Part II Statistical Mechanics, Lent 2005

Prof. R.R. Horgan

February 14, 2005

1 Statistical Mechanics

1.1 Introduction

The classical and quantum dynamics of simple systems with few degrees of freedom is well understood in that we know the equations of motion which govern their time evolution and, in most cases, we can solve these equations exactly or, in any case, to the required degree of accuracy. Even systems with many degrees of freedom can be analyzed, for example, many problems in electromagnetic theory, where the field degrees of freedom are large, can be solved using Maxwell's equations. However, it remains that there are physical systems, which are typified by possessing a large number of degrees of freedom, for which analysis of this kind in inappropriate or, in practice, impossible, Whilst these systems are made up of identifiable constituents whose interactions are well known, different concepts are needed to describe the evolution of the system and the way it interacts with an observer. In particular, the concepts of heat, temperature, pressure, entropy, etc. must arise from a proper analysis – they are certainly not evident in the equations of motion, even though knowledge of the equations of motion is indispensable. These are ideas which come from a notion of statistical averaging over the detailed properties of the system. This is because they only make sense when thought of as properties which apply on the large scales typical of the observer, rather than the microscopic scales of the individual constituents. Statistical mechanics was developed to address this conceptual problem. It enables one to answer questions like "how do you calculate the physical properties of a system in thermal contact with large reservoir of heat (known as a heat bath) at temperature T?"

Statistical mechanics deals with **macroscopic**[†] systems with very many degrees of freedom. For example, a gas with a large number of particles or a complex quantum system with a large number of quantum states. A **thermally isolated** or **closed** system is one whose boundary neither lets heat in or out. Consider such a system, $\mathcal{S}$, that is not subject to any external forces. No matter how it is prepared, it is an experimental fact that after sufficient time $\mathcal{S}$ reaches a steady state in that it can be specified by a set, Σ, of time-independent **macroscopic** or **thermodynamic**, variables which usefully describe all the large-scale properties of the system for all practical purposes. The system is then said to be in a state of **thermal equilibrium** specified

---

[†] By macroscopic we mean that the system is large on the scale of the typical lengths associated with the microscopic description, such as the inter-atomic spacing in a solid or mean particle separation in a gas. That being said, it can be the case that systems with as few as a hundred particles can be successfully described by statistical mechanics.

by Σ. For example, for a gas isolated in this manner, Σ includes the pressure P, volume V, temperature T, energy E, and entropy S.

Although the state of thermal equilibrium described by Σ is unique and time-independent in this sense, the individual degrees of freedom such as particle positions and momenta are changing with time. It is certainly not practical to know the details of how each degree of freedom behaves, and it is a basic tenet of statistical mechanics that it is not necessary to know them in order to fully specify the equilibrium state: they are not of any practical importance to calculating the accessible experimental properties of the equilibrium system. The task of statistical mechanics is to use averaging techniques and statistical methods to predict the variables Σ characterizing the equilibrium state.

We are generally able to describe the dynamics of such systems at the **microscopic** level. For instance, we can specify the interaction between neighbouring atoms in a crystal or between pairs of particles in a gas. This description arises ultimately from our knowledge of quantum mechanics (and relativity), but it may also be legitimate and useful to rely on classical mechanics which we view as derivable from quantum mechanics by the correspondence principle. We shall develop the ideas of statistical mechanics using quantum mechanics in what follows. We define

(a) The **microscopic** states of $\mathcal{S}$ are the stationary quantum states $|i\rangle$.

[*This is the Dirac 'bra' ($\langle i|$) and 'ket' ($|i\rangle$) notation for states. The label i stands implicitly for the complete set of quantum numbers which specify the state uniquely. For example, $H|i\rangle = E_i|i\rangle$. The matrix element of an operator X between states $|i\rangle$ and $|j\rangle$ is $\langle i|X|j\rangle$ which, in wave-function notation, is the same as $\int dx\, \phi_i^* X \phi_j$. Then $\langle i|X|i\rangle$ is the expectation value, $\langle X\rangle$, of X.*]

(b) The **macroscopic** states of $\mathcal{S}$ are the possible states of thermodynamic equilibrium and are described by the corresponding set of thermodynamic variables, Σ. These are not states in the quantum mechanical sense but involve a vast number of microstates.

An important idea is the **ergodic hypothesis** (which has only been proved for a few systems[†]) which states that:

A system $\mathcal{S}$ evolves in time through all accessible microstates compatible with its state of thermodynamic equilibrium.

This means that time averages for a single system $\mathcal{S}$ can be replaced by averages at a **fixed** time over a suitable ensemble $\mathcal{E}$ of systems, each member identical to $\mathcal{S}$ in its macroscopic properties.

An important feature is that the ensemble, $\mathcal{E}$, can be realized in different ways whilst giving the same results for the thermodynamic properties of the original system. The major realizations of $\mathcal{E}$ are:

- **the microcanonical ensemble:** each member of $\mathcal{E}$ has the same value of E and N. This is appropriate to an **isolated** system for which the energy is fixed.

[†] There are many examples in one dimension where the ergodic hypothesis does not hold. This is a current research area

- **the canonical ensemble:** each member of $\mathcal{E}$ has N particles but E is not fixed but the average energy over $\mathcal{E}$ **is** fixed. This is appropriate to a system of a fixed number of particles but which is in thermal contact with a heat bath with which it can exchange energy. The heat bath is all the other systems of the ensemble.

- **the grand canonical ensemble:** neither E nor N is fixed for each member but the ensemble averaged values are fixed. This describes a system in thermal contact with a heat bath with which particles can also be exchanged.

Members of the last two ensembles represent the original system in the average. The reason why they all give the same thermal physics is that the fluctuations of E and N about their average values are very small because the number of particles involved is very large. For example, consider the probability p_m of obtaining $N/2+m$ heads from N coin tosses:

$$p_m \;=\; 2^{-N}\binom{N}{N/2+m} \;\sim\; \sqrt{\frac{2}{\pi N}}e^{-2m^2/N} \;,$$

using Stirling's formula ($\log n! \;=\; n\log n - n + \frac{1}{2}\log 2\pi n$) for $N \gg m \gg 1$. We see that only for $m \ll \sqrt{N}$ is p_m appreciable in size. Hence the mean fluctuation, $m/N \sim 1/\sqrt{N}$, is very small.

The density of a volume of gas V, $\rho = N/V$, is a macroscopic quantity on which thermodynamic variables depend. If the fluctuations in N are $O(\sqrt{N})$ then

$$\delta\rho \;\sim\; \sqrt{\frac{\rho}{V}} \;,$$

which vanishes as $V \longrightarrow \infty$. Hence, we expect the results from the different ensembles to be the same in the infinite volume limit but to differ by corrections of $O(1/\sqrt{V})$. The large volume limit is always assumed in what follows.

In this course we mainly study the canonical ensemble, but for fermions and bosons it will be simpler to use the grand canonical ensemble.

1.2 The Canonical Ensemble

We study a system $\mathcal{S}$ of N particles in a volume V with N and V fixed. The microstates of $\mathcal{S}$ are the complete orthogonal set $|i\rangle$. We assume that the spectrum is discrete with possible degeneracy.

We associate with $\mathcal{S}$ a **canonical ensemble** $\mathcal{E}$. This is a very large number A of distinguishable replicas $\mathcal{S}_1, \mathcal{S}_2, \mathcal{S}_3, \ldots$ of $\mathcal{S}$. Suppose we were to measure a complete set of commuting observables for all members of $\mathcal{E}$ simultaneously. Then let a_i be the number of members found in state $|i\rangle$. Then

$$\sum_i a_i \;=\; A \;, \qquad \text{and} \qquad \sum_i a_i E_i \;=\; AE \;, \tag{1.2.1}$$

where E is the fixed, time-independent ensemble average energy of the members of $\mathcal{E}$. In other words a_i/A is the (quantum mechanical) probability for finding the system in microstate $|i\rangle$. The set $\{a_i\}$ is called a **configuration** of $\mathcal{E}$, and the number of ways of partitioning A into the set $\{a_i\}$ is

$$W(\boldsymbol{a}) \;=\; \frac{A!}{\prod_i a_i!} \;. \tag{1.2.2}$$

The important principle underlying statistical mechanics is that we assign equal **a priori** probability to **each way** of realizing **each** allowed configuration. This means that the probability of finding configuration $\{a_i\}$ is proportional to $W(\boldsymbol{a})$. In the light of the earlier discussion, the probability distribution for the $\{a_i\}$ will be sharply peaked and the fluctuations about the peak value will be very small: $\sim O(1/\sqrt{A})$. Since we can take A as large as we like, the average of any variable over the ensemble will be overwhelmingly dominated by the value it takes in the most probable configuration, $\{\bar{a}_i\}$. For fixed N, V and E, we shall associate the state of thermodynamic equilibrium of S with the most probable configuration of $\mathcal{E}$ that is consistent with the constraints (1.2.1).

We find $\{\bar{a}_i\}$ by maximizing $\log W(\boldsymbol{a})$ w.r.t. $\{a_i\}$. Using Stirling's formula, ($\log n! = n \log n - n + \frac{1}{2}\log 2\pi n$), we have

$$\log W \sim A\log A - A - \sum_i a_i(\log a_i - 1) = A\log A - \sum_i a_i \log a_i \ , \tag{1.2.3}$$

since $\sum_i a_i = A$. We maximize this expression subject to the constraints (1.2.1) by using Lagrange multipliers. Then we have

$$\frac{\partial}{\partial a_j}\left(A\log A - \sum_i a_i \log a_i - \alpha\sum_i a_i - \beta\sum_i a_i E_i\right) \ = \ 0 \ , \quad \forall i \ . \tag{1.2.4}$$

Thus

$$\begin{aligned} &\log a_j + 1 + \alpha + \beta E_j = 0 \ , \\ \Longrightarrow \quad & \\ &a_j = e^{-1-\alpha-\beta E_j} \ . \end{aligned} \tag{1.2.5}$$

We eliminate α using (1.2.1), and define Z, the **canonical partition function**:

$$\begin{aligned} A \ &= \ \sum_i a_i \ = \ e^{-1-\alpha} Z \ , \\ \text{with} \quad & \\ Z \ &= \ \sum_i e^{-\beta E_i} \ = \ \sum_E \Omega(E) e^{-\beta E} \ , \end{aligned} \tag{1.2.6}$$

where $\Omega(E)$ is the **degeneracy** of levels with energy E, and E runs over all distinct values in the spectrum of the system.

The fraction of members of $\mathcal{E}$ found in microstate $|i\rangle$ in the macrostate of thermodynamic equilibrium is

$$\rho_i \ = \ \frac{a_i}{A} \ = \ \frac{1}{Z} e^{-\beta E_i} \ . \tag{1.2.7}$$

This is the **Boltzmann distribution** and is thought of as the (quantum mechanical) probability of finding $|i\rangle$ in the state of thermodynamic equilibrium. The average $\langle X\rangle$ of a physical observable is then

$$\langle X\rangle \ = \ \sum_i \langle i|X|i\rangle \rho_i \ . \tag{1.2.8}$$

For example,

$$\langle E\rangle \ = \ \sum_i \langle i|E|i\rangle \rho_i \ = \ \frac{1}{A}\sum_i a_i E_i \ = \ E \ , \tag{1.2.9}$$

from (1.2.1), as we expect it should.

Z is very important since we shall see that it allows us to calculate the thermodynamic and large-scale properties of a system in equilibrium from the quantum mechanical description of the system. An important example is that

$$\langle E\rangle \;=\; -\left(\frac{\partial \log Z}{\partial \beta}\right)_V \ . \qquad (1.2.10)$$

Holding V fixed means holding the E_i fixed since these latter depend on V.

As emphasized earlier for the canonical ensemble we can think of each individual system as being in contact with a heat bath made up of the rest of the ensemble. What are the **fluctuations** in energy of a system? Using above results, we have

$$\begin{aligned}\frac{\partial E}{\partial \beta} \;&=\; -\frac{1}{Z}\frac{\partial^2 Z}{\partial \beta^2} + \frac{1}{Z^2}\left(\frac{\partial Z}{\partial \beta}\right)^2 \\ &=\; -\langle E^2\rangle + \langle E\rangle^2 \;=\; -\,(\Delta E)^2 \ . \qquad (1.2.11)\end{aligned}$$

For large systems typically $E \sim N$ and, as E depends smoothly on β we expect

$$\frac{\partial E}{\partial \beta} \sim N \ , \qquad \Longrightarrow \qquad \frac{|\Delta E|}{E} \;\sim\; \frac{\sqrt{N}}{N} \;=\; N^{-\frac{1}{2}} \;\propto\; V^{-\frac{1}{2}} \ . \qquad (1.2.12)$$

We have found that energy plays a vitally central rôle in determining the probability for finding a system in a given state. In principle, the ρ_i can depend on other variables too. For example, the particle number N_i, corresponding to the grand ensemble, or even angular momenta where relevant. In general, such quantities must be be conserved and observable on the large scale.

1.3 Temperature

Consider two systems S_a and S_b with volumes V_a, V_b and particle numbers N_a, N_b, respectively. Form a joint ensemble $\mathcal{E}_{ab}$ from the separate ensembles $\mathcal{E}_a$ and $\mathcal{E}_b$ by allowing all members of both to exchange energy whilst keeping the overall total energy fixed. By allowing S_a and S_b to exchange energy we are allowing for interactions between the particles of the two systems. These will be sufficient to establish equilibrium of S_a with S_b but otherwise be negligible. For instance, they take place across a boundary common to the systems and so they contribute energy corrections which scale with the boundary area and so are therefore negligible compared with the energy, which scales with volume. Now form a composite system S_{ab} by joining S_a and S_b together. Let S_a have microstates $|i\rangle_a$ with energy E^a_i and S_b have microstates $|i\rangle_b$ with energy E^b_i. Because the extra interactions can be neglected in the bulk, then S_{ab} will have microstates $|ij>_{ab} = |i>_a\,|j>_b$ with energy $E^{ab}_{ij} = E^a_i + E^b_j$.

Now, $\mathcal{E}_a$ and $\mathcal{E}_b$ are **separately** in equilibrium characterized by Lagrange multipliers β_a and β_b, respectively. Also, $\mathcal{E}_{ab}$ is an equilibrium ensemble which is characterized by β_{ab}, and so we must have

$$\rho^{ab}_{ij}(E^a_i + E^b_j) \;=\; \rho^a_i(E^a_i)\rho^b_j(E^b_j) \ , \quad \forall\ E^a_i \text{ and } E^b_j \ . \qquad (1.3.1)$$

Then

$$\frac{e^{-\beta_{ab}(E^a_i+E^b_j)}}{Z_{ab}} \;=\; \frac{e^{-(\beta_a E^a_i + \beta_b E^b_j)}}{Z_a Z_b} \ , \quad \forall\ E^a_i \text{ and } E^b_j \ . \qquad (1.3.2)$$

This can only be satisfied if $\beta_{ab} = \beta_a = \beta_b$ and hence $Z_{ab} = Z_a Z_b$.

It is interesting to note that the result (1.3.1) could be given as the **definition** of what it means for S_a to be in thermal equilibrium with S_b. This is because we expect that in equilibrium these probabilities depend only on the energies of the systems. Then relation (1.3.1) must hold since energies add and probabilities multiply. The only solution to (1.3.1) is the Boltzmann distribution (1.2.7).

The relation (1.3.1) can be established directly by finding the most probable distribution for the ensemble $\mathcal{E}_{ab}$. Suppose $|i\rangle_a$ occurs a_i times, $|j\rangle_b$ occurs b_j times and $|ij\rangle_{ab}$ occurs c_{ij} times in $\mathcal{E}_{ab}$. Then we have

$$\sum_i a_i = A \ , \qquad \sum_i b_i = A \tag{1.3.3}$$

$$a_i = \sum_j c_{ij} \ , \quad b_j = \sum_i c_{ij} \ , \quad \sum_{ij} c_{ij} = A \ . \tag{1.3.4}$$

We only have **one** constraint on the energy which is that the total energy of $\mathcal{E}_{ab}$ is fixed:

$$AE = \sum_i a_i E_i^a + \sum_j b_j E_j^b \ , \tag{1.3.5}$$

$$AE = \sum_{ij} c_{ij}(E_i^a + E_j^b) \ . \tag{1.3.6}$$

The configuration $\{a_i, b_j\}$ arises in

$$W(\boldsymbol{a})W(\boldsymbol{b}) \propto \frac{A!}{\prod_i a_i!}\frac{A!}{\prod_j b_j!} \ , \tag{1.3.7}$$

ways and the configuration $\{c_{ij}\}$ arises in

$$W(\boldsymbol{c}) \propto \frac{A!}{\prod_{ij} c_{ij}!} \ , \tag{1.3.8}$$

ways. The most probable distribution of $\{a_i, b_j\}$ is obtained by maximizing $W(\boldsymbol{a})W(\boldsymbol{b})$ subject to (1.3.4) and (1.3.6) to get

$$\begin{aligned} a_i &= e^{-1-\alpha_a-\beta E_i^a} \ , \quad \rho_i^a = \frac{a_i}{A} \ , \\ b_j &= e^{-1-\alpha_b-\beta E_j^b} \ , \quad \rho_j^b = \frac{b_j}{A} \ . \end{aligned}$$

Note that since there is only **one** energy constraint the most probable $\{a_i\}$ and $\{b_j\}$ depend on the **same** value of β. The temperature, T, is **defined** by the statement that two systems in thermal equilibrium have the same temperature. We therefore conclude that

two systems in thermal equilibrium are described by Boltzmann distributions with a common value of β which can therefore be thought of as defining the temperature

The most probable distribution of $\{c_{ij}\}$ is obtained by maximizing $W(\boldsymbol{c})$ subject to the constraints (1.3.4) and (1.3.6) to get

$$c_{ij} = e^{-1-\alpha_{ab}-\beta'(E_i^a+E_j^b)} \ , \quad \rho_{ij}^{ab} = \frac{c_{ij}}{A} \ . \tag{1.3.9}$$

So far β and β' do not necessarily have the same value since they are just Lagrange multipliers imposing the energy constraint in two separate maximizations. However, the relations between $\{c_{ij}\}$ and $\{a_i, b_j\}$ given in (1.3.4) can be only be satisfied if $\beta' = \beta$. It then follows that (c.f. (1.3.1))

$$Z_{ab}(\beta) \;=\; Z_a(\beta) Z_b(\beta) \;, \quad \text{and} \quad \rho^{ab}_{ij} \;=\; \rho^a_i \rho^b_j \;\; . \tag{1.3.10}$$

The temperature can be **defined** by

$$\beta \;=\; \frac{1}{kT} \;, \tag{1.3.11}$$

where k is Boltzmann's constant and T is in degrees Kelvin. Because the analysis applies to any pair of systems k is a universal constant: $k = 1.38\ 10^{-23}$ Joules/^0K. This definition ensures that the average energy increases as T increases. For a dilute gas with N particles this implies Boyle's law, $PV = NkT$, as we shall see.

1.4 Thermodynamic ideas

A system undergoes an adiabatic change if

(i) the system is thermally isolated so that no energy in the form of heat can cross the system boundary: $\delta Q = 0$,

(ii) the change is caused by purely by external forces acting on the system,

(iii) the change takes place arbitrarily slowly.

By heat Q we mean energy that is transfered in a disordered way by random processes. In contrast, work W is energy transfered through the agency of external forces in an ordered way.

Work can be done on the system by moving one of the walls (like a piston) with an external force, which corresponds to changing the volume. If the change is adiabatic, no energy is permitted to enter or leave the system through the walls. Because the motion is slow the applied force just balances the pressure force PA (A is the area of the wall), and the work done **on** the system is then $\delta W = -PA\delta L = -P\delta V$. If the system is in a given microstate with energy E_i the change is slow enough that it remains in this state – there is no induced quantum-mechanical transition to another level. This is the Ehrenfest principle. The change in energy balances the work done and so

$$\delta W_i \;=\; \delta E_i \;=\; \frac{\partial E_i}{\partial V}\delta V \;. \tag{1.4.1}$$

For a system in TE the work done in an adiabatic change is the ensemble average of this result

$$-P\delta V \;=\; \delta W \;=\; \sum_i \rho_i \frac{\partial E_i}{\partial V}\delta V \qquad \Longrightarrow \qquad P \;=\; -\sum_i \rho_i \frac{\partial E_i}{\partial V} \;. \tag{1.4.2}$$

This is reasonable since if P exists then it is a thermodynamic variable and so should be expressible as an ensemble average.

These ideas are most easily understood in the context of an ideal gas. Consider the quantum mechanics of a gas of N non-interacting particles in a cube of side L. The single particle energy levels are $E_{\boldsymbol{n}} = \hbar^2 \boldsymbol{n}^2/2mL^2$ with $\boldsymbol{n} \in \mathcal{Z}^3$. Note that for given $\boldsymbol{n}$, $E_{\boldsymbol{n}} \propto V^{-2/3}$ – generally, we expect $E = E(V)$. Then

$$\begin{aligned} \frac{\partial E_i}{\partial V} &= -\frac{2}{3}\frac{E_i}{V} \\ \Longrightarrow \quad P &= \frac{2}{3V}\sum_i \rho_i E_i \;=\; \tfrac{2}{3}\epsilon, \end{aligned} \tag{1.4.3}$$

where $\epsilon = E/V$ is the energy density. This is the formula relating the pressure to the energy density for a non-relativistic system of particles. We also deduce the energy conservation equation for an adiabatic change ($\delta Q = 0$):

$$P\delta V \;+\; \delta E \;=\; 0\ . \tag{1.4.4}$$

For the ideal gas an adiabatic change **does not** change the occupation probabilities ρ_i of the energy levels – just the **energies** of the levels change. This is the Ehrenfest principle. For systems with $E_i(V) \propto V^{-s}$ the ρ_i are **still** Boltzmann equilibrium probabilities in $V + \delta V$ but with a different temperature – $s = 2/3$ for the ideal gas.

If the change is not adiabatic then heat energy can enter or leave the system. Since $E = \sum_i \rho_i E_i$, in the most general case we have

$$\delta E \;=\; \sum_i \rho_i \delta E_i \;+\; \sum_i E_i \delta\rho_i$$

or

$$\delta E \;=\; -P\delta V \;+\; \sum_i E_i\delta\rho_i\ . \qquad \text{(note that} \sum_i \delta\rho_i = 0.) \tag{1.4.5}$$

To deal with this eventuality we introduce the concept of entropy. Define the entropy, S, of a system by

$$S \;=\; \frac{k}{A}\log W(\boldsymbol{a})\ , \tag{1.4.6}$$

where, as before, $W(\boldsymbol{a})$ is the number of ways of realizing the ensemble configuration $\{a_i\}$. Since there are A members in the ensemble this is the entropy per system. Note that the ensemble is not necessarily an equilibrium one. However, with this definition we see that for an **equilibrium** ensemble

$$S_{\text{eq}} \;=\; \frac{k}{A}\log W_{\text{max}}\ , \tag{1.4.7}$$

and from (1.4.6) and (1.4.7) we deduce that S is maximized for the an **equilibrium** ensemble. Using (1.2.2) we get

$$\begin{aligned} S &= \frac{k}{A}\left(A\log A - \sum_i a_i \log a_i\right) \\ &= \frac{k}{A}\left(A\log A - \sum_i a_i\log\rho_i - \log A\sum_i a_i\right)\ , \\ \Longrightarrow & \end{aligned}$$

$$\boxed{S \;=\; -\,k\textstyle\sum_i \rho_i\log\rho_i} \tag{1.4.8}$$

where we have used $\rho_i = a_i/A$, $\sum_i a_i = A$. This is an important formula which can be derived in many different ways from various reasonable requirements. However, for complex systems it needs some interpretation but for the free gas there is no difficulty. More on entropy shortly. Let us apply the idea to a general change between nearby states of TE, e.g., an isothermal change ($dT = 0$).

Change the volume of the system arbitrarily slowly and allow heat energy to be absorbed/emitted, e.g., by contact with a heat bath. At every stage the system will have time to equilibrate and so although the ρ_i will be time-dependent they will take their equilibrium values appropriate to the conditions prevailing (i.e., values of V, T). In this case we have (S will stand for the equilibrium value unless otherwise stated)

$$\begin{aligned}
\rho_i &= \frac{1}{Z(V,T)} e^{-\beta E_i(V)} \;, \\
\Longrightarrow & \\
\delta S &= -k \sum_i \delta\rho_i (\log \rho_i + 1) \;, \\
&= -k \sum_i \delta\rho_i (-\beta E_i - \log Z + 1) \;, \\
&= \frac{1}{T} \sum_i E_i \delta\rho_i \\
\Longrightarrow & \\
\sum_i E_i \delta\rho_i &= T\delta S \;. \qquad (1.4.9)
\end{aligned}$$

So, for a slow change, using (1.4.5), we get in general

$$\delta E \;=\; T\delta S \;-\; P\delta V \;,$$

or, in exact differentials

$$dE \;=\; TdS \;-\; PdV \;. \qquad (1.4.10)$$

We can think of this as a comparison between two infinitesimally different states of thermodynamic equilibrium. This equation is an expression of the **First Law of Thermodynamics** which is the statement of energy conservation:

The First Law of Thermodynamics states that

$$dE \;=\; \delta Q \;+\; \delta W \;. \qquad (1.4.11)$$

Here dE is the differential change in the **internal energy** of $\mathcal{S}$, δQ is the heat energy supplied **to** $\mathcal{S}$, and δW is the work done **on** $\mathcal{S}$ (e.g., $\delta W = -PdV$).

For a slow change where the system is instantaneously **always** in thermal equilibrium the change is said to be **reversible**. This is because no work done is dissipated as unrecoverable energy such as through frictional forces or stirring of the system. In principle, the work done **on** the system can be recovered by allowing it to do work **on** another system, for example, a slowly compressed spring. From above, comparing (1.4.10) and (1.4.11), for a **reversible** change we have

reversible change $\Longrightarrow$

$$\delta Q \;=\; TdS \;. \qquad (1.4.12)$$

Note that neither δQ or δW are exact derivatives even though dE, dS and dV are. For an **adiabatic** change we have $dS = \delta Q/T = 0$. So the definition of an adiabatic change may be taken to be $dS = 0$.

From (1.4.10) we see that

$$T \;=\; \left(\frac{\partial E}{\partial S}\right)_V \;, \qquad P \;=\; -\left(\frac{\partial E}{\partial V}\right)_S \;. \tag{1.4.13}$$

For a system in equilibrium we can derive $S \equiv S_{\rm eq}$ from the partition function. We have

$$\begin{aligned} S &= -k\sum_i \rho_i \log \rho_i \\ &= -k\sum_i \rho_i(-\beta E_i \;-\; \log Z) \\ &= \frac{1}{T}E \;+\; k\log Z \;. \end{aligned} \tag{1.4.14}$$

We define the **free energy**, F, to be

$$F \;=\; -\,kT\log Z \quad\Longrightarrow\quad F \;=\; E - TS \;. \tag{1.4.15}$$

In differential form we then have

$$dF \;=\; dE - TdS - SdT \;,$$

or, using the first law in the form (1.4.10) for nearby equilibrium states

$$dF \;=\; -\,PdV - SdT \;, \quad\Longrightarrow\quad P \;=\; -\left(\frac{\partial F}{\partial V}\right)_T \;, \quad S \;=\; -\left(\frac{\partial F}{\partial T}\right)_V \;. \tag{1.4.16}$$

We shall see that the idea of the free energy, F, is a very important one in thermodynamics. From (1.4.16) we deduce that F is a function of the independent variables V and T. Similarly, from (1.4.16),

$$F \;=\; F(V,T) \;, \quad\Longrightarrow\quad P \;=\; P(V,T) \;, \quad S \;=\; S(V,T) \;. \tag{1.4.17}$$

Note also that $E = E(S,V)$ since S,V are the independent variables. Thus the thermodynamic state of equilibrium characterized by P,V,T,S can be determined from the partition function, Z, through the idea of the free energy, F. The equation which relates P,V and T , in either of the forms shown in (1.4.16) and (1.4.17), is known as the **equation of state**. Remember, all these quantities depend implicitly on the fixed value of N and hence on the particle density $n = N/V$.

1.5 Extensive and intensive variables

Suppose an equilibrium system is placed next to an exact copy and the partition between them is withdrawn to make the system twice the size. Variables that double in this process are called **extensive** and those whose value is the same as the original are called **intensive**. For example, N, V, S, E are extensive and P, T are intensive.

To elaborate. Consider two equilibrium systems at temperature T, $\mathcal{S}_1$ and $\mathcal{S}_2$, which are identical in every way except that they have volumes V_1 and V_2, respectively. Put them next to each other but do not remove the partition. Then for the whole system, $\mathcal{S}_{12}$, we have $V_{12} = V_1 + V_2,\ N_{12} = N_1 + N_2$. Since, $Z_{12} = Z_1 Z_2$, we see that $E = E_1 + E_2,\ S = S_1 + S_2$ etc. If these variables are **extensive** as claimed then these equalities will hold **after** the partition is removed. We expect this to be the case if $\mathcal{S}_1, \mathcal{S}_2$ are homogeneous (which means $\mathcal{S}_{12}$ will be, too). The particles in each system are of the same nature and are not **distinguishable** and so by removing the partition we do not mix the systems in a **detectable** way. Hence, the entropy S_{12} should remain unchanged. Then we expect that V, E, S are proportional to N. Clearly, $\mathcal{S}_1$ and $\mathcal{S}_2$ have the same T, P and so T, P are **intensive**.

This expectation can fail, for example,

(a) if surface energies are not negligible compared with volume energies – the partition matters too much,

(b) there are long-range forces such as gravity. The systems are no longer homogeneous. Think of the atmosphere – different parcels of air at different heights may have the same temperature but the pressure and density are functions of height.

1.6 General Remarks

(i) Let $\mathcal{S}$ be the sum of two systems, $\mathcal{S} = \mathcal{S}_1 + \mathcal{S}_2$, neither of which is **necessarily** in equilibrium. Although the partition function is not useful we can still show that the entropy defined by (1.4.8) is additive. Let the level occupation probabilities for $\mathcal{S}_1, \mathcal{S}_2$ be ρ_i, σ_i, respectively. Then

$$\begin{aligned} S &= -k \sum_{ij} \rho_i \sigma_j \log \rho_i \sigma_j \\ &= -k \sum_{ij} \rho_i \sigma_j (\log \rho_i + \log \sigma_j) \\ &= -k \sum_{i} \rho_i \log \rho_i + \sum_{j} \sigma_j \log \sigma_j \\ &= S_1 \ + S_2 \ , \end{aligned} \tag{1.6.1}$$

where we have used $\sum_i \rho_i = \sum_j \sigma_j = 1$.

(ii) From the definition of entropy (1.4.6) we see that the entropy is a **maximum** for a system in equilibrium. Indeed, it was the entropy that we maximized to determine the equilibrium state. More on this later.

(iii) Of the four variables, P, V, T, S, any two can be treated as independent on which the others depend. We have naturally found above V, T to be independent, but we can define $G = F + PV$ which leads to

$$dG \ = \ dF + PdV + VdP \ , \quad \Longrightarrow \quad dG \ = \ VdP - SdT \ . \tag{1.6.2}$$

G is called the **Gibbs function**. Clearly, $G = G(P, T)$ and now V and S are functions of P, T. This merely means that the experimenter chooses to control P, T and see how V, S vary as they are changed. The transformation $G = F + PV$, which changes the set of independent variables, is called a **Legendre transformation**.

1.7 The example of a perfect gas

A perfect gas consists of a system of free particles which therefore do not interact with each other. The gas is in a cubical box with walls at temperature T, and energy will be exchanged between the particles and the walls when the particle collide (inelastically) with the walls. The gas will come into equilibrium with the box and will have temperature T, also.

1.7.1 the density of states

Let the number of particles be N, the box be of side L with volume $V = L^3$, and the mass of each particle be m. We use periodic boundary conditions to impose the condition of finite volume so that the single particle wavefunction in the box satisfies

$$\psi(\boldsymbol{x} + L\boldsymbol{e}_i) \;=\; \psi(\boldsymbol{x}) \;, \tag{1.7.1.1}$$

where the $\boldsymbol{e}_i,\ i = 1, 2, 3$ are unit vectors spanning R^3. For free non-relativistic particles we then have energy eigenstates

$$\psi_{\boldsymbol{k}}(\boldsymbol{x}) \;=\; \frac{1}{\sqrt{V}}\, e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \quad \text{with} \quad \boldsymbol{k} \;=\; \frac{2\pi}{L}\boldsymbol{n} \;, \quad \boldsymbol{n} \in \mathcal{Z}^3 \;, \tag{1.7.1.2}$$

$$\boldsymbol{p} \;=\; \hbar\boldsymbol{k} \;, \qquad \epsilon(\boldsymbol{p}) \;=\; \frac{\hbar^2\boldsymbol{k}^2}{2m} \;. \tag{1.7.1.3}$$

We are interested in the limit $L \longrightarrow \infty$ and so for fixed momentum, $\boldsymbol{p}$, we have that $|\boldsymbol{n}| \longrightarrow \infty \;, \quad |\boldsymbol{n}|/L$ fixed. Since, $|\boldsymbol{n}|$ is large we can treat $\boldsymbol{n}$ as a continuous variable. The number of states in $\boldsymbol{n} \longrightarrow \boldsymbol{n} + d\boldsymbol{n}$ is d^3n which is therefore the number of states in $\boldsymbol{k} \longrightarrow \boldsymbol{k} + d\boldsymbol{k}$ with $d\boldsymbol{k} = (2\pi/L)d\boldsymbol{n}$. Thus,

$$\text{The number of states in} \quad \boldsymbol{k} \longrightarrow \boldsymbol{k} + d\boldsymbol{k} \quad \text{is} \quad \left(\frac{L}{2\pi}\right)^3 d^3k \;. \tag{1.7.1.4}$$

Because the state label can be treated as continuous we have

$$\begin{aligned}
\sum_i \;\longrightarrow\; & \left(\frac{L}{2\pi}\right)^3 \int d^3k \;=\; V\int \frac{d^3k}{(2\pi)^3} \;=\; V\int \frac{d^3p}{(2\pi\hbar)^3} \\
=\; & \frac{V}{(2\pi\hbar)^3}\int 4\pi p^2 dp \;=\; \frac{V}{2\pi^2\hbar^3}\int p^2\frac{dp}{d\epsilon}\, d\epsilon \\
\equiv\; & \int g(\epsilon) d\epsilon
\end{aligned}$$

$$\text{where} \quad g(\epsilon) \;=\; \frac{V}{2\pi^2\hbar^3}p^2\,\frac{dp}{d\epsilon} \;. \tag{1.7.1.5}$$

For particles with spin there is an extra factor of $g_s = 2s + 1$. Then, using (1.7.1.3), we have

$$\begin{aligned}
g(\epsilon) \;&=\; g_s\frac{V}{2\pi^2\hbar^3}\,(2m\epsilon)\,\frac{m}{\sqrt{2m\epsilon}} \;, \\
\Longrightarrow \qquad g(\epsilon) \;&=\; g_s\frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\epsilon^{\frac{1}{2}} \;.
\end{aligned} \tag{1.7.1.6}$$

The function $g(\epsilon)$ is the **density of states**. A general analysis in D-dimensions gives

$$g(\epsilon) \;=\; g_s \frac{V S_D}{(2\pi\hbar)^D}\, p^{D-1} \frac{dp}{d\epsilon} \;, \tag{1.7.1.7}$$

where S_D is the surface area of the unit sphere in D-dimensions, and $V = L^D$. Again for non-relativistic particles we can use (1.7.1.3) to get

$$\begin{aligned} D=1 \qquad & g(\epsilon) \;=\; g_s \frac{V}{2\pi}\left(\frac{2m}{\hbar^2}\right)^{\frac{1}{2}} \epsilon^{-\frac{1}{2}} \\ D=2 \qquad & g(\epsilon) \;=\; g_s \frac{V}{4\pi}\left(\frac{2m}{\hbar^2}\right) \;. \end{aligned} \tag{1.7.1.8}$$

Note that for $D = 2$ $g(\epsilon)$ is a constant independent of ϵ.

For relativistic particles of rest mass m, the analysis is identical and we can use the general formula (1.7.1.7) with the relativistic dispersion relation $\epsilon(\boldsymbol{p}) = \sqrt{p^2c^2 + m^2c^4}$. In particular, photons have $m = 0$ and $g_s = 2$ since there are two distinct polarization states. In this case, and for $D = 3$ we get

$$g(\epsilon) \;=\; V \frac{\epsilon^2}{\pi^2 \hbar^3 c^3} \;. \tag{1.7.1.9}$$

Since $\epsilon = \hbar\omega$ we get the density of states in $\omega \longrightarrow \omega + d\omega$, $g(\omega)$, to be

$$g(\omega) \;=\; g(\epsilon)\frac{d\epsilon}{d\omega} \;=\; V\frac{\omega^2}{\pi^2 c^3} \;. \tag{1.7.1.10}$$

1.7.2 the partition function for free spinless particles

Consider N particles in volume $V = L^3$. Then the single particle partition function is

$$z \;=\; \sum_i e^{-\beta\epsilon_i} \;\longrightarrow\; \int_0^\infty d\epsilon\, g(\epsilon) e^{-\beta\epsilon} \;. \tag{1.7.2.1}$$

From (1.7.1.6), and setting $y = \beta\epsilon$, this gives

$$z \;=\; \frac{V}{4\pi^2}\left(\frac{2mkT}{\hbar^2}\right)^{\frac{3}{2}} \int_0^\infty dy\, y^{\frac{1}{2}} e^{-y} \;=\; V\left(\frac{2m\pi kT}{h^2}\right)^{\frac{3}{2}} \;. \tag{1.7.2.2}$$

In D-dimensions a similar results holds and we find from (1.7.1.7)

$$z \;\propto\; (kT)^{D/2} \;. \tag{1.7.2.3}$$

Now, from before, we have $Z = z^N$ and so $\log Z = N \log z$. Then

$$E \;=\; -N\frac{\partial}{\partial\beta}\log z \;=\; -N\frac{\partial}{\partial\beta}(-\frac{D}{2}\log\beta + \text{const.}) \;=\; \frac{D}{2}NkT \;. \tag{1.7.2.4}$$

This is an important result of classical statistical mechanics and is called the **equipartition of energy** which states that

The average energy per particle degree of freedom is $\frac{1}{2}kT$

Combining this result with $PV = \frac{2}{3}E$ from (1.4.3) we get the equation of state for a perfect gas.

$$PV \;=\; NkT \qquad\qquad \textbf{Boyle's Law.} \tag{1.7.2.5}$$

Note that this result holds in D-dimensions since the generalization of (1.4.3) is $P = \frac{2}{D}\epsilon$.

1.7.3 entropy and the Gibbs paradox

Using (1.4.15) and (1.4.16), which states that

$$F \;=\; -\,kT\log Z\ , \qquad\qquad S \;=\; -\left(\frac{\partial F}{\partial T}\right)_V\ ,$$

we get

$$\begin{aligned} S &= Nk\log z + \frac{3}{2}Nk \\ &= Nk\log V + \frac{3}{2}Nk\log\left(\frac{2m\pi kT}{h^2}\right) + \frac{3}{2}Nk\ , \end{aligned} \tag{1.7.3.1}$$

which is **not** extensive since $V \propto N$. This is **Gibbs paradox**. The resolution is to realize that using $Z = z^N$ treats the particles as **distinguishable** since it corresponds to summing **independently** over the states of each particle. In reality the particles are **indistinguishable**, and we should not count states as different if they can be transformed into each other merely by shuffling the particles. The paradox is resolved by accounting for this fact and taking $Z = z^N/N!$. Using Stirling's approximation, we find

$$S \;=\; Nk\log\left(\frac{V}{N}\right) + \frac{3}{2}Nk\log\left(\frac{2m\pi kT}{h^2}\right) + \frac{3}{2}Nk\ , \tag{1.7.3.2}$$

which **is** extensive.

We should note at this stage that dividing by $N!$ correctly removes the overcounting only in the case where the probability that two particles occupy the same level is negligible. This is the situation where Boltzmann statistics is valid and certainly it works for T large enough or the number density is low enough. In other cases we must be more careful and we shall see that the consequences are very important.

1.8 The Increase of Entropy

We saw in section 1.4 from the definition of entropy (1.4.6) and (1.4.7), that S is maximized for an **equilibrium** ensemble. This is a general principle which states that

For a system with fixed external conditions the entropy, S, is a maximum in the state of thermodynamic equilibrium.

By "fixed external conditions" we mean variables like V, E, N for an isolated system or V, T, N for a system in contact with a heat bath at temperature T.

For example, let $\mathcal{S}_1$ be a perfect gas of volume V_1 with N_1 molecules of species 1, and $\mathcal{S}_2$ be a perfect gas of volume V_2 with N_2 molecules of species 2. Both systems are in equilibrium at temperature T. Place them next to each other with the partition in place. The total entropy is $S_1 + S_2$ with

$$S_i \;=\; kN_i \log\left(\frac{V_i}{N_i}\right) + kN_i \underbrace{\left(\frac{3}{2}\log\left(\frac{2m_i\pi kT}{h^2}\right) + \frac{3}{2}\right)}_{B_i(T)} \; . \qquad (1.8.1)$$

Now remove the partition. Because the species are **distinguishable**, the new system, $\mathcal{S}_{12}$, is **not** initially in equilibrium but initially has entropy $S_1 + S_2$. The gases mix and after equilibrium has been attained the partition function of the new system, of volume $V = V_1 + V_2$, is

$$Z_{12} \;=\; \frac{{z_1}^{N_1}}{N_1!}\,\frac{{z_2}^{N_2}}{N_2!} \; . \qquad (1.8.2)$$

Note, **not** $(N_1 + N_2)!$ in the denominator. This gives entropy

$$\begin{aligned} S_{12} \;&=\; \\ & k\left[(N_1+N_2)\log(V_1+V_2) - N_1\log N_1 - N_2\log N_2 + N_1B_1(T) + N_2B_2(T)\right] \\ &=\; k\left[N_1\log\left(\frac{V_1+V_2}{N_1}\right) + N_2\log\left(\frac{V_1+V_2}{N_2}\right) + N_1B_1(T) + N_2B_2(T)\right] \; . \qquad (1.8.3) \end{aligned}$$

So

$$S_{12} - S_1 - S_2 \;=\; k\left[N_1\log\left(\frac{V_1+V_2}{V_1}\right) + N_2\log\left(\frac{V_1+V_2}{V_2}\right)\right] \;>\; 0 \; . \qquad (1.8.4)$$

The entropy has **increased** because the gases have mixed. To separate the two species requires to system to become more ordered – all the boys to the left and all the girls to the right. This is unlikely to happen spontaneously. From the definition of the entropy (1.4.6) the **more ordered** a system is, the **smaller** is $W(\boldsymbol{a})$ and the **lower** is the entropy.

Notice that S_{12} is the sum of the entropies that each gas would have if they each occupied the volume $V_1 + V_2$ separately. This is the law of **partial entropies** which holds for ideal gases.

Remarks:

(i) Entropy is a measure of information. When a system is in equilibrium we know the least about it since all available microstates are equally likely. In contrast, the entropy is low when we know that the system is in one or very few microstates. Information theory is based on this observation.

(ii) The microscopic definition of entropy $S = -k\sum_i \rho_i \log \rho_i$, for a complex system is subtle. In particular, reconciling the microscopic and thermodynamic definitions of an adiabatic change ($\delta S = 0$) is not simple but the perfect gas is an ideal system to examine these ideas.

(iii) When $\mathcal{S}_1$ and $\mathcal{S}_2$ separately come into equilibrium we maximize S_1 and S_2 **separately** each with two constraints (for E_1, N_1 and E_2, N_2), making four in all. When $\mathcal{S}_1$ and $\mathcal{S}_2$ are combined there are only two constraints (for $E_1 + E_2$ and $N_1 + N_2$) which are a subset of the original four. We expect that the maximum of $S_1 + S_2$ subject to these two constraints will be greater than the maximum attained when it is subject to the four constraints (it *may* be unchanged c.f. Gibbs paradox). Hence, the entropy will generally **increase** when the systems are combined. Mixing just corresponds to **removal** of some constraints.

(iv) We have not ever discussed **how** a system reaches equilibrium. In practice, it takes a certain time to evolve to equilibrium and some parts of the system may reach equilibrium faster than others. Indeed, the time constants characterizing the approach to equilibrium may differ markedly. This can mean that although equilibrium will be established it might not occur on time scales relevant to observation. For example, some kinds of naturally occurring sugar molecules are found in one isomeric form or handedness. In equilibrium, there would be equal of each isomer but this is not relevant to nature.

The description of the system at equilibrium need not include those forces which are negligible for most purposes. For example, exactly how a gas molecule interacts with the vessel walls is a surface effect and does not affect any thermodynamic property in the infinite volume limit. However, it is generally assumed that it is such small randomizing interactions which allow energy to be transfered between levels, allowing the most probable state to be achieved. They are vital to the working of the **ergodic hypothesis**. This is by no means certain to hold and in one dimension it is quite hard to establish that it does.

The point is, that if a closed system is not in equilibrium then, given enough time, the most probable consequence is an evolution to more probable macrostates and a consequent steady increase in the entropy. Given ergodicity, the probability of transition to a state of higher entropy is **enormously** larger than to a state of lower entropy. Chance fluctuations to a state of lower entropy will never be observed.

The law of increase of entropy states that:

If at some instant the entropy of a closed system does not have its maximum value, then at subsequent instants the entropy will not decrease. It will increase or at least remain constant.

2 Thermodynamics

Consider a volume V of gas with fixed number N of particles. From the work on statistical mechanics we have found a number of concepts which enable us to describe the thermodynamic properties of systems:

(i) E: the total energy of the system. This arose as the average ensemble energy and which has negligible fluctuations. It is extensive.

(ii) S: the entropy. This is a new idea which measures the number of microstates contributing to the system macrostate. S is a maximum at equilibrium and is an extensive property of the system. It has meaning for systems not in equilibrium.

(iii) For nearby equilibrium states $dE = TdS - PdV$.

(iv) F: the free energy. It is naturally derived from the partition function and, like E, is a property of the system. For nearby equilibrium states $dF = -SdT - PdV$. It is extensive.

(v) All thermodynamic variables are functions of two suitably chosen independent variables: $F = F(T, V),\ E = E(S, T)$. Note, if X is extensive we cannot have $X = X(P, T)$.

(vi) Equations of state express the relationship between thermodynamic variables. For example, $PV = NkT$ for a perfect gas.

It is because E, S, P, V, T, F etc. are properties of the equilibrium state that we are able to write infinitesimal changes in their values as exact derivatives. The changes in their values are **independent** of the path taken from an initial to final equilibrium state. This is not true of the heat δQ absorbed or emitted in a change between nearby equilibrium states – it does depend on the path: δQ is not an exact derivative. However, $dS = \delta Q/T$ **is** an exact derivative, and $1/T$ is the integrating factor just like in the theory of first-order o.d.es.

Energy conservation is expressed by the fist law:

$$\begin{aligned} \delta E &= \delta Q + \delta W \quad \text{for infinitesimal change} \\ \Delta E &= \Delta Q + \Delta W \quad \text{for finite change} \end{aligned} \tag{2.1}$$

where $\delta Q(\Delta Q)$ is the heat supplied **to** the system and $\delta W(\Delta W)$ is the work done **on** the system. Suppose the state changes from $E(S_1, V_1)$ to $E(S_2, V_2)$ then in general we expect

$$\Delta W \ \geq \ -\int_{V_1}^{V_2} PdV \tag{2.2}$$

because additional work is done against friction, generating turbulence etc. However, $dE = TdS - PdV$ which implies

$$\Delta Q \ \leq \ \int_{S_1}^{S_2} TdS\ . \tag{2.3}$$

2.1 reversible and irreversible changes

These inequalities hold as equalities for **reversible** changes. This will hold if the change is non-dissipative and **quasistatic**. **Quasistatic** means a change so slow that the system is always arbitrarily close to a state of equilibrium. The path between initial and final states can then be plotted in the space of relevant thermodynamic variables. This is necessary because otherwise thermodynamic variables such as P, T etc. are not defined on the path and we cannot express $\delta W = -P\delta V$ as must be the case for reversibility. In this case, no energy is dissipated in a way that cannot be recovered by reversing the process.

For an **irreversible** change the extra increase in entropy is due to the extra randomizing effects due to friction, turbulence etc. A simple example is rapid piston movement as well as mixing of distinguishable gases and free expansion of gas into a vacuum. For an isolated, or closed, system $\Delta Q = 0$ but $\Delta S \geq 0$, the equality holding for a **reversible** change. In real life there is always some dissipation and so even if $\delta Q = 0$ we expect $\delta S > 0$.

These ideas are compatible with common sense notions. For reversible transitions the system can be returned to its original state by manipulating the external conditions.

E.g., a spring that has been quasistatically compressed. Since the change is through a sequence of equilibrium configurations there is no thermodynamic reason why it could not be traversed in the opposite direction, thus reversing the change. For an irreversible change, such as mixing of different gases, no such reverse transition can occur. In the case of mixing it would require the gases to be separated **within** a closed system. It must be emphasized that the entropies of the individual parts of the closed system need not all increase or remain constant, but the total for the whole system must. Of course, chemical means can be used to separate mixed gases but only at the cost of **increasing** the entropy elsewhere since the closed system must include all parts of the experiment. If this were not the case then, overall, entropy would have decreased.

2.2 Applications

2.2.1 The Maxwell Relations

The variables we have considered so far, F, G, E, P, V, T, S, are all properties of the equilibrium state. They each depend on an appropriate pair of variables chosen as independent and are differentiable. These properties are encoded in differential relations between the variables. Using $dE = TdS - PdV$ we have

$$\begin{array}{lrcrlrcl} \text{Free energy} & F & = & E - TS : & & & & \\ \Longrightarrow & dF & = & -SdT - PdV & \Longrightarrow & F & = & F(T,V) \\[1em] \text{Enthalpy} & H & = & E + PV : & & & & \\ \Longrightarrow & dH & = & TdS + VdP & \Longrightarrow & H & = & H(S,P) \\[1em] \text{Gibbs function} & G & = & E - TS + PV : & & & & \\ \Longrightarrow & dG & = & -SdT + VdP & \Longrightarrow & G & = & G(T,P) \end{array} \tag{2.2.1.1}$$

Because these are all exact differentials we have, for example,

$$\left(\frac{\partial^2 E}{\partial S \partial V}\right) = \left(\frac{\partial^2 E}{\partial V \partial S}\right) . \tag{2.2.1.2}$$

Since

$$T = \left(\frac{\partial E}{\partial S}\right)_V , \quad P = -\left(\frac{\partial E}{\partial V}\right)_S , \tag{2.2.1.3}$$

we have

$$-\left(\frac{\partial P}{\partial S}\right)_V = \left(\frac{\partial T}{\partial V}\right)_S . \tag{2.2.1.4}$$

This is a **Maxwell relation**. Another such Maxwell relation arises from

$$S = \left(\frac{\partial F}{\partial T}\right)_V , \quad P = -\left(\frac{\partial F}{\partial V}\right)_T , \tag{2.2.1.5}$$

we find

$$\left(\frac{\partial S}{\partial V}\right)_T = \left(\frac{\partial P}{\partial T}\right)_V . \tag{2.2.1.6}$$

Two more Maxwell relations can be derived from G and H giving four in all. However, they are not all independent. This can be seen by considering $z = z(x, y)$. Then

$$\delta z = \left(\frac{\partial z}{\partial x}\right)_y \delta x + \left(\frac{\partial z}{\partial y}\right)_x \delta y \quad \Longrightarrow$$

$$
\begin{aligned}
1 &= \left(\frac{\partial z}{\partial x}\right)_y \left(\frac{\partial x}{\partial z}\right)_y \quad \text{and} \\
0 &= \left(\frac{\partial z}{\partial x}\right)_y + \left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z \ .
\end{aligned} \tag{2.2.1.7}
$$

Multiply the last equation by $(\partial x/\partial z)_y$ and use the second equation to get

$$
\left(\frac{\partial z}{\partial y}\right)_x \left(\frac{\partial y}{\partial x}\right)_z \left(\frac{\partial x}{\partial z}\right)_y = -1 \ . \tag{2.2.1.8}
$$

Substitute any three of P, V, T, S for x, y, z to relate the terms in Maxwell's relations.

We note that E, F, H, G are all extensive. Note that G explicitly only depends on the intensive quantities T, P and so its extensive character is due to N and we can write $G = \mu(P,T)N$, where μ is called the chemical potential. More on this later.

Consider $E = E(V,S)$ and change variables $(V,S) \longrightarrow (V,T)$. Then we have

$$
\left(\frac{\partial E}{\partial V}\right)_T = \left(\frac{\partial E}{\partial V}\right)_S + \left(\frac{\partial E}{\partial S}\right)_V \left(\frac{\partial S}{\partial V}\right)_T \ .
$$

Using $dE = TdS - PdV$ this gives

$$
\left(\frac{\partial E}{\partial V}\right)_T = -P + T\left(\frac{\partial S}{\partial V}\right)_T \ ,
$$

and using (2.2.1.6) (from differentiability of F) we have

$$
\left(\frac{\partial E}{\partial V}\right)_T = -P + T\left(\frac{\partial P}{\partial T}\right)_V \ . \tag{2.2.1.9}
$$

For a perfect gas the equation of state is $PV = NkT$ which implies

$$
\left(\frac{\partial E}{\partial V}\right)_T = 0 \ . \tag{2.2.1.10}
$$

Since we only need two independent variables we have in this case that $E = E(T)$ – a function of T only. This is an example of equipartition where $E = \frac{3}{2}NkT$.

2.2.2 specific heats

Because δQ is not an exact derivative there is no concept of the amount of heat stored in a system. Instead δQ is the amount of energy absorbed or emitted by the system by random processes and excluding work done by mechanical or other means. If we specify the conditions under which the heat is supplied then we can determine δQ using

$$
\delta Q = dE + \delta W \ , \tag{2.2.2.1}
$$

where dE **is** an exact derivative. It may be the case that absorption/emission of heat is **not** accompanied by a change in T. An example, is the latent heat of evaporation of water to a gas. However, this is a phase change which we exclude from the current discussion. We can write $\delta Q = CdT$ where C is the **heat capacity** of the system. Clearly, to give C meaning we need to specify the conditions under which the change takes place. Important cases are where the change is reversible and at fixed V or fixed P. Then $\delta Q = TdS$ and we define

$$
C_V = T\left(\frac{\partial S}{\partial T}\right)_V \ , \qquad C_P = T\left(\frac{\partial S}{\partial T}\right)_P \ , \tag{2.2.2.2}
$$

where C_V is the **heat capacity at constant volume** and C_P is **heat capacity at constant pressure**. Using

$$dE \;=\; TdS - PdV \quad \text{and} \quad dH \;=\; TdS + VdP \;. \tag{2.2.2.3}$$

we have also

$$C_V \;=\; \left(\frac{\partial E}{\partial T}\right)_V \;, \quad C_P \;=\; \left(\frac{\partial H}{\partial T}\right)_P \;. \tag{2.2.2.4}$$

Changing variables from T, P to T, V we have

$$\left(\frac{\partial S}{\partial T}\right)_P \;=\; \left(\frac{\partial S}{\partial T}\right)_V + \left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P \;, \tag{2.2.2.5}$$

and using the Maxwell relation (2.2.1.6) and the definitions of C_P and C_V we have

$$\begin{aligned} C_P \;&=\; C_V + T\left(\frac{\partial S}{\partial V}\right)_T \left(\frac{\partial V}{\partial T}\right)_P \\ \Longrightarrow& \\ C_P \;&=\; C_V + T\left(\frac{\partial V}{\partial T}\right)_P \left(\frac{\partial P}{\partial T}\right)_V \;. \end{aligned} \tag{2.2.2.6}$$

For any physical system we expect

$$\left(\frac{\partial V}{\partial T}\right)_P \;>\; 0 \;, \quad \left(\frac{\partial P}{\partial T}\right)_V \;>\; 0 \;, \tag{2.2.2.7}$$

and hence that $C_P > C_V$. This makes sense because for a given increment dT the system at fixed P expands and does work PdV on its surroundings whereas the system at fixed V does no work. The extra energy expended in the former case must be supplied by δQ which is correspondingly larger than in the latter case.

An important example is that of a perfect gas. An amount of gas with N_A (Avogadro's number) of molecules is called a **mole**. The ideal gas law tells us that at fixed T, P a mole of any **ideal** gas occupies the same volume. The ideal gas law for n moles is

$$PV \;=\; nN_AkT \;=\; nRT \;, \tag{2.2.2.8}$$

where R is the **gas constant**, and is the same for all ideal gases. Note that at sufficiently low density all gases are essentially ideal. The heat capacity per mole, c, is called the **specific heat**: $c = C/n$. Then from (2.2.2.6) we have

$$c_p - c_v \;=\; R \;. \tag{2.2.2.9}$$

We define

$$\gamma \;=\; \frac{c_p}{c_v} \;=\; \frac{c_v + R}{c_v} \quad \Longrightarrow \quad \gamma \;>\; 1 \;. \tag{2.2.2.10}$$

If the number of degrees of freedom per molecule is denoted N_F then equipartition of energy gives that the energy of 1 mole is

$$\begin{aligned} E \;&=\; \frac{N_F}{2}RT \\ \Longrightarrow& \\ c_v \;&=\; \frac{N_F}{2}R \;, \quad c_p \;=\; \left(\frac{N_F}{2} + 1\right)R \;, \quad \gamma \;=\; \left(\frac{N_F + 2}{N_F}\right). \end{aligned} \tag{2.2.2.11}$$

For a monatomic gas $N_F = 3$, for a diatomic gas $N_F = 5$ (3 position, 2 orientation), etc, and so $\gamma = 5/3, 7/5, \ldots$.

2.2.3 adiabatic changes

$\delta Q = 0$ for an adiabatic change. For n moles of an ideal gas $E = nc_vT,\ PV = nRT$ and then

$$0 \ = \ RdE + RPdV \qquad \text{(extra } R \text{ for convenience).} \tag{2.2.3.12}$$

As a function of (V, T) we have

$$dE \ = \ \left(\frac{\partial E}{\partial T}\right)_V dT + \left(\frac{\partial E}{\partial V}\right)_T dV \ = \ nc_v dT \ , \tag{2.2.3.13}$$

since $(\partial E/\partial V)_T = 0$ (eqns. (2.2.1.9),(2.2.1.10)) for a perfect gas. (Note, this will be true for all systems where E does not depend on V at fixed N.) Then

$$\begin{aligned} 0 &= Rnc_v dT + RPdV \\ &= c_v(PdV + VdP) + RPdV \quad \text{(using } PV = nRT\text{)} \\ &= c_p PdV + c_v VdP \ . \end{aligned} \tag{2.2.3.14}$$

Thus

$$c_p PdV + c_v VdP \ = \ 0 \ . \tag{2.2.3.15}$$

Thus

$$\boxed{PV^\gamma \ \textbf{ is constant on adiabatics} \qquad (2.2.3.16)}$$

Note that adiabatics, PV^γ constant, are steeper than isothermals, PV constant.

We can use these results to calculate the entropy of n moles of ideal gas. We have

$$\begin{aligned} TdS &= dE + PdV \\ &= nc_v dT + nRT\frac{dV}{V} \\ \Longrightarrow \\ dS &= nc_v\frac{dT}{T} + nR\frac{dV}{V} \\ \Longrightarrow \\ S &= nc_v \log T + nR\log V + S_0 \ . \end{aligned}$$

S_0 is the unknown constant of integration. Thermodynamics cannot determine S_0 and does not care whether S is extensive. This is resolved by the more complete approach of statistical mechanics given earlier. This result is the same as we found before (1.7.3.2) if all the constant parts are lumped into S_0. Using $PV = nRT$ we get

$$S \ = \ nc_v \log PV^\gamma + S_0' \ , \tag{2.2.3.17}$$

and so, as expected, S is constant on adiabatics (isentropics).

2.2.4 van der Waals' equation

In reality even a low density gas is not ideal because there are interactions between the gas molecules. Ultimately, at temperatures low enough these interactions cause a **phase transition** to occur and the gas liquefies. We can account approximately for the interactions by correcting the ideal gas law to give van der Waals' equation. The form of the inter-particle potential is shown in figure 1. There is a repulsive, hard-core, part at small separations and an attractive part for intermediate separations. The effect of these two regions is qualitatively modeled as follows:

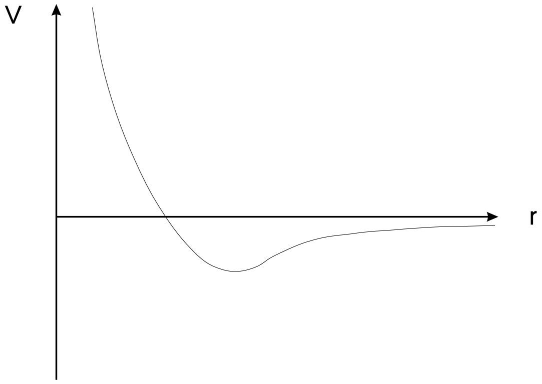


Figure 1:

(i) The hard-core is represented as an excluded volume since the volume available is reduced. We replace $V \longrightarrow V - b$.

(ii) The pressure measured by an observer, P, is due to the particles on the surface hitting the walls. Whilst in the bulk the attractive forces roughly cancel out because the particles are in homogeneous environment, this is not true at the surface since they are pulled back into the bulk. Consequently, P is less than the bulk pressure, P_B. Now, P is given by $P = d(\text{momentum/unit-area})/dt \propto nmv^2$ where n is the number density. The reduction in K.E. (and hence in v^2) due to the unbalanced attractions near the surface is also proportional to n. Hence the reduction in pressure takes the form

$$P_B - P \;\propto\; n^2 \quad\Longrightarrow\quad P_B - P \;=\; \frac{a}{V^2}\;, \tag{2.2.4.18}$$

for fixed N since $n = N/V$.

The ideal gas equation corrected for the excluded volume applies in the bulk so that

$$P_B(V-b) \;=\; NkT \quad\Longrightarrow$$

$$\left(P + \frac{a}{V^2}\right)(V-b) \;=\; RT\;. \qquad \textbf{van der Waals' equation}$$
$$a \propto N^2\;,\quad b \propto N\;,\quad R = Nk\;. \tag{2.2.4.19}$$

2.3 the Second Law of Thermodynamics

Kelvin-Planck: *It is impossible to construct an engine that, operating in a cycle, will produce no other effect than the extraction of heat from a reservoir and the performance of an equivalent amount of work.*

Clausius: *It is impossible to construct a device that, operating in a cycle, will produce no other effect than the transfer of heat from a cooler to a hotter body.*

These two versions can be shown to be equivalent.

By a cycle we mean a series of processes operating on a system for which the initial and final states are the same. Clearly, the most efficient cycle is a reversible one for

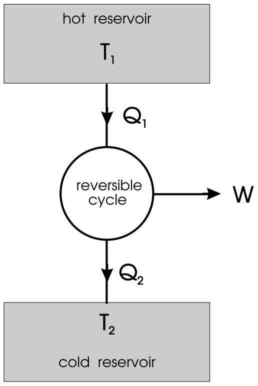


Figure 2:

which all the constituent processes are reversible. Such a cycle can be drawn as a closed curve in the 2D space of the independent thermodynamic variables (see figure 3).

Consider system shown in figure 2 which, acting in a reversible cycle, extracts heat Q_1 from the hot reservoir at temperature T_1 and delivers heat Q_2 to the cold reservoir at temperature T_2 and does work $W = Q_1 - Q_2$. Because the cycle is reversible the **total** change in entropy is zero. Thus,

$$\begin{aligned} \Delta S &= \frac{Q_2}{T_2} - \frac{Q_1}{T_1} = 0 \,, \\ \Longrightarrow \qquad Q_2 &= \frac{Q_1 T_2}{T_1}. \end{aligned} \tag{2.3.1}$$

The law of increase of entropy tells us that for realistic systems which have frictional losses etc.

$$Q_2 \geq \frac{Q_1 T_2}{T_1} \,. \tag{2.3.2}$$

This clearly agrees with the Kelvin-Planck statement of the second law which says that $Q_2 > 0$. The efficiency of the cycle is defined as

$$\eta = \frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1} \leq \frac{T_1 - T_2}{T_1} \,. \tag{2.3.3}$$

and we see that the reversible cycle (or engine) has the greatest efficiency.

The work done by the cycle is the areaof the cycle: $W = \int PdV$.

This important example of a reversible cycle is the Carnot cycle shown in figure 3 in which a sample $\mathcal{S}$ of perfect gas is taken around a closed curve in the PV plane by a sequence of adiabatic and isothermal reversible processes with $T_1 > T_2$. On AB the heat bath supplies heat Q_1 to $\mathcal{S}$ at temperature T_1. On CD the heat bath extracts heat Q_2 from $\mathcal{S}$ at temperature T_2. On the adiabatics, BC and DA, $TV^{\gamma-1}$ is constant so that

$$\begin{aligned} T_1 V_A^{\gamma-1} &= T_2 V_D^{\gamma-1} \,, \qquad T_1 V_B^{\gamma-1} = T_2 V_C^{\gamma-1} \qquad \Longrightarrow \\ \frac{V_A}{V_B} &= \frac{V_D}{V_C} \,. \end{aligned} \tag{2.3.4}$$

The direct connection of the second law in the form of Clausius with the law of increase of entropy can be seen as follows. Suppose there exists an engine I which is

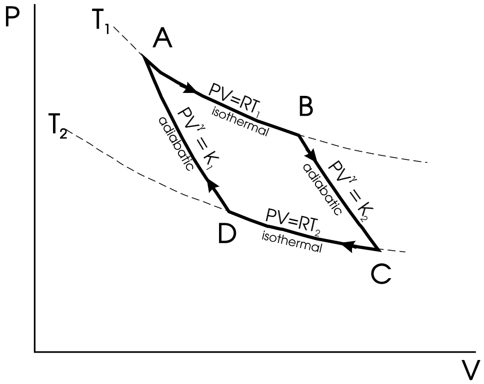


Figure 3:

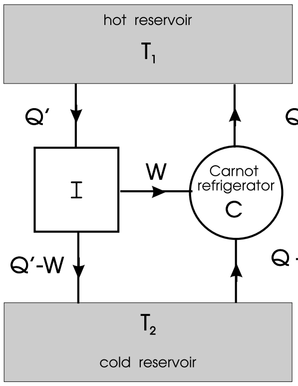


Figure 4:

more efficient that a Carnot engine, C, and so violates the law of increase of entropy. Let I drive C backwards as a Carnot refrigerator in the manner shown in figure 4. Note that no work is supplied or absorbed by the external surroundings since the output, W, of I is absorbed by C. Then

$$\eta_I \ > \ \eta_C \quad \Longrightarrow \quad \frac{W}{Q'} > \frac{W}{Q} \quad \Longrightarrow \quad Q' \ < \ Q \ . \tag{2.3.5}$$

However, the heat supplied **to** the hot heat bath is then $\Delta Q = Q - Q' > 0$ and this has been wholly transfered from the cold heat bath **without** the external input of work. This violates the Clausius statement of the second law. It also shows that the most efficient cycle is the Carnot cycle. Remember W and Q are path-dependent **even** for reversible paths and hence η also depends on the path used to define the cycle.

3 The Grand Canonical Ensemble

3.1 formalism

We follow an identical procedure to before. We consider a grand (canonical) ensemble, $\mathcal{G}$ of A replicas of the system $\mathcal{S}$ in microstates $|i\rangle$ of $\mathcal{S}$ with N_i particles and energies

E_i. Suppose a_i members of $\mathcal{G}$ are in microstate $|i\rangle$ so that

$$\sum_i a_i \;=\; A \;, \qquad \sum_i a_i E_i \;=\; AE \;, \qquad \sum_i a_i N_i \;=\; AN \;. \tag{3.1.1}$$

As before we associate thermodynamic equilibrium of $\mathcal{S}$ with the most probable configuration of $\mathcal{G}$ subject to the given constraints. We maximize the same $\log W$ as before:

$$\frac{\delta}{\delta a_i}\left(\log W - \alpha \sum_i a_i - \beta \sum_i a_i E_i - \gamma \sum_i a_i N_i\right) \;=\; 0 \;. \tag{3.1.2}$$

This gives us

$$a_i \;=\; e^{-(1+\alpha)} e^{-\beta(E_i - \mu N_i)} \;, \tag{3.1.3}$$

where we have defined the **chemical potential** μ by $\beta\mu = -\gamma$. The grand partition function is then

$$Z \;=\; \sum_i e^{-\beta(E_i - \mu N_i)} \;. \tag{3.1.4}$$

The fraction of members in microstate $|i\rangle$ is

$$\rho_i \;=\; \frac{a_i}{A} \;=\; \frac{e^{-\beta(E_i - \mu N_i)}}{Z} \;. \tag{3.1.5}$$

The grand ensemble average is

$$\hat{O} \;=\; \sum_i \rho_i O_i \;, \tag{3.1.6}$$

which gives $\hat{E} = E,\ \hat{N} = N$.

What we have done is exchange a **fixed** value of N for a value that is maintained in the **average** by choosing the corresponding value for μ – just as we did for E by choosing β. This has big advantages since it is often difficult to impose fixed N as a constraint on a system. It is something that we have not needed to explicitly encode hitherto. By introducing μ the ability to tune $\hat{N}$ to the right value using a variable that is explicitly part of the formalism is a huge advantage, as we shall see.

Then we have

$$\begin{aligned} N \;&=\; \frac{1}{\beta}\left(\frac{\partial \log Z}{\partial \mu}\right)_{\beta,V} \;, \\ E - \mu N \;&=\; -\left(\frac{\partial \log Z}{\partial \beta}\right)_{\mu,V} \;. \end{aligned} \tag{3.1.7}$$

Remember, the V dependence is in $E_i \equiv E_i(V)$.

As before, we considering adiabatic changes in V (e.g., constant ρ_i for the ideal gas). Then, as before, $\delta E = -P\delta V$ and so

$$P \;=\; -\sum \rho_i \frac{\delta E}{\delta V} \;=\; \frac{1}{\beta}\left(\frac{\partial \log Z}{\partial V}\right)_{\beta,\mu} \;. \tag{3.1.8}$$

Following the familiar argument, more general changes $\delta\rho_i \neq 0$, in which heat enters or leaves the system, obey

$$\sum_i \delta\rho_i \;=\; 0 \;, \quad \sum_i \delta\rho_i N_i \;=\; \delta N \;, \quad \delta E \;=\; -P\delta V + \sum_i \delta\rho_i E_i \;. \tag{3.1.9}$$

The entropy is defined in general, as before, to be

$$S \;=\; -\frac{k}{A}\log W \;=\; -\,k\sum_i \rho_i \log \rho_i \;. \tag{3.1.10}$$

For an equilibrium ensemble we then find

$$\begin{aligned} \delta S \;&=\; -k\sum_i \delta\rho_i \log \rho_i \;-\; \sum_i \rho_i \frac{1}{\rho_i}\delta\rho_i \\ &=\; -k\sum_i \delta\rho_i(-\beta(E_i-\mu N_i) - \log Z + 1)\;. \end{aligned} \tag{3.1.11}$$

This gives the fundamental thermodynamic relation

$$TdS \;=\; dE + PdV - \mu dN\;, \tag{3.1.12}$$

which corresponds to the first law $dE = \delta Q + \delta W_{\text{mech}} + \delta W_{\text{chem}}$, where $\delta W_{\text{chem}} = \mu dN$ is the work done in adding dN particles to the system. Note, we have conformed to the convention that an exact derivative is denoted as e.g., dE, whereas an inexact derivative is denoted as δQ. The chemical potential is then given by

$$\mu \;=\; \left(\frac{\partial E}{\partial N}\right)_{S,V}\;. \tag{3.1.13}$$

From the definition the entropy in equilibrium is

$$S \;=\; k\beta(E-\mu N) \;+\; k\log Z\;, \tag{3.1.14}$$

which can be rewritten as

$$S \;=\; \left(\frac{\partial\, kT\log Z}{\partial T}\right)_{\mu,V}\;. \tag{3.1.15}$$

A similar argument to the one given earlier in section 1.3 shows that two systems placed in thermal and diffusive contact (i.e., they can exchange particles through, say, a permeable membrane) will reach a thermal equilibrium characterized by a common value of β, μ. The argument is simply that when the systems are **separate** there are six Lagrange multipliers leading to equilibrium characterized by $(\alpha_1, \beta_1, \mu_1)$, $(\alpha_2, \beta_2, \mu_2)$, but that when they are in contact there are only three Lagrange multipliers and hence common values for these two sets.

The first law states that $dE = TdS - PdV + \mu dN$ and so E can be viewed as a function of (S, V, N) , and when E, S, V and N are extensive (as is the case for many large systems) we get

$$\begin{aligned} E(\lambda S, \lambda V, \lambda N) \;&=\; \lambda E(S, V, N) \\ \Longrightarrow \quad E \;&=\; S\left(\frac{\partial E}{\partial S}\right)_{V,N} \;+\; V\left(\frac{\partial E}{\partial V}\right)_{S,N} \;+\; N\left(\frac{\partial E}{\partial N}\right)_{S,V} \\ \Longrightarrow \quad E \;&=\; TS \;-\; PV \;+\; \mu N\;. \end{aligned} \tag{3.1.16}$$

We define the **grand potential** Ω for a state of thermodynamic equilibrium by the Legendre transformation

$$\begin{aligned} \Omega \;&=\; E \;-\; TS \;-\; \mu N \\ &=\; E \;-\; \mu N \;-\; (E-\mu N \;+\; kT\log Z) \\ &=\; -kT\log Z\;, \qquad \Longrightarrow \qquad Z \;=\; e^{-\beta\Omega}\;. \end{aligned} \tag{3.1.17}$$

We also have that

$$d\Omega \;=\; -\,SdT - Nd\mu - PdV \;, \tag{3.1.18}$$

which means $\Omega \;=\; \Omega(\mu, T, V)$.

In cases where Ω is extensive we get from (3.1.16) that $\Omega \;=\; -\,PV$. Hence, we find that in these cases

$$\boxed{PV \;=\; kT \log Z \;. \qquad (3.1.19)}$$

This allows the calculation of the equation of state from the grand partition function.

Note, that as for fluctuations in the energy (1.2.12), we find

$$(\Delta N)^2 \;=\; \frac{1}{\beta}\left(\frac{\partial N}{\partial \mu}\right)_\beta \;\sim\; N. \tag{3.1.20}$$

Thus

$$\frac{|\Delta N|}{N} \;\sim\; N^{-\frac{1}{2}} \;. \tag{3.1.21}$$

Although, the particle number is no longer **fixed** to be N, the effect of fluctuations around the most probable ensemble vanish like $N^{-\frac{1}{2}}$. In all practical circumstances where N is very large we should therefore expect that the results from the grand canonical and canonical ensembles are indistinguishable. However, we have the advantage in the grand ensemble that N can be made a function of external variable under the experimenters control: μ, β. We see also that the equation of state is given explicitly in terms of the grand partition function.

3.2 Systems of Non-interacting Identical Particles

Suppose the system consists of N identical particles. Then labelling them $1, 2, \ldots$etc. wrongly distinguishes them. For example, there is no way of distinguishing

$$\psi(\boldsymbol{x}_1, \alpha_1, \boldsymbol{x}_2, \alpha_2, t) \qquad \text{from} \qquad \psi(\boldsymbol{x}_2, \alpha_2, \boldsymbol{x}_1, \alpha_1, t) \;, \tag{3.2.1}$$

experimentally. They must be the **same** state with the **same** energy. They may differ by a phase since this does not affect probabilities or energies. Somehow we must construct the states to take into account the indistinguishability of the particles. This is done in one of the two following ways which have far reaching consequences for most physical systems:

(i) **Fermi-Dirac (FD) statistics** states that the wavefunction must be **antisymmetric** under interchange of any two particles. In the example above this means $1 \leftrightarrow 2$ with a phase factor of -1. Particles which obey Fermi-Dirac statistics are called **fermions** and include all spin-1/2 particles and in particular the electron.

(ii) **Bose-Einstein (BE) statistics** states that the wavefunction for identical particles is **symmetric** under interchange of any pair of particles. Particles which obey Bose-Einstein statistics are called **bosons** which includes the photon.

Thus we have:

$$\begin{aligned} \text{Fermi-Dirac:} \qquad \psi(\boldsymbol{x}_1,\alpha_1,\boldsymbol{x}_2,\alpha_2,t) &= -\,\psi(\boldsymbol{x}_2,\alpha_2,\boldsymbol{x}_1,\alpha_1,t)\ , \\ \text{Bose-Einstein:} \qquad \psi(\boldsymbol{x}_1,\alpha_1,\boldsymbol{x}_2,\alpha_2,t) &= +\,\psi(\boldsymbol{x}_2,\alpha_2,\boldsymbol{x}_1,\alpha_1,t)\ . \end{aligned} \tag{3.2.2}$$

In the separable case for two particles

$$H \;=\; -\,\frac{\hbar^2}{2m}\nabla_1^2 \;-\; \frac{\hbar^2}{2m}\nabla_2^2 \;+\; V(\boldsymbol{x}_1) \;+\; V(\boldsymbol{x}_2)\ , \tag{3.2.3}$$

and the wavefunction becomes

$$\Psi \;=\; \sqrt{\tfrac{1}{2}}\,\left(\psi_1(\boldsymbol{x}_1,\alpha_1,t)\psi_2(\boldsymbol{x}_2,\alpha_2,t) \;\mp\; \psi_1(\boldsymbol{x}_2,\alpha_2,t)\psi_2(\boldsymbol{x}_1,\alpha_1,t)\right)\ , \tag{3.2.4}$$

with $(-)$ for Fermi-Dirac and $(+)$ for Bose-Einstein.

Consider two electrons and neglect the interaction between them. The Hamiltonian is then separable. Put one in single particle state $\psi_1(\boldsymbol{x},s,t)$ and the other in $\psi_2(\boldsymbol{x},s,t)$. The total energy is $E \;=\; E_1+E_2$ and the correct wavefunction is

$$\Psi(\boldsymbol{x}_1,s_1,\boldsymbol{x}_2,s_2,t) \;=\; \sqrt{\tfrac{1}{2}}\,\left(\psi_1(\boldsymbol{x}_1,s_1,t)\psi_2(\boldsymbol{x}_2,s_2,t) \;-\; \psi_1(\boldsymbol{x}_2,s_2,t)\psi_2(\boldsymbol{x}_1,s_1,t)\right)\ . \tag{3.2.5}$$

If $\psi_1 \;=\; \psi_2$ then $\Psi \;=\; 0$. This is the **Exclusion Principle** which states that

"No two electrons may occupy identical single particle states."

For a system of N electrons with separable Hamiltonian the energy eigenfunctions take the form

$$\Psi(\boldsymbol{x}_1,s_1,\ldots,\boldsymbol{x}_N,s_N,t) \;=\; \Phi(\boldsymbol{x}_1,s_1,\ldots,\boldsymbol{x}_N,s_N,t)\,e^{-iEt/\hbar}\ , \tag{3.2.6}$$

where Φ can be written most conveniently as

$$\Phi \;=\; \sqrt{\frac{1}{N!}}\;\left|\begin{array}{cccc} \phi_1(\boldsymbol{x}_1,s_1) & \phi_1(\boldsymbol{x}_2,s_2) & \ldots\ldots & \phi_1(\boldsymbol{x}_N,s_N) \\ \phi_2(\boldsymbol{x}_1,s_1) & \phi_2(\boldsymbol{x}_2,s_2) & \ldots\ldots & \phi_2(\boldsymbol{x}_N,s_N) \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \phi_N(\boldsymbol{x}_1,s_1) & \phi_N(\boldsymbol{x}_2,s_2) & \ldots\ldots & \phi_N(\boldsymbol{x}_N,s_N) \end{array}\right|\ , \tag{3.2.7}$$

with $E \;=\; \sum_{i=1}^{N} E_i$. Fermi-Dirac statistics are automatically satisfied.

This is the **Slater determinant** and eigenfunctions of this kind form the complete basis set of functions for all atomic physics calculations.

E.g., for $N=3$ we have $E=E_1+E_2+E_3$ and

$$\Phi(\boldsymbol{x}_1,s_1,\boldsymbol{x}_2,s_2,\boldsymbol{x}_3,s_3) \;=\; \sqrt{\frac{1}{6}}\;\left|\begin{array}{ccc} \phi_1(\boldsymbol{x}_1,s_1) & \phi_1(\boldsymbol{x}_2,s_2) & \phi_1(\boldsymbol{x}_3,s_3) \\ \phi_2(\boldsymbol{x}_1,s_1) & \phi_2(\boldsymbol{x}_2,s_2) & \phi_2(\boldsymbol{x}_3,s_3) \\ \phi_3(\boldsymbol{x}_1,s_1) & \phi_3(\boldsymbol{x}_2,s_2) & \phi_3(\boldsymbol{x}_3,s_3) \end{array}\right|\ , \tag{3.2.8}$$

In the case that the Hamiltonian does not depend on the spin of the electron the energy eigenstates which differ just by spin orientation have the **same** energy and **same** spatial wavefunction. In this circumstance we may ignore the presence of spin in many of our calculations except to remember that any **given** state may be occupied by at most two electrons: one with spin "up" $(+1/2)$ and one with spin "down" $(-1/2)$.

3.2.1 the partition functions

$\mathcal{S}$ consists of non-interacting identical particles. Let the one particle states of $\mathcal{S}$ be $\phi_r(\boldsymbol{x})$ with energies ϵ_r. Note that our notation is that r stands for the set of quantum numbers that uniquely labels the state. Suppose that in microstate $|i\rangle$, n_r of the particles have the wavefunction ϕ_r. Then

$$N_i \;=\; \sum_r n_r\ , \qquad \text{and} \qquad E_i \;=\; \sum_r n_r\epsilon_r\ . \tag{3.2.1.1}$$

We see that $|i\rangle$ is completely determined by the set $\{n_r\}$. We obtain the full set of microstates by allowing the n_r to vary without restriction over their allowed range of values. Thus we cannot **fix** N_i or E_i directly – this is why we use the grand ensemble method. The particular values of N and E for a given system are obtained through choosing the appropriate values of μ and β. This is a strength of the approach.

We then write the grand partition function as

$$\begin{aligned} Z &= \sum_i e^{-\beta(E_i-\mu N_i)} \\ &= \sum_{n_1,n_2,\ldots} e^{-\beta(n_1\epsilon_1+n_2\epsilon_2+\ldots-\mu n_1-\mu n_2-\ldots)} \\ &= \prod_r \sum_{n_r} e^{-\beta n_r(\epsilon_r-\mu)}\ . \end{aligned} \tag{3.2.1.2}$$

For **fermions** we have $n_r = 0,1$ only. Hence,

$$Z \;=\; \prod_r \left(1 \;+\; e^{-\beta(\epsilon_r-\mu)}\right)\ . \tag{3.2.1.3}$$

Then we find

$$N \;=\; \frac{1}{\beta}\left(\frac{\partial \log Z}{\partial \mu}\right)_{\beta,V} \;=\; \sum_r \frac{1}{e^{\beta(\epsilon_r-\mu)}+1}\ . \tag{3.2.1.4}$$

Also,

$$\begin{aligned} \hat{n}_r \;\equiv\; \bar{n}_r &= \frac{\sum_{n_1,n_2,\ldots} n_r e^{-\beta\sum_s n_s(\epsilon_s-\mu)}}{Z} \\ &= -\frac{1}{\beta}\left(\frac{\partial \log Z}{\partial \epsilon_r}\right)_{\beta,\mu,\epsilon_s} \qquad s\neq r \end{aligned} \tag{3.2.1.5}$$

$$\Longrightarrow$$

$$\boxed{\bar{n}_r \;=\; \frac{1}{e^{\beta(\epsilon_r-\mu)}+1}\ , \qquad \textbf{Fermi-Dirac distribution.}} \tag{3.2.1.6}$$

We see that $N = \sum_r \bar{n}_r$ where $\bar{n}_r$ is the **mean occupation number** of the r-th single particle state in thermal equilibrium.

For **bosons** the sum on the n_r is unrestricted: $0 \le n_r < \infty$ and

$$Z \;=\; \prod_r \frac{1}{1-e^{-\beta(\epsilon_r-\mu)}}\ . \tag{3.2.1.7}$$

The same formula (3.2.1.5) applies to calculate $\bar{n}$ and we get

$$\boxed{\bar{n}_r \;=\; \frac{1}{e^{\beta(\epsilon_r-\mu)}-1}\ , \qquad \textbf{Bose-Einstein distribution}} \tag{3.2.1.8}$$

If there are g_r states of energy ϵ_r then the average number of particles with energy ϵ_r is

$$\bar{n}(\epsilon_r) \;=\; \frac{g_r}{e^{\beta(\epsilon_r-\mu)} \pm 1} \;. \tag{3.2.1.9}$$

For large V we can, as before, pass to the continuum limit for free particles confined to a box and we have

$$\sum_r \;\rightarrow\; \int d\epsilon g(\epsilon) \;. \tag{3.2.1.10}$$

The general result for $g(\epsilon)$ was given in (1.7.1.7). Here we remember that in D=3

$$g(\epsilon) \;=\; \begin{cases} g_s \frac{V}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \epsilon^{\frac{1}{2}} \;\equiv\; KV\epsilon^{\frac{1}{2}} \;, & \text{non-relativistic particles} \\ \\ g_s V \frac{\epsilon^2}{2\pi^2\hbar^3c^3} & \text{massless particles e.g., photons,} \end{cases} \tag{3.2.1.11}$$

where g_s is the spin/polarization degeneracy and $K = \frac{g_s}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}$.

The average number of particles in the range $\epsilon \rightarrow \epsilon + d\epsilon$ is

$$n(\epsilon)d\epsilon \;=\; \frac{g(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1} \;. \tag{3.2.1.12}$$

From the definitions for Z in (3.2.1.3) and (3.2.1.7), we find in the continuum limit that the grand partition function is given by

$$\log Z \;=\; \pm \int_0^\infty d\epsilon\; g(\epsilon) \log(1 \pm e^{-\beta(\epsilon-\mu)}) \;, \tag{3.2.1.13}$$

with + for fermions and − for bosons. We can now use the results of section 3.1 to derive expressions for N and E. Of course, these expressions can be straightforwardly derived from the appropriate averages using (3.2.1.12). We find

$$N \;=\; \int_0^\infty \frac{g(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1} \;, \qquad E \;=\; \int_0^\infty \frac{\epsilon\, g(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1} \;. \tag{3.2.1.14}$$

The expression for N determines μ **in terms of** N. Generally this is complicated but μ must be chosen (as must β) to realize the physical value of N. For bosons (− sign) we must have $\mu \leq 0$ otherwise the integrand is singular. For fermions there is no restriction on μ but we shall see that in many applications $\mu > 0$ is natural.

For non-relativistic particles we get

$$\log Z \;=\; \pm KV \int_0^\infty d\epsilon\, \epsilon^{\frac{1}{2}} \log(1 \pm e^{-\beta(\epsilon-\mu)})$$

by parts we find

$$= \; KV \int_0^\infty \tfrac{2}{3}\beta \frac{\epsilon^{3/2}\, d\epsilon}{e^{\beta(\epsilon_r-\mu)} \pm 1}$$

$$\Longrightarrow$$

$$\log Z \;=\; \tfrac{2}{3}\beta E \;. \tag{3.2.1.15}$$

Combining this with the equation of state formula (3.1.19) we get $PV = \frac{2}{3}E$ – remember (1.4.3). The 2/3 comes from the $\epsilon^{1/2}$ factor. In D dimensions the exponent is $(D-2)/2$, see (1.7.1.7), and so we find $PV = (2/D)E$ in general.

3.2.2 the classical limit

When the number density is low and/or the temperature is very high we would not expect that there would be more than one particle at most in any given quantum state. In this case the restrictions because of particles being identical should become irrelevant and we should recover the classical results. Now

$$N = \int_0^\infty \frac{g(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1} \, , \qquad E = \int_0^\infty \frac{\epsilon\, g(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)} \pm 1} \, , \tag{3.2.2.1}$$

and, putting $z = \beta\epsilon$ and $y = -\beta\mu$, we find $N = KV(kT)^{3/2}I_{1/2}(-\beta\mu)$ and $E = KV(kT)^{5/2}I_{3/2}(-\beta\mu)$, where

$$I_n(y) = \int_0^\infty \frac{z^n\, dz}{e^{z+y} \pm 1} \, . \tag{3.2.2.2}$$

When $e^y \gg 1$ we can ignore the 1 in the denominator. We then find that

$$n(\epsilon) \sim e^{\beta\mu}e^{-\beta\epsilon}g(\epsilon) \quad \forall\epsilon \, , \tag{3.2.2.3}$$

which is the classical result for **all** values of ϵ since μ is just chosen to impose that the total number of particles is $N = \int n(\epsilon)d\epsilon$. Note that for ϵ such that $\beta(\epsilon-\mu) \gg 1$ the classical approximation will also hold. This is true, for example, for the conduction electrons in a semi-conductor **even though** e^y is not large.

Under what circumstances is e^y large? Suppose this to be the case, then

$$\frac{1}{e^{z+y} \pm 1} \sim e^{-(z+y)}\left(1 \mp e^{-(z+y)} + e^{-2(z+y)} \mp \ldots\right) \, . \tag{3.2.2.4}$$

Substituting in the expression for N above gives

$$N = KV(kT)^{3/2}e^{\beta\mu}\left[\underbrace{\int_0^\infty dz\, z^{1/2}e^{-z}}_{\sqrt{\pi}/2} \mp e^{\beta\mu}\underbrace{\int_0^\infty dz\, z^{1/2}e^{-2z}}_{\sqrt{\pi}/4\sqrt{2}} + \ldots\right] \Longrightarrow$$

$$N = KV\frac{\sqrt{\pi}}{2}(kT)^{\frac{3}{2}}e^{\beta\mu}\left(1 \mp \frac{e^{\beta\mu}}{2\sqrt{2}} + \ldots\right) \, . \tag{3.2.2.5}$$

The condition that we require is then

$$e^{-\beta\mu} \gg 1 \qquad \Longrightarrow \qquad \left(\frac{2m\pi kT}{h^2}\right)^{\frac{3}{2}}\frac{V}{N} \gg 1 \, . \tag{3.2.2.6}$$

This means large T and low density N/V . These conditions hold for most real gases under normal conditions and most gases liquefy before quantum-mechanical effects become significant. Of course, in order to liquefy, the gas deviates from ideal behaviour at low enough temperatures since liquefication is due to attractive intermolecular forces. The density of states calculation must be modified to account for these forces. An exception is helium for which these forces are small and so the liquefication temperature is very low.

By similar means the energy of the gas is then

$$E \sim \frac{3}{2}NkT\left(1 \pm \frac{e^{\beta\mu}}{4\sqrt{2}} + \ldots\right) \, . \tag{3.2.2.7}$$

We have $PV = \frac{2}{3}E$ and so the lowest order correction to the equation of state is

$$PV \;=\; NkT \;\pm\; \frac{NP}{(kT)^{\frac{3}{2}}}\frac{1}{K\sqrt{8\pi}} \;. \qquad (3.2.2.8)$$

3.3 Black-Body Radiation

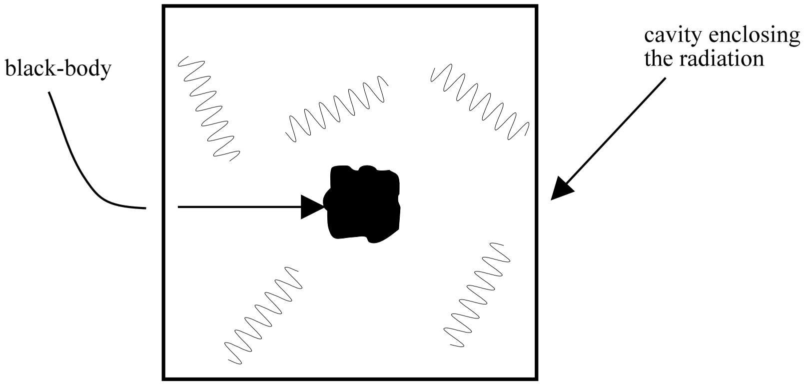
Any body with temperature $T > 0$ emits radiation in the form of photons. For example:

(i) a light bulb filament is heated by the electric current to a high temperature and it glows;

(ii) iron heated in a forge glows red hot meaning that it emits visible red light. As the iron gets hotter the colour changes toward the blue end of the spectrum, i.e., there are more photons of shorter wavelength than before, which means they have higher energy;

(iii) bodies at room temperature typically emit infra-red radiation, e.g., people. This is often referred to as "heat" but really it is only distinguished by being of longer wavelength than visible light.

Consider radiation trapped in a box with perfectly reflecting walls. The different energy states for the photons will each be occupied by different numbers of photons. Suppose, a body of temperature T is introduced into the box.



This body will emit and absorb photons and so change the occupation numbers of the levels. The box will thus contain radiation in the form of many photons of varying frequency which are being continually absorbed and emitted by the body, and these photons will come into equilibrium with it and so will conform to the **Bose-Einstein distribution**. Now, the number of photons is not constrained since they are continually being emitted from (or absorbed by) the body, and hence there is **no need** for the Lagrange multiplier γ in the derivation in section (3.1) above. The effect is to set $\mu = 0$ for photons.

The body is a **black-body** if it can absorb and emit radiation of **all** frequencies (i.e. energies) across a wide spectral range. If it is not a black-body then photons of some frequencies **cannot** be absorbed or emitted by the body and so their number will be fixed at some non-equilibrium value. The corollary is that if radiation is in contact with a black-body at temperature T, then photons of **all** frequencies will conform to the Bose-Einstein distribution with $\mu = 0$. Thus:

Radiation in equilibrium with a **black-body** at temperature T is characterized **only** by the temperature T. Such radiation is called **black-body radiation**.

3.3.1 The Black-Body Distribution

We use the relativistic formula for the density of states (3.2.1.11) with $\epsilon = \hbar\omega$. Each photon has two polarization states, $g_s = 2$, and so the total number of states in $(\omega, \omega + d\omega)$ is:

$$g(\omega)d\omega \;=\; \frac{\omega^2 V}{\pi^2 c^3} d\omega \; . \tag{3.3.1.1}$$

where $g(\omega)$ is the density of states.

Thus the number density of photons, $N(\omega)$, is given using the Bose-Einstein distribution to be $(\mu = 0)$:

$$N(\omega) \;=\; g(\omega)\frac{1}{e^{\beta\hbar\omega} \;-\; 1} \; , \tag{3.3.1.2}$$

which, using eqn. (3.3.1.1), becomes

$$\boxed{N(\omega) \;=\; \frac{V\omega^2}{\pi^2 c^3}\,\frac{1}{e^{\beta\hbar\omega} - 1} \qquad \textbf{Black-body radiation distribution}.} \tag{3.3.1.3}$$

(i) The **energy density**, $E(\omega)$, is given by

$$\begin{aligned} E(\omega) \;&=\; \hbar\omega\, N(\omega) \\ &=\; \frac{V\hbar\omega^3}{\pi^2 c^3}\,\frac{1}{e^{\beta\hbar\omega} - 1} \end{aligned} \tag{3.3.1.4}$$

This is the **Planck radiation law**. For high T or small ω we can approximate

$$e^{\beta\hbar\omega} - 1 \;\sim\; \beta\hbar\omega \qquad \Longrightarrow \qquad E(\omega) \;\sim\; g(\omega)kT \; . \tag{3.3.1.5}$$

This is the classical equipartition of energy – $\frac{1}{2}kT$ per degree of freedom. Since $g(\omega) \sim \omega^2$ the **total** energy, $E = \int E(\omega)d\omega$, diverges if we use this result. This is the **ultraviolet** catastrophe of the classical approach which treats the photon gas as ideal.

(ii) The total energy is then

$$\begin{aligned} E(T) \;&=\; \int_0^\infty E(\omega)d\omega \\ &=\; \frac{V}{\hbar^3 c^3 \pi^2}\int_0^\infty (\hbar\omega)^3\,\frac{1}{e^{\beta\hbar\omega} - 1}\,d(\hbar\omega) \; . \end{aligned} \tag{3.3.1.6}$$

Setting $x = \beta\hbar\omega$ we get

$$E(T) \;=\; \frac{V}{\beta^4}\frac{1}{\hbar^3 c^3 \pi^2}\int_0^\infty x^3 \frac{1}{e^x - 1}\,dx \; .$$

Hence, (we look up the integral, it is $\pi^4/15$)

$$E(T) \;=\; CVT^4 \; , \qquad C \;=\; \frac{\pi^2 k}{15}\left(\frac{k}{\hbar c}\right)^3 \; . \tag{3.3.1.7}$$

(iii) The energy flux, $\mathcal{E}(T)$, radiated from a black-body is defined as the energy per second leaving a small hole of unit area in the wall of the box, assuming no inward flux. The speed of the photons is c, and the number in interval $(\omega, \omega + d\omega)$ crossing unit area per second is

$$df(\omega) \;=\; \frac{1}{V}\frac{c}{4}N(\omega)\,d\omega \;,$$

where the factor 1/4 comes from an angular integration over velocity directions. Thus,

$$\mathcal{E}(T) \;=\; \frac{1}{4}\frac{E(T)}{V}c \;,$$

or

$$\boxed{\mathcal{E}(T) \;=\; \sigma T^4 \;, \qquad \sigma \;=\; \frac{\pi^2 kc}{60}\left(\frac{k}{\hbar c}\right)^3 \qquad \textbf{Stefan-Boltzmann law}. \qquad (3.3.1.8)}$$

Here $\sigma = 5.67 \cdot 10^{-8} Js^{-1}m^{-2}K^{-4}$ is Stefan's constant.

From (3.1.14) with $\mu = 0$

$$S \;=\; k\log Z \;+\; \frac{E}{T}\;. \tag{3.3.1.9}$$

From (3.2.1.13) we have

$$\log Z \;=\; -\int_0^\infty d\omega\; g(\omega)\log\left(1 - e^{-\beta\hbar\omega}\right)\;. \tag{3.3.1.10}$$

Using (3.3.1.1), $g(\omega) = BV\omega^2$ and by parts we find $k\log Z = E/3T$ and hence

$$S \;=\; \frac{4E}{3T} \;=\; \frac{4V}{3}CT^3\;. \tag{3.3.1.11}$$

Then

$$E \;=\; \left(\frac{3S}{4}\right)^{\frac{4}{3}}\left(\frac{1}{VC}\right)^{\frac{1}{3}} \;\sim\; V^{-\frac{1}{3}} \qquad \text{at constant entropy}\;. \tag{3.3.1.12}$$

Note: $|\boldsymbol{k}| \propto L^{-1}, \;\; E = c|\boldsymbol{k}| \Rightarrow E \propto V^{-1/3}$.

The free energy is

$$F \;=\; E \;-\; TS \;=\; -\tfrac{1}{3}E\;. \tag{3.3.1.13}$$

The radiation pressure, P, can be calculated from

$$P \;=\; -\left(\frac{\partial F}{\partial V}\right)_T \;=\; \tfrac{1}{3}\frac{E}{V} \;=\; \tfrac{1}{3}CT^4\;. \tag{3.3.1.14}$$

(Remember $F = F(T,V)$ and $E \propto V$ since it is extensive.)

By similar calculation the number density of photons is

$$n \;=\; \frac{N}{V} \;\propto\; T^3 \qquad \text{(see example sheet)}. \tag{3.3.1.15}$$

3.4 Debye theory of specific heat

3.4.1 the harmonic oscillator

We begin by discussing a single quantum harmonic oscillator of frequency ω at temperature T. Think of it as being in thermal contact with a heat bath. The partition function is then

$$z(\omega) \;=\; \sum_n e^{-\beta(n+\frac{1}{2})\hbar\omega} \quad=\quad \frac{e^{-\frac{1}{2}\beta\hbar\omega}}{1-e^{-\beta\hbar\omega}} \quad=\quad \frac{1}{2\mathrm{sinh}\,(\frac{1}{2}\beta\hbar\omega)}. \qquad (3.4.1.1)$$

Then

$$E \;=\; -\,\frac{\partial \log z}{\partial\beta} \;=\; \tfrac{1}{2}\hbar\omega \;+\; \frac{\hbar\omega}{e^{\beta\hbar\omega}-1}\;. \qquad (3.4.1.2)$$

The average excitation number $\bar{n}$ is

$$\bar{n} \;=\; \frac{E-\frac{1}{2}\hbar\omega}{\hbar\omega} \;=\; \frac{1}{e^{\beta\hbar\omega}-1}\;. \qquad (3.4.1.3)$$

This is the Bose-Einstein distribution if we interpret each excitation of $\hbar\omega$ as putting another boson into an energy level $\epsilon = \hbar\omega$. In otherwords, the excitations of the oscillator act like bosons for all practical purposes. (This is how the quantum field theory of bose particles works – it starts from the oscillator.)

For $\hbar\omega \ll kT$

$$E \;\sim\; kT\;, \qquad (3.4.1.4)$$

equipartition of energy for two degrees of freedom. [$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2x^2$ has two degrees of freedom classically – p and x.]

Clearly, for $\hbar\omega \gg kT$ we recover classical Boltzmann statistics since the "1" in the denominator can be ignored.

3.4.2 vibrational specific heat of a solid

We shall consider the contribution to the specific heat of a solid from its elastic or vibrational modes. Other contributions come from the electrons which, in a metal, behave like a gas and also from internal atomic degrees of freedom such as rotational modes.

The long wavelength elastic modes are sound waves of which there are three kinds – two where the displacement is transverse to the velocity and one where it is longitudinal. We shall not distinguish them and shall assume that they all have the same velocity c_s which is constant and thus independent of frequency. To a good approximation these modes are harmonic and each sound wave is viewed as an excitation of the corresponding quantum oscillator. It is labelled by its frequency and wavevector $(\omega, \boldsymbol{k})$ with $\omega = c_s k$ ($k = |\boldsymbol{k}|$), where the wave solution is identical to that already given in (1.7.1.3):

$$\psi_{\boldsymbol{k}}(\boldsymbol{x}) \;=\; \frac{1}{\sqrt{V}}e^{i\boldsymbol{k}\cdot\boldsymbol{x}} \quad \text{with} \quad \boldsymbol{k} \;=\; \frac{2\pi}{L}\boldsymbol{n}\;,\quad \boldsymbol{n}\in\mathcal{Z}^3\;, \qquad (3.4.2.1)$$

$$\boldsymbol{p} \;=\; \hbar\boldsymbol{k}\;, \qquad \epsilon(\boldsymbol{p}) \;=\; \hbar\omega \;=\; pc_s\;. \qquad (3.4.2.2)$$

This is the same dispersion relation as for massless particles (e.g., photons) and in the alternative description of the previous section, each excitation of the oscillator can be viewed as a boson of energy $\hbar\omega$ with momentum $\hbar\boldsymbol{k}$. These "vibrational" bosons are called **phonons**, and the elastic excitations of the solid can be described as a gas of free phonons.

The density of modes is given by the formula for massless particles with $g_s = 3$

$$g(\omega) \;=\; \frac{3V\omega^2}{2\pi^2 c_s^3}\;. \qquad (3.4.2.3)$$

The solid contains N atoms and so there are $3N$ oscillator degrees of freedom. This gives a maximum frequency $\omega_{\rm max}$ determined by

$$\int_0^{\omega_{\rm max}} d\omega\; g(\omega) \;=\; 3N \qquad \Longrightarrow \qquad \omega_{\rm max} \;=\; \left(\frac{6\pi^2 N}{V}\right)^{\frac{1}{3}} c_s\;. \qquad (3.4.2.4)$$

Thus $\omega_{\rm max}$ gives an **ultraviolet cutoff** on the energy: $\hbar\omega_{\rm max}$.

Calculations are very similar to those of blackbody radiation. In thermal equilibrium at temperature T we have the Planck distribution

$$E \;=\; \int_0^{\omega_{\rm max}} d\omega\; \frac{\hbar\omega\, g(\omega)}{e^{\beta\hbar\omega} - 1}\;. \qquad (3.4.2.5)$$

Note: E is **relative** to the zero–point energy $3N\hbar\omega/2$ which is **unobservable** and can be subtracted off.

Define the **Debye temperature** by $kT_D = \hbar\omega_{\rm max}$ and set $x = \hbar\omega/kT$ to get

$$\begin{aligned} E \;&=\; \frac{3V}{2\pi^2(\hbar c_s)^3}\,(kT)^4 \int_0^{T_D/T} dx\; \frac{x^3}{e^x - 1} \\ &=\; 3NkT\, D\left(\frac{T_D}{T}\right)\;, \end{aligned} \qquad (3.4.2.6)$$

where $D(z)$ is the Debye function

$$D(z) \;=\; \frac{3}{z^3}\int_0^z dx\; \frac{x^3}{e^x - 1}\;. \qquad (3.4.2.7)$$

Now

(i) $T \gg T_D\;, \quad \Longrightarrow \quad z \ll 1$

$$D(z) \;=\; 1 \;-\; \frac{3z}{8} \;+\; O(z^2)\;. \qquad (3.4.2.8)$$

Then

$$E \;\sim\; 3NkT \qquad \Longrightarrow \qquad C_V \;\sim\; 3Nk\;. \qquad (3.4.2.9)$$

This is the classical regime – note there is no ultraviolet catastrophe since there is an explicit ultraviolet cutoff $\omega_{\rm max}$.

(ii) $T \ll T_D\;, \quad \Longrightarrow \quad z \gg 1$

$$D(z) \;=\; \frac{\pi^4}{5z^3},\; z \gg 1. \qquad (3.4.2.10)$$

then

$$E \;\sim\; \frac{3\pi^4 NkT^4}{5T_D^3} \qquad \Longrightarrow \qquad C_V \;\sim\; \frac{12\pi^4}{5}Nk\left(\frac{T}{T_D}\right)^3 \;=\; \frac{2\pi^2 Vk^4T^3}{5(\hbar c_s)^3}\;. \qquad (3.4.2.11)$$

This is the quantum regime.

As in most cases this theory is very approximate since the vibrational properties of solids are much more complex in general.

3.5 Bose-Einstein Condensation

For a free boson gas the chemical potential μ is given implicitly in terms of N and T via

$$\begin{aligned} N &= KV \int_0^\infty d\epsilon\, \frac{\epsilon^{\frac{1}{2}}}{e^{\beta(\epsilon-\mu)}-1} \\ &= KV(kT)^{\frac{3}{2}} \int_0^\infty dx\, \frac{x^{\frac{1}{2}}}{e^{(x-\beta\mu)}-1} \\ &= KV(kT)^{\frac{3}{2}} \frac{\sqrt{\pi}}{2} f(-\beta\mu) \ , \end{aligned} \tag{3.5.1}$$

where

$$f(-\beta\mu) = \sum_{n=1}^{\infty} \frac{e^{n\beta\mu}}{n^{\frac{3}{2}}} \ . \tag{3.5.2}$$

[This is derived by writing the integral as

$$\int_0^\infty dx\, x^{1/2} \sum_1^\infty e^{-nx}\, e^{n\beta\mu} \ ,$$

and using

$$\int_0^\infty dx\, x^{1/2}\, e^{-nx} = \frac{\sqrt{\pi}}{2n^{\frac{3}{2}}} \ .$$

]

This is a power series expansion in the **fugacity** $\kappa = e^{\beta\mu}$, and it converges for all $\beta\mu < 0$. The maximum allowed value of μ is thus $\mu = 0$, for which case the denominator in the BE distribution vanishes at $\epsilon = 0$ but the singularity is integrable. It is easy to see that f has a maximum at $\beta\mu = 0$ and decreases monotonically as $\beta\mu$ decreases from 0.

Now suppose the density, N/V, is fixed. As T decreases both sides of the equation for N, (3.5.1), balance since f can **increase** by **decreasing** $|\beta\mu|$ (remember $\mu < 0$). However, this is accomplished by decreasing $|\mu|$ since β is increasing. Thus there must be some $T = T_C$ at which $\mu = 0$ and f has reached its maximum value. A further decrease in T cannot be accommodated by the equation and we seemingly have a problem with the approach. From (3.5.1) T_C is given by

$$\frac{N}{V} = \hat{K}(kT_C)^{\frac{3}{2}} \ , \qquad \hat{K} = \frac{K\sqrt{\pi}}{2} f(0) \ . \tag{3.5.3}$$

Note that $f(0) = \zeta(3/2) = 2.612$.

Why does the equation for N fail for $T < T_C$? The reason is that we must take more care in passing to the continuum limit when calculating the density of states. The $\epsilon^{\frac{1}{2}}$ factor in $g(\epsilon)$ assigns zero weight to the $\epsilon = 0$ state. There is only one state with $\epsilon = 0$ however large V is, whereas for $\epsilon > 0$ the number of states is proportional to V. Thus, in the limit $V \to \infty$ this state at $\epsilon = 0$ is neglected in the approximation

$$\sum_r \to \int d\epsilon\, g(\epsilon) \ . \tag{3.5.4}$$

For fermions this is always good since no more than g_s fermions can occupy the $\epsilon = 0$ state. However, for bosons the number that can occupy this state is unrestricted. The resolution of the problem for $T < T_C$ is to write $N = N_0 + N_C$ where there are N_0 particles in the $\epsilon = 0$ state and N_C in the $\epsilon > 0$ states. The total energy, E, receives **no** contribution from the N_0 particles with $\epsilon = 0$. For $T < T_C$ we have (using (3.5.3)) and $\mu = 0$

$$\begin{aligned} \frac{N_C}{V} &= \hat{K}(kT)^{\frac{3}{2}} , \\ \frac{N_C}{N} &= \left(\frac{T}{T_C}\right)^{\frac{3}{2}} \quad \Longrightarrow \quad \frac{N_0}{N} = \left[1 - \left(\frac{T}{T_C}\right)^{\frac{3}{2}}\right] . \end{aligned} \tag{3.5.5}$$

This phenomenon is called **Bose-Einstein condensation** and is a phase transition between the phase where $N_0 = 0$ and the phase where $N_0 > 0$. Formulae of this kind are typical a such transitions (c.f. the Curie point separating the ferromagnetic and paramagnetic phases of iron). BE condensation occurs in nature – for example, superconductivity and superfluidity are related effects. However, it is only very recently (search the web for "bose condensation") that low temperature experiments have been done with bosonic-type atoms and the effect directly observed.

For $T < T_C$ we have

$$E = KV \int_0^\infty d\epsilon \, \frac{\epsilon^{\frac{3}{2}}}{e^{\beta\epsilon} - 1} = KV(kT)^{\frac{5}{2}} \int_0^\infty dx \, \frac{x^{\frac{3}{2}}}{e^x - 1} . \tag{3.5.6}$$

As before, the integral is

$$\frac{3\sqrt{\pi}}{4}\zeta(5/2) \quad \text{using} \quad \sum_{n=1}^{\infty} \frac{1}{n^{\frac{5}{2}}} = \zeta(5/2) = 1.341 . \tag{3.5.7}$$

From (3.5.3) we have

$$KV = \frac{2}{\sqrt{\pi}\zeta(3/2)} \frac{N}{(kT_C)^{\frac{3}{2}}} , \tag{3.5.8}$$

and so

$$E = \underbrace{\frac{3\zeta(5/2)}{2\zeta(3/2)}}_{A} NkT \left(\frac{T}{T_C}\right)^{\frac{3}{2}} . \tag{3.5.9}$$

Using (3.5.5) we get

$$E = AN_C kT . \tag{3.5.10}$$

For $T < T_C$:

(i) The heat capacity is

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{5E}{2T} \propto T^{\frac{3}{2}} . \tag{3.5.11}$$

The specific heat gets a contribution from the condensed bosons since as T changes N_C changes, and so the energy is affected by particles entering or vacating the $\epsilon > 0$ states.

(ii) The entropy is

$$S = \int dT \, \frac{C_V}{T} = \frac{5E}{3T} . \tag{3.5.12}$$

(iii) The free energy is

$$F = E - TS = -\tfrac{2}{3}E . \tag{3.5.13}$$

(iv) The pressure is (remember $F = F(T, V)$)

$$P = -\left(\frac{\partial F}{\partial V}\right)_T = \frac{2E}{3V} \quad \Longrightarrow \quad PV = \tfrac{2}{3}E \ . \tag{3.5.14}$$

This follows since $E \propto V$ since it is extensive (see also (3.5.10)).

Finally we note that although C_V is continuous at $T = T_C$ there is a **discontinuity** in its derivative $(\partial C_V/\partial T)_V$. This is again typical of a phase transition. This is because $(\partial^2 \mu/\partial T^2)$ has a discontinuity $T = T_C$ since it turns out that

$$-\mu \sim (T - T_C)^2 \quad T = T_C + 0 \ , \quad \mu = 0 \quad T < T_C \ . \tag{3.5.15}$$

We have $E = E(V, T, \mu)$ and for given N we have

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{V,N} = \left(\frac{\partial E}{\partial T}\right)_{V,\mu} + \left(\frac{\partial \mu}{\partial T}\right)_{V,N}\left(\frac{\partial E}{\partial \mu}\right)_{V,T} \tag{3.5.16}$$

Clearly, $(\partial C_V/\partial T)_V$ will contain $(\partial^2 \mu/\partial T^2)$ and so will also be **discontinuous** at $T = T_C$.

4 The Degenerate Fermi Gas

For values of T and μ for which the Pauli exclusion principle dominates, the fermi gas is degenerate. It is in the extreme QM limit and cannot be described by a classical approach. Examples of systems where this is the case are

Electrons in metals at room temperature;
Electrons in white dwarf stars;
Neutrons in neutron stars;
Nucleons in nuclear matter.

For N electrons in volume V at temperature T:

$$\begin{aligned} N &= \int_0^\infty g(\epsilon) F(\epsilon)\, d\epsilon \ , \\ E &= \int_0^\infty \epsilon\, g(\epsilon) F(\epsilon)\, d\epsilon \ , \\ F(\epsilon) &= \frac{1}{e^{\beta(\epsilon-\mu)} + 1} \ . \end{aligned} \tag{4.1}$$

As discussed $e^{-\beta\mu} \gg 1$ gives the classical limit (e.g., high T, $\mu < 0$). We expect degenerate QM behaviour for $e^{\beta\mu} \gg 1$, i.e., low T, $\mu > 0$. Clearly, low T means $T \ll \mu$.

4.1 Properties at $T = 0$

At $T = 0$ write $\mu = E_F = kT_F$. Then as $T \to 0$:

$$\begin{aligned} e^{\beta(\epsilon-\mu)} &\longrightarrow \begin{cases} 0 & \text{for} \quad \epsilon < E_F \\ \infty & \text{for} \quad \epsilon > E_F \end{cases} \\ \Rightarrow \quad F(\epsilon) &\longrightarrow \begin{cases} 1 & \text{for} \quad \epsilon < E_F \\ 0 & \text{for} \quad \epsilon > E_F \end{cases} \end{aligned} \tag{4.1.1}$$

It can be shown (see later section) that $\mu = E_F + O(T^2)$

Thus at $T = 0$, $F(\epsilon) = \theta(E_F - \epsilon)$. This is a consequence of the Pauli principle since at $T = 0$ the electrons will fill the lowest available one-particle states (in pairs with spin ↑, ↓) until the Fermi energy E_F is reached. E_F is the highest energy of an occupied state at $T = 0$. Then

$$\begin{aligned} N &= \int_0^{E_F} KV\epsilon^{1/2}\, d\epsilon \;=\; \frac{2}{3}KVE_F^{3/2} \;, \\ E_F &= \left(\frac{3N}{2KV}\right)^{2/3} \;, \\ E &= \int_0^{E_F} KV\epsilon^{3/2}\, d\epsilon \;=\; \frac{2}{5}KVE_F^{5/2} \;=\; \frac{3}{5}NE_F \;\propto\; V^{-2/3} \;. \end{aligned} \tag{4.1.2}$$

The equation of state is then

$$PV \;=\; \tfrac{2}{3}E \quad\Longrightarrow\quad PV^{5/3} \;=\; \text{const} \;. \tag{4.1.3}$$

Classical theory gives $P \to 0$ as $T \to 0$, but the Pauli principle requires that electrons have non-zero momenta **even** at $T = 0$ and so P is non-zero. With all the factors we get for the pressure

$$P \;=\; \frac{\hbar^2}{5m}\left(3\pi^2\right)^{2/3}\left(\frac{N}{V}\right)^{5/3} \;. \tag{4.1.4}$$

So $P \propto 1/m$ and we see that the pressure at $T = 0$ is much greater for electrons than for protons.

4.2 White Dwarf Stars

When the hydrogen fuel of a star has been used up, the gravitational forces are no longer balanced by the pressure of the plasma and the star collapses. The matter becomes very dense and cold and the electrons form a degenerate fermi gas. For stars below a certain mass, known as the Chandrasakar limit, the collapse is halted when the pressure due to the degenerate electrons balances the gravitational pressure. The nucleons will also contribute but, as found in the previous section, their effect is negligible compared with that of the electrons. We can estimate the Chandrasakar limit by minimizing the total energy of the star with respect to its radius.

Since for relativistic particles $PV^{4/3} =$ const. we have $P \propto \rho^{4/3}$. [*] Using the equations for hydrodynamic equilibrium we find that the gravitational potential energy can easily be calculated and is given by

$$E_G \;=\; -\frac{3}{2}\frac{GM^2}{R} \;, \tag{4.2.1}$$

where $M = Nm_p$ is the mass of the star, N is the number of electrons and protons and m_p is the proton mass.

As the star collapses the value of E_F will rise and the temperature T is low enough that $T \ll T_F$, justifying the subsequent calculation using the degenerate $T = 0$ formalism. Then

$$\frac{N}{V} \;=\; \frac{8\pi}{(2\pi\hbar c)^3}\int_{mc^2}^{E_F} d\epsilon\; \epsilon\sqrt{\epsilon^2 - m^2c^4} \;, \tag{4.2.2}$$

---

[*] For relativistic particles $PV = E/3$ and at constant entropy $E \propto V^{-1/3}$. Hence, for adiabatic changes $PV^{4/3} =$ constant. This gives $\gamma = 4/3$.

where the general formula for the density of states (1.7.1.7) has been used and m is the **electron** mass. This integral can be done using the substitution

$$mc^2x = \sqrt{\epsilon^2 - m^2c^4} \quad \Longrightarrow \quad \epsilon \;=\; mc^2\sqrt{1+x^2}\ ,$$

to give

$$\frac{N}{V} \;=\; 8\pi\left(\frac{mc}{h}\right)^3\frac{x_F^3}{3}\ . \tag{4.2.3}$$

Thus

$$x_F \;=\; \left(\frac{h}{mc}\right)\left(\frac{9}{32\pi^2}\right)^{1/3}\frac{N^{1/3}}{R}\ . \tag{4.2.4}$$

Similarly, the energy is

$$\frac{E_e}{V} \;=\; 8\pi\left(\frac{mc}{h}\right)^3 mc^2\int_0^{x_F} dx\; x^2\sqrt{1+x^2}\ . \tag{4.2.5}$$

This integral can be done exactly, but since $x_F \gg 1$ we can read off the leading terms by expanding the $\sqrt{1+x^2}$ to get

$$\frac{E_e}{V} \;=\; \frac{mc^2}{4}\,8\pi\left(\frac{mc}{h}\right)^3 x_F^3\,\left(x_F + \frac{1}{x_F} + O\left(\frac{1}{x_F^2}\right)\right)\ . \tag{4.2.6}$$

Thus

$$\frac{E_e}{N} \;=\; \frac{3}{4}mc^2\,\left(x_F + \frac{1}{x_F} + \ldots\right)\ . \tag{4.2.7}$$

Then the total energy per nucleon is

$$\begin{aligned}\frac{E}{N} \;&=\; \frac{E_e}{N} \;+\; \frac{E_G}{N}\\ &=\; \frac{a}{R} \;+\; bR \;-\; \frac{c}{R}\ ,\end{aligned} \tag{4.2.8}$$

where

$$a \;=\; \frac{3hc}{4}\left(\frac{9}{32\pi^2}\right)^{1/3}N^{1/3}\ , \quad c \;=\; \frac{3}{2}GNm_p^2\ . \tag{4.2.9}$$

The a/R term is from the **relativistic** electrons. The value of b can be read from the above calculation but we do not need the explicit value in what follows. We note that E only has a minimum w.r.t R if $c \leq a$. In this case the star forms a dense cold body called a **white dwarf**. If $c > a$ the collapse cannot be halted and the star becomes so dense that inverse beta decay occurs, $e^- + p \rightarrow \nu_e + n$, and the electrons and protons coalesce into neutrons with the violent release of energy in the form of neutrinos. This is a supernova. The energy burst causes material to be blown off in a spectacular way and the star's brightness increases hugely for some days. What is left is a star which is a degenerate gas of neutrons — a **neutron star**. This can be rapidly rotating to form a **pulsar**. If the star is too massive the collapse cannot be halted even by the pressure of the degenerate neutrons and a **black hole** results.

The limit given above for a white dwarf to be stable is

$$N \;\leq\; \sqrt{\frac{2}{\pi}}\,\frac{3}{64\pi^2}\,\left(\frac{\lambda_p}{l_P}\right)^3\ , \tag{4.2.10}$$

where $\lambda_p = 1.32\ 10^{-15}$ metres is the Compton wavelength of the proton and l_P is the Planck length defined by

$$l_P \;=\; \sqrt{\frac{G\hbar}{c^3}} \;=\; 1.616\ 10^{-35}\ \text{metres}\ . \tag{4.2.11}$$

Thus

$$N \;\leq\; 2\ 10^{57} \;=\; 1.7\ N_{\text{sun}}\ . \tag{4.2.12}$$

A more general calculation gives $\sim 1.5\ N_{\text{sun}}$.

4.3 Heat Capacity of Fermi Gas at Low T

We want to calculate the energy E for given N as a function of T for low T. The meaning of "low" will be clear in what follows. We will need to calculate $\mu(T)$ as well. We use eq. (4.1) and to see how to expand around $T = 0$ we consider the generic form for these integrals:

$$I(T) \;=\; \int_0^\infty \frac{f(\epsilon)d\epsilon}{e^{\beta(\epsilon-\mu)}+1}\ . \tag{4.3.1}$$

Make the transformation $\beta(\epsilon - \mu) = z$ to get

$$\begin{aligned} I(T) \;&=\; kT\int_{-\beta\mu}^{\infty} \frac{f(\mu+kTz)}{e^z+1}\,dz \\ &=\; kT\int_0^{\beta\mu} \frac{f(\mu-kTz)}{e^{-z}+1}\,dz \;+\; kT\int_0^\infty \frac{f(\mu+kTz)}{e^z+1}\,dz\ . \end{aligned} \tag{4.3.2}$$

Now, in the first integral we can write

$$\frac{1}{e^{-z}+1} \;=\; 1 - \frac{1}{e^z+1}$$

to give

$$I(T) \;=\; \int_0^\mu f(\epsilon)\,d\epsilon \;-\; kT\int_0^{\beta\mu} \frac{f(\mu-kTz)}{e^z+1}\,dz \;+\; kT\int_0^\infty \frac{f(\mu+kTz)}{e^z+1}\,dz\ . \tag{4.3.3}$$

Note that in the first term the variable of integration has been transformed back to ϵ. This term is just the result for $T = 0$ **but** with upper limit $\mu \equiv \mu(T)$. In the second term we replace the upper limit by ∞ since we certainly require $\mu/T \gg 1$. This amounts to neglecting exponentially small terms which behave like $\exp(-\mu/T)$. Note that terms of this kind are **not** expansible around $T = 0$. [*For interest, we are deriving a power series expansion in T for I(T). This series is an asymptotic, not convergent, series. Given the series, resummation techniques designed to recover the original function naturally give results ambiguous up to non-expansible functions of this kind*].

We thus find

$$I(T) \;=\; \int_0^\mu f(\epsilon)\,d\epsilon \;+\; kT\int_0^\infty \frac{f(\mu+kTz)-f(\mu-kTz)}{e^z+1}\,dz\ . \tag{4.3.4}$$

The numerator of the second integral can now be expanded as a Taylor series in z which can be integrated term by term. This gives

$$\begin{aligned} I(T) \;&=\; \int_0^\mu f(\epsilon)\,d\epsilon \;+\; 2(kT)^2 f'(\mu)\int_0^\infty \frac{z}{e^z+1}\,dz \\ &+\; \frac{1}{3}(kT)^4 f'''(\mu)\int_0^\infty \frac{z^3}{e^z+1}\,dz \;+\; \ldots \end{aligned} \tag{4.3.5}$$

The integrals are given by

$$\int_0^\infty \frac{z^{x-1}}{e^z+1} \;=\; \left(1-\frac{1}{2^{x-1}}\right)\Gamma(x)\zeta(x)\;, \qquad x>0\;. \tag{4.3.6}$$

where $\Gamma(x)$ is the Gamma function and $\zeta(x)$ the Riemann zeta function. This is derived by expanding the denominator in the integrand in (4.3.6) (see section 3.5) and using standard representations for $\Gamma(x)$ and $\zeta(x)$:

$$\begin{aligned} \Gamma(x) \;&=\; \int_0^\infty t^{x-1}e^{-t}\,dt\;, \qquad \zeta(x) \;=\; \sum_{n=1}^\infty \frac{1}{n^x}\;, \\ \sum_{n=1}^\infty \frac{(-1)^{n-1}}{n^x} \;&=\; \left(1-\frac{1}{2^{x-1}}\right)\zeta(x)\;. \end{aligned} \tag{4.3.7}$$

For integer $x>0$ we have $\Gamma(x) \;=\; (x-1)!$ For the values of x of interest the sums defining $\zeta(x)$ can be done using half-range sine and cosine series. For example,

$$\zeta(2) \;=\; \frac{\pi^2}{6}\;, \qquad \zeta(4) \;=\; \frac{\pi^4}{90}\;. \tag{4.3.8}$$

(Note, integrals for Bose-Einstein statistics are done in a similar way.)

The outcome is

$$I(T) \;=\; \int_0^\mu f(\epsilon)\,d\epsilon \;+\; \frac{\pi^2}{6}(kT)^2 f'(\mu) \;+\; \frac{7\pi^4}{360}(kT)^4 f'''(\mu) \;+\; \ldots.. \tag{4.3.9}$$

Then

$$N \;=\; KV\left(\frac{2}{3}\mu^{3/2} \;+\; \frac{\pi^2}{12}(kT)^2\mu^{-1/2} \;+\; O(T^4)\right), \quad \text{with} \quad f(z) \;=\; KVz^{1/2}\;, \tag{4.3.10}$$

$$E(T) \;=\; KV\left(\frac{2}{5}\mu^{5/2} \;+\; \frac{\pi^2}{4}(kT)^2\mu^{1/2} \;+\; O(T^4)\right), \quad \text{with} \quad f(z) \;=\; KVz^{3/2}\;. \tag{4.3.11}$$

We must rewrite (4.3.11) in terms of N using (4.3.10) to express μ as a function of N. This will give a series in T:

$$\mu(T,N) \;=\; E_F \;+\; AT^2 \;+\; O(T^4)\;. \tag{4.3.12}$$

The coefficient A can be determined by substitution into (4.3.10). However, since we want the heat capacity C, we can write

$$C \;=\; \left(\frac{\partial E}{\partial T}\right)_{N,V} \;=\; KV\left(\mu^{3/2}\left(\frac{\partial \mu}{\partial T}\right)_{N,V} \;+\; \frac{\pi^2}{2}k^2T\mu^{1/2} \;+\; O(T^2)\right)\;. \tag{4.3.13}$$

From (4.3.10) we find

$$\left(\frac{\partial \mu}{\partial T}\right)_{N,V} \;=\; -\frac{\pi^2}{6}\frac{k^2T}{\mu} \;+\; O(T^2)\;, \qquad \text{since} \qquad \left(\frac{\partial N}{\partial T}\right)_{N,V} = 0\;,$$

and so the sum of the first two terms gives

$$C \;=\; k^2T\,\frac{\pi^2}{3}KV\,\mu^{1/2} \;+\; O(T^2)\;. \tag{4.3.14}$$

To $O(T)$ we can use the zeroth-order approximation $\mu = E_F$ from (4.3.10). To $O(T)$ we get

$$C \;=\; Nk\,\frac{\pi^2}{2}\,\frac{kT}{E_F} \;=\; Nk\,\frac{\pi^2}{2}\,\frac{T}{T_F}\;, \tag{4.3.15}$$

where $kT_F = E_F$. The expansion is thus good for $T \ll T_F$.

5 Reminder of Classical Mechanics

Let a system be described by M generalized coordinates q_α, with generalized velocities $\dot{q}_\alpha$ and momenta p_α, $1 \leq \alpha \leq M$. For example, a system of N point particles in D dimensions where the q_α are the particle positions and $M = ND$. The forces acting on all parts of the system are derived from a potential which is a function of the q_α. The **Lagrangian** is $L(q, \dot{q}) = T - V$ where T is the kinetic energy term, $T = \frac{1}{2}\dot{q}_\alpha\, T_{\alpha\beta}\, \dot{q}_\beta$ (summation convention applies), and $V = V(q_\alpha)$ is the potential energy. The $T_{\alpha\beta}$ may depend on the q_α but not on the $\dot{q}_\alpha$. It is a non-singular matrix.

Lagrange's equations are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_\alpha} \;-\; \frac{\partial L}{\partial q_\alpha} \;=\; 0\;, \qquad \forall \alpha\;. \tag{5.1}$$

They imply and are implied by Newton's equations. E.g.,

$$\begin{aligned} \text{if} \qquad T \;&=\; \tfrac{1}{2} m \dot{x}_i \dot{x}_i\;, \quad V \;=\; V(x_i) \quad \text{then} \\ \frac{d}{dt}(m\dot{x}_i) \;&=\; -\frac{\partial V}{\partial x_i} \qquad \Rightarrow \qquad m\ddot{\boldsymbol{x}} \;=\; \boldsymbol{F} \;=\; -\nabla V\;. \end{aligned}$$

Lagranges's method can be applied to any choice of coordinates, for example, angular coordinates. In plane polar coordinates $\boldsymbol{x} = r\hat{\mathbf{r}}$, $\dot{\boldsymbol{x}} = \dot{r}\hat{\mathbf{r}} + r\dot{\theta}\hat{\boldsymbol{\theta}}$. Then for the central force $V(\boldsymbol{x}) = -\mu/r$ we have

$$\begin{aligned} L \;=\; T - V \;&=\; \tfrac{1}{2} m\dot{r}^2 \;+\; \tfrac{1}{2} m r^2 \dot{\theta}^2 \;+\; \frac{\mu}{r}\;, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{r}}\right) \;-\; \frac{\partial L}{\partial r} \;&=\; m\ddot{r} \;-\; m r \dot{\theta}^2 \;+\; \frac{\mu}{r^2} \;=\; 0\;, \\ \frac{d}{dt}\left(\frac{\partial L}{\partial \dot{\theta}}\right) \;-\; \frac{\partial L}{\partial \theta} \;&=\; \frac{d}{dt}\left(m r^2 \dot{\theta}\right) \;=\; 0\;. \end{aligned}$$

These are well-known results with conserved angular momentum $h = mr^2\dot{\theta}$.

Define the **generalized momenta** by

$$p_\alpha \;=\; \frac{\partial L}{\partial \dot{q}_\alpha} \;=\; T_{\alpha\beta}\,\dot{q}_\beta \qquad \Rightarrow \qquad \dot{q}_\alpha \;=\; (T^{-1})_{\alpha\beta}\,p_\beta\;.$$

Note that if L does not depend **explicitly** on q_γ, for some γ, then Lagrange's equations (5.1) imply the equation of motion $\dot{p}_\gamma = 0$. This means that p_γ is a **conserved** momentum. In the case of the central potential above p_θ is conserved (we denoted it h there).

The Hamiltonian H is defined by

$$H \;=\; \dot{q}_\alpha p_\alpha \;-\; L\;, \tag{5.2}$$

and is viewed as a function $H(p,q)$ of the coordinates and momenta by elimination of the $\dot{q}_\alpha$ in favour of the p_α. In fact, this is just a Legendre transformation of L to change the independent variables from $(q_\alpha, \dot{q}_a)$ to (p_α, q_α). Then

$$H(p,q) \;=\; \tfrac{1}{2} p_a \,(T^{-1})_{\alpha\beta}\, p_\beta \;+\; V \;. \tag{5.3}$$

Remember $T_{\alpha\beta}$ can depend on the q_α as does V. H gives the total energy of the system since we can write (5.3) as

$$H \;=\; \dot{q}_\alpha \, T_{\alpha\beta}\, \dot{q}_\beta \;-\; L \;=\; T \;+\; V \;. \tag{5.4}$$

Now consider the response of $H(p,q)$ to a small change of the independent variables:

$$\delta H \;=\; \frac{\partial H}{\partial q_\alpha}\delta q_\alpha \;+\; \frac{\partial H}{\partial p_\alpha}\delta p_\alpha \;=\; \delta\dot{q}_\alpha p_\alpha \;+\; \dot{q}_\alpha \delta p_\alpha \;-\; \frac{\partial L}{\partial q_\alpha}\delta q_\alpha \;-\; \frac{\partial L}{\partial \dot{q}_\alpha}\delta \dot{q}_\alpha \;. \tag{5.5}$$

Using the definition of p_α the $\delta\dot{q}_\alpha$ terms cancel and we recover Hamilton's equations:

$$\frac{\partial H}{\partial q_\alpha} \;=\; -\,\frac{\partial L}{\partial q_\alpha} \;=\; -\,\dot{p}_\alpha \;, \qquad \frac{\partial H}{\partial p_\alpha} \;=\; \dot{q}_\alpha \;. \tag{5.6}$$

There are $2M$ such first-order equations which are equivalent to the M second-order Lagrange's equations above.

As an example consider a light rod of length $2a$ with equal masses m fixed at each of its ends, and which rotates about its fixed centre-of-mass G. Each of the particles has kinetic energy $\frac{1}{2}m\dot{\boldsymbol{x}}^2$. Write $\boldsymbol{x} = a(\sin\theta\cos\phi,\ \sin\theta\sin\phi,\ \cos\theta)$ in spherical polars since each particle moves on the surface of a sphere of radius a. Then

$$\dot{\boldsymbol{x}}^2 \;=\; a^2(\dot{\theta}^2 \;+\; \sin^2\theta\,\dot{\phi}^2) \;, \qquad \Rightarrow \qquad T \;=\; \tfrac{1}{2}I(\dot{\theta}^2 \;+\; \sin^2\theta\,\dot{\phi}^2) \;, \tag{5.7}$$

where $I = 2ma^2$ is the moment of inertia of the system. Then

$$p_\theta \;=\; \frac{\partial L}{\partial \dot{\theta}} \;=\; I\dot{\theta} \;, \quad p_\phi \;=\; \frac{\partial L}{\partial \dot{\phi}} \;=\; I\sin^2\theta\,\dot{\phi} \;. \tag{5.8}$$

If $V = 0$ then $L = T$ and so

$$H \;=\; \frac{1}{2I}\left(p_\theta^2 \;+\; \frac{p_\phi^2}{\sin^2\theta}\right) \;. \tag{5.9}$$

Also have

$$H \;=\; \frac{1}{2I}\,\boldsymbol{L}^2 \;,$$

where $\boldsymbol{L} \;=\; \boldsymbol{x}\wedge\boldsymbol{p} \;=\; m\boldsymbol{x}\wedge\dot{\boldsymbol{x}}$ is the angular momentum vector. Note that since the Lagrangian does not depend explicitly on ϕ the corresponding equation of motion is $\dot{p}_\phi = 0$.

6 Classical Statistical Physics

The classical expression for the partition function of a system of N interacting particles in D dimensions in TE at temperature T is

$$Z \;=\; \int e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}\, d\boldsymbol{p}d\boldsymbol{q} \;, \tag{6.1}$$

where $\boldsymbol{p}$ stands for the set of particle momenta $\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_N$ and likewise $\boldsymbol{q}$ stands for the set of particle positions $\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_N$. The Hamiltonian for the whole system including inter-particle interactions and external potential fields is $H(\boldsymbol{p}, \boldsymbol{q})$. The measure is defined to be

$$d\boldsymbol{p}d\boldsymbol{q} \;=\; \prod_{i=1}^{N} d^D p_i \, d^D q_i \ . \tag{6.2}$$

We recognize the Boltzmann factor with the energy evaluated by the classical Hamiltonian. This form can be argued from the quantum formalism as follows.

6.1 The Classical limit

We consider the example of one particle in $D = 1$. The result is easily generalized. The partition function is

$$Z \;=\; \sum_r e^{-\beta\epsilon_r} \;=\; \sum_r \langle r|e^{-\beta\hat{H}}|r\rangle \ , \tag{6.1.1}$$

where $\hat{H}|r\rangle = \epsilon_r|r\rangle$. Clearly, $|r\rangle$ is an energy eigenstate. We can use resolution of the identity to express Z in the form

$$\begin{aligned} Z \;&=\; \sum_r \langle r| \left(\int dq \ |q\rangle\langle q| \right) \ e^{-\beta\hat{H}} \ \left(\int dq' \ |q'\rangle\langle q'| \right) |r\rangle \ , \\ &=\; \int dq dq' \langle q|e^{-\beta\hat{H}}|q'\rangle \ \sum_r \ \langle q'|r\rangle\langle r|q\rangle \ . \end{aligned} \tag{6.1.2}$$

But

$$\sum_r \ \langle q'|r\rangle\langle r|q\rangle \;=\; \langle q|q'\rangle \;=\; \delta(q-q') \ , \tag{6.1.3}$$

and so

$$Z \;=\; \int dq \ \langle q|e^{-\beta\hat{H}}|q\rangle \ . \tag{6.1.4}$$

[Comment: both equations (6.1.1) and (6.1.4) are explicit realizations of the expression

$$Z \;=\; \text{trace}\left(e^{-\beta\hat{H}} \right) \ . \tag{6.1.5}$$

which treats the operator as the generalization of a finite-dimensional matrix. E.g., the operator $\hat{O}$ has matrix elements $\langle r|\hat{O}|r'\rangle$ or $\langle q|\hat{O}|q'\rangle$. It is well known that the trace is invariant under a unitary change of basis such as we have here, namely, the change from the complete set $|r\rangle$ to the complete set $|q\rangle$. It is the sum over **diagonal** elements in **all** bases.]

Then use $\hat{H}(\hat{p}, \hat{q}) = K(\hat{p}) + V(\hat{q})$ and write

$$e^{-\beta\hat{H}} \;=\; e^{-\beta K(\hat{p})} e^{-\beta V(\hat{q})} \;+\; O(\hbar) \ . \tag{6.1.6}$$

Ignore the extra, $\hbar$-dependent, terms to get the classical limit. These arise because $\hat{p}$ and $\hat{q}$ do not commute.

Then

$$Z \;=\; \int dq \ e^{-\beta V(q)} \ \langle q|e^{-\beta K(\hat{p})}|q\rangle \ . \tag{6.1.7}$$

But

$$\begin{aligned}\langle q|e^{-\beta K(\hat{p})}|q\rangle &= \langle q|\int dpdp'|p\rangle \underbrace{\langle p|e^{-\beta K(\hat{p})}|p'\rangle}_{e^{-\beta K(p)}\delta(p-p')}\langle p'|q\rangle \\ &= \int dp\ |\langle q|p\rangle|^2\ e^{-\beta K(p)} \\ &= \frac{1}{h}\int dp\ e^{-\beta K(p)}\ , \qquad \text{since}\ \ \langle q|p\rangle \ = \ \frac{e^{ipq/\hbar}}{\sqrt{2\pi\hbar}}\ . \end{aligned} \tag{6.1.8}$$

Thus

$$Z \ = \ \frac{1}{h}\int dpdq\ e^{-\beta H(p,q)}\ , \tag{6.1.9}$$

where $H(p,q) = K(p) + V(q)$ is the classical Hamitonian.

From now on we shall omit the factor of $1/h$ although it is the correct normalization for Z which comes from the quantum theory and which a classical derivation can never give. In particular, it correctly sets the origin of the entropy.

The average of a physical variable then takes the obvious form:

$$\langle f(\boldsymbol{p},\boldsymbol{q})\rangle \ = \ \frac{\int\ d\boldsymbol{p}d\boldsymbol{q}\ f(\boldsymbol{p},\boldsymbol{q})\ e^{-\beta H(\boldsymbol{p},\boldsymbol{q})}}{\int\ d\boldsymbol{p}d\boldsymbol{q}\ e^{\beta H(\boldsymbol{p},\boldsymbol{q})}}\ , \tag{6.1.10}$$

so that as before, for example,

$$E \ = \ \langle H(\boldsymbol{p},\boldsymbol{q})\rangle \ = \ -\left(\frac{\partial \log Z}{\partial\beta}\right)\ .$$

Note, that in general the momentum integrals in the partition function are always Gaussian integrals which can be done in the standard way.

6.2 The Monatomic Gas

Consider a gas of N non-interacting particles, each of mass m, in volume V in D dimensions. The Hamiltonian H is separable and is given by

$$H \ = \ \sum_{i=1}^{N}\frac{\boldsymbol{p}_i^2}{2m}\ , \tag{6.2.1}$$

and $Z = z^N$ where z is the single-particle partition function

$$z \ = \ \int\ d^Dpd^Dx\ e^{-\beta\,\boldsymbol{p}^2/2m} \ = \ V\left(\frac{2\pi m}{\beta}\right)^{D/2} \ = \ V(2\pi mkT)^{D/2}\ . \tag{6.2.2}$$

Then

$$E \ = \ \frac{D}{2}NkT\ , \tag{6.2.3}$$

which demonstrates equipartition of energy.

Let $f(v)dv$ be the number of particles in $(v, v+dv)$. Using symmetry and $p = mv$ we have the measure $d^Dpd^Dx = CVv^{D-1}dv$ and hence from above

$$f(v) \ = \ CVv^{D-1}\ e^{-mv^2/2kT}\ . \tag{6.2.4}$$

This is the **Maxwell distribution** for velocities.

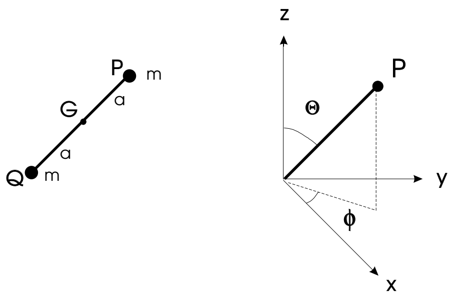


Figure 5:

6.3 The Diatomic Molecule

We consider the rigid “dumbbell” model shown in figure 5, where the atoms, each of mass m, are a fixed distance $2a$ apart. The internal coordinate degrees of freedom are then (θ, ϕ), and the centre-of-mass G has coordinate $\boldsymbol{x}$. There is no potential energy and so the Lagrangian is given by

$$L \;=\; T \;=\; \tfrac{1}{2}M\dot{\boldsymbol{x}}^2 \;+\; \tfrac{1}{2}I\left(\dot{\theta}^2 \;+\; \sin^2\theta\,\dot{\phi}^2\right) \;, \tag{6.3.1}$$

where $M = 2m$, and the kinetic energy due to the internal motion was derived in section 5 with the moment of inertia $I = 2ma^2$. Then get the momenta from $p_\alpha = \partial L/\partial \dot{q}_\alpha$ and find

$$p_i \;=\; \frac{\partial L}{\partial \dot{x}_i} \;=\; M\dot{x}_i \;, \quad p_\theta \;=\; \frac{\partial L}{\partial \dot{\theta}} \;=\; I\dot{\theta} \;, \quad p_\phi \;=\; \frac{\partial L}{\partial \dot{\phi}} \;=\; I\sin^2\theta\,\dot{\phi} \;. \tag{6.3.2}$$

Then

$$\begin{aligned} H \;&=\; \dot{x}_i p_i \;+\; \dot{\theta}p_\theta \;+\; \dot{\phi}p_\phi \;-\; T \;, \\ &=\; \frac{\boldsymbol{p}^2}{2M} \;+\; \frac{p_\theta^2}{2I} \;+\; \frac{p_\phi^2}{2I\sin^2\theta} \;. \end{aligned} \tag{6.3.3}$$

The partition function for one molecule is then (omitting $1/h^5$ factor)

$$z \;=\; \int d^3p d^3x \, dp_\theta d\theta \, dp_\phi d\phi \; e^{-\beta H} \;. \tag{6.3.4}$$

We can write $z = z_t z_r$ where

$$z_t \;=\; \int d^3p d^3x \; e^{-\beta \boldsymbol{p}^2/2M} \;=\; V\left(\frac{2\pi M}{\beta}\right)^{3/2} \;,$$

like the monatomic gas, and

$$z_r \;=\; \int dp_\theta d\theta \, dp_\phi d\phi \; e^{-\beta\left(p_\theta^2/2I \;+\; p_\phi^2/2I\sin^2\theta\right)} \;.$$

Then have

$$z_r \;=\; \sqrt{\frac{2\pi I}{\beta}}\int_0^\pi d\theta\,\sqrt{\frac{2\pi I\sin^2\theta}{\beta}}\int_0^{2\pi} d\phi$$

$$= \frac{2\pi I}{\beta} \int_0^{\pi} d\theta \sin\theta \int_0^{2\pi} d\phi \;=\; \frac{8\pi^2 I}{\beta} \;. \tag{6.3.5}$$

Thus

$$z \;\propto\; (kT)^{5/2} \;, \qquad Z \;=\; z^N \;. \tag{6.3.6}$$

We then find $E = \frac{5}{2}NkT$ which is **equipartition of energy** with 5 degrees of freedom per particle – 3 translational and 2 rotational.

6.4 Equipartition in General

Consider a system described by a set of momenta and coordinates p_α, q_α $1 \le \alpha \le r$, with Hamiltonian $H(\boldsymbol{p}, \boldsymbol{q})$. Take H to be of the form

$$H(\boldsymbol{p}, \boldsymbol{q}) \;=\; p_\alpha\, A_{\alpha\beta}\, p_\beta \;+\; \underbrace{q_\alpha\, B_{\alpha\beta}\, q_\beta}_{V(\boldsymbol{q})} \;, \quad 1 \le \alpha, \beta \le r \;, \tag{6.4.1}$$

where A and B may depend on the q_α. Let H be a homogeneous function of degree 2 in the p_α and a **subset** of the q_α. Denote these variables by $\{l_i\} = (p_1, p_2, \ldots, p_r, q_1, q_2, \ldots, q_s)$ where $s \le r$, $1 \le i \le r + s$. Then by Euler's theorem

$$2H \;=\; \sum_{i=1}^{r+s} l_i \, \frac{\partial H}{\partial l_i} \;. \tag{6.4.2}$$

Thus, under the change of variables $t_i = \sqrt{\beta}\, l_i$ we have

$$H(\boldsymbol{l}, q_{s+1} \ldots q_r) \;=\; H(\frac{1}{\sqrt{\beta}}\boldsymbol{t}, q_{s+1} \ldots q_r) \;=\; \frac{1}{\beta}\; H(\boldsymbol{t}, q_{s+1} \ldots q_r) \;. \tag{6.4.3}$$

The partition function is

$$Z \;=\; = \int \left(\prod_i dl_i\right) dq_{s+1} \ldots dq_r \; e^{-\beta H(\boldsymbol{l}, q_{s+1} \ldots q_r)} \;, \tag{6.4.4}$$

and under the change of variables we find

$$Z \;=\; \left(\frac{1}{\beta}\right)^{(r+s)/2} \int \left(\prod_i dt_i\right) dq_{s+1} \ldots dq_r \; e^{-H(\boldsymbol{t}, q_{s+1} \ldots q_r)} \;. \tag{6.4.5}$$

The integral no longer has any dependence on β and so can be done to give an overall constant. Then we find

$$Z \;\propto\; \beta^{-(r+s)/2} \;, \tag{6.4.6}$$

and so

$$E \;=\; -\left(\frac{\partial \log Z}{\partial \beta}\right) \;=\; \frac{1}{2}(r+s)kT \;. \tag{6.4.7}$$

This is general equipartition of energy for $(r + s)$ "degrees of freedom". For example, the harmonic oscillator in $D = 3$ has

$$H \;=\; \frac{1}{2m}\boldsymbol{p}^2 \;+\; \frac{1}{2}m\omega^2\boldsymbol{q}^2 \qquad \Longrightarrow \qquad E \;=\; 3NkT \;, \tag{6.4.8}$$

which has 6 degrees of freedom.

6.5 Heat Capacities of Gases

Monatomic Molecules. The classical treatment of $H = \boldsymbol{p}^2/2m$ gives $E = \frac{3}{2}NkT \Rightarrow C_V = \frac{3}{2}Nk$, which agrees with experiment.

Diatomic Molecules. We found $E = \frac{5}{2}NkT \Rightarrow C_V = \frac{5}{2}Nk$. However, this only applies at high enough temperature. Find $E = \frac{3}{2}NkT$ when T is low: the rotational degrees of freedom are 'cut out'. This is explained by the full QM analysis. We have $Z = z^N, \;\; z = z_t z_r$. Look at z_r using

$$H_r \;=\; \frac{1}{2I}\left(p_\theta^2 \;+\; \frac{p_\phi^2}{\sin^2\theta}\right) \;=\; \frac{\boldsymbol{L}^2}{2I}\;, \tag{6.5.1}$$

where $\boldsymbol{L} = \boldsymbol{x}\wedge\boldsymbol{p}$ is the angular momentum which we must now treat as quantized. Use eigenstates $|l,m\rangle$:

$$H_r|l,m\rangle \;=\; \frac{\hbar^2}{2I}l(l+1)|l,m\rangle\;, \quad l \;=\; 0,1,2,\ldots \qquad -l \le m \le l\;. \tag{6.5.2}$$

Then

$$z_r \;=\; \sum_r e^{-\beta E_r} \;=\; \sum_{l=0}^{\infty}(2l+1)\,e^{-l(l+1)T_r/T}\;, \quad T_r \;=\; \frac{\hbar^2}{2Ik}\;, \tag{6.5.3}$$

where $(2l+1)$ is the level degeneracy and T_r, typically $\sim 50^oK$, gives the critical temperature. For $T \gg T_r$ have

$$\sum_l \longrightarrow \int_0^\infty dl\;.$$

Then let $x = l(l+1) \quad\Rightarrow\quad dx = (2l+1)dl$ and get

$$z_r \;=\; \int_0^\infty dx\, e^{-xT_r/T} \;=\; \frac{T}{T_r} \;=\; \frac{2I}{\hbar^2}kT \;=\; \frac{8\pi^2 I}{h^2}kT\;, \tag{6.5.4}$$

which agrees with the earlier classical result.

For $T \ll T_r$ only the $l = 0$ term contributes since the other exponents are very large and negative. Then $z_r = 1$ and there is **no** contribution to the specific heat.

Note that for T very large the vibrational modes **along** the axis of the dumbbell will be excited. These are described accurately by the one-dimensional harmonic oscillator which has two degrees of freedom and so contributes NkT to give $E = \frac{7}{2}NkT$.

6.6 Paramagnetism

Each molecule of N molecules in a solid is a little magnet fixed at the lattice site and constrained only to rotate about it. The magnetic dipole has energy $-\boldsymbol{m}\cdot\boldsymbol{B}$ in an external magnetic field $\boldsymbol{B} = (0,0,B)$ as shown in figure 6. Then

$$H \;=\; \frac{1}{2I}\left(p_\theta^2 \;+\; \frac{p_\phi^2}{\sin^2\theta}\right) \;-\; mB\cos\theta\;, \tag{6.6.1}$$

and so

$$z \;=\; \left(\frac{2\pi I}{\beta}\right)\,2\pi\int_0^\pi \underbrace{\sin\theta}_{\text{from } p_\phi}\; d\theta\, e^{\,\beta mB\cos\theta}\;. \tag{6.6.2}$$

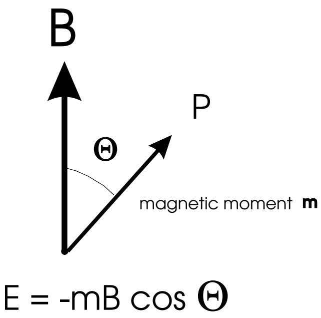


Figure 6:

$$= \underbrace{\left(\frac{8\pi^2 I}{\beta}\right)}_{z_r} \left(\frac{\sinh\ y}{y}\right) \ , \quad y \ = \ m\beta B \ . \tag{6.6.3}$$

The magnetic moment of the solid, M, is defined by $(0,0,M) = N\langle \boldsymbol{m}\rangle$. Only the z-component survives and we find

$$\begin{aligned} M &= N\langle m\cos\theta\rangle \ = \ N\frac{1}{\beta}\,\frac{\partial \log Z}{\partial B} \\ &= N\frac{1}{\beta}\frac{\partial}{\partial B}\left[\log\sinh\left(m\beta B\right) \ - \ \log\ m\beta B\right] \\ &= N\frac{1}{\beta}\left[m\beta\coth(m\beta B) \ - \ \frac{1}{B}\right] \\ &= Nm\left[\coth\ y \ - \ \frac{1}{y}\right] \ . \end{aligned} \tag{6.6.4}$$

For small y (e.g., high T) we find **Curie's law**

$$M \ = \ Nm\frac{y}{3} \ = \ \frac{Nm^2B}{3kT} \ . \tag{6.6.5}$$

6.7 Weak Interparticle Forces

A classical gas of N monatomic molecules has

$$H \ = \ \sum_r \frac{\boldsymbol{p}_r^2}{2m} \ + \ \sum_{r<s} \phi(u_{rs}) \ , \quad u_{rs} \ = \ |\boldsymbol{x}_r \ - \ \boldsymbol{x}_s| \ . \tag{6.7.1}$$

Then

$$Z \ = \ (2\pi mkT)^{3N/2}\, V^N W \ \equiv \ Z_0 W \ , \tag{6.7.2}$$

where $Z_0 = z_0^N$ is the partition function of a perfect gas of N spinless particles, and

$$W \ = \ \frac{1}{V^N}\int d\boldsymbol{x}\ e^{-\beta\sum_{r<s}\ \phi(u_{rs})} \ . \tag{6.7.3}$$

Suppose the interaction ϕ is weak with a possible hard core. Then W can be calculated in perturbation theory. Let

$$\lambda_{rs} \ = \ e^{-\beta\phi(u_{rs})} \ - \ 1 \ , \tag{6.7.4}$$

and so

$$W \;=\; \frac{1}{V^N}\int d\boldsymbol{x}\,\prod_{r<s}(1\;+\;\lambda_{rs}) \;=\; \frac{1}{V^N}\int d\boldsymbol{x}\,(1\;+\;\sum_{r<s}\lambda_{rs}\;+\;\ldots)\;. \qquad (6.7.5)$$

The reason for expanding in λ_{rs} rather than in $\beta\phi(u_{rs})$ is that ϕ might be very large for small ranges of u_{rs} which would render such an expansion invalid. However, keeping the exponential form has no difficulty in accommodating such behaviour (such as a hard core). Then

$$W \;=\; 1\;+\;\frac{1}{V^2}\frac{N(N-1)}{2}\int d^3x_a d^3x_b\,\lambda_{ab}\;.\;\;\text{(no sum on }\;a,b) \qquad (6.7.6)$$

By a change of variables $\boldsymbol{u}=(\boldsymbol{x}_a-\boldsymbol{x}_b),\;\boldsymbol{w}=(\boldsymbol{x}_a+\boldsymbol{x}_b)/2$ we have

$$\begin{aligned}\int d^3x_a d^3x_b\,\lambda_{ab} \;&=\; Vf(T)\;,\quad f(T)\;=\;\int d^3u\,[e^{-\beta\phi(u)}\;-\;1]\;.\\ \Longrightarrow\qquad & \\ W \;&=\; 1\;+\;\frac{N^2}{2V}f(T)\;.\end{aligned} \qquad (6.7.7)$$

Thus

$$\log Z \;=\; \ldots\;+\;N\log V\;+\;\underbrace{\log W}_{\approx\,\frac{N^2}{2V}f(T)}\;, \qquad (6.7.8)$$

and then have equation of state

$$\begin{aligned}\frac{P}{kT}\;&=\;\left(\frac{\partial\log Z}{\partial V}\right)_T\;=\;\frac{N}{V}\;-\;\frac{N^2}{2V^2}f(T)\;,\\ \Longrightarrow\qquad & \\ PV\;&=\;NkT\left[\,1\;-\;\frac{N}{2V}f(T)\,\right]\;.\end{aligned} \qquad (6.7.9)$$

This is the beginning of the virial expansion which is an expansion in the density:

$$P \;=\; \frac{N}{V}kT\left(1+\frac{N}{V}B(T)+\left(\frac{N}{V}\right)^2 C(T)+\ldots\right)\;. \qquad (6.7.10)$$

An example is the van der Waals' equation which arises from the potential reproduced in figure 7.

(I) $\phi=\infty,\;\; e^{-\beta\phi}=0$, for $u<d$. The contribution to $f(T)$ is

$$\frac{4\pi d^3}{3}\cdot(-1)\;\equiv\;-\,2b\;. \qquad (6.7.11)$$

(II) For high T and ϕ attractive and weak for $u>d$ we get the rest of $f(T)$ to be

$$\approx\;\int_d^\infty 4\pi u^2 du\,[\,-\beta\phi(u)\,]\;\equiv\;2a\beta\;. \qquad (6.7.12)$$

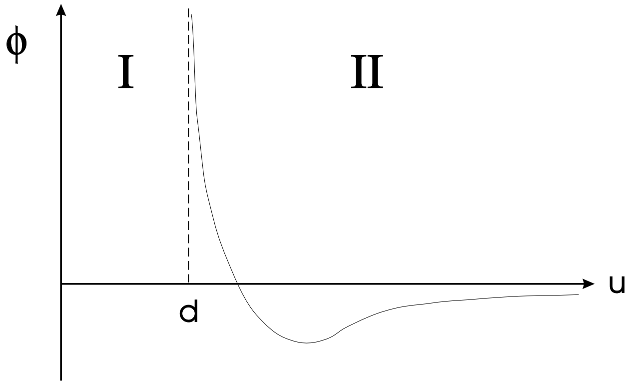


Figure 7:

Here a and b are positive constants. Then $f(T) = -2(b - a/kT)$. Putting this all together we get

$$\begin{aligned} PV \;&=\; NkT\left[1 \,+\, \frac{N}{V}b \,-\, \frac{N}{V}\frac{a}{kT} \,+\, \ldots\right] \;\approx\; NkT\left(1 \,+\, \frac{N}{V}b\right)\left(1 \,-\, \frac{N}{V}\frac{a}{kT}\right) \\ \Longrightarrow \qquad & P\left(1 \,+\, \frac{N}{V}\frac{a}{kT}\right)V\left(1 \,-\, \frac{N}{V}b\right) \;\approx\; NkT \\ \Longrightarrow \qquad & \left(P \,+\, \frac{N^2 a}{V^2}\right)(\,V \,-\, Nb\,) \;\approx\; NkT\;, \end{aligned} \tag{6.7.13}$$

where, in the last manipulation, we have used $PV \approx NkT + O(1/V)$ to eliminate T in favour of P in the first bracket. This is correct to the order in $1/V$ to which we are working. This is the van der Waals' equation of state argued earlier.