4. Canonical ensemble

In this chapter we will formulate statistical physics for subsystems that are held at constant temperature. In the next chapter we will work at constant temperature and constant chemical potential. The calculations are different from those we have done so far, and often far simpler. For systems at fixed temperature there are very useful numerical techniques, Monte Carlo methods, that we will explain.

Ensembles: Gibbs gave odd names to methods involving subsystems in contact with reservoirs.

- In closed system, with fixed energy we estimate time averages by sampling uniformly over the energy shell, as we have done until now. Gibbs called this the *microcanonical* ensemble.

- In a system which can exchange energy at fixed temperature we average using the Boltzmann factor of Eq. (3.49). This is called the *canonical* ensemble.

- In a system which can exchange both energy and number with reservoirs, we have a different method of averaging that we will derive below. This is called the *grand canonical* ensemble.

4.1. Averages and the partition function

In this section we study systems which are subsystems of a larger system: they are held at constant temperature by contact with a (much larger) heat bath. For a system in contact with a heat bath we have already derived the probability of finding a given energy or a given state, Eq. (3.49). To review, we say:

$$\begin{aligned}
\mathsf{W}_{tot} &= e^{S_r(E-E_s)/k_B}\mathsf{W}_s(E_s) \propto e^{-\beta E_s}\,\mathsf{W}_s(E_s), \\
P(E_i) &\propto \exp(-\beta E_i).
\end{aligned} \qquad (4.1)$$

Here E_i labels a particular quantum state, and $\beta = 1/k_B T$. Note that this formula holds in general: classical systems and quantum systems

obey this law[1].

Any probability satisfies $\sum_j P(j) = 1$, where the sum is over all the alternatives. Thus we can write, for the canonical ensemble:

$$P(E_i) = \frac{e^{-\beta E_i}}{Z}; \quad Z = \sum_i e^{-\beta E_i}. \tag{4.2}$$

The normalization, Z, is called the partition function. The notation, Z, comes from the German *Zustandsumme* = sum over states. In the chemical literature Z is called Q. For a classical system we replace the sum over states by an integral in the same way as for the microcanonical ensemble:

$$P(E) = \frac{1}{N!h^{3N}} \frac{e^{-\beta \mathcal{H}(\mathbf{r},\mathbf{p})}}{Z}, \quad Z = \frac{1}{N!h^{3N}} \int d^N\mathbf{r} \ d^N\mathbf{p} \ e^{-\beta \mathcal{H}(\mathbf{r},\mathbf{p})}. \tag{4.3}$$

4.2. Canonical averages and equipartition

There is a very useful special case of the formalism above.

4.2.1. Distributions of additive degrees of freedom

Let us look at the canonical probability more closely. Suppose that we have a classical Hamiltonian in which certain degrees of freedom (coordinates or momenta) appear additively. For example, for an interacting fluid,

$$\mathcal{H} = p_1^2/2m + \sum_2^N p_i^2/2m + \Phi(\{\mathbf{r}\}).$$

We can find the probability distribution for that variable (the distribution for $\mathbf{p}_1$ in this case) by considering it as a *marginal* distribution. We will use the relation from probability theory that if $P(AB)$ is the probability of A and B, then:

$$P(A) = \sum_B P(AB).$$

The distribution, $P(A)$, is the marginal distribution, the distribution of A ignoring B. We can find $P(A)$ by summing over all the possibilities of B.

---

[1]This may seem to contradict the form of the Fermi and Bose distributions of Eq. (2.41), Eq. (2.43). It does not, as we will see below.

For our case, we integrate the normalized probability over $\mathbf{p}_j, j = 2, 3, \ldots, N$ and $\mathbf{r}_j, j = 1, 2, \ldots, N$.

$$
\begin{aligned}
P(\mathbf{p}_1) &= \frac{e^{-\beta p_1^2/2m} \int d^{N-1}\mathbf{p} \; e^{-\beta \sum_2^N p_i^2/2m} \int d^N\mathbf{r} \; e^{-\beta\Phi}}{\int d^N\mathbf{p} \; e^{-\beta \sum_1^N p_i^2/2m} \int d^N\mathbf{r} e^{-\beta\Phi}} \\
&= \frac{e^{-\beta p_1^2/2m}}{\int d\mathbf{p}_1 \; e^{-\beta p_1^2/2m}} \quad\quad\quad\quad\quad\quad (4.4)
\end{aligned}
$$

This is the Maxwell-Boltzmann distribution which we have seen above for the ideal gas. Note that the interactions cancel out in numerator and demominator. Any classical system, gas, liquid, or solid (or polymer, glass, etc.) has this distribution for the momentum of any particle.

4.2.2. Boltzmann equipartition theorem

Continuing in the same vein, consider the average of any quantity that appears in the Hamiltonian as an additive term which is quadratic in a degree of freedom:

$$
\mathcal{H} = \mathcal{H}_\circ(\{q,p\}) + CQ^2.
$$

Here, $\{q,p\}$ denotes all the other variables. Then, in the same way as in Eq. (4.4) we have:

$$
\begin{aligned}
\langle CQ^2 \rangle &= \frac{\int dQ \; CQ^2 \; e^{-\beta CQ^2} \int dq \; dp \; e^{-\beta\mathcal{H}_\circ}}{\int dQ e^{-\beta CQ^2} \int dq \; dp \; e^{-\beta\mathcal{H}_\circ}} \\
&= \frac{\int dQ \; CQ^2 \; e^{-\beta CQ^2}}{\int dQ e^{-\beta CQ^2}} \\
&= -\frac{\partial}{\partial\beta} \ln[\int dQ e^{-\beta CQ^2}] \\
&= \frac{k_B T}{2}. \quad\quad\quad\quad\quad\quad\quad\quad\quad\quad (4.5)
\end{aligned}
$$

This theorem quickly gives many results. For example, the average internal energy of the ideal gas is:

$$
E = \left\langle \sum_1^N p_i^2/2m \right\rangle = 3Nk_B T/2.
$$

The heat capacity follows: $C_V = \partial E/\partial T_V = 3Nk_B/2$.

As another example, air is composed, almost entirely, of the diatomic molecules O_2 and N_2. We can model them as rigid rotors with moment

of intertia I. Then the kinetic energy has extra degrees of freedom. For each type of molecule we can write:

$$E = \sum_j [p_j^2/2m + \mathbf{L}_j^2/2I]. \qquad (4.6)$$

The angular momentum vector $\mathbf{L}$ should be regarded as having two components – in our classical picture we neglect the moment of inertia around the molecular axis. Thus we have two more degrees of freedom, and

$$\langle E \rangle = 3N(k_B T/2) + 2N(k_B T/2) = (5/2)Nk_B T.$$

The first term is the translational part and the second rotational. Thus $C_V = (5/2)Nk_B$, $C_p = (7/2)Nk_B$. (The last identity comes from the thermodynamic relation for ideal gases, $C_p = C_V + Nk_B$. The proof is left as an exercise.)

This is simple, and gives answers in agreement with experiment. However, it is quite puzzling in one respect: if we think of diatomic molecules classically, we should consider that they can vibrate. The simplest model for a diatomic molecule with a chemical bond would be two masses connected by a spring. The spring should give another quadratic degree of freedom, and another $k_B T/2$ per molecule in the energy. This is not observed. This fact was extremely troubling to the founders of statistical thermodynamics; see Gibbs (1902). The explanation is that the vibrations must be treated in quantum mechanics, and equipartition does not hold in quantum theory.

We can apply these ideas to a solid as well. The simplest model for the vibrations of a monatomic solid is due to A. Einstein. Einstein noted that in a solid in equilibrium each atom sits at the minimum of the potential well due to all of the other atoms. Expanding about the minimum means that each atom is in a quadratic potential, and if an atom moves away from equilibrium, the force is of the form $-K\mathbf{r}$ (assuming isotropy). Then the energy is $Kr^2/2m$, where K is a spring constant and m is the atomic mass, so that the frequency of the vibration is $\omega_E = \sqrt{K/m}$. There are three more quadratic degrees of freedom. We find, as before:

$$C_V = \frac{3Nk_B}{2}(\text{kinetic}) + \frac{3Nk_B}{2}(\text{vibrational}) = 3Nk_B. \qquad (4.7)$$

This is called the law of Dulong and Petit; it was discovered empirically by P. L. Dulong and A. T. Petit in 1819.

The law is quite a lot more general than this derivation seems to imply. In fact, if we let N denote all the atoms in a molecular solid, it
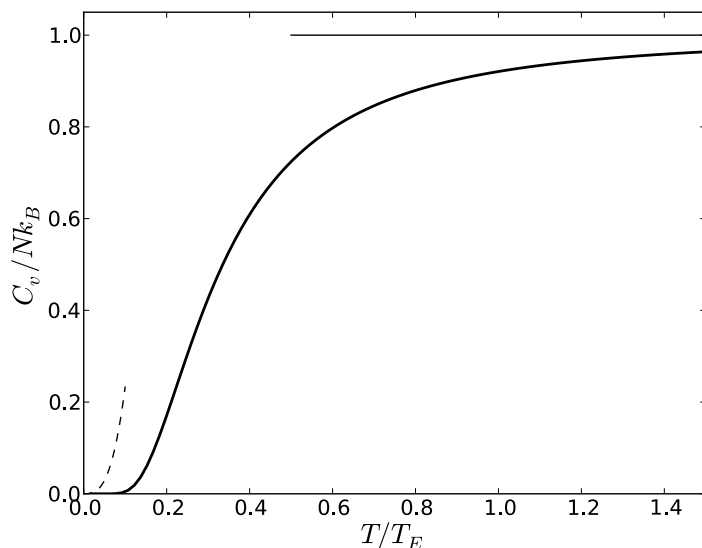
Figure 4.1.:   The Einstein model for C_V as a function of T/T_E where $k_B T_E = \hbar\omega_E$. The upper line is the Dulong-Petit law, and the lower dotted line schematically depicts the deviation of the theory from experiment at very low T.

applies as well. The reason is this: in the Einstein model each atom is thought to vibrate independently. In fact, all of them are coupled so that the real excitations are waves of vibration, e.g. sound waves. The Hamiltonian is quadratic in the amplitude and momentum of the waves, and we have as many different waves as we have degrees of freedom. How this works out – the normal mode transformation – will be treated in Chapter 6.

However, well below room temperature, the Dulong-Petit law fails, and

$$\lim_{T \to 0} C_V = 0,$$

as required by the third law. A quantum mechanical treatment of the Einstein model gives this result; see Figure 4.1. We will do this derivation below. The real situation in solids is qualitatively like the Einstein model with a deviation at very low T.

4.3. The partition function and the free energy

The partition function is of physical significance itself, as shown by the following theorem.

$$Z = e^{-\beta F}; \quad F = -k_B T \ln Z. \tag{4.8}$$

4.3.1. First proof: average of F

We will show this in two ways. First we need to interpret F in the equation. Remember $F = E - TS(E)$, and E varies as heat flows in and out of the system from the heat bath. In fact, the quantity of interest is $\langle F \rangle = \langle E \rangle - T_r \langle S \rangle$. In order to emphasize that the temperature is fixed by the reservoir, we write T_r. Now from Eq. (4.2) and Eq. (3.5) we have:

$$
\begin{aligned}
\langle F \rangle = \langle E \rangle - T_r \langle S \rangle \;\; &= \;\; \frac{1}{Z} \sum_i (E_i + k_B T_r \ln P(E_i)) e^{-\beta E_i} \\
&= \;\; \frac{1}{Z} \sum_i (E_i - k_B T_r [\beta E_i + \ln Z]) e^{-\beta E_i} \\
&= \;\; -k_B T_r \ln Z. \tag{4.9}
\end{aligned}
$$

4.3.2. Second proof: energy fluctuations

It is interesting to derive Eq. (4.8) another way. Note that for a large system we can expect that the energy levels will be close together (or, in a classical system, the energy is continuous). We can convert the sum in Z into an integral by introducing the many-particle density of states, i.e. the number of states per energy interval:

$$\omega(E) = \sum_i \delta(E - E_i). \tag{4.10}$$

Clearly:

$$Z = \sum_i e^{-\beta \epsilon_i} = \int dE \, \omega(E) e^{-\beta E}. \tag{4.11}$$

However, W is the number of states for which $E < \epsilon_i < E + \Delta$, i.e.

$$\mathsf{W} = \sum_i [\theta(E + \Delta - \epsilon_i) - \theta(E - \epsilon_i)],$$

where θ is the step function: $\theta(x) = 1, x \geq 0, \theta(x) = 0, x < 0; d\theta/dx = \delta(x)$. For $\Delta \ll E$ we can expand:

$$\mathsf{W} \approx \Delta \sum_i \delta(E - \epsilon_i) = \Delta\omega. \tag{4.12}$$

Thus:

$$S = k_B \ln(\Delta\omega) \tag{4.13}$$

Now we use this result to deal with the integral in Eq. (4.11).

$$Z = \int dE e^{-\beta(E - T_r S(E))}/\Delta = \int dE e^{-\beta F(E)}/\Delta. \tag{4.14}$$

Note that F is extensive, so the argument of the exponential is very large. We can use the Laplace method; see Appendix A. We need to find the energy at which $F(E)$ is a minimum, and the second derivative at that point.

$$
\begin{aligned}
\frac{dF}{dE} &= 0 \quad = 1 - T_r \frac{dS}{dE}; \quad \frac{dS}{dE} = \frac{1}{T_r} \\
\frac{d^2 F}{dE^2} &= -T_r \frac{d^2 S}{dE^2}\Big|_{T=T_r} = -T_r \frac{d(1/T)}{dE}\Big|_{T=T_r} = \frac{1}{C_V T_r} \\
F(E) &= F(\overline{E}) + \frac{1}{2C_V T_r}(E - \overline{E})^2 \ldots
\end{aligned}
\tag{4.15}
$$

The first equation determines an energy, $\bar{E}$, at which F is a minimum. At that point the system temperature is equal to the reservoir temperature, and we can write $F(\overline{E}) = F(T_r)$. The second derivative must be positive in order that F be a minimum. This gives a *stability condition*: $C_V > 0$. If this does not hold, we are dealing with a system that cannot come to equilibrium.

Using Eq. (A.9), we have:

$$
\begin{aligned}
Z &\rightarrow e^{-\beta F(T_r)} \int e^{-(E-\bar{E})^2/2k_B C_V T_r^2} dE/\Delta \\
&= e^{-\beta F(T_r)} \sqrt{2\pi k_B T^2 C_V}/\Delta.
\end{aligned}
\tag{4.16}
$$

The correction to Eq. (4.8) is now explicit

$$F = -k_B T \ln Z - k_B T \ln \sqrt{2\pi k_B T^2 C_V} + \mathcal{O}(1).$$

The leading correction is of order $\ln(N)$.

We have actually proved more: the integrand in Eq. (4.16) is the probability distribution for energy fluctuations. It a gaussian centered

on $\bar{E}$; the width is $\delta E = \sqrt{2k_B T^2 C_V}$. The gaussian is very narrow: $\delta E / \bar{E} = \mathcal{O}(N^{-1/2}) \approx 10^{-10}$ for a macroscopic system. This shows that the canonical ensemble will give the same results as the microcanonical: the energy fluctuations are very small.

4.4. The canonical recipe

With these results in hand, we can write down a procedure for deriving a fundamental relation from statistical physics. This is the most popular method in statistical thermodynamics.

4.4.1. Recipe

Here are the steps to do a calculation:

1. For a quantum system calculate:

$$Z(T, \mathbf{X}, N) = \sum_i e^{-\beta E_i}.$$

    For a classical system calculate

$$Z(T, \mathbf{X}, N) = \frac{1}{N! h^{3N}} \int d^{3N}q \ d^{3N}p \ e^{-\beta \mathcal{H}(q,p)}.$$

2. Calculate $F(T, \mathbf{X}, N) = -k_B T \ln Z$.

3. Calculate $S = -(\partial F/\partial T)_{\mathbf{X},N}$, $\mathbf{f} = (\partial F/\partial \mathbf{X})_{T,N}$. For example, for a fluid:

$$S = -(\partial F/\partial T)_{V,N}, \quad p = -(\partial F/\partial V)_{T,N}.$$

    These are the equations of state.

4. To average any quantity of interest, $R(q, p)$ in a classical system compute:

$$\langle R \rangle = \frac{\int d^{3N}q \ d^{3N}p \ e^{-\beta \mathcal{H}(q,p)} R(q,p)}{\int d^{3N}q \ d^{3N}p \ e^{-\beta \mathcal{H}(q,p)}}.$$

    In a quantum system, for any operator, $\hat{R}$ we need its diagonal matrix elements, $\langle i|\hat{R}|i \rangle$. Then the average is:

$$\langle \hat{R} \rangle = \frac{\sum_i \langle i|\hat{R}|i \rangle e^{-\beta E_i}}{\sum_i e^{-\beta E_i}}.$$

4.4.2. Phase space density and density matrix

We can rewrite our results in terms of the phase space density. For the canonical ensemble, our results amount to saying:

$$\rho_c(q,p) = \frac{1}{N!h^{3N}} \frac{e^{-\beta\mathcal{H}(\mathbf{r},\mathbf{p})}}{Z}. \tag{4.17}$$

The density matrix is:

$$\rho_c^{j,k} = \delta_{j,k} \frac{e^{-\beta E_j}}{Z}. \tag{4.18}$$

The density operator is:

$$\hat{\rho} = \frac{e^{-\beta\hat{\mathcal{H}}}}{\mathrm{Tr}(e^{-\beta\hat{\mathcal{H}}})} \tag{4.19}$$

Averages are, as usual:

$$\left\langle \hat{R} \right\rangle = \mathrm{Tr}(\hat{\rho}R). \tag{4.20}$$

4.5. Examples

With the recipe we can do a large number of very interesting computations. Here are a few examples.

4.5.1. Ideal paramagnet

For the paramagnet $E = -\sum_j \sigma_j h$. The first step of the recipe is:

$$
\begin{aligned}
Z &= \sum_{\{\sigma_i = \pm 1\}} e^{\beta h \sum_j \sigma_j} \\
&= \sum_{\sigma_1 = \pm 1} e^{\beta h \sigma_1} \sum_{\sigma_2 = \pm 1} e^{\beta h \sigma_2} \ldots \\
&= \left(e^{\beta h} + e^{-\beta h} \right)^N. \tag{4.21}
\end{aligned}
$$

The last equation is important: the partition function is the Nth power of a partition function for each spin: $Z = (Z_1)^N$. This always happens for non-interacting systems.

Now, the second step:

$$F = -Nk_B T \ln(e^{\beta h} + e^{-\beta h}). \tag{4.22}$$

The equation of state is:

$$M = - \left(\frac{\partial F}{\partial h} \right)_T = N \tanh \beta h, \tag{4.23}$$

as in Eq. (3.7).

For small field we have:

$$M \to N\beta h; \quad N\chi = \frac{M}{h} = \frac{N}{k_B T}. \tag{4.24}$$

This is the Curie law. In more conventional units the Curie law reads (for spin-1/2 particles):

$$\bar{M} \to \frac{N\mu^2}{k_B T} H. \tag{4.25}$$

Here $\bar{M}$ is the magnetic moment, and μ the magnetic moment per particle.

4.5.2. Ideal gas

For the ideal gas $\mathcal{H} = \sum p_j^2/2m$. The first step is:

$$
\begin{aligned}
Z &= \frac{V^N}{N! h^{3N}} \int d\mathbf{p}_1 d\mathbf{p}_2 \ldots e^{-\beta \sum p_j^2/2m} \\
&= \frac{V^N}{N! h^{3N}} \int d\mathbf{p}_1 \; e^{-\beta p_1^2/2m} \int d\mathbf{p}_2 \; e^{-\beta p_2^2/2m} \int \ldots \\
&= \frac{1}{N!} \left[\frac{V}{h^3} \int d\mathbf{p} \; e^{-\beta p^2/2m} \right]^N \\
&= \left[\frac{Ve}{N\lambda^3} \right]^N, \quad \lambda = (h^2/2\pi m k_B T)^{1/2}. \tag{4.26}
\end{aligned}
$$

In the last step Stirling's approximation, $N! \approx N^N e^{-N}$ was used. Again, note the factorization of the partition function into the Nth power of a one particle quantities.

The second step is to find the free energy:

$$F = -k_B T \ln Z = N k_B T(\ln(n\lambda^3) - 1), \quad n = N/V. \tag{4.27}$$

The third step:

$$
\begin{aligned}
S &= -\left(\frac{\partial F}{\partial T} \right)_{V,N} = N k_B \left(\frac{5}{2} + \ln(V/N) - \ln(\lambda^3) \right) \\
p &= -\left(\frac{\partial F}{\partial V} \right)_{T,N} = N k_B T/V. \tag{4.28}
\end{aligned}
$$

The first line is the Sackur-Tetrode equation, Eq. (3.12) expressed as a function of T.

4.5.3. Ideal gas with internal degrees of freedom: diatomic molecules

We now return to the gas of molecules with internal degrees of freedom (such as rotations) which was discussed in 4.2. Consider, again, diatomic molecules. Now the Hamiltonian has several parts:

$$\mathcal{H} = \mathcal{H}_{trans} + \mathcal{H}_{rot} + \mathcal{H}_{vib},$$

corresponding to translation of the molecule as a whole, rotation about two axes with moment of inertia I, and vibrations, i.e. stretching of the chemical bond between the molecules. As usual when we have sums of independent parts of the Hamiltonian, the partition function factors:

$$Z = Z_1^N; \quad Z_1 = Z_{trans}Z_{rot}Z_{vib}.$$

The first factor is the familiar integral arising from the momentum terms: $Z_{trans} = Ve/N\lambda^3$. Correspondingly, the free energy is a sum:

$$
\begin{aligned}
F &= F_{trans} + F_{rot} + F_{vib} \\
&= -Nk_BT(\ln Z_{trans} + \ln Z_{rot} + \ln Z_{vib}).
\end{aligned}
\tag{4.29}
$$

And the specific heat, $-T(\partial^2 F/\partial T^2)$ has three parts as well.

We now discuss the last two factors.

Rotations

The rotations arise from the rotational kinetic energy $\mathbf{L}^2/2I$. It is an elementary result of quantum theory (Messiah (1968), Sakurai & Napolitano (2010)) that for such a rotation the energy levels and degeneracies are:

$$\epsilon_J = \hbar^2 J(J+1)/2I, \quad g(J) = 2J+1, \quad J = 0, 1, \ldots \tag{4.30}$$

To give a scale for the energies here we define the *rotational temperature*, T_R, by $k_BT_R = \hbar^2/2I$. For O_2 $T_R = 2.1$K, N_2, 2.9K, well below the boiling point of these gases and very far below room temperature. The largest T_R's are for H_2, 85.4K, and HD (D stands for Deuterium), 64K.

For a diatomic molecule of different species such as HCl we can write the rotational partition function at once:

$$Z_{rot} = \sum_{0}^{\infty}(2J+1)e^{-\beta\hbar^2 J(J+1)/2I} = \sum(2J+1)e^{-J(J+1)T_R/T}. \quad (4.31)$$

At room temperature, we note that $T_R/T << 1$, and the terms in the sum differ very little on increasing J by unity. It is a very good approximation to replace the sum by an integral:

$$Z_{rot} \approx \int_{0}^{\infty} 2Je^{-T_R J^2/T} dJ = 2k_B TI/\hbar^2.$$

Now the free energy is $F_{rot} = -Nk_B T \ln(2k_B TI/\hbar^2)$, and the heat capacity is:

$$C_v^{rot} = -T(\partial^2 F_{rot}/\partial T^2) = Nk_B,$$

as above.

For $T < T_R$ we have a different situation. As a practical matter this only applies to HD. Now the sum in Eq. (4.31) is dominated by the first few terms, $Z \approx 1 + 3e^{-2T_R/T}$, and the heat capacity vanishes exponentially.

For a homo-nuclear molecule all this isn't quite right because of symmetry requirements. There results an interesting behavior for the heat capacity of H_2 for low temperatures, an effect first explained by D. Dennison; see Problem 7.

Vibrations:

Diatomic molecules are attached by a chemical bond whose length we call r. The equilibrium value, $r = L$, is set by the atoms sitting at the minimum of the mutual potential energy well; compare Figure 1.1. Near the bottom of the well, we can expand, and replace the well with a quadratic approximation: $V(r) \approx V(L) + K(r-L)^2/2 \ldots$. If we set $u = r - L$ we have a quadratic degree of freedom, $\mathcal{H} = \mathcal{H}_\circ + Ku^2/2$ which is the Hamiltonian for a simple harmonic oscillator.

We know how quantize a simple harmonic oscillator, (Messiah (1968), Sakurai & Napolitano (2010)). The energy levels are:

$$\epsilon_n = \hbar\omega(n+1/2), \quad \omega = \sqrt{K/m}, \quad n = 0, 1, \ldots \quad (4.32)$$

To give an idea of the numbers here, we define a *vibrational temperature*, $T_v = \hbar\omega/k_B$. For O_2, $T_v = 2230$K. Since T_v is far above

room temperature, we will always considering small n for which the quadratic approximation of the potential is adequate.

Now we can write:

$$Z_{vib} = e^{-\beta\hbar\omega/2} \sum_{0}^{\infty} e^{-\beta n\hbar\omega} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}}. \tag{4.33}$$

Differentiating, we find the heat capacity:

$$C_{vib} = Nk_B \left(\frac{T_v}{T}\right)^2 \frac{e^{T_v/T}}{(e^{T_v/T} - 1)^2}. \tag{4.34}$$

This function is very small at room temperature since it is proportional to $e^{-T_v/T}$ for $T_v/T >> 1$. The missing degree of freedom which troubled Gibbs (see section 4.2) is frozen out by the large spacing of the vibrational energy levels.

4.5.4. Low temperature solids in the Einstein model

We now have the analysis we need to see how the Einstein model works at low T. We take over the result of Eq. (4.34) for N oscillators of frequency $\omega_E = k_B T_E/\hbar$. This is the result plotted in Figure 4.1.

The deviations from the Einstein model at low temperatures are due to the fact that in a real solid the frequencies are not all the same. This theory (the Debye theory) will be given in Chapter 6.

4.5.5. Interacting fluid

For the interacting fluid we have $\mathcal{H} = \sum_j p_j^2/2m + \Phi(\{\mathbf{r}\})$. The partition function divides into two factors:

$$
\begin{aligned}
Z &= \frac{1}{N!h^{3N}} \int d^{3N}p \; e^{-\beta\sum p_j^2/2m} \int d^{3N}r \; e^{-\beta\Phi} \\
&= \left[\frac{V^N}{N!\lambda^{3N}}\right] \left[\frac{1}{V^N} \int d^{3N}r e^{-\beta\Phi}\right] \\
&= \frac{V^N}{N!\lambda^{3N}} Q_N \approx \left[\frac{Ve}{N\lambda^3}\right]^N Q_N
\end{aligned}
\tag{4.35}
$$

The first factor is the ideal gas result, and the second is called[2] the configuration integral,

$$Q_N = \left[\frac{1}{V^N} \int d^{3N}r e^{-\beta\Phi}\right]. \tag{4.36}$$

---

[2] Our definition differs from the more usual one by the factor V^{-N}.

The configuration integral cannot be evaluated exactly. There are a number of ways to approximate it analytically. For numerical methods, we see that it would be useful to evaluate it alone since the translational degrees of freedom are treated exactly. One goal in evaluating Q_N would be to get insight into the non-trivial parts of the phase diagram such as the critical point.

Virial expansion

In the 1930's considerable effort was expanded in finding an expansion of Q_N around the ideal gas result. When converted to an equation of state this takes the form of an expansion in density:

$$p/k_B T = n(1 + B_2(T)n + B_3(T)n^2 \dots). \tag{4.37}$$

The B_i are called virial coefficients.

A clever method of deriving entire series of virial coefficients was devised by H. Ursell and J. E. Mayer. It is summarized in most modern books on statistical physics, and in Mayer's own book, Mayer & Mayer (1940). Unfortunately, the series is quite useless since its sum fails to get the critical point correctly; see Mayer & Mayer (1940). We will limit ourselves here to show how to get the first correction, B_2 which is useful at low density.

We follow the method of Landau & Lifshitz (1980). First we recall the form of Φ:

$$\Phi = \phi(r_{1,2}) + \phi(r_{1,3}) + \phi(r_{2,3}) \dots \tag{4.38}$$

The free energy can be written:

$$
\begin{aligned}
F &= F_{trans} - k_B T \ln \frac{1}{V^N} \int d^{3N}r \, e^{-\beta \Phi} \\
&= F_{trans} - k_B T \ln \frac{1}{V^N} \int d^{3N}r \left[(e^{-\beta \Phi} - 1) + 1 \right] \\
&= F_{trans} - k_B T \ln \left[1 + \frac{1}{V^N} \int d^{3N}r (e^{-\beta \Phi} - 1) \right]. \tag{4.39}
\end{aligned}
$$

Now we use the fact that we are considering a density expansion. We suppose that at a given time in the *whole* sample there is only one pair that is close together, and that there are no higher-order collisions at all. This means that we are looking at a relatively small system. In this case only one of the terms in Eq. (4.38) is large enough to make the integrand much different from zero. However, we must sum over

all the possible pairs that could be close, and take $e^{-\beta\phi(r_{j,k})} = 1$ for all the rest. Thus:

$$
\frac{1}{V^N} \int d^{3N}r(e^{-\beta\Phi} - 1) \approx \frac{N(N-1)}{2V^2} \int d\mathbf{r_1}d\mathbf{r_2}(e^{-\beta\phi(r_{1,2})} - 1)
$$

$$
\approx \frac{N^2}{2V} \int d\mathbf{s}(e^{-\beta\phi(s)} - 1). \tag{4.40}
$$

In the last line we have gone to the center of mass frame of the two atoms.

Now we can expand the logarithm, assuming the correction is small:

$$
F = F_{trans} - k_B T \frac{N^2}{2V} \int d\mathbf{s}(e^{-\beta\phi(s)} - 1) \tag{4.41}
$$

To get the pressure we differentiate:

$$
p = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{Nk_BT}{V} + \frac{N^2 k_B T}{2V^2} \int d\mathbf{s}(1 - e^{-\beta\phi(s)}) \tag{4.42}
$$

Comparing to Eq. (4.37) gives:

$$
B_2 = \frac{1}{2} \int d\mathbf{s}(1 - e^{-\beta\phi(s)}). \tag{4.43}
$$

Now we can try to approximate the integral. We can write:

$$
B_2 = 2\pi \int_0^\infty r^2 dr f(r); \quad f = 1 - e^{-\beta\phi}.
$$

Recall that $\phi = 4\epsilon[(\sigma/r)^{12} - (\sigma/r)^6]$ for the Lennard-Jones case. Note that the function f is approximately unity for $r < \sigma$ because of the repulsive part of the potential. For $r > \sigma$ we have only the small attractive part. We assume that $\phi_{att}/k_BT < 1$. So, crudely, we put:

$$
B_2 \approx 2\pi \left[\int_0^\sigma r^2 dr - 4\epsilon\beta\sigma^6 \int_\sigma^\infty r^2 dr/r^6\right] = \frac{2\pi\sigma^3}{3}\left[1 - \frac{4\epsilon}{k_B T}\right] \tag{4.44}
$$

Now B_2 is of the form $b - a/k_BT$. Thus, from Eq. (4.37):

$$
\begin{aligned}
p &= nk_BT(1 + [b - a/k_B/T]n) \\
&= nk_BT(1 + bn) - an^2 \approx \frac{nk_BT}{1 - nb} - an^2.
\end{aligned} \tag{4.45}
$$

This is the van der Waals' equation, Eq. (2.18) with an approximate microscopic identification of a, b:

$$
a = 8\pi\epsilon\sigma^3/3; \quad b = 2\pi\sigma^3/3.
$$

The temperature dependence of B_2 is qualitatively correct. The temperature at which $B_2 = 0$ is called the Boyle temperature.

4.6. Simulations in the canonical ensemble

Numerical solutions for systems at fixed temperature are widely used. In fact, the most common method in numerical computation for statistical systems is the one that we will describe here, the Metropolis Monte Carlo technique of Metropolis, Rosenbluth, Rosenbluth, Teller & Teller (1953). It was developed at Los Alamos in the late 1940's and early 50's, and arose out of the work of Stanaslaw Ulam and John von Neumann. The term "Monte Carlo" was coined by Ulam: it refers to the famous gambling casino in Monaco: the method makes copious use of random sampling and random numbers.

There are numerous references on this subject, e.g. Newman & Barkema (1999), Gould et al. (2006) and chapters in Chandler (1987), Gould & Tobochnik (2010), and Pathria & Beale (2011), Peliti (2011). For some of the mathematics of Markov chains see Sethna (2006) and Van Kampen (2007).

4.6.1. Statistical sampling at fixed T

Suppose we want to compute the canonical average of a quantity, R, for example, the magnetic moment for an Ising system. According to Eq. (4.2) we need to sum over the states of the system which we label by j:

$$\langle R \rangle = \frac{\sum_j R(j) e^{-\beta E_j}}{\sum_j e^{-\beta E_j}}. \tag{4.46}$$

In this expression, $R(j)$ is the value of R in state j. For a small system we could simply sum over all the states. For 9 Ising spins we have to sum over $2^9 = 512$ states which is perfectly doable. However, for 256 spins, there are 1.2×10^{77} states – far beyond the capacity of any computer. We must sample, i.e., pick a very much smaller subset of states, $\{J\}$ and estimate the result. If we pick the J at random we estimate:

$$\langle R \rangle = \frac{\sum_J R(J) e^{-\beta E_J}}{\sum_j e^{-\beta E_J}}.$$

However, this is likely to give very bad estimates. For example, at low T most of the states we pick will have very small Boltzmann factors and will be essentially discarded from the sum.

The trick which Metropolis et al. devised is to use a set $\{\mu\}$ which prefers more important states. As a result, the method is called importance sampling. We will see, in a moment, how to generate the μ's.

First we need to see how to use them to estimate averages. Suppose the μ's generate a biased sample so that a state μ has a probability P_μ. Then more probable states will occur more often than they should in Eq. (4.46), and we need to correct for this: if in a sample $R(\mu)$ occurs twice as often as it should, we need to divide its contribution by 2. In general, to get a good estimate, we should put:

$$\langle R \rangle = \frac{\sum_\mu R(\mu) P^{-1}(\mu) e^{-\beta E_\mu}}{\sum_\mu P^{-1}(\mu) e^{-\beta E_\mu}}. \tag{4.47}$$

Now the $P(\mu)$'s are at our disposal. The best choice is to make every term in the sum equally important: presumably it will take computer resources to generate a term in the sum, so we want to use all of them. A nice choice will be: $P(\mu) \propto e^{-\beta E_\mu}$. In this case we have, for a sample of size $\mathcal{N}$:

$$\langle R \rangle = \frac{\sum_\mu R(\mu)}{\mathcal{N}}. \tag{4.48}$$

Our sample is the simple numerical average over the sequence of μ's.

This is not the only choice for P which has proven useful. We may want to emphasize regions of the state space where R has values we are interested in. Methods like this are called umbrella sampling: see Chandler (1987). We will only discuss the simplest method.

4.6.2. Markov processes

We need to pick the set $\{\mu\}$. This sounds hard since the important states might be a very small fraction of all the possibilities. However, we know that the problem of finding the high probability regions of the state space cannot be too hard, since real dynamics does go to equilibrium in the real world. The strategy is to impose a simplified dynamics on the system. That is, $\{\mu\}$ will be generated as a sequence of states arising from a fictitious dynamics. The equilibrium attained will be real; the way we get there need not be.

We will choose the μ as states of a *Markov chain*. A Markov chain is a sequence that is generated by picking states one after another at discrete time steps according to prescribed time-independent transition probabilities, $\mathbb{W}$. If we are in μ at step t the probability of being in state ν at $t + 1$ is $\mathbb{W}(\nu \leftarrow \mu)$. (We will usually abbreviate this to $\mathbb{W}(\nu, \mu)$.) Since we pick new states according to a probability, if we rerun the process we will, in general, get a different sequence. These different sequences are called sample paths.

We need to require that the $\mathbb{W}$ be interpreted as a probability, and that something happens at each step. This means that we require:

$$\mathbb{W}(\nu, \mu) \geq 0; \quad \sum_{\nu} \mathbb{W}(\nu, \mu) = 1. \tag{4.49}$$

It is quite possible for the state to stay the same at a step: that is $\mathbb{W}(\mu, \mu) \neq 0$. We do not require that the matrix of the $\mathbb{W}$ be symmetric. Note that the transition probabilities at a given step do not depend on how the path got to state μ. Markov processes have no memory.

Master equation

Since the dynamics we are representing is ergodic, we require our process to be ergodic: that is, we must choose the $\mathbb{W}$ so that every state in our finite set of states can be reached by a finite number of steps from any other. If we start in some state, at time t later there will be a a probability, $P(\mu, t)$ for being in state μ at time t. At the next time step:

$$P(\mu, t+1) = \sum_{\nu} \mathbb{W}(\mu, \nu) P(\nu, t). \tag{4.50}$$

This is called the *master equation*. Subtracting $P(\mu, t)$ from both sides, and using Eq. (4.49) gives:

$$P(\mu, t+1) - P(\mu, t) = \sum_{\nu} [\mathbb{W}(\mu, \nu) P(\nu, t) - \mathbb{W}(\nu, \mu) P(\mu, t)]. \tag{4.51}$$

Sometimes it is useful to consider a continuous time process. In this case we replace the unit time step in Eq. (4.51) by a time step τ, divide both sides of of the equation by τ, and replace the $\mathbb{W}$'s by transition rates $\overline{\mathbb{W}}$. This gives:

$$dP(\mu)/dt = \sum_{\nu} [\overline{\mathbb{W}}(\mu, \nu) P(\nu, t) - \overline{\mathbb{W}}(\nu, \mu) P(\mu, t)].$$

We will continue here with the discrete time version.

Approach to equilibrium

Our first goal must be to find out under what conditions the process we invent actually does go to equilibrium. This discussion is a bit technical.

If we write the P's as a column vector, $\mathbb{P}$, whose components are $P(\mu)$, then we can read Eq. (4.50) as a matrix equation:

$$\mathbb{P}(t+1) = \mathbb{W}\mathbb{P}(t), \mathbb{P}(t+2) = \mathbb{W}^2\mathbb{P}(t), \ldots, \mathbb{P}(t+n) = \mathbb{W}^n\mathbb{P}(t). \quad (4.52)$$

We are looking for equilibrium:

$$\mathbb{P}_e(t+1) = \mathbb{P}_e(t) = \mathbb{W}\mathbb{P}_e(t).$$

The equilibrium vector, if it exists, is an eigenvector of $\mathbb{W}$ with eigenvalue 1:

$$[\mathbb{W} - \mathbb{I}]\mathbb{P}_e = 0; \quad \det[\mathbb{W} - \mathbb{I}] = 0, \quad\quad\quad (4.53)$$

Where $\mathbb{I}$ is the unit matrix.

The existence of $\mathbb{P}_e$ is easy to prove. The transpose of $\mathbb{W}$ has an eigenvector with unit eigenvalue, namely the vector with all ones as components: this is just Eq. (4.49). The eigenvector of the transpose matrix is called a left eigenvector. Left eigenvectors and right eigenvectors do not need to be equal. But, since the determinant of $\mathbb{W}^T$ same as the determinant of $\mathbb{W}$, a right eigenvector with the same eigenvalue must exist.

Further, no other eigenvalue is larger than 1 by the triangle inequality:

$$
\begin{aligned}
|\lambda P(\mu)| &= |\lambda||P(\mu)| \le \sum_{\nu} W(\nu,\mu)|P(\nu)| \\
|\lambda|\sum_{\mu}|P(\mu)| &\le \sum_{\nu,\mu} W(\nu,\mu)|P(\nu)| = \sum_{\nu}|P(\nu)|.
\end{aligned}
$$

In the last line we summed over μ and used Eq. (4.49). Thus $|\lambda| \le 1$.

Thus equilibrium exists, the remaining question is whether it is unique, and whether we approach it properly. Ergodicity guarantees uniqueness. However, in the general case there is the possibility of *limit cycles*, namely states where we cycle through some set:

$$\mu_1 \to \mu_2 \to \ldots \mu_k \to \mu_1. \quad\quad\quad (4.54)$$

If we get caught in a limit cycle, we will not go to $\mathbb{P}_e$.

If we have no cycles, then there is a famous theorem of Perron and Frobenius that guarantees approach to equilibrium. We do not need this strong result because we will impose a further condition on the process, that of *detailed balance*.

Detailed balance

Let us return to the master equation, Eq. (4.51), and note that for $\mathbb{P}_e$ we must have:

$$\sum_\nu \mathbb{W}(\mu,\nu)P_e(\nu,t) = \sum_\nu \mathbb{W}(\nu,\mu)P_e(\mu,t)$$

This is certainly guaranteed if we make another assumption: we postulate that each term in the two sums are equal, and that we have a probability distribution such that:

$$\mathbb{W}(\mu,\nu)P_e(\nu,t) = \mathbb{W}(\nu,\mu)P_e(\mu,t). \tag{4.55}$$

This is called detailed balance. Real mechanics has this property, as do quantum transitions Van Kampen (2007). If we impose it on our artificial dynamics, it makes things easier. Of course, this $\mathbb{P}_e$ has unit eigenvalue.

We can see that detailed balance eliminates cycles automatically. In equilibrium we cannot have Eq. (4.54) because we would be equally likely to traverse the sequence backwards.

Also, since $P_e(\mu) > 0$ we can symmetrize $\mathbb{W}$. Set:

$$\mathbb{V}(\nu,\mu) = [P_e(\nu)]^{-1/2}\mathbb{W}(\nu,\mu)[P_e(\mu)]^{1/2}.$$

A problem shows that $\mathbb{V}$ is symmetric and that its eigenvectors obey:

$$S_i(\mu) = [P_e(\mu)]^{1/2}P_i(\mu); \quad \mathbb{V}\mathbb{S}_i = \lambda_i\mathbb{S}_i; \quad \mathbb{W}\mathbb{P}_i = \lambda_i\mathbb{P}_i. \tag{4.56}$$

Since the eigenvectors of a real symmetric matrix span the space, we can expand any vector in eigenvectors of $\mathbb{V}$, and use this result to expand in eigenvectors of $\mathbb{W}$.

Result: approach to equilibrium

We have shown that if $\mathbb{W}$ is ergodic and obeys detailed balance then for any starting point we can write:

$$\mathbb{P}(t) = c_1\mathbb{P}_e + \sum_i c_i\mathbb{P}_i; \quad \mathbb{W}\mathbb{P}_e = \mathbb{P}_e; \quad \mathbb{W}\mathbb{P}_i = \lambda_i\mathbb{P}_i. \tag{4.57}$$

From Eq. (4.52)

$$\mathbb{P}(t+n) = c_1\mathbb{P}_e + \sum_i c_i\lambda_i^n\mathbb{P}_i. \tag{4.58}$$

Since the eigenvalues obey $|\lambda_i| < 1$ the sum becomes negligible for large n, and we must have $c_1 = 1, \mathbb{P}(t+n) \to \mathbb{P}_e$. If we write $\lambda_i = e^{-\Lambda_i}$ we identify the Λ's as relaxation rates.

4.6.3. Metropolis algorithm

We need to devise a Markov process that is ergodic and obeys detailed balance. Detailed balance allows us to build in the equilibrium we want, namely the Boltzmann distribution. Thus we need, from Eq. (4.55)

$$\frac{P_e(\mu)}{P_e(\nu)} = \frac{e^{-\beta E_\mu}}{e^{-\beta E_\nu}} = \frac{W(\mu \leftarrow \nu)}{W(\nu \leftarrow \mu)}. \tag{4.59}$$

There are many algorithms that satisfy this equation. The most popular and most efficient (see Newman & Barkema (1999)) is that of Metropolis et al. (1953). It involves several steps.

1. Start in state n, and try a move to state m.

2. Compute $\Delta = E_m - E_n$.

3. If $\Delta \leq 0$ accept the move, i.e. put the system in state m.

4. If $\Delta > 0$ accept the move with probability $e^{-\beta\Delta}$. That is, choose a random number $u \in [0,1]$ and if $u < e^{-\beta\Delta}$ move to state m, otherwise do nothing.

This does satisfy Eq. (4.59) because if $\Delta > 0$, $E_m > E_n$ the probability of moving $n \to m$ is $e^{-\beta(E_m - E_n)}$ and the probability of moving $m \to n$ is 1.

4.6.4. Monte Carlo simulations

Here is a very simple example to illustrate the technique. Consider a two level system, $n = 0, 1$ with energies $E_0 = 0$ and $E_1 = \Delta$. An elementary application of the canonical recipe gives $\langle n \rangle = 1/(e^{\beta\Delta} + 1)$.

Now we implement Monte Carlo: if you start in $n = 0$ you go to $n = 1$ with probability $e^{-\beta\Delta}$. If you start in $n = 1$ you go to $n = 0$ with probability 1. Here is the matrix $\mathbb{W}$:

$$\begin{pmatrix} 1 - e^{-\beta\Delta} & 1 \\ e^{-\beta\Delta} & 0 \end{pmatrix}$$

It is easy to see that this vector is an eigenvector with eigenvalue unity:

$$\begin{pmatrix} 1 \\ e^{-\beta\Delta} \end{pmatrix}$$

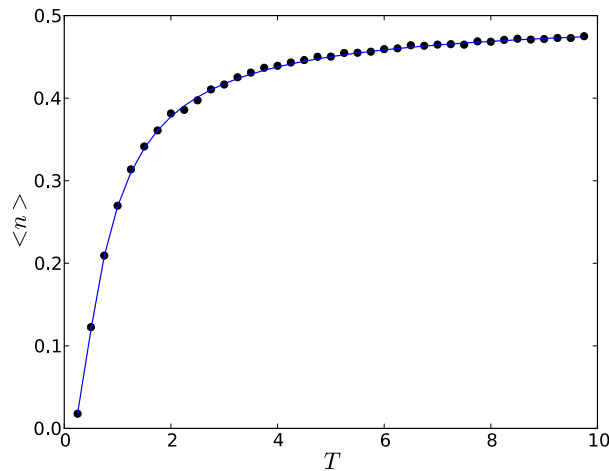Here is a code fragment to calculate $\langle n \rangle$:

Figure 4.2.: Monte Carlo results for a two level system: $\langle n \rangle$ as a function of T. The solid line is the exact result.

```
n=1
nsum=0.
w=exp(-1./T)
for j in arange(maxsteps):
    if n==1: nn=0
    if n==0:
        u=random()
        if u <= w:
            nn=1
        else:
            nn=0
    n=nn
    nsum=nsum+n
nave=nsum/maxsteps
```

Results for various T are plotted in Figure 4.2.

There is now an enormous industry in doing Monte Carlo simulations of the type we have described. The first application (Metropolis et al. (1953)) was to the hard-sphere gas in 2d and the second (Rosenbluth & Rosenbluth (1954)) was to 3d hard spheres, and to the Lennard-Jones gas.

A modern version of the second computation goes as follows: From above we see that we need only compute the configuration integral, Q_N. To do this, start with some configuration of atoms, and try a

move, in this case, simply pick another place for one of the atoms, e.g. at random. Then compute Δ and proceed as above. It is wise to monitor the energy to guess if the system has approached equilibrium, and then start to average. To get the pressure use the virial equation, Eq. (2.11). A program that does this is in Appendix B. We will use this method in Chapter 8.

For the Ising model the procedure is still simpler. Start with an arbitrary arrangement of spins. Pick a spin at random, and flip it according to the Metropolis scheme. After relaxation has occurred monitor the energy and magnetic moment. From problem 2, below, compute the heat capacity and magnetic susceptibility. This is the subject of problem 5.

Suggested reading

The canonical ensemble is treated in all the general references, e.g.:

Huang (1987)

The rotational and vibrational heat capacities and the low temperature behavior of H_2 see:

Pathria & Beale (2011)

Landau & Lifshitz (1980)

For the Mayer expansion the original source is

Mayer & Mayer (1940)

The modern treatment in

Pathria & Beale (2011)

is particularly detailed. For a quick treatment of B_2 alone, see

Landau & Lifshitz (1980)

Monte Carlo methods are in:

Newman & Barkema (1999)

Gould et al. (2006)

Chandler (1987)

Gould & Tobochnik (2010)

For Markov chains see:

Sethna (2006)

Van Kampen (2007)

Problems

1. Consider a tall cylinder with a classical ideal gas of N parti-
cles inside. Because of gravity, the gas is denser at the bottom.
Figure out the partition function, the free energy, and the heat
capacity.

2. Derive the formula for the vibrational heat capacity, Eq. (4.34).
Plot it as a function of T/T_v. It should resemble Figure 4.1.

3. a). Prove that $Var(E) = k_B T^2 C$, where C is the heat capacity.

b.) Prove for the Ising system that the magnetic susceptibility,
$(1/N)\partial M/\partial h|_T$ is given by:

$$N\chi = \frac{1}{k_B T} Var(M).$$

c.) Show that for the zero-field susceptibility for the Ising model
in the disordered phase you can write:

$$\chi = \frac{1}{k_B T} \sum_j \langle \sigma_j \sigma_0 \rangle_0.$$

The average uses the Hamiltonian with $h = 0$. Use Eq. (1.27).

d.) For the Ising case express χ in terms of the spin-spin cor-
relation function, Eq. (1.33). This is a simple example of a
fluctuation-dissipation relation.

4. Consider the Joule-Thomson process, pushing a gas through a
porous plug; see Figure 4.3. a.) It is a standard result that the
enthalpy is conserved in the process. Prove this by figuring out
the work done on both sides of the plug to move a parcel of gas
from left to right.

b.) Prove that the temperature change can be written:

$$(N/C_p)(TdB_2/dT - B_2),$$

where B_2 is the second viral coefficient, and we neglect higher
order terms.

Hints: You should compute $\partial T/\partial p|_H$. It will help to prove

$$\left(\frac{\partial H}{\partial T}\right)_p = C_p, \quad \left(\frac{\partial H}{\partial p}\right)_T = V + T\left(\frac{\partial S}{\partial p}\right)_T.$$
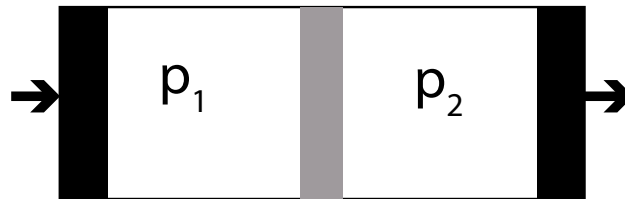
Figure 4.3.: The Joule-Thomson effect. Gas is forced through a porous plug
              (center); $p_1 > p_2$. The flow is slow enough that both sides of
              the system can be considered to be in equilibrium.

Also prove that, for small density, from the virial expansion

$$V \approx NkBT/p + NB_2.$$

c.) Show that for $B_2 = b - a/k_B T$, Eq. (4.44), the gas cools at
low T and heats for high T.

5. a.) Write Metropolis Monte Carlo code for the 1d Ising model,
and verify the result in Eq. (4.23).

b.) Write a 2d code and use the result of Problem (4.3)a to find
the heat capacity as a function of T. Be sure you go through the
phase transition.

6. Use detailed balance to prove that $\mathbb{V}$ is symmetric and that
Eq. (4.56) is correct.

7. An interesting application of quantum statistics is to the rota-
tional states of H_2 gas (see Landau & Lifshitz (1980)).

a.) The two nuclei– protons in this case– are identical and have
spin 1/2. Quantum mechanics says that if the total nuclear spin
is 1 (othohydrogen) then only odd J are allowed in the rotations,
and for spin 0 (parahydrogen) then only even J are allowed.
Explain this from quantum theory.

b.) Show that the equilibrium ratio of othro (o) to para (p) is:

$$x = N_o/N_p = 3Z_{odd}/Z_{even}$$

$$Z_{even,odd} = \sum (2J + 1)e^{-\beta(\hbar^2/2I)J(J+1)}$$

and the sum is over even and odd J's and I is the moment of
inertia.

c.) What is the low and high temperature limit of x?

d.) David Dennison, late of the University of Michigan physics department, pointed out that the heat capacity of H_2 gas at T's low compared to $\hbar^2/2Ik_B$ could be understood by assuming that x takes a long time to equilibrate so that when you cool the gas it retains the high-T value. Show that this means that C^{rot} for H_2 is much smaller than it would be in equilibrium. Find the leading temperature dependence for C^{rot} at low T for the two cases. (This calculation was the first solid evidence that the proton has spin.)

8. a.) Consider a system of N 1d classical anharmonic oscillators for which the potential is $V(x) = cx^2 + fx^4$. Suppose that the extra term is small. Show that the approximate heat capacity is $C \propto Nk_B(1 - [3fk_BT/2c^2])$.

b.) Suppose two molecules are bound by the Lennard-Jones potential. Find the equilibrium distance between the molecules, r_o. Expand the potential around r_o to third order, and figure out the shift of the equilibrium with temperature. Assume the shift is small. (Use the same technique as in part a.) This is the origin of thermal expansion of most solids.

9. Classical theory of paramagnetism: suppose a molecule has magnetic moment $\mathbf{m}$ and is in a magnetic field $\mathbf{B}$ which is in the z-direction. Use the Hamiltonian, $\mathcal{H} = -\mathbf{m} \cdot \mathbf{B} = -mB\cos(\theta)$ where θ is the polar angle. Figure out the average magnetic moment along the z direction as a function of T, i.e., $\langle m\cos(\theta)\rangle$. Expand the result for high temperature, and find the magnetic susceptibility, $\lim_{B \to 0} m/B$.

10. Classical theory of diamagnetism (Bohr - van Leeuven theorem): consider a classical gas of charged particles in a constant, uniform external field in the z-direction, B_z. We might expect that the field would induce eddy currents by Lenz's law, and give rise to a diamagnet. Show that this is false: the magnetic moment of the gas is zero for any field.

Hint: Show that the external field can be taken to arise from a vector potential, $\mathbf{A} = (0, xB_z, 0)$. Write the Hamiltonian from Eq. (1.2). Then, by a change of variables, show that Z is independent of B_z.