

The temperature is obtained from the relation

$$\begin{aligned}\frac{1}{T} &= \left(\frac{\partial S}{\partial E}\right)_M = -\frac{1}{W}\left(\frac{\partial s}{\partial q}\right)_m \\ &= -\frac{k_{\rm B}}{W}\log(1-q) + \frac{k_{\rm B}}{W}\log\left[\tfrac{1}{2}\sqrt{q^2-m^2}\right] \quad .\end{aligned} \tag{4.327}$$

Thus,

$$T = -\frac{W/k_{\rm B}}{\log\left[2(1-q)/\sqrt{q^2-m^2}\right]} \quad . \tag{4.328}$$

(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 $T = -W/k_{\rm B}\log(2q^{-1}-2)$ and $T < 0$ for all $q \in [0, \frac{2}{3})$. The reason for this strange state of affairs is that the entropy S is bounded, and is not a monotonically increasing function of the energy E (or the dimensionless quantity Q). The entropy is maximized for $N_\uparrow = N_0 = N_\downarrow = \frac{1}{3}N$, which says $m = 0$ and $q = \frac{2}{3}$. Decreasing q below $q = \frac{2}{3}$ at fixed $m = 0$ results in an increased energy $E = -NWq$ and a decreased entropy, thus $(\partial S/\partial E) < 0$ which says $T < 0$. 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.

4.11.4 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 of the interface between the fluids, relative to equilibrium. The potential energy is a sum of gravitational and surface tension terms, with

$$U_{\rm grav} = \int d^2x \int\limits_0^z dz' \, \Delta\rho \, g \, z' \qquad , \qquad U_{\rm surf} = \tfrac{1}{2}\sigma \int d^2x \, (\boldsymbol{\nabla} z)^2 \quad . \tag{4.329}$$

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

$$T = \int d^2x \int d^2x' \, \tfrac{1}{2}\mu(\boldsymbol{x},\boldsymbol{x}') \, \frac{\partial z(\boldsymbol{x},t)}{\partial t} \, \frac{\partial z(\boldsymbol{x}',t)}{\partial t} \quad . \tag{4.330}$$

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

$$z(\boldsymbol{x}) = \big(L_x \, L_y\big)^{-1/2} \sum_{\boldsymbol{k}} z_{\boldsymbol{k}} \, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \quad , \tag{4.331}$$

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

$$\boldsymbol{k} = \frac{2\pi n_x}{L_x}\,\hat{\boldsymbol{x}} + \frac{2\pi n_y}{L_y}\,\hat{\boldsymbol{y}} \quad , \tag{4.332}$$

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

$$L = \frac{1}{2}\sum_{\boldsymbol{k}} \Big[\mu_{\boldsymbol{k}} \, \big|\dot{z}_{\boldsymbol{k}}\big|^2 - \big(g\,\Delta\rho + \sigma\boldsymbol{k}^2\big)\,\big|z_{\boldsymbol{k}}\big|^2 \Big] \quad , \tag{4.333}$$

where

$$\mu_{\boldsymbol{k}} = \int d^2x \, \mu\big(|\boldsymbol{x}|\big) \, e^{-i\boldsymbol{k}\cdot\boldsymbol{x}} \quad . \tag{4.334}$$

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

$$p_{\boldsymbol{k}} = \frac{\partial L}{\partial \dot{z}^*_{\boldsymbol{k}}} = \mu_{\boldsymbol{k}}\,\dot{z}_{\boldsymbol{k}} \qquad , \qquad p^*_{\boldsymbol{k}} = \frac{\partial L}{\partial \dot{z}_{\boldsymbol{k}}} = \mu_{\boldsymbol{k}}\,\dot{z}^*_{\boldsymbol{k}} \tag{4.335}$$

The Hamiltonian is then

$$\begin{aligned} \hat{H} &= {\sum_{\boldsymbol{k}}}' \Big[p_{\boldsymbol{k}}\,\dot{z}^*_{\boldsymbol{k}} + p^*_{\boldsymbol{k}}\,\dot{z}_{\boldsymbol{k}} \Big] - L \\ &= {\sum_{\boldsymbol{k}}}' \left[\frac{|p_{\boldsymbol{k}}|^2}{\mu_{\boldsymbol{k}}} + \big(g\,\Delta\rho + \sigma\boldsymbol{k}^2\big)\,|z_{\boldsymbol{k}}|^2 \right] \quad , \end{aligned} \tag{4.336}$$

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:

$$\begin{aligned} Z &= \prod_{\boldsymbol{k}}{}' \int \frac{d^2p_{\boldsymbol{k}}\, d^2z_{\boldsymbol{k}}}{(2\pi\hbar)^2}\, e^{-|p_{\boldsymbol{k}}|^2/\mu_{\boldsymbol{k}} k_{\rm B}T}\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= \prod_{\boldsymbol{k}}{}' \left(\frac{k_{\rm B}T}{2\hbar}\right)^2 \left(\frac{\mu_{\boldsymbol{k}}}{g\,\Delta\rho+\sigma\boldsymbol{k}^2}\right) \quad . \end{aligned} \tag{4.337}$$

Thus,

$$F = -k_{\rm B}T \sum_{\boldsymbol{k}} \log\left(\frac{k_{\rm B}T}{2\hbar\Omega_{\boldsymbol{k}}}\right) \quad , \tag{4.338}$$

where[15]

$$\Omega_{\boldsymbol{k}} = \left(\frac{g\,\Delta\rho+\sigma\boldsymbol{k}^2}{\mu_{\boldsymbol{k}}}\right)^{1/2} \quad . \tag{4.339}$$

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.

It is now easy to compute the thermal average

$$\begin{aligned} \big\langle\, |z_{\boldsymbol{k}}|^2 \,\big\rangle &= \int d^2z_{\boldsymbol{k}}\, |z_{\boldsymbol{k}}|^2\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \Big/ \int d^2z_{\boldsymbol{k}}\, e^{-(g\,\Delta\rho+\sigma\boldsymbol{k}^2)\,|z_{\boldsymbol{k}}|^2/k_{\rm B}T} \\ &= \frac{k_{\rm B}T}{g\,\Delta\rho+\sigma\boldsymbol{k}^2} \quad . \end{aligned} \tag{4.340}$$

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

$$\begin{aligned} C(\boldsymbol{x}) \equiv \big\langle z(\boldsymbol{x})\, z(0)\big\rangle &= \frac{1}{L_xL_y} \sum_{\boldsymbol{k}} \big\langle\, |z_{\boldsymbol{k}}|^2 \,\big\rangle\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} = \int \frac{d^2k}{(2\pi)^2} \left(\frac{k_{\rm B}T}{g\,\Delta\rho+\sigma\boldsymbol{k}^2}\right) e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \\ &= \frac{k_{\rm B}T}{2\pi\sigma} \int_0^\infty dk\, k\, \frac{J_0\big(k\,|\boldsymbol{x}|\big)}{k^2+\xi^{-2}} = \frac{k_{\rm B}T}{2\pi\sigma}\, K_0\big(|\boldsymbol{x}|/\xi\big) \quad , \end{aligned} \tag{4.341}$$

where $\xi = \sqrt{\sigma/g\,\Delta\rho}$ is the correlation length, and where $K_0(z)$ is the modified Bessel function of the second kind. The asymptotic behavior of $K_0(z)$ for small z is $K_0(z) \sim \log(2/z)$, 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 be cured by imposing a cutoff on the integral, or by insisting that the shortest distance scale is a molecular diameter.

[15]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.

4.11.5 Dissociation of atomic hydrogen

Suppose we fill a volume V with a number density n of hydrogen atoms. Now let the reaction

$$\mathrm{H} \rightleftharpoons \mathrm{p}^+ + \mathrm{e}^- \tag{4.342}$$

run until thermal equilibrium is attained. Under equilibrium conditions, some fraction x of the original nV hydrogen atoms will have dissociated, and we will have

$$n_{\mathrm{H}} = (1-x)\,n \qquad , \qquad n_{\mathrm{p}} = xn \qquad , \qquad n_{\mathrm{e}} = xn \quad . \tag{4.343}$$

What is the equilibrium relationship between the temperature T, the dissociation fraction x, and the original number density n? To find this, we must invoke the machinery of statistical mechanics, invoking the key relation (see eqn. 2.379)

$$\mu_{\mathrm{H}} = \mu_{\mathrm{p}} + \mu_{\mathrm{e}} \quad . \tag{4.344}$$

The single particle partition function for each species is

$$\zeta = \left(\frac{gV}{\lambda_T^3}\right) e^{-\varepsilon_{\mathrm{int}}/k_{\mathrm{B}}T} \quad , \tag{4.345}$$

where g is the degeneracy and $\varepsilon_{\mathrm{int}}$ the internal energy for a given species. We have $\varepsilon_{\mathrm{int}} = 0$ for p and e, and $\varepsilon_{\mathrm{int}} = -\Delta$ for H, where $\Delta = e^2/2a_{\mathrm{B}} = 13.6\,\mathrm{eV}$, the binding energy of hydrogen. Neglecting hyperfine splittings[16], we have $g_{\mathrm{H}} = 4$, while $g_{\mathrm{e}} = g_{\mathrm{p}} = 2$ because each has spin $S = \frac{1}{2}$. Thus, the associated grand potentials are

$$\begin{aligned}
\Omega_{\mathrm{H}}(T,V,\mu_{\mathrm{H}}) &= -g_{\mathrm{H}}\,Vk_{\mathrm{B}}T\,\lambda_{T,\mathrm{H}}^{-3}\,e^{(\mu_{\mathrm{H}}+\Delta)/k_{\mathrm{B}}T} \\
\Omega_{\mathrm{p}}(T,V,\mu_{\mathrm{p}}) &= -g_{\mathrm{p}}\,Vk_{\mathrm{B}}T\,\lambda_{T,\mathrm{p}}^{-3}\,e^{\mu_{\mathrm{p}}/k_{\mathrm{B}}T} \\
\Omega_{\mathrm{e}}(T,V,\mu_{\mathrm{e}}) &= -g_{\mathrm{e}}\,Vk_{\mathrm{B}}T\,\lambda_{T,\mathrm{e}}^{-3}\,e^{\mu_{\mathrm{e}}/k_{\mathrm{B}}T} \quad ,
\end{aligned} \tag{4.346}$$

where

$$\lambda_{T,a} = \sqrt{\frac{2\pi\hbar^2}{m_a k_{\mathrm{B}} T}} \tag{4.347}$$

for species a. The corresponding number densities are

$$n = -\frac{1}{V}\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V} = g\,\lambda_T^{-3}\,e^{(\mu-\varepsilon_{\mathrm{int}})/k_{\mathrm{B}}T} \quad , \tag{4.348}$$

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

$$z = g^{-1} n\,\lambda_T^3\,e^{\varepsilon_{\mathrm{int}}/k_{\mathrm{B}}T} \quad . \tag{4.349}$$

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

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

$$g_{\rm H}^{-1} n_{\rm H}\, \lambda_{T,{\rm H}}^3\, e^{-\Delta/k_{\rm B}T} = \big(g_{\rm p}^{-1} n_{\rm p}\, \lambda_{T,{\rm p}}^3\big)\big(g_{\rm e}^{-1} n_{\rm e}\, \lambda_{T,{\rm e}}^3\big) \quad , \tag{4.350}$$

which yields

$$\left(\frac{x^2}{1-x}\right) n \tilde{\lambda}_T^3 = e^{-\Delta/k_{\rm B}T} \quad , \tag{4.351}$$

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

$$\tilde{\lambda}_T = a_{\rm B} \sqrt{\frac{4\pi m_{\rm H}}{m_{\rm p}}} \sqrt{\frac{\Delta}{k_{\rm B}T}} \quad , \tag{4.352}$$

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

$$\left(\frac{x^2}{1-x}\right) \cdot (4\pi)^{3/2}\, \nu = \left(\frac{T}{T_0}\right)^{3/2} e^{-T_0/T} \quad , \tag{4.353}$$

where $T_0 = \Delta/k_{\rm B} = 1.578 \times 10^5\,{\rm K}$ 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 $\nu = 1.69 \times 10^{-27}$, corresponding to a density of $n = 1.14 \times 10^{-2}\,{\rm cm}^{-3}$. At this temperature, the fraction of hydrogen atoms in their first excited (2s) state is $x' \sim e^{-T_0/2T} = 3.8 \times 10^{-12}$. 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.