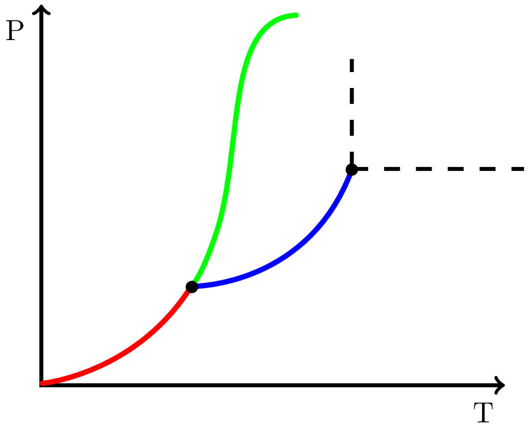
Lecture Notes

Statistical Physics and Thermodynamics

Based on the lecture held by Prof. Roland Netz in winter 2017
Authors: Reyk Börner, Martin Borchert, Philip Loche, Jan Daldrop, Douwe Bonthuis
partly revised, as of June 2019



Send comments to `rnetz@physik.fu-berlin.de`

Contents

1 Introduction

One mole of a substance contains $N_A \approx 6.022 \cdot 10^{23}$ particles[1]. This corresponds to 12 grams of carbon ^{12}C or about 2 to 100 grams of a gas, depending on its type. At room temperature and atmospheric pressure, 1 mol of a gas takes up a volume of approximately 24.6 liters. Knowing that even the three-body problem does not possess a general closed-form solution, it seems hopelessly complicated to calculate properties of a system of this size. However, we will see in the course of this lecture that large systems are particularly amenable to statistical treatment. In the regime of the thermodynamic limit, where the particle number $N \to \infty$, things can in fact become easier – as long as we are only interested in statistical properties of the system as a whole. This is the central motivation behind Statistical Physics and Thermodynamics.

1.1 Objective of Statistical Physics and Thermodynamics

The fields of Statistical Physics and Thermodynamics are concerned with physical systems containing a large number of particles. Examples include gases, liquids, solids, and photon gases. In fact, most systems are large; isolated particles rarely occur.

The key objective of Statistical Physics (or Statistical Mechanics, SM) is to predict the behavior of a large system based on known properties of the individual particles and their interactions. In principle, we can write down the equations of motion of an N-particle system. Solving this problem, however, is difficult or impossible. Furthermore, the solution would not be very interesting because we couldn't test its details experimentally. From an experimental perspective, we are interested in *macroscopic* parameters like temperature, pressure, heat capacity, etc. Statistical Mechanics aims at deriving these macroscopic parameters from *microscopic* properties of the system.

Conversely, Thermodynamics (ThD) states relations between macroscopic quantities without considering microscopic details. These relations result from the *laws of Thermodynamics*, a set of heuristic rules that have been deduced and generalized from experimental findings. While these laws are very general, allowing various predictions of a system's behavior regardless of its particular composition, they fail to yield substance-specific properties like heat capacity.

1.2 A brief history

Historically, Thermodynamics was developed earlier than Statistical Mechanics and even before scientists understood the atomic structure of matter. The idea of the equivalence of heat and energy (Mayer 1842, Joule 1849) marked an important milestone. The laws of Thermodynamics were formulated by Clausius and Kelvin around 1850; Gibbs completed their work in 1878. The theory of Statistical Mechanics was developed mainly by Boltzmann and Gibbs between 1860 and 1900. With the emergence of Quantum Mechanics, many results of Statistical Mechanics were modified; yet the original framework remains valid today.

In this lecture, we will not follow the historical path but instead focus on deriving thermodynamic results from the theory of Statistical Mechanics.

---

[1] 1 mol is defined as the amount of substance containing exactly N_A elementary entities (atoms, molecules, ...), where N_A is the Avogadro constant.

1.3 Systems and equilibrium

We distinguish between three types of idealized systems.

- **Isolated system.** A system that is completely isolated from the surrounding environment, i.e. it exchanges neither energy nor particles with the environment. (Example: hot coffee in a closed thermos bottle.)
- **Closed system.** A system that may exchange energy but not particles with its environment. (Example: hot coffee in a closed glass jar.)
- **Open system.** A system that may exchange both energy and particles with the surrounding environment. (Example: hot coffee in an open cup.)

In reality, even the best thermos bottle exchanges some energy and some particles with the outside world. The isolated system and the closed system are idealizations that serve merely as approximations of real-world systems.

Equilibrium. A system is said to be in *equilibrium* when none of the macroscopic parameters of the system change in time. Systems that are not in equilibrium will change their macroscopic properties until eventually reaching an equilibrium state.

Example 1.1. Imagine you add a cube of sugar to your hot tea. When the cube dissolves, it initially forms a layer at the bottom of the tea cup, leading to an inhomogeneous sugar concentration. After a while, the sugar is completely dissolved and homogeneously distributed in the tea.

Equilibrium states are significantly easier to describe than non-equilibrium states. This lecture will mainly restrict itself to the former case. Transitions between two equilibrium states may be interpreted as a quasi-continuous concatenation of equilibrium states, as long as the transition process is sufficiently slow.

Example 1.2. Consider the compression of a gas in a cylinder with a moveable piston. As long as the piston motion is slow enough, the density distribution of the gas inside the cylinder is homogenous and the gas pressure corresponds to the equilibrium pressure at all times.

1.4 Introductory example: The Boyle-Mariotte law

Before diving into the theory, let us look at a simple example that illuminates how Statistical Mechanics works. To do this we only need Newton's laws and a bit of statistics.

Consider an ideal gas composed of N non-interacting identical mass points (particles), confined in a container of volume V. The container is sealed by a lid of surface area A that may move up and down in the z direction due to an external force (see fig. 1.1). The lid, with a weight of mass m_l, experiences a downwards-oriented gravitational force F. In equilibrium, this force is compensated by the pressure $P = F/A$ the gas exerts on the lid from below. The pressure results from the elastic scattering of particles against the bottom surface of the lid.

During each elastic collision, momentum and energy must be conserved. We assume that before the collision, the lid is at rest. Let m be the mass of a particle, v its velocity before the collision, and v' its velocity after the collision. Furthermore, we denote by ${v_l}'$ the velocity (in the z direction) of the lid after the collision. Conservation of momentum and energy requires

$$\begin{aligned} m v_z &= m {v_z}' + m_l {v_l}' \\ \frac{m {v_z}^2}{2} &= \frac{m {v_z}'^2}{2} + \frac{m_l {v_l}'^2}{2} \ , \end{aligned} \tag{1.1}$$

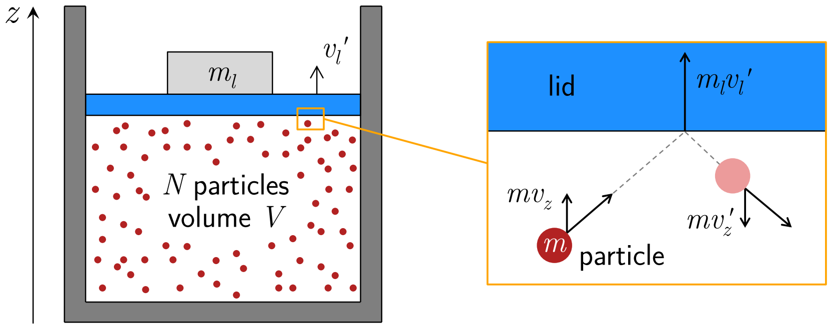


Figure 1.1: **Gas in a container with moveable lid.** Left: N particles of an ideal gas are confined in a container of volume V which is sealed with a moveable lid of mass m_l. Right: Illustration of the scattering process. A particle with mass m and vertical velocity v_z collides elastically with the lid ($m \ll m_l$). After the collision, the particle has a vertical velocity of $v_z' = -v_z$; the momentum transfer to the lid is $m_l v_l' = 2mv_z$.

where v_z represents the z component of the particle velocity (normal to the lid surface A; components parallel to the lid surface are not relevant). The first line of (1.1) can be rewritten as

$$m{v_z}'^2 = \frac{1}{m}\left(mv_z - m_l{v_l}'\right)^2 \ .$$

Inserting this into the second line of (1.1) yields

$$m_l{v_l}' = \frac{2mv_z}{1+\frac{m}{m_l}} \ . \tag{1.2}$$

If the lid is much heavier than the particle ($\frac{m}{m_l} \to 0$), the momentum transfer from the particle to the lid becomes

$$m_l{v_l}' = 2mv_z \ . \tag{1.3}$$

According to Newtonian mechanics, the force acting on the lid equals the time derivative of its momentum. To compute the total force F exerted by the gas, we must find out how many particles collide with the lid in a certain time interval Δt. The vertical distance a particle travels during this time is $\Delta z = v_z \Delta t$. Assuming that half of the particles move upwards ($v_z > 0$) while the other half moves downwards ($v_z < 0$), the probability that a particle in the volume $V_{\text{coll}} = Av_z\Delta t$ hits the lid is $p = 1/2$. Particles outside V_{coll} will not reach the lid during the time interval Δt. Thus, the number N_{coll} of particles colliding with the lid during Δt is

$$N_{\text{coll}} = \frac{1}{2}\rho A v_z \Delta t \ ,$$

where $\rho := N/V$ denotes the particle density. The force F, given by the total momentum transfer per time, reads

$$\begin{aligned} F &= \frac{N_{\text{coll}} \cdot m_l{v_l}'}{\Delta t} \\ &= \rho A m v_z^2 \ , \end{aligned} \tag{1.4}$$

where we have used (1.3). This leads to a gas pressure $P = F/A$ of

$$P = \frac{N}{V} m v_z^2 \ . \tag{1.5}$$

Actually, particles in a gas follow a continuous velocity *distribution* and the pressure P should depend on the average of the squared velocity $\langle v_z^2 \rangle$ (more on statistical distributions follows in chapter 2). Thermal particle motion in a gas has no preferred direction, i.e. $\langle v_x^2 \rangle = \langle v_y^2 \rangle = \langle v_z^2 \rangle$ and the average kinetic energy $\langle E_{\text{kin}} \rangle$ is given by

$$\begin{aligned} \langle E_{\text{kin}} \rangle &= \frac{m}{2} \left\langle v_x^2 + v_y^2 + v_z^2 \right\rangle \\ &= \frac{3m}{2} \left\langle v_z^2 \right\rangle \ . \end{aligned} \tag{1.6}$$

With (1.6), we may write the pressure (1.5) as

$$P = \frac{2}{3} \frac{N}{V} \langle E_{\text{kin}} \rangle \ . \tag{1.7}$$

This agrees with the experimentally known result

$$\boxed{P \propto \frac{1}{V}} \tag{1.8}$$

which is referred to as the Boyle-Mariotte law.

Experiments have shown that the product PV is constant at constant temperature T (for sufficiently low pressure, see fig. 1.2). This holds independently of the type of gas and is thus independent of the particle mass m; we will see later why this is the case. At the same time, the temperature relates closely to the average kinetic energy $\langle E_{\text{kin}} \rangle$. Indeed we may define the temperature of a mono-atomic gas as

$$T := \frac{2}{3} \frac{\langle E_{\text{kin}} \rangle}{k_B} \ . \tag{1.9}$$

Here, k_B denotes the Boltzmann constant. Its value $k_B \approx 1.38 \cdot 10^{-23}$ JK^{-1} is chosen such that water freezes at $T = 273.15$K (0°C) and boils at $T = 373.15$K (100°C). The Boltzmann constant serves the purpose of aligning the thermodynamic temperature scale (in units of Kelvin) with the Celsius scale; it is *not* a natural constant.

Now, we may invoke (1.9) to replace the average kinetic energy in (1.7) with the temperature, leading to the ideal gas law:

$$\boxed{PV = Nk_BT} \tag{1.10}$$

Fig. 1.2 shows that even real gases with particle-particle interaction follow the ideal gas law for pressures up to 10 bar. We will return to the ideal gas in sec. 3.6.

Velocity of gas particles. In the derivation above, the average squared velocity $\langle v^2 \rangle = \langle v_x^2 + v_y^2 + v_z^2 \rangle$ of gas particles appeared. How fast are gas molecules actually, say, of gaseous Hydrogen H_2 or Oxygen O_2? Equations (1.6) and (1.9) yield an expression for the average velocity $\bar{v}$ in terms of temperature as well as the particle mass m:

$$\bar{v} := \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3k_BT}{m}} \ . \tag{1.11}$$

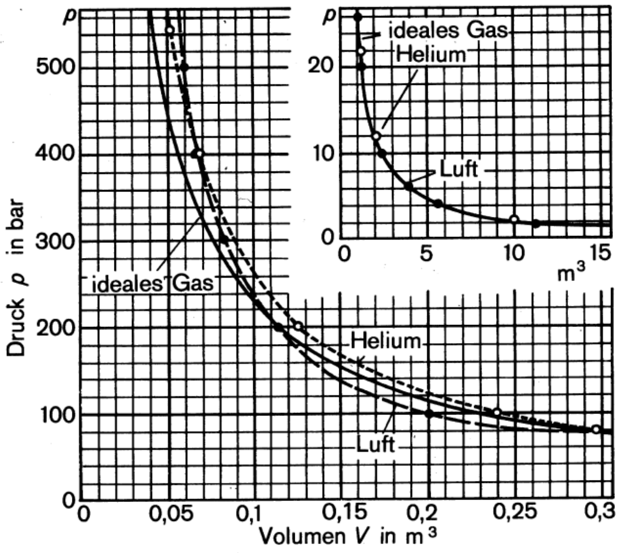
At $T = 273$K, we find

$$\begin{array}{lllll} \text{for Hydrogen:} & m_{H_2} \approx & 2 \cdot 1.661 \cdot 10^{-27} \text{ kg} & \Rightarrow & \bar{v} \approx 1800 \ \frac{\text{m}}{\text{s}} \\ \text{for Oxygen:} & m_{O_2} \approx & 32 \cdot 1.661 \cdot 10^{-27} \text{ kg} & \Rightarrow & \bar{v} \approx 460 \ \frac{\text{m}}{\text{s}} \ . \end{array}$$

Hence gas molecules move quite fast! Furthermore, the relation between temperature and average kinetic energy,

$$\langle E_{\text{kin}} \rangle = \frac{3}{2} k_BT \ ,$$

implies the existence of *absolute zero*: a minimum value ($T = 0$K) of the thermodynamic temperature scale where particles do not move (according to the classical description, in disagreement with quantum mechanics).



Luft bei 0 °C

p in bar	pV in J
1	1,0000
2	0,9994
4	0,9982
6	0,9971
8	0,9960
10	0,9949
20	0,9898
50	0,9780
100	0,9710
200	1,0072
300	1,0952
400	1,2096
600	1,4663

Figure 1.2: Pressure of ideal and real (air and helium) gases at constant temperature 0°C. Data in the figure are for one kmol of gas. The table is for an amount of air that at 0°C and normal pressure of 1 bar has a volume of 10cm^3. Up to pressures of ca. 10 bar the ideal gas law works well. For higher pressures interactions between gas molecules give rise to corrections that can be treated by the virial expansion.

2 Mathematical Statistics

2.1 Probability

We demonstrate the key characteristics of a statistical description with the help of an *experiment* which we conduct N times. As an *outcome*, each experiment yields an integer m. After N repetitions the *overall outcome*is the set $\{m_i\}$ with $i = 1, \ldots, N$. The *absolute frequency* of the *event* m, denoted my $n(m)$, states how often the event m occurred. We define the *relative frequency* h of the event as

$$h(m) := \frac{n(m)}{N} \ ,$$

which is a normalized quantity such that $\sum_m h(m) = 1$.

In the limit of a large number of experiments ($N \to \infty$), the relative frequency converges to the *probability* p of an event,

$$p(m) := \lim_{N\to\infty} h(m) \qquad \text{where } \sum_m p(m) = 1 \ . \tag{2.1}$$

Addition theorem

For mutually exclusive events, the probability of observing any of these events in an experiment is given by the sum of individual probabilities:

$$p(m_1 \vee m_2 \vee m_3 \vee \cdots) = p(m_1) + p(m_2) + p(m_3) + \cdots \ . \tag{2.2}$$

Here, the symbol $\vee$ means “or.”

Example 2.1. Imagine we roll a single die. The probability of obtaining either a 1 or a 2 is the sum of individual probabilities,

$$p(1 \vee 2) = p(1) + p(2) = \frac{1}{6} + \frac{1}{6} = \frac{1}{3} \ .$$

Multiplication theorem

The probability of observing two independent events simultaneously is given by the product of individual probabilities:

$$p(m_1 \wedge m_2) = p(m_1) \cdot p(m_2) \ , \tag{2.3}$$

where the symbol $\wedge$ means “and.”

Example 2.2. Now we roll two dice at the same time. The probability of obtaining 1 two times is the square of the probability of observing one 1 alone,

$$p(1 \wedge 1) = p(1) \cdot p(1) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36} \ .$$

Example 2.3. As a slightly more complicated example, consider the chance of rolling a straight (1,2,3,4) with four dice. If we roll the dice in a specific order, the probability of first rolling 1, then 2, then 3, and then 4 is

$$p(1 \wedge 2 \wedge 3 \wedge 4) = \left(\frac{1}{6}\right)^4 = \frac{1}{1296} \approx 0.0008 \ ,$$

according to the multiplication theorem. However, also (1243) or (3142) qualify as a straight if the order of the dice doesn't matter. In total there exist $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$ possibilities to arrange the numbers 1, 2, 3, and 4. Therefore, according to the addition theorem, the probability of observing any of the 24 possible arrangements of a straight is

$$p([1 \wedge 2 \wedge 3 \wedge 4] \vee [1 \wedge 2 \wedge 4 \wedge 3] \vee \cdots) = \frac{4!}{6^4} \approx 0.019 \ .$$

In the above example, we introduced the *factorial* of a positive integer n,

$$n! := n(n-1)(n-2) \cdots 2 \cdot 1 \ ,$$

which states the number of permutations (possible distinct arrangements) of n distinct items.

2.2 Expectation value, variance, and deviation

Let x be an arbitrary system variable which takes the value $x(m)$ when the system is in the state m. In terms of the normalized probability distribution $p(m)$, we define the *expectation value* or *mean value* $\langle x \rangle$ as

$$\langle x \rangle := \sum_m x(m) p(m) \ . \tag{2.4}$$

Example 2.4. A die has a uniform probability distribution $p(m) = \frac{1}{6}$ for $m = 1, 2, \ldots, 6$. Furthermore, $x(m) = m$. The expectation value is thus

$$\langle x \rangle = \sum_{m=1}^{6} \frac{m}{6} = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5 \ .$$

To quantify deviations from the expectation value, we define the *variance* Δx^2,

$$\Delta x^2 := \left\langle \left(x - \langle x \rangle \right)^2 \right\rangle \ , \tag{2.5}$$

which states the mean squared deviation of x from the expectation value $\langle x \rangle$. Expanding the square, we can rewrite the variance as

$$\begin{aligned} \Delta x^2 &= \left\langle x^2 - 2x \langle x \rangle + \langle x \rangle^2 \right\rangle \\ &= \langle x^2 \rangle - 2 \langle x \rangle \langle x \rangle + \langle x \rangle^2 \\ &= \langle x^2 \rangle - \langle x \rangle^2 \ . \end{aligned} \tag{2.6}$$

Thus, the variance is given in terms of the expectation values of x and x^2. The *deviation* Δx is defined as the square root of the variance,

$$\Delta x := \sqrt{\langle x^2 \rangle - \langle x \rangle^2} \ . \tag{2.7}$$

Example 2.5. What's the deviation of the number of pips when rolling a die? In example 2.4, we calculated the expectation value of the number of pips, $\langle x \rangle = 7/2$. Moreover, we have

$$\langle x^2 \rangle = \sum_{m=1}^{6} \frac{m^2}{6} = \frac{1+4+9+16+25+36}{6} = \frac{91}{6} \ .$$

Thus, the deviation is

$$\Delta x = \sqrt{\frac{91}{6} - \left(\frac{7}{2} \right)^2} \approx 1.7 \ .$$

2.3 Binomial distribution

Imagine a random walker in one dimension who, per unit time, takes a step upwards with probability p or a step downwards with probability $q = 1 - p$. Random walks are used to model the food search of animals, fluctuations in protein configuration, diffusion processes, and many other applications in science. What is the probability that the walker, after taking $N = 5$ steps, will end up one step higher than its original starting point? The *binomial distribution* provides the answer.

To arrive one step higher after five steps, three of the walker's steps must go up, while two must go down. According to the multiplication theorem (sec. 2.1), the probability of such a path is $p^3 q^2$. However, since we disregard the order in which the walker goes up or down, there exist ten distinct paths to reach the desired final position. The number of possible paths can be found by calculating

$$10 = \frac{5!}{3! \cdot 2!} ,$$

where $n!$ again denotes the faculty of n, stating the number of possible ways to arrange n distinguishable objects in a line. The denominator accounts for the fact that not all five steps are distinct: the three steps upwards (as well as the two steps downwards) are indistinguishable among each other, and exchanging them will not change the path. Consequently, the probability of taking three steps up and two steps down in an arbitrary order is $P_5(3) = 10p^3(1-p)^2$.

Example 2.6. If $p = 1/2$, i.e. a step up is as probable as a step down, the probability of a net upwards movement of one step after $N = 5$ steps is

$$P_5(3) = 10 \cdot \left(\frac{1}{2}\right)^5 \approx 0.3 .$$

Generally, if a random walker takes N steps of which m go upwards (and $N - m$ downwards), the total number of different paths is given by the *binomial coefficient*,

$$\binom{N}{m} := \frac{N!}{m!(N-m)!} . \tag{2.8}$$

We may think of the binomial coefficient as the number of possibilities to allocate m objects to N places, or to choose m objects from among N objects. The probability of taking m steps upwards among N steps in total is thus given by the *binomial distribution*

$$\boxed{P_N(m) = \binom{N}{m} p^m (1-p)^{N-m} .} \tag{2.9}$$

Ultimately, the binomial distribution states the probability of observing precisely m single events of probability p in a total of N experiments. It is suitable for repeated experiments with "digital" outcome (e.g. yes/no, up/down, success/failure, one/zero).

Example 2.7. The probability of a newborn child being a girl is approximately $p \approx 1/2$. Therefore, the probability that a family with three children has two girls is $P_3(2) \approx 0.375$.

Example 2.8. The probability of being born on New Year's Eve is $p = 1/365$ (except in leap years). This implies that the probability of having precisely one student with birth date December 31 in a Stat Mech lecture of 50 students is $P_{50}(1) \approx 0.12$.

Expectation value

First, let us prove that the binomial distribution is a normalized probability distribution. This isn't simple when applying brute force, but invoking the N-th order binomial formula,

$$(p+q)^N = \sum_{m=0}^{N} \binom{N}{m} p^m q^{N-m} ,$$

brings us directly to the result. With $q = 1 - p$, the proof reads

$$\begin{aligned}\sum_{m=0}^{N} P_N(m) &= \sum_{m=0}^{N} \binom{N}{m} p^m (1-p)^{N-m} \\ &= (p + (1-p))^N \\ &= 1 \qquad \text{q.e.d.}\end{aligned} \tag{2.10}$$

Since the binomial distribution $P_N(m)$ is a normalized probability distribution, its expectation value (or mean) is, according to the definition in (2.4),

$$\langle m \rangle = \sum_{m=0}^{N} m P_N(m)\ . \tag{2.11}$$

Now, we use the binomial formula again to obtain a simple expression for $\langle m \rangle$ in terms of N and p:

$$\begin{aligned}\langle m \rangle &= \sum_m m P_N(m) \\ &= \sum_m m \frac{N!}{m!(N-m)!} p^m q^{N-m} \\ &\overset{*}{=} p \frac{\partial}{\partial p} \sum_m \frac{N!}{m!(N-m)!} p^m q^{N-m} \\ &\overset{**}{=} p \frac{\partial}{\partial p} (p+q)^N \\ &= pN(p+q)^{N-1} \\ &= pN\ ,\end{aligned}$$

where we have used $p+q = 1$ in the final step. Note that at $*$ we rewrote the expression in terms of a partial derivative, $\langle m \rangle = p\frac{\partial}{\partial p}\sum_m P_N(m)$, to get rid of the factor m in the sum. This is a trick worthy to remember, as we will use it more often in the course of this lecture. At $**$, we inserted the binomial formula. In summary, the expectation value of the binomial distribution is given by

$$\boxed{\langle m \rangle = Np\ .} \tag{2.12}$$

This result agrees with the addition theorem for independent events (sec. 2.1).

Variance

To find the variance, we must additionally calculate the expectation value of m^2. The procedure is similar to the calculation above, though with two partial derivatives this time:

$$\begin{aligned}\left\langle m^2 \right\rangle &= \sum_m m^2 P_N(m) \\ &= \sum_m m^2 \frac{N!}{m!(N-m)!} p^m q^{N-m} \\ &= p\frac{\partial}{\partial p} p \frac{\partial}{\partial p} \sum_m \frac{N!}{m!(N-m)!} p^m q^{N-m} \\ &= p\frac{\partial}{\partial p} p \frac{\partial}{\partial p} (p+q)^N \\ &= p\frac{\partial}{\partial p} \left(pN(p+q)^{N-1}\right) \\ &= pN(p+q)^{N-1} + p^2 N(N-1)(p+q)^{N-2} \\ &= pN + p^2 N(N-1)\ ,\end{aligned}$$

where $p+q=1$. Together with (2.12), we have

$$\begin{aligned}\Delta m^2 &= \langle x^2\rangle - \langle x\rangle^2 \\ &= pN + p^2N(N-1) - p^2N^2 \\ &= pN + p^2N\ .\end{aligned}$$

Thus the variance of the binomial distribution is given by

$$\boxed{\Delta m^2 = pN(1-p)\ .} \tag{2.13}$$

Relative Deviation

By taking the square root of the variance (2.13), we obtain the deviation Δm from the mean for the binomial distribution:

$$\boxed{\Delta m = \sqrt{pN(1-p)}\ .} \tag{2.14}$$

When comparing probability distributions characterized by different expectation values, the comparison of deviations is not very meaningful unless they are put in relation to their respective mean. A standardized measure which takes this into account is the *relative deviation* σ_{rel} defined by

$$\sigma_{\text{rel}} := \frac{\Delta m}{\langle m\rangle}\ . \tag{2.15}$$

The relative deviation quantifies the dispersion of a probability distribution relative to its mean value $\langle m\rangle$. For the binomial distribution, the relative deviation is given by

$$\boxed{\frac{\Delta m}{\langle m\rangle} = \sqrt{\frac{1-p}{pN}} \quad \sim \frac{1}{\sqrt{N}}\ .} \tag{2.16}$$

Since the relative deviation is proportional to the reciprocal of the square root of N, it approaches zero as N gets very large:

$$\lim_{N\to\infty} \frac{\Delta m}{\langle m\rangle} = 0\ . \tag{2.17}$$

This important result is known as the *law of large numbers.* It justifies describing the properties of statistical systems in terms of mean values, as long as the number of experiments N is sufficiently large.

Example 2.9 (Compound system). Consider $N = 10^{24}$ molecules of a gas (that's roughly 1 mol) in a box of volume V. We divide the box into two equally large sub-volumes. Then, the probability of finding m out of N molecules in one of the sub-volumes is given by the binomial distribution. Let's assume that a molecule is equally likely to be in either of the sub-volumes, such that $p = q = 1/2$. On average, there are $\langle m\rangle = Np = 0.5\cdot 10^{24}$ molecules in a sub-volume. The deviation of this mean occupation number is $\Delta m = \sqrt{Npq} = \sqrt{N/4} = 0.5\cdot 10^{12}$ (half a trillion molecules!). However, in relation to the expectation value $\langle m\rangle$, we obtain a negligible relative deviation of $\frac{\Delta m}{\langle m\rangle} = 10^{-12}$. This illustrates why Thermodynamics works: since most physical systems are very large, the probability distributions of statistical system properties have very sharp peaks (relative to the mean value).

Rare events. There exist situations where $N \gg 1$ is very large but the probability $p \ll 1$ is very small, such that the product $\langle m\rangle = Np$ is finite[1]. This characterizes *rare events* for which the relative deviation may in fact be large. To demonstrate this, let us consider the following example.

---

[1] Here we mean *finite* in the sense that Np isn't a very large number like N.

Example 2.10 (Rare events). The number of car trips a driver undertakes in a lifetime can arguably be very large, whereas the probability of having an accident per car ride is pretty small. For $p \ll 1$, the deviation Δm is approximately

$$\Delta m = \sqrt{Np(1-p)} \approx \sqrt{Np} = \sqrt{\langle m \rangle} \ ,$$

where $\langle m \rangle$ gives the average number of accidents in a lifetime. Then, the relative deviation,

$$\frac{\Delta m}{\langle m \rangle} \approx \frac{1}{\sqrt{\langle m \rangle}} \ ,$$

does not become negligible for large N ($\langle m \rangle$ remains finite). For example, if the mean of a rare event is $\langle m \rangle = 0.1$, the relative deviation is approximately $\frac{\Delta m}{\langle m \rangle} \approx 3$. Thus, we may expect significant deviations from the mean, which is why it's good to insure yourself against events with large relative deviations!

2.4 Normal distribution

We can simplify the binomial distribution (2.9) in the limit of large $N \gg 1$, while p remains a finite number, such that $Np \to \infty$. Let us take the logarithm of the binomial distribution,

$$\begin{aligned} \ln P_N(m) &= \ln\left(\frac{N!}{m!(N-m)!} p^m (1-p)^{N-m}\right) \\ &= m \ln p + (N-m)\ln(1-p) + \ln(N!) - \ln(m!) - \ln((N-m)!) \ . \end{aligned} \tag{2.18}$$

This expression may be simplified using the *Stirling formula*,

$$\boxed{\ln(N!) = N \ln N - N + \mathcal{O}(\ln N)} \ . \tag{2.19}$$

To derive the formula (2.19), we first rewrite

$$\ln(N!) = \ln\left(\prod_{j=1}^{N} j\right) = \sum_{j=1}^{N} \ln j \ . \tag{2.20}$$

Since $N \gg 1$, we may approximate the sum by an integral:

$$\begin{aligned} \sum_{j=1}^{N} \ln j &\approx \int_1^N \mathrm{d}x \ln x \\ &= x \ln x - x \Big|_1^N \\ &= N \ln N - N + 1 \\ \Rightarrow \quad \ln(N!) &= N \ln N - N + \mathcal{O}(\ln N) \ . \end{aligned} \tag{2.21}$$

Choosing 1 as the lower boundary of the integral is merely an approximation, such that the formula contains an error in the order of $\ln N$. Inserting the Stirling formula (2.19) into (2.18), we now obtain

$$\ln P_N(m) = m \ln p + (N-m)\ln(1-p) + N \ln N - m \ln m - (N-m)\ln(N-m) \ . \tag{2.22}$$

$P_N(m)$ has a sharp peak at $\langle m \rangle$ for large N, as we know from the law of large numbers. This also holds for the logarithm, $\ln P_N(m)$. We perform a Taylor expansion around the maximum m^*:

$$\ln P_N(m) = \ln P_N(m^*) + \frac{1}{2}(m-m^*)^2 \left.\frac{\mathrm{d}^2}{\mathrm{d}m^2} \ln P_N(m)\right|_{m=m^*} + \dots \ , \tag{2.23}$$

where m^* is defined by

$$0 = \frac{\mathrm{d}}{\mathrm{d}m} \ln P_N(m) \bigg|_{m=m^*} \quad ,$$

such that the linear term vanishes in the expansion. At the maximum, we have

$$\begin{aligned}
0 = \frac{\mathrm{d}}{\mathrm{d}m} \ln P_N(m) &= \ln \frac{p}{1-p} + \ln \frac{N-m}{m} \\
\Rightarrow \quad \ln \frac{1-p}{p} &= \ln \frac{N-m}{m} \\
\frac{1-p}{p} &= \frac{N-m}{m} \\
\frac{1}{p} &= \frac{N}{m} \quad \Rightarrow \quad m^* = Np \ .
\end{aligned} \tag{2.24}$$

Thus, the maximum m^* of our approximation for $\ln P_N(m)$ coincides with the mean $\langle m \rangle = Np$. As we might expect, the mean is equivalent to the most probable value in the limit of $N \to \infty$. Using this result, the second derivative reads

$$\begin{aligned}
\frac{\mathrm{d}^2}{\mathrm{d}m^2} \ln P_N(m) \bigg|_{m=m^*} &= \left(-\frac{1}{m} - \frac{1}{N-m} \right) \bigg|_{m=m^*} \\
&= -\frac{1}{Np} - \frac{1}{N-Np} \\
&= -\frac{1}{Np} - \frac{1}{Nq} \\
&= -\frac{q+p}{Npq} \\
&= -\frac{1}{Npq} = -\frac{1}{\Delta m^2} \ ,
\end{aligned} \tag{2.25}$$

where $q = 1-p$ and the last line follows from the definition of the variance Δm^2 of the binomial distribution (2.13). Finally, it is easily verified by inserting $m^* = Np$ into (2.22) that the constant term of the Taylor expansion vanishes, $\ln P_N(m^*) = 0$. The expansion up to second order therefore reads

$$\ln P_N(m) = -\frac{(m-m^*)^2}{2\Delta m^2} + \mathcal{O}\left(m^3\right) \ . \tag{2.26}$$

Exponentiating both sides yields an approximation $W(m)$ of the binomial distribution,

$$W(m) \propto \exp\left(-\frac{(m - \langle m \rangle)^2}{2\Delta m^2} \right) \ , \tag{2.27}$$

where we replaced m^* by $\langle m \rangle$. Note that by Taylor expanding the logarithm of the binomial distribution instead of $P_N(m)$ itself, we obtained an exponential solution for $W(m)$ that is normalizable (such that $\int_{-\infty}^{\infty} \mathrm{d}m\, W(m) = 1$), as required for a probability distribution. The normalized *normal distribution* $W(m)$ is given by

$$\boxed{W(m) = \frac{1}{\sqrt{2\pi\Delta m^2}} \exp\left(-\frac{(m - \langle m \rangle)^2}{2\Delta m^2} \right)} \ . \tag{2.28}$$

As a simplification of the binomial distribution, the normal distribution is valid in the limit of $N \to \infty$ while p remains finite. It is fully determined by the mean $\langle m \rangle$ and the variance Δm^2.

2.5 Poisson distribution

For very large $N \gg 1$ but simultaneously $p \ll 1$, such that the product Np is finite, we can simplify the binomial distribution in a different way. This case leads us to the Poisson distribution. Let us break down the binomial distribution (2.9) into parts,

$$P_N(m) = \frac{N!}{(N-m)!}\,\frac{1}{m!}\left(\frac{p}{1-p}\right)^m (1-p)^N \ . \tag{2.29}$$

First, we approximate the factor $(1-p)^N$ by writing

$$(1-p)^N = \exp\left[N\ln(1-p)\right] \approx e^{-Np} \ , \tag{2.30}$$

where we have Taylor expanded the logarithm up to first order, $\ln(1+x) = x + \mathcal{O}(x^2)$. Secondly, we use $p \ll 1$ to approximate

$$\left(\frac{p}{1-p}\right)^m \approx p^m \ . \tag{2.31}$$

Lastly, we tackle the binomial coefficient:

$$\begin{aligned}
\frac{N!}{(N-m)!} &= \exp\left[\ln(N!) - \ln((N-m)!)\right] \\
&\overset{*}{\approx} \exp\left[N\ln(N) - N - (N-m)\ln(N-m) + N - m\right] \\
&= \exp\left[N\ln\left(\frac{N}{N-m}\right) + m\ln(N-m) - m\right] \\
&= \exp\left[-N\ln\left(\frac{N-m}{N}\right) + m\ln\left(\frac{N-m}{N}\right) + m\ln N - m\right] \\
&= \exp\left[-N\ln\left(1-\frac{m}{N}\right) + m\ln\left(1-\frac{m}{N}\right) + m\ln N - m\right] \\
&\overset{\dagger}{\approx} \exp\left[m + \mathcal{O}\left(\frac{m^2}{N}\right) + \mathcal{O}\left(\frac{m^2}{N}\right) + m\ln N - m\right] \\
&\approx \exp\left[m\ln N\right] = N^m \ .
\end{aligned} \tag{2.32}$$

Here we applied the Stirling formula at $*$ and used the Taylor series up to first order, $\ln(1+x) = x + \mathcal{O}(x^2)$, in the step marked by †. Putting things together, we arrive at an approximation of the binomial distribution for finite Np,

$$P_N(m) \approx N^m \frac{1}{m!} p^m e^{-Np} \ . \tag{2.33}$$

For simplicity we define $\lambda := Np$. Then, the *Poisson distribution* is given by

$$\boxed{W(m) = \frac{\lambda^m}{m!}e^{-\lambda}} \ . \tag{2.34}$$

This result turns out to be normalized: using $N \to \infty$ we show

$$\sum_{m=0}^{N} W(m) \approx \sum_{m=0}^{\infty} W(m) = \sum_{m=0}^{\infty} \frac{\lambda^m}{m!}e^{-\lambda} = e^{\lambda}e^{-\lambda} = 1 \ . \tag{2.35}$$

Expectation value. Let us calculate the expectation value of the Poisson distribution:

$$\begin{aligned}
\langle m \rangle &= \sum_{m=0}^{\infty} m W(m) \\
&= e^{-\lambda} \sum_{m=0}^{\infty} \frac{m\lambda^m}{m!} \\
&= e^{-\lambda}\, \lambda \frac{\partial}{\partial \lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \\
&= e^{-\lambda}\, \lambda \frac{\partial}{\partial \lambda} e^{\lambda} \\
&= e^{-\lambda} \lambda e^{\lambda} \\
&= \lambda \; .
\end{aligned} \tag{2.36}$$

As expected, the expectation value of the Poisson distribution, $\langle m \rangle = \lambda = Np$, is identical to the expectation value of the binomial distribution.

2.6 Continuous density distributions

Previously, we have seen two different ways of normalizing a probability distribution. The normalization of the normal distribution was carried out via integration from $-\infty$ to ∞, whereas we showed the normalization of the Poisson distribution by means of a discrete sum over all m. In other words, we interpreted the normal distribution as a continuous distribution but the Poisson distribution as a discrete distribution. In this section, let us briefly address how discrete and continuous probability distributions relate.

In the continuous case, we have a continuous random variable x controlled by a density distribution $p(x)$ (e.g. the normal distribution). The normalization requirement reads

$$\int_{-\infty}^{\infty} \mathrm{d}x\, p(x) = 1 \; .$$

The *moments*[2] of this density distribution are then calculated via

$$\langle x^n \rangle := \int_{-\infty}^{\infty} \mathrm{d}x\, x^n p(x) \; .$$

More generally, the expectation value of any function $f(x)$ is given by

$$\langle f \rangle := \int_{-\infty}^{\infty} \mathrm{d}x\, f(x) p(x) \; . \tag{2.37}$$

A discrete probability distribution p_m may be written as a continuous density distribution $p(x)$ by summing over δ-distributions representing the outcomes x_m, such that

$$p(x) = \sum_m p_m \delta(x - x_m) \; . \tag{2.38}$$

The δ-distribution, as a reminder, is defined by

$$\int_{-\infty}^{\infty} \mathrm{d}x\, f(x) \delta(x - y) = f(y) \; .$$

---

[2] The n-th moment of the distribution $p(x)$ is defined as the expectation value $\langle x^n \rangle$. For more about moments, see section 2.7.

We may switch between the continuous and the discrete description of the expectation value $\langle f\rangle$ by inserting (2.38) into (2.37), which yields

$$\begin{aligned}\langle f\rangle &= \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)p(x)\\ &= \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\sum_m p_m\delta(x-x_m)\\ &= \sum_m p_m \int_{-\infty}^{\infty} \mathrm{d}x\, f(x)\delta(x-x_m)\\ &= \sum_m p_m f(x_m)\ . \end{aligned} \tag{2.39}$$

2.7 Characteristic function

Let us introduce the *characteristic function* G,

$$G(k) := \left\langle e^{-ikx}\right\rangle = \int \mathrm{d}x\, p(x)\, e^{-ikx}\ . \tag{2.40}$$

This is nothing but the Fourier transform of $p(x)$. As we will see, the characteristic function serves as a tool to quickly calculate moments and cumulants of a distribution.

Moments. The n-th derivative of $G(k)$ is

$$\frac{\mathrm{d}^n G(k)}{\mathrm{d}k^n} = \int \mathrm{d}x\, p(x)\,(-ix)^n\, e^{-ikx}\ .$$

Note that n-fold differentiation leads to a factor $(-ix)^n$ within the integral. Evaluating the derivative at $k=0$, we find

$$\left.\frac{\mathrm{d}^n G(k)}{\mathrm{d}k^n}\right|_{k=0} = \langle(-ix)^n\rangle = -i^n\,\langle x^n\rangle\ .$$

Thus, the n-th derivative of $G(k)$ is directly linked to the n-th moment of the distribution $p(x)$. The characteristic function allows us to easily calculate all moments:

$$\boxed{\langle x^n\rangle = i^n\left.\frac{\mathrm{d}^n G(k)}{\mathrm{d}k^n}\right|_{k=0}}\ . \tag{2.41}$$

In turn, moments of $p(x)$ correspond to coefficients of the Taylor series of $G(k)$,

$$G(k) = \sum_{n=0}^{\infty}\frac{k^n}{n!}\left.\frac{\mathrm{d}^n G(k)}{\mathrm{d}k^n}\right|_{k=0} = \sum_{n=0}^{\infty}\frac{(-ik)^n}{n!}\langle x^n\rangle\ . \tag{2.42}$$

In fact, it turns out that the logarithm of the characteristic function is much better for characterizing probability distributions. Taking the logarithm has already proven to be useful when we derived the normal distribution (see section 2.4). In many cases, the logarithm is easier to work with. For example, the normal distribution is exponential, $p(x)\propto\exp\left[-(x-x^*)^2\right]$. Its logarithm, however, is simply a second-order polynomial: $\ln p(x)\propto(x-x^*)^2$.

Cumulants. We define the n-th *cumulant* $\langle x^n\rangle_c$ as

$$\boxed{\langle x^n\rangle_c := i^n\left.\frac{\mathrm{d}^n \ln G(k)}{\mathrm{d}k^n}\right|_{k=0}}\ , \tag{2.43}$$

which implies that

$$\ln G(k) = \sum_{n=1}^{\infty} \frac{(-ik)^n}{n!} \left\langle x^n \right\rangle_c \ . \tag{2.44}$$

How do moments and cumulants relate? We show this with the following derivation sketch. First, we explicitly write the Taylor series in eq. (2.42),

$$G(k) = 1 \underbrace{- ik \left\langle x \right\rangle - \frac{k^2}{2} \left\langle x^2 \right\rangle + \frac{ik^3}{6} \left\langle x^3 \right\rangle + \dots}_{=:z} \ , \tag{2.45}$$

where we label the underbraced term z and use the Taylor expansion of $\ln(1+z) = z - \frac{1}{2}z^2 + \frac{1}{3}z^3 - \frac{1}{4}z^4 + \dots$ to find

$$\ln G(k) = -ik \left\langle x \right\rangle - \frac{k^2}{2} \left\langle x^2 \right\rangle + \frac{k^2}{2} \left\langle x \right\rangle^2 + \dots \ . \tag{2.46}$$

Comparing this with eq. (2.44), that is, $\ln G(k) = -ik \left\langle x \right\rangle_c - \frac{k^2}{2} \langle x^2 \rangle_c + \dots$, we obtain the cumulants in terms of moments:

$$\begin{aligned} \left\langle x \right\rangle_c &= \left\langle x \right\rangle \\ \langle x^2 \rangle_c &= \langle x^2 \rangle - \left\langle x \right\rangle^2 \\ \langle x^3 \rangle_c &= \langle x^3 \rangle - 3\langle x^2 \rangle \langle x \rangle + 2 \left\langle x \right\rangle^3 \\ &\dots \end{aligned} \tag{2.47}$$

Note that the first cumulant corresponds to the mean, while the second cumulant corresponds to the variance Δx^2. Generally, cumulants are great for characterizing probability distributions!

Example 2.11 (Cumulants of the normal distribution). The normal distribution is given by eq. (2.28). According to (2.40), the associated characteristic function is

$$G(k) = \int_{-\infty}^{\infty} \mathrm{d}x \, W(x) e^{-ikx} = \int_{-\infty}^{\infty} \frac{\mathrm{d}x}{\sqrt{2\pi\Delta^2}} \exp\left[-ikx - \frac{(x-x^*)^2}{2\Delta^2} \right] \ ,$$

where x^* denotes the mean and Δ^2 represents the variance. We solve the integral by applying the technique *completing the square.* In the first step, we shift the integration variable by the mean, $x = \tilde{x} + x^*$. (The infinite integration limits remain unchanged; x^* is finite.) This yields

$$G(k) = \int_{-\infty}^{\infty} \frac{\mathrm{d}\tilde{x}}{\sqrt{2\pi\Delta^2}} \exp\left[-ikx^* - ik\tilde{x} - \frac{\tilde{x}^2}{2\Delta^2} \right] \ .$$

Now we complete the square in the square brackets by rewriting

$$-ik\tilde{x} - \frac{\tilde{x}^2}{2\Delta^2} = -\frac{1}{2\Delta^2} \left(\tilde{x} + ik\Delta^2 \right)^2 - \frac{k^2\Delta^2}{2} \ .$$

We transform variables again, such that $\hat{x} = \tilde{x} + ik\Delta^2$. This results in

$$\begin{aligned} G(k) &= \int_{-\infty}^{\infty} \frac{\mathrm{d}\hat{x}}{\sqrt{2\pi\Delta^2}} \exp\left[-ikx^* - \frac{1}{2\Delta^2}\hat{x}^2 - \frac{k^2\Delta^2}{2} \right] \\ &= e^{-ikx^* - \frac{k^2\Delta^2}{2}} \int_{-\infty}^{\infty} \frac{\mathrm{d}\hat{x}}{\sqrt{2\pi\Delta^2}} e^{-\frac{\hat{x}^2}{2\Delta^2}} \\ &= \exp\left[-ikx^* - \frac{k^2\Delta^2}{2} \right] \ . \end{aligned} \tag{2.48}$$

Based on eq. (2.43), the cumulants of the normal distribution are found by calculating derivatives of $\ln G(k) = -ikx^* - \frac{k^2\Delta^2}{2}$, evaluated at $k = 0$. Specifically, the first and second cumulants are

$$\begin{aligned}
\frac{\mathrm{d}\ln G(k)}{\mathrm{d}k} &= -ix^* - k\Delta^2 \quad \Rightarrow \quad \langle x \rangle_c = \langle x \rangle = i \cdot (-ix^*) = x^* \\
\frac{\mathrm{d}^2 \ln G(k)}{\mathrm{d}k^2} &= -\Delta^2 \quad \Rightarrow \quad \langle x^2 \rangle_c = \langle x^2 \rangle - \langle x \rangle^2 = -1 \cdot (-\Delta^2) = \Delta^2 \ ,
\end{aligned} \tag{2.49}$$

as expected. Since $\ln G(k)$ is a quadratic function in this example (note that the logarithm is easy to work with), we immediately know that all higher derivatives vanish. Thus, all higher-order cumulants ($n > 2$) are zero; the normal distribution is characterized entirely by the first two cumulants. This implies that non-zero higher-order cumulants indicate deviations from the normal distribution.

2.8 Multi-dimensional probability distributions

A joint distribution of several random variables $x_1, x_2, \ldots, x_n$ is determined by the multi-dimensional probability density $p(x_1, x_2, \ldots, x_n)$. Normalization requires that

$$\int \mathrm{d}x_1 \int \mathrm{d}x_2 \cdots \int \mathrm{d}x_n \ p(x_1, x_2, \ldots, x_n) \ . \tag{2.50}$$

To obtain the probability density for one of the variables, we must integrate over all other random variables:

$$p(x_1) = \int \mathrm{d}x_2 \cdots \int \mathrm{d}x_n \ p(x_1, x_2, \ldots, x_n) \ . \tag{2.51}$$

Some refer to this projection process as the *marginalization* of the variables $x_2, \ldots, x_n$: we are interested in the probability of x_1 regardless of the values of the other variables. Furthermore, moments are generally of the form $\langle x_1^{m_1} \ldots x_n^{m_n} \rangle$. In this context, the *covariance*,

$$\mathrm{cov}(x_j, x_k) := \langle x_j x_k \rangle - \langle x_j \rangle \langle x_k \rangle \tag{2.52}$$

plays an important role as a measure of the correlation of the two random variables x_j and x_k. In the case that their joint distribution factorizes, i.e. $p(x_j, x_k) = p(x_j)p(x_k)$, the variables are independent of one another and the covariance vanishes:

$$\begin{aligned}
\langle x_j x_k \rangle &= \int \mathrm{d}x_j \mathrm{d}x_k \, x_j x_k p(x_j) p(x_k) \\
&= \int \mathrm{d}x_j \, x_j p(x_j) \int \mathrm{d}x_k \, x_k p(x_k) \\
&= \langle x_j \rangle \langle x_k \rangle \qquad \Rightarrow \quad \mathrm{cov}(x_j, x_k) = 0 \ .
\end{aligned}$$

2.9 Central limit theorem

Consider a sum of random variables $x_1, x_2, \ldots, x_m$,

$$y = \sum_{i=1}^{m} \frac{x_i}{m} \ , \tag{2.53}$$

where each random number x_i follows the same probability distribution $p(x_i)$. In the limit $m \to \infty$, y represents the mean of $p(x_i)$. What is the distribution of y? The central limit theorem provides the answer. Let us illuminate this theorem by means of a concrete example.

Example 2.12. Let $p(x)$ be the weight distribution of a single person. Then, what is the weight distribution of 100 people? This question is important, for instance, when estimating before a flight how much fuel the airplane must carry. The average weight y of 100 individuals obeys the distribution $W(y)$, given by

$$W(y) = \int \mathrm{d}x_1 \dots \mathrm{d}x_m \, p(x_1) \cdots p(x_m) \, \delta\left(y - \sum_{i=1}^{m} \frac{x_i}{m}\right) \; . \tag{2.54}$$

Here δ denotes the delta distribution. Note that we assume individual weights to be uncorrelated, such that the multiplication theorem (see section 2.1) holds. For the n-th moment we find

$$\begin{aligned}
\langle y^n \rangle &= \int \mathrm{d}y \, y^n W(y) \\
&= \int \mathrm{d}y \, y^n \int \mathrm{d}x_1 \dots \mathrm{d}x_m \, p(x_1) \cdots p(x_m) \, \delta\left(y - \sum_{i=1}^{m} \frac{x_i}{m}\right) \\
&= \int \mathrm{d}x_1 \dots \mathrm{d}x_m \, p(x_1) \dots p(x_m) \int \mathrm{d}y \, y^n \delta\left(y - \sum_{i=1}^{m} \frac{x_i}{m}\right) \\
&= \int \mathrm{d}x_1 \dots \mathrm{d}x_m \, p(x_1) \dots p(x_m) \left(\sum_{i=1}^{m} \frac{x_i}{m}\right)^n \\
&= \left\langle \left(\sum_{i=1}^{m} \frac{x_i}{m}\right)^n \right\rangle \; .
\end{aligned} \tag{2.55}$$

The characteristic function of $W(y)$ reads

$$\begin{aligned}
G(k) &= \int \mathrm{d}y \, e^{-iky} W(y) \\
&= \int \mathrm{d}x_1 \dots \mathrm{d}x_m \, p(x_1) \dots p(x_m) \int \mathrm{d}y \, e^{-iky} \delta\left(y - \sum_{i=1}^{m} \frac{x_i}{m}\right) \\
&= \int \mathrm{d}x_1 \dots \mathrm{d}x_m \, p(x_1) \dots p(x_m) \, \exp\left[-ik \sum_{i} \frac{x_i}{m}\right] \\
&= \int \mathrm{d}x_1 \, p(x_1) e^{-ikx_1/m} \cdot \int \mathrm{d}x_2 \, p(x_2) e^{-ikx_2/m} \cdots \int \mathrm{d}x_m \, p(x_m) e^{-ikx_m/m} \\
&= \left[g\left(\frac{k}{m}\right)\right]^m \; ,
\end{aligned} \tag{2.56}$$

where $g(k) = \int \mathrm{d}x \, p(x) e^{-ikx}$ represents the characteristic function of $p(x)$. This implies that

$$\ln G(k) = m \ln \left[g\left(\frac{k}{m}\right)\right] \; . \tag{2.57}$$

Consequently, the cumulants are given by

$$\begin{aligned}
\langle y^n \rangle_c &= i^n \left. \frac{\mathrm{d}^n \ln G(k)}{\mathrm{d}k^n} \right|_{k=0} \\
&= i^n m \left. \frac{\mathrm{d}^n \ln g(k/m)}{\mathrm{d}k^n} \right|_{k=0} \\
&= i^n m^{1-n} \left. \frac{\mathrm{d}^n \ln g(q)}{\mathrm{d}q^n} \right|_{q=0}
\end{aligned} \tag{2.58}$$

where we have substituted $q = \frac{k}{m}$ in the final step. Now, since

$$\langle x^n \rangle_c = i^n \left. \frac{\mathrm{d}^n \ln g(q)}{\mathrm{d}q^n} \right|_{q=0}$$

by definition, we may write

$$\boxed{\langle y^n \rangle_c = m^{1-n} \langle x^n \rangle_c} \,. \tag{2.59}$$

This is the *central limit theorem.* The only assumption made here is that the n-th moments must exist. Importantly, the central limit theorem has the following implications.

- $n = 1 \quad \Rightarrow \quad \langle y^n \rangle_c = \langle x^n \rangle_c$,
  In terms of our example, the mean of the average weight of 100 people equals the mean of the weight of an individual.

- $n = 2 \quad \Rightarrow \quad \langle y^n \rangle_c = \frac{1}{m} \langle x^n \rangle_c$.
  This underpins the law of large numbers: for the distribution $W(y)$ of the sum, the mean deviation becomes much smaller than for the individual distribution $p(x)$. Thinking back to the application to fuel planning on airplanes, this presents a relieving result. While individual body weights might vary quite significantly among airplane passengers, the total weight of all $m = 100$ passengers will not deviate much from its mean, making safe estimates of needed fuel possible.

- $n = 3 \quad \Rightarrow \quad \langle y^n \rangle_c = \frac{1}{m^2} \langle x^n \rangle_c$,
  In words, deviations from the normal distribution (which are characterized by the third or higher-order cumulants, as shown in example 2.11) are very small for large m – this states the key result of the central limit theorem!

3 Statistical Mechanics

3.1 Recap: Classical mechanics

Before introducing key concepts of Statistical Mechanics, let us recapitulate some fundamentals of classical mechanics. Newton's equation of motion in one dimension is

$$m\ddot{x}(t) = F(x,t) \ , \tag{3.1}$$

where the conservative force F can be extracted from a potential V according to

$$F(x,t) = -\frac{\partial V(x,t)}{\partial x} \ . \tag{3.2}$$

In terms of the momentum $p = m\dot{x}$, eqs. (3.1) and (3.2) yield $\dot{p} + \frac{\partial V}{\partial x} = 0$ as well as

$$\begin{aligned}\dot{p} &= \frac{\mathrm{d}}{\mathrm{d}t}(m\dot{x}(t)) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial}{\partial \dot{x}}\left(\frac{m\dot{x}^2}{2}\right) \\ &= \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{x}} \ ,\end{aligned}$$

where $T = m\dot{x}^2/2$ represents the kinetic energy. Thus, we may rewrite the equation of motion (3.1) as

$$\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial T}{\partial \dot{x}} + \frac{\partial V}{\partial x} = 0 \ . \tag{3.3}$$

We define the Lagrange function L,

$$L(x,\dot{x}) := T(\dot{x}) - V(x) \ , \tag{3.4}$$

and rewrite the equation of motion in terms of L to obtain the *Euler-Lagrange equation*,

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L(x,\dot{x})}{\partial \dot{x}} - \frac{\partial L(x,\dot{x})}{\partial x} = 0} \ . \tag{3.5}$$

Moreover, let us define the action S,

$$S := \int_{t_0}^{t_1} \mathrm{d}t \, L(x(t),\dot{x}(t)) \ . \tag{3.6}$$

The action is a functional, i.e. a function of a function. For a functional $F[x(t)]$, its functional derivative[1] is defined as

$$\frac{\delta F[x(\cdot)]}{\delta x(\tilde{t})} := \left.\frac{F[x(t) + \varepsilon\delta(t-\tilde{t})] - F[x(t)]}{\varepsilon}\right|_{\varepsilon\to 0} \ . \tag{3.7}$$

---

[1] For more about functional derivatives, see the Appendix.

Thus the functional derivative of the action reads

$$\begin{aligned}\frac{\delta S[x(\cdot),\dot{x}(\cdot)]}{\delta x(\tilde{t})} &= \int_{t_0}^{t_1} \mathrm{d}t \left[\frac{\partial L}{\partial x}\delta(t-\tilde{t}) + \frac{\partial L}{\partial \dot{x}}\frac{\mathrm{d}}{\mathrm{d}t}\delta(t-\tilde{t})\right] \\ &\overset{*}{=} \int_{t_0}^{t_1} \mathrm{d}t \left[\frac{\partial L}{\partial x} - \frac{\mathrm{d}}{\mathrm{d}t}\frac{\partial L}{\partial \dot{x}}\right]\delta(t-\tilde{t}) \\ &= \frac{\partial L}{\partial x(\tilde{t})} - \frac{\mathrm{d}}{\mathrm{d}\tilde{t}}\frac{\partial L}{\partial \dot{x}(\tilde{t})} \overset{!}{=} 0 \ . \end{aligned} \tag{3.8}$$

Here the step marked by $*$ results from partial integration; in the last line we inserted the Euler-Lagrange equation. This calculation accords with Hamilton's principle which states that a solution of the Euler-Lagrange equation minimizes the action.

The Lagrangian formalism has proven to be very useful for problems with constraints. In Statistical Mechanics, however, the Hamiltonian formalism turns out more suitable. To switch from Lagrange to Hamilton, we perform a *Legendre transform* which replaces $\dot{x}$ in L by the momentum p,

$$p(t) = m\dot{x}(t) = \frac{\partial T(\dot{x})}{\partial \dot{x}} = \frac{\partial L(x,\dot{x})}{\partial \dot{x}} \ . \tag{3.9}$$

The Legendre transform yields the Hamilton function $\mathcal{H}$,

$$\boxed{\mathcal{H}(x,p,t) = \dot{x}p - L(x,\dot{x},t)} \ . \tag{3.10}$$

Note that $\mathcal{H}$ does not depend on $\dot{x}$. To see this explicitly, let us write down the total differential[2] of the Hamilton function,

$$\begin{aligned}\mathrm{d}\mathcal{H}(x,\dot{x},p,t) &= \left.\frac{\partial H}{\partial x}\right|_{\dot{x},p,t}\mathrm{d}x + \left.\frac{\partial H}{\partial \dot{x}}\right|_{x,p,t}\mathrm{d}\dot{x} + \left.\frac{\partial H}{\partial p}\right|_{x,\dot{x},t}\mathrm{d}p + \left.\frac{\partial H}{\partial t}\right|_{x,\dot{x},p}\mathrm{d}t \\ &= -\left.\frac{\partial L}{\partial x}\right|_{\dot{x},p,t}\mathrm{d}x + \underbrace{p\,\mathrm{d}\dot{x} - \left.\frac{\partial L}{\partial \dot{x}}\right|_{x,p,t}\mathrm{d}\dot{x}}_{=0} + \dot{x}\,\mathrm{d}p - \left.\frac{\partial L}{\partial t}\right|_{x,\dot{x},p}\mathrm{d}t \\ &= \mathrm{d}\mathcal{H}(x,p,t) \ . \end{aligned} \tag{3.11}$$

Due to the Legendre transform, the $\dot{x}$-dependence cancels[3] (the underbraced term vanishes according to eq. (3.9)) and the three remaining variables are (x,p,t). In general, Legendre transforms can be used to change independent variables of functions. Eq. (3.11) now leads to

$$\mathrm{d}\mathcal{H}(x,p,t) = -\dot{p}\,\mathrm{d}x + \dot{x}\,\mathrm{d}p - \frac{\partial L}{\partial t}\mathrm{d}t \ . \tag{3.12}$$

A comparison with the general form for the total differential,

$$\mathrm{d}\mathcal{H}(x,p,t) = \left.\frac{\partial H}{\partial x}\right|_{p,t}\mathrm{d}x + \left.\frac{\partial H}{\partial p}\right|_{x,t}\mathrm{d}p + \left.\frac{\partial H}{\partial t}\right|_{x,p}\mathrm{d}t \ , \tag{3.13}$$

yields the Hamilton equations,

$$\boxed{\frac{\partial \mathcal{H}}{\partial p} = \dot{x} \ , \qquad \frac{\partial \mathcal{H}}{\partial x} = -\dot{p} \ , \qquad \frac{\partial \mathcal{H}}{\partial t} = -\frac{\partial L}{\partial t}} \ . \tag{3.14}$$

---

[2] The total differential of a function $f(x,y)$ is defined as $\mathrm{d}f(x,y) = \left.\frac{\partial f}{\partial x}\right|_y \mathrm{d}x + \left.\frac{\partial f}{\partial y}\right|_x \mathrm{d}y$.

[3] Even though $\dot{x}$ still appears in the total differential (3.11) in the term $\dot{x}\,\mathrm{d}p$, it is *not* an independent variable of $\mathcal{H}$ anymore.

These equations completely specify the dynamics of a system. We can see this by calculating the total time derivative of the Hamilton function using the chain rule,

$$\begin{aligned}\frac{\mathrm{d}\mathcal{H}(x(t),p(t),t)}{\mathrm{d}t} &= \frac{\partial\mathcal{H}}{\partial x}\frac{\mathrm{d}x}{\mathrm{d}t}+\frac{\partial\mathcal{H}}{\partial p}\frac{\mathrm{d}p}{\mathrm{d}t}+\frac{\partial\mathcal{H}}{\partial t}\\ &= -\dot{p}\dot{x}+\dot{x}\dot{p}+\frac{\partial\mathcal{H}}{\partial t}\\ \Rightarrow\quad \frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t} &= \frac{\partial\mathcal{H}}{\partial t}\ . \end{aligned}\tag{3.15}$$

If the potential is constant in time such that the Hamiltonian $\mathcal{H}$ is not explicitly time-dependent, i.e. $\frac{\partial\mathcal{H}}{\partial t}=0$, then $\mathcal{H}$ is a conserved quantity. But what is $\mathcal{H}$ physically? Returning to the Legendre transform, we find

$$\begin{aligned}\mathcal{H} &= p\dot{x}-L\\ &= m\dot{x}^2-(T-V)\\ &= 2T-T+V\\ &= T+V\ .\end{aligned}\tag{3.16}$$

Thus, the Hamiltonian represents the total energy! In a time-independent potential, the total energy is conserved.

After this brief summary, let us get started with Statistical Mechanics.

3.2 Liouville's theorem

A gas of N atoms is completely specified by $3N$ position coordinates $q_1,\ldots,q_{3N}$ and $3N$ momentum coordinates $p_1,\ldots,p_{3N}$. These variables constitute a $6N$-dimensional coordinate space termed *phase space*, or Γ-space. A point in phase space defines a *microstate*. It moves according to the canonical Hamilton equations,

$$\dot{q}_i(t)=\frac{\partial\mathcal{H}(q^{3N},p^{3N})}{\partial p_i}\tag{3.17}$$

$$\dot{p}_i(t)=-\frac{\partial\mathcal{H}(q^{3N},p^{3N})}{\partial q_i}\ ,\tag{3.18}$$

where $i=1,\ldots,3N$ ($6N$ equations) and q^{3N} abbreviates $q_1,q_2,\ldots,q_{3N}$. Typically, we assume that $\mathcal{H}$ depends only on q^{3N} and p^{3N}, not on time t, $\dot{p}$, etc. Then,

$$\frac{\mathrm{d}\mathcal{H}}{\mathrm{d}t}=\frac{\partial\mathcal{H}}{\partial t}=0\ .$$

As shown in the previous section, this implies conservation of total energy $\mathcal{H}$. Therefore, with given initial conditions, eqs. (3.17) and (3.18) uniquely determine $q^{3N}(t)$ and $p^{3N}(t)$ for all times t. This also means that $\dot{q}^{3N}(t)$ and $\dot{p}^{3N}(t)$ are unique, which brings about the consequence that trajectories in phase space never cross. Indeed, if a trajectory would cross itself in a certain point, there would exist two different ways to move on from that point – in contradiction with the uniqueness of solutions. Closed curves, however, are possible and describe periodic motion.

In Statistical Mechanics, where we deal with large systems, following $6N$ coordinates around is neither feasible nor worthwhile; it's just too much information. But we don't need all details because not all microstates have distinct physical characteristics. Imagine, for example, a gas composed of N identical particles. Then all $N!$ possible permutations describe identical systems but represent different points in phase space! Many microstates are equivalent and lead to the samemacroscopic properties such as temperature, pressure, total energy, etc. In other words, one *macrostate* characterized by these "coarse-grain" quantities can be realized by a vast number of microstates. This motivates a probabilistic description of the problem. We assign each microstate a probability $\rho(q^{3N}, p^{3N}, t)$. The probability density can be normalized such that

$$\int \mathrm{d}^{3N}q\, \mathrm{d}^{3N}p\ \rho(q^{3N}, p^{3N}, t) = 1\ , \tag{3.19}$$

where, in short notation, $\mathrm{d}^{3N}q = \mathrm{d}q_1 \cdots \mathrm{d}q_{3N}$. Thus, we may calculate the expectation value of any observable A, e.g. pressure or internal energy, according to

$$\langle A(t)\rangle = \int \mathrm{d}^{3N}q\, \mathrm{d}^{3N}p\ A(q^{3N}, p^{3N})\, \rho(q^{3N}, p^{3N}, t)\ . \tag{3.20}$$

A probability distribution ρ in phase space is also called an *ensemble*, and the expectation value in eq. (3.20) is sometimes termed *ensemble average*. We may think of an ensemble as a collection of "mental copies" of the given system, each of which represents a possible microstate the system could be in under the constraints of a given macrostate. This powerful interpretation, introduced by Gibbs, sets the conceptual foundation for much of what will follow in this lecture. We will discuss different types of ensembles soon.

What are the properties of the density distribution ρ, and how does it evolve in time? A microstate (q^{3N}, p^{3N}), with probability $\rho(q^{3N}, p^{3N}, t)$ at time t, moves along a completely determined trajectory according to the canonical Hamilton equations (3.17), (3.18). Since probability is conserved over time when integrating over the entire phase space (there are no sources or sinks of microstates), ρ must obey a so-called conservation or balance equation. Consider a fixed volume Ω in phase space: the change in probability ρ integrated over Ω must be balanced by a probability flux through the surface $S(\Omega)$ of the volume Ω. In mathematical terms, we write this as

$$-\frac{\mathrm{d}}{\mathrm{d}t}\int_\Omega \mathrm{d}^{3N}q\, \mathrm{d}^{3N}p\ \rho(q^{3N}, p^{3N}, t) = \int_{S(\Omega)} \mathrm{d}s\ \vec{n}(s)\cdot\vec{v}(s)\rho(s,t)\ , \tag{3.21}$$

where $\vec{v}(s) = (\dot{q}_1, \ldots, \dot{q}_{3N}, \dot{p}_1, \ldots, \dot{p}_{3N})$ denotes the $6N$-dimensional velocity in phase space, s is the surface element, and $\vec{n}(s)$ denotes the surface normal on S (pointing outwards). Using Gauss's theorem, sometimes referred to as the divergence theorem, we rewrite eq. (3.21) as

$$-\int_\Omega \mathrm{d}^{3N}q\, \mathrm{d}^{3N}p\ \frac{\partial}{\partial t}\rho(q^{3N}, p^{3N}, t) = \int_\Omega \mathrm{d}^{3N}q\, \mathrm{d}^{3N}p\ \vec{\nabla}\cdot\Big(\vec{v}(q^{3N}, p^{3N})\, \rho(q^{3N}, p^{3N}, t)\Big)\ . \tag{3.22}$$

Here $\vec{\nabla}$ represents the $6N$-dimensional gradient operator,

$$\vec{\nabla} = \left(\frac{\partial}{\partial q_1}, \ldots, \frac{\partial}{\partial q_{3N}}, \frac{\partial}{\partial p_1}, \ldots, \frac{\partial}{\partial p_{3N}}\right)\ .$$

Note that we could pull the time derivative inside the volume integral since we assume that Ω remains fixed in time. Rearranging eq. (3.22), we get

$$\int_\Omega \mathrm{d}^{3N}q\, \mathrm{d}^{3N}p\ \left(\frac{\partial}{\partial t}\rho(q^{3N}, p^{3N}, t) + \vec{\nabla}\cdot\Big(\vec{v}(q^{3N}, p^{3N})\, \rho(q^{3N}, p^{3N}, t)\Big)\right) = 0\ . \tag{3.23}$$

We may choose the time-independent volume Ω arbitrarily, which implies that eq. (3.23) is only true for any test volume Ω if the integrand vanishes. This requirement results in an equation of continuity for the ρ,

$$\begin{aligned}
-\frac{\partial \rho}{\partial t} &= \vec{\nabla}\cdot(\vec{v}\rho) \\
&= \sum_{i=1}^{3N}\left[\frac{\partial}{\partial q_i}(\dot{q}_i\rho) + \frac{\partial}{\partial p_i}(\dot{p}_i\rho)\right] \\
&= \sum_{i=1}^{3N}\left[\dot{q}_i\frac{\partial\rho}{\partial q_i} + \dot{p}_i\frac{\partial\rho}{\partial p_i} + \rho\underbrace{\left(\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\dot{p}_i}{\partial p_i}\right)}_{=0}\right]\ ,
\end{aligned} \tag{3.24}$$

where we know from the Hamilton equations (3.17), (3.18) that the underbraced term must vanish:

$$\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i} = \frac{\partial}{\partial q_i}\frac{\partial \mathcal{H}}{\partial p_i} - \frac{\partial}{\partial p_i}\frac{\partial \mathcal{H}}{\partial q_i} = 0 \ ,$$

according to Schwarz's theorem about the symmetry of partial derivatives. This leads to the *Liouville equation*,

$$-\frac{\partial \rho(q^{3N}, p^{3N}, t)}{\partial t} = \sum_{i=1}^{3N} \left[\dot{q}_i \frac{\partial}{\partial q_i} + \dot{p}_i \frac{\partial}{\partial p_i}\right] \rho \ . \tag{3.25}$$

Defining the Liouville operator $\hat{L} := \sum_{i=1}^{3N} \left[\dot{q}_i \frac{\partial}{\partial q_i} + \dot{p}_i \frac{\partial}{\partial p_i}\right]$, we may write this compactly as[4]

$$\boxed{-\frac{\partial \rho}{\partial t} = \hat{L}\rho} \ . \tag{3.26}$$

The Liouville equation describes how ρ changes in time at a fixed position in phase space $(q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N})$ (Eulerian description). We might ask ourselves: how does ρ evolve in time when we "go with the flow" by moving along with the trajectory as it moves through phase space (Lagrangian description)? Using the chain rule, we find that

$$\begin{aligned} \frac{\mathrm{d}\rho(q^{3N}(t), p^{3N}(t), t)}{\mathrm{d}t} &= \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i}\frac{\mathrm{d}q_i}{\mathrm{d}t} + \frac{\partial \rho}{\partial p_i}\frac{\mathrm{d}p_i}{\mathrm{d}t}\right) \\ &= \frac{\partial \rho}{\partial t} + \hat{L}\rho \\ &\Rightarrow \quad \boxed{\frac{\mathrm{d}\rho}{\mathrm{d}t} = 0} \ . \end{aligned} \tag{3.27}$$

This is *Liouville's theorem.* It states that the density ρ is (locally) constant when moving along a trajectory through phase space. Essentially, ρ propagates through phase space like an incompressible fluid.

3.3 Postulate of equal *a priori* probability

If we imagine that a trajectory visits every point in phase space compatible with energy conservation ($\mathcal{H} = \text{const.}$), then the Liouville theorem tells us that

$$\rho(\vec{q}, \vec{p}, t) = \begin{cases} \text{const.} & U \leq \mathcal{H}(\vec{q}, \vec{p}) \leq U + \Delta \\ 0 & \text{otherwise.} \end{cases} \tag{3.28}$$

Here we allow the total energy $\mathcal{H}$ to lie in a small interval $[U, U + \Delta]$ to comply with the uncertainty principle. According to this argument – which only assumes that a trajectory may, in principle, visit every point in phase space without crossing itself –, every possible microstate which results in total energy U is equally probable. In other words, a system with total energy U is equally likely to be in any of the compatible microstates. This defines the *microcanonical ensemble*, which we will discuss in further detail soon. Since the total energy is fixed, the microcanonical ensemble describes an isolated system.

---

[4] As a side remark, we mention that the Liouville operator may be used to formally solve the Liouville equation (3.25). Given an initial distribution $\rho(\vec{q}, \vec{p}, 0)$, the solution is $\rho(\vec{q}, \vec{p}, t) = \exp\big[-\hat{L}t\big]\rho(\vec{q}, \vec{p}, 0)$. This is an extremely powerful method, since it provides an exact solution of the many-body problem!

Remark. If we more generally consider an ensemble where different energies are populated with different probabilities, that is,

$$\rho(\vec{q},\vec{p},t) = \rho(\mathcal{H}(\vec{q},\vec{p}),t) \ , \tag{3.29}$$

then the Liouville equation tells us that

$$\begin{aligned}
-\frac{\partial\rho(\vec{q},\vec{p},t)}{\partial t} &= \sum_{i=1}^{3N}\left[\dot{q}_i\frac{\partial\rho}{\partial q_i}+\dot{p}_i\frac{\partial\rho}{\partial p_i}\right] \\
&= \sum_{i=1}^{3N}\left[\frac{\partial\mathcal{H}}{\partial p_i}\frac{\partial\rho}{\partial\mathcal{H}}\frac{\partial\mathcal{H}}{\partial q_i}-\frac{\partial\mathcal{H}}{\partial q_i}\frac{\partial\rho}{\partial\mathcal{H}}\frac{\partial\mathcal{H}}{\partial p_i}\right] \\
&= 0 \ ,
\end{aligned} \tag{3.30}$$

where we have used the chain rule. This implies that ρ cannot be explicitly time-dependent:

$$\rho(\mathcal{H}(\vec{q},\vec{p}),t) = \rho(\mathcal{H}(\vec{q},\vec{p})) \ . \tag{3.31}$$

Consequently, a distribution that depends on the Hamiltonian only and not on phase space variables directly is stationary. This statement is true for arbitrary systems and for a general Hamiltonians $\mathcal{H}(\vec{q},\vec{p})$.

3.4 Concepts of entropy and temperature

The "number" of microstates within the energy range $U < \mathcal{H}(\vec{q},\vec{p}) < U+\Delta$ is given by the volume integral

$$\Gamma(U,V,N) = \int_{U<\mathcal{H}<U+\Delta} \mathrm{d}^{3N}q\,\mathrm{d}^{3N}p\ 1 \ , \tag{3.32}$$

which quantifies the allowed volume in phase space of a system with that energy. This relates to the *density of states*: the total number of states with any energy $\mathcal{H}$ below U is

$$\Sigma(U) = \int_{\mathcal{H}<U} \mathrm{d}^{3N}q\,\mathrm{d}^{3N}p\ 1 \ . \tag{3.33}$$

Thus, we may write the number of allowed microstates Γ as

$$\begin{aligned}
\Gamma(U) &= \Sigma(U+\Delta)-\Sigma(U) \\
&\approx \Sigma(U)+\Delta\Sigma'(U)-\Sigma(U) \\
&= \Delta\Sigma'(U) \ ,
\end{aligned} \tag{3.34}$$

where we used the Taylor approximation up to first order, assuming that $\Delta \ll U$. Based on this, we define the density of states $\omega(U)$ for a certain energy U as

$$\omega(U) := \lim_{\Delta\to 0}\frac{\Gamma(U)}{\Delta} = \Sigma'(U) \ , \tag{3.35}$$

$$\Sigma(U) = \int_{-\infty}^{U} \mathrm{d}U'\,\omega(U') \ . \tag{3.36}$$

Two Coupled Systems. Consider now an isolated system, characterized by fixed total energy U, particle number N, and volume V. Imagine that we insert a dividing plane which divides the system into two subsystems. The dividing plane lets energy pass from one side to the other; however, neither particles nor volume may pass through. Thus, subsystems 1 and 2 have fixed particle numbers N_1, $N - N_1$ as well as fixed volumes V_1, $V - V_1$ but fluctuating energies $U_1, U - U_1$ under the constraint $U_1 \leq U$ (see fig. 3.1).

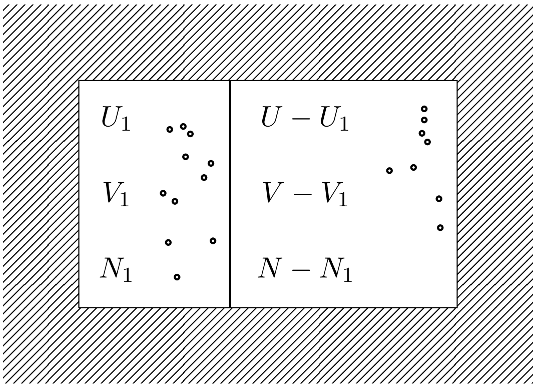


Figure 3.1: Two coupled subsystems isolated from the environment

The total number of microstates Γ for the compound system at energy U is calculated as

$$\Gamma(U, V, N) = \int \mathrm{d}U_1 \, \Gamma_1(U_1) \cdot \Gamma_2(U - U_1) \ . \tag{3.37}$$

Note that for a fixed energy U_1, the number of microstates Γ of the total system equals the product $\Gamma_1 \cdot \Gamma_2$ of the number of microstates of the subsystems. This is founded upon the multiplication theorem (see eq. (2.3)). Since the subsystems may exchange energy, we must furthermore integrate over all possible energies U_1 (from 0 to U).

Let us now introduce a new function which we label $S(U)$. The reason why we do this will become clear later; it will turn out that S represents the *entropy* in Thermodynamics. We define

$$\boxed{\frac{S(U)}{k_B} := \ln(\Gamma(U))} \ . \tag{3.38}$$

Let us use this function to explore our example of a compound system further. Inserting eq. (3.37) for $\Gamma(U)$, we obtain

$$\frac{S(U)}{k_B} = \ln\left(\int \mathrm{d}U_1 \, e^{[S_1(U_1) + S_2(U - U_1)]/k_B} \right) \ . \tag{3.39}$$

Assuming that the integrand has a maximum at energy U_1^*, the sum $S_1 + S_2$ is also maximal at U_1^* because the exponential is a monotonic equation. Therefore, we may Taylor expand the sum around its maximum,

$$\begin{aligned} S_1(U_1) + S_2(U - U_1) \approx {} & S_1(U_1^*) + S_2(U - U_1^*) \\ & + (U_1 - U_1^*) \underbrace{\left[S_1'(U_1^*) - S_2'(U - U_1^*)\right]}_{=0} \\ & + \frac{(U_1 - U_1^*)^2}{2} \left[S_1''(U_1^*) + S_2''(U - U_1^*)\right] + \dots \ , \end{aligned} \tag{3.40}$$

where the first derivative vanishes at the maximum. Re-inserting the Taylor series up to second order into eq. (3.39) yields

$$\frac{S(U)}{k_B} = \ln\left(e^{[(S_1(U_1^*) + S_2(U - U_1^*)]/k_B} \int \mathrm{d}U_1 \, e^{(U_1 - U_1^*)^2 [S_1''(U - U_1^*) + S_2''(U - U_1^*)]/(2k_B)} \right) \ .$$

Here we have pulled the U_1-independent part in front of the integral. The integral itself is a Gaussian integral for which the solution is known: $\int \mathrm{d}x\, e^{-a(x+b)^2} = \sqrt{\pi/a}$. Performing the integral, we find

$$S(U) = S_1(U_1^*) + S_2(U - U_1^*) + k_B \ln \sqrt{-\frac{2\pi k_B}{S_1''(U_1^*) + S_2''(U - U_1^*)}} \ , \tag{3.41}$$

where we assume that $S_1''(U_1*) + S_2''(U - U_1^*) < 0$ (which will be shown to hold true). From this result we may draw several fundamental conclusions:

- **The function $S(U)$ is extensive.** This means that the entropy of the total system equals the sum of the individual entropies of the subsystems. In other words, $S(U)$ is proportional to the system size N. This statement is true because in fact the square root term in eq. (3.41) becomes negligible for large N. To see this, imagine that each particle in a system contributes an equal share of entropy to the entropy S of the whole system, such that $S = Ns$, where the lower-case s denotes the per-particle entropy. We can similarly define a per-particle energy u such that $U = Nu$ (s and u are independent of N). This implies

  $$\frac{\mathrm{d}^2 S(U)}{\mathrm{d}U^2} = \frac{\mathrm{d}^2 Ns(u)}{\mathrm{d}(Nu)^2} = \frac{1}{N}\frac{\mathrm{d}^2 s(u)}{\mathrm{d}u^2} \quad \Rightarrow \quad S''(U) \sim \frac{1}{N} \ .$$

  Thus, the second derivative of $S(U)$ scales like $1/N$, and the square root term consequently scales according to

  $$k_B \ln \sqrt{-\frac{2\pi k_B}{S_1''(U_1^*) + S_2''(U - U_1^*)}} \sim \ln \sqrt{N} \ .$$

  For large N, we have $N \gg \ln \sqrt{N}$, such that the square root term approaches zero and $S(U)$ is an extensive function, i.e.,

  $$S(U) = S_1(U_1^*) + S_2(U - U_1^*) \ . \tag{3.42}$$

  Eq. (3.42) is exact in the thermodynamic limit ($N \to \infty$).

  In Thermodynamics, we generally distinguish between *extensive* quantities which scale with the system size N and *intensive* quantities which are independent of N. Other examples of extensive variables include the total system the volume V, energy U, and obviously N. Contrarily, the temperature T or pressure P belong to the intensive variables.

- **In equilibrium, $S(U)$ is maximal.** The function $S(U) \approx S_1(U_1^*) + S_2(U - U_1^*)$ is maximized with respect to the free variable U_1. Since S relates to the phase space volume via (3.38), maximizing S also maximizes Γ. Therefore, the composite system automatically goes to the state where $U_1 = U_1^*$ because this is the most probable one in equilibrium. This corresponds to the *Second Law of Thermodynamics*: in equilibrium, the entropy S is maximal with respect to all "free" variables (under all constraints imposed by the macrostate).

- **Systems in contact exchange energy until equilibrium.** In equilibrium, where the energy of subsystem 1 has reached $U_1 = U_1^*$, we find that $S'(U)$ is the same for both subsystems:

  $$\frac{\mathrm{d}S_1(U_1^*)}{\mathrm{d}U_1^*} = \frac{\mathrm{d}S_2(U_2^*)}{\mathrm{d}U_2^*} \ ,$$

  where $U_2^* = U - U_1*$. This suggests that in equilibrium the derivative S' takes a constant value, which we define as

  $$\boxed{\frac{1}{T} := \frac{\mathrm{d}S(U)}{\mathrm{d}U}} \ . \tag{3.43}$$

You may have heard of T; in Thermodynamics it is called *temperature*. We deduce that systems in thermal contact must have the same temperature (after sufficient time, when equilibrium is reached). Otherwise, if two systems in thermal contact would not have the same temperature, they would exchange energy until the two derivatives $\mathrm{d}S_1/\mathrm{d}U_1^*$ and $\mathrm{d}S_2/\mathrm{d}U_2^*$ are the same. This corresponds to the *Zeroth Law of Thermodynamics.*

3.5 Canonical ensemble

As in the previous section, let us consider two systems in thermal contact (fig. (3.1)). Their combined energy U, particle number N, and volume V are fixed. The particle number and volume of system 1, N_1 and V_1, are also fixed. However, the systems may exchange energy, such that U_1 may fluctuate. Once again, the "number" of microstates for a given energy U_1 is given by the multiplication theorem,

$$\Gamma(U, U_1) = \Gamma_1(U_1) \cdot \Gamma_2(U - U_1) \ . \tag{3.44}$$

We now argue the following. If a microstate of system 1 allows for relatively many microstates of the composite system, then that microstate of system 1 will have a relatively high probability. In other words, the number of microstates of the composite system per microstate of system 1 equals the (non-normalized) probability of that microstate of system 1:

$$\begin{aligned}
\rho(U_1, N_1, V_1) &= \frac{\Gamma(U_1, V_1, N_1) \cdot \Gamma(U - U_1, V - V_1, N - N_1)}{\Gamma(U_1, V_1, N_1)} \\
&= \Gamma(U - U_1, V - V_1, N - N_1) \\
&= \exp\left[\frac{1}{k_B} S(U - U_1, V - V_1, N - N_1)\right] \\
&= \exp\left[\frac{1}{k_B} S(U) - \frac{U_1}{k_B} S'(U) + \frac{U_1^2}{2k_B} S''(U) + \ldots\right] \ ,
\end{aligned} \tag{3.45}$$

where ρ is the non-normalized probability distribution of system 1. Here we used eq. (3.38) and expanded $S(U - U_1, V - V_1, N - N_1)$ around $U_1 = 0$. Let us now assume that system 2 is much larger than system 1, i.e. $N_2 = N - N_1 \gg N_1$, $V_2 = V - V_1 \gg V_1$, and $U_2 = U - U_1 \gg U_1$. Then, the second-order term of the Taylor expansion is negligible because

$$\frac{U_1^2}{2k_B} \frac{\partial^2 S(U)}{\partial U^2} \sim \frac{N_1^2}{N} \to 0 \quad (N_1 \ll N) \ .$$

Similar arguments hold for higher-order terms of the expansion. Furthermore, we know from the previous section that

$$\frac{dS(U)}{\mathrm{d}U} = \frac{1}{T} \ ,$$

where T is the temperature. Inserting this into eq. (3.45), we obtain the *Boltzmann distribution*,

$$\boxed{\rho(U_1) \propto \exp\left[-\frac{U_1}{k_B T}\right]} \ . \tag{3.46}$$

Here we simply discarded the constant factor $e^{S(U)/k_B}$ coming from the zeroth order of the expansion in eq. (3.45), as the distribution is not normalized anyway.

Above we assumed that system 2 is much larger than system 1, such that $N - N_1 \approx N$. In that case, we call system 2 a *reservoir* or heat bath – a system which is so large that its extensive properties like N, V, U essentially do not change when put in contact with a small system like system 1. The Boltzmann distribution, sometimes termed canonical distribution, thus gives the probability distribution for a closed system that may exchange energy with a reservoir. It defines the so-called *canonical ensemble* which is characterized by the macroscopic variables N, V, T. The fixed temperature T of the reservoir controls the energy of the closed system. All states are possible, even states with extremely high energy U_1, but they will be extremely unlikely unless the heat bath is very, very hot.

Expectation values and variances. We now want to calculate the mean energy and its variance for a system governed by the canonical distribution $\rho(U_i)$. Let us suppose the system has discrete states i with energy U_i. Then the expectation value reads

$$\langle U \rangle = \frac{\sum_i U_i \, \rho(U_i)}{\sum_i \rho(U_i)} = \frac{\sum_i U_i e^{-\frac{U_i}{k_B T}}}{\sum_i e^{-\frac{U_i}{k_B T}}} \ . \tag{3.47}$$

It is convenient to define $\beta := \frac{1}{k_B T}$ for increased simplicity. In terms of β, we write

$$\begin{aligned} \langle U \rangle &= \frac{\sum_i U_i e^{-\beta U_i}}{\sum_i e^{-\beta U_i}} \\ &= \frac{-\frac{\partial}{\partial \beta} \sum_i e^{-\beta U_i}}{\sum_i e^{-\beta U_i}} \\ &= -\frac{\partial}{\partial \beta} \ln \left(\sum_i e^{-\beta U_i} \right) \ , \end{aligned} \tag{3.48}$$

where, in the final step, we have used the chain rule “backwards” to arrive at the logarithm. Thus, the expectation value of the energy in the canonical ensemble is given by

$$\boxed{\langle U \rangle = -\frac{\partial}{\partial \beta} \ln Z \ , \qquad Z = \sum_i e^{-\beta U_i}} \ . \tag{3.49}$$

Here we introduced the *partition function* Z. It states the sum of Boltzmann weights $e^{-\beta U_i}$ over all microstates. The partition function is the central object in statistical mechanics; it allows to calculate any expectation values, variances, etc. by means of suitable derivatives. To obtain the variance of the energy, for example, we calculate

$$\begin{aligned} \frac{\partial^2}{\partial \beta^2} \ln(Z) &= -\frac{\partial}{\partial \beta} \langle U \rangle \\ &= -\frac{\partial}{\partial \beta} \left(\frac{\sum_i U_i e^{-\beta U_i}}{\sum_i e^{-\beta U_i}} \right) \\ &= \frac{\sum_i U_i^2 e^{-\beta U_i}}{\sum_i e^{-\beta U_i}} + \frac{\sum_i U_i e^{-\beta U_i}}{(\sum_i e^{-\beta U_i})^2} \frac{\partial}{\partial \beta} \sum_i e^{-\beta U_i} \\ &= \frac{\sum_i U_i^2 e^{-\beta U_i}}{\sum_i e^{-\beta U_i}} - \frac{(\sum_i U_i e^{-\beta U_i})^2}{(\sum_i e^{-\beta U_i})^2} \\ &= \langle U^2 \rangle - \langle U \rangle^2 = \Delta U^2 \ . \end{aligned} \tag{3.50}$$

Connection to Thermodynamics. The expectation value of the energy $\langle U \rangle$, as derived here from Statistical Mechanics, is a macroscopic observable that we would also deal with in Thermodynamics. How does the variance, which describes deviations from the mean, connect to Thermodynamics?

Remembering that β depends on the temperature T, the variance can be written as

$$\begin{aligned}\langle U^2\rangle - \langle U\rangle^2 &= \frac{\partial}{\partial\beta}\langle U\rangle \\ &= -\frac{\partial T}{\partial\beta}\frac{\partial}{\partial T}\langle U\rangle \\ &= -\frac{\partial}{\partial\beta}\frac{1}{k_B\beta}\frac{\partial}{\partial T}\langle U\rangle \\ &= k_B T^2\frac{\partial}{\partial T}\langle U\rangle \ . \end{aligned} \tag{3.51}$$

Moreover, we introduce the *heat capacity* C_V at constant volume as the derivative of energy by temperature,

$$C_V = \left.\frac{\partial\langle U\rangle}{\partial T}\right|_{V,N} \ . \tag{3.52}$$

The heat capacity quantifies how the system's energy increases with increasing temperature. More specifically, it states the amount of heat needed to increase the temperature by one Kelvin. C_V is always positive and extensive; we need to double the heat to warm up double the amount!

Example 3.1 (Heating water)*.* At 15°C, the specific heat capacity of water is $C_V = 4.2\frac{\text{kJ}}{\text{kg K}}\cdot m$, where m denotes the water mass in kilograms. So to heat up 1 kg of water from $T = 0$°C (273 K) to $T = 100$°C (373 K), we need roughly 420 kJ of energy (C_V is temperature-dependent). A 1000 W heater adds 1 kJ per second and would thus take 7 minutes to boil freezing water. In comparison, the same amount of energy would suffice to lift up a weight of 1 kg by 42 kilometers (1 kg excerts 10 N, 1 J=1 N·m, 420 kJ=10 N · 42 km). This demonstrates that heat is costly!

Coming back to the variance, we now combine eqs. (3.51) and (3.52) to obtain

$$\langle U^2\rangle - \langle U\rangle^2 = k_B T^2 C_V \ . \tag{3.53}$$

Accordingly, energy deviations are given by

$$\Delta U = \sqrt{\langle U^2\rangle - \langle U\rangle^2} = T\sqrt{k_B C_V} \ , \tag{3.54}$$

and the relative energy deviations become

$$\frac{\Delta U}{U} = \frac{T\sqrt{k_B C_v}}{U} \sim \frac{\sqrt{N}}{N} \sim \frac{1}{\sqrt{N}} \to 0 \quad (N\to\infty) \ . \tag{3.55}$$

Once again, this result reflects the law of large numbers: as $N \to \infty$, relative deviations in the energy approach zero. In the thermodynamic limit, the canonical ensemble (where the energy may fluctuate) and the microcanonical ensemble (where the energy is fixed) are equivalent because fluctuations around the mean vanish. Thermodynamics does not know about these fluctuations, but Statistical Mechanics does! Variances are determined by response functions (Fluctuation-dissipation theorem).

Concept of free energy. We have introduced the canonical partition function $Z(T, V, N) = \sum_i e^{-\beta\mathcal{H}_i}$, where $\beta = \frac{1}{k_B T}$ and the index i sums over all microstates. Calculating or approximating this sum is difficult. Therefore, we look for an expression of the partition function which circumvents the explicit summation over microstates. Let us insert unity, i.e.

$$1 = \int_{-\infty}^{\infty} \mathrm{d}U\,\delta(U - \mathcal{H}_i) \ , \tag{3.56}$$

into the expression of the partition function:

$$\begin{aligned} Z &= \sum_i \int_{-\infty}^{\infty} \mathrm{d}U \, \delta(U - \mathcal{H}_i) e^{-\beta \mathcal{H}_i} \\ &= \int_{-\infty}^{\infty} \mathrm{d}U \sum_i \delta(U - \mathcal{H}_i) e^{-\beta U} \\ &= \int_{-\infty}^{\infty} \mathrm{d}U \, \omega(U) e^{-\beta U} \\ &= \int_{-\infty}^{\infty} \mathrm{d}U \, e^{\ln \omega(U) - \beta U} \ . \end{aligned} \tag{3.57}$$

Here $\omega(U) := \sum_i \delta(U - \mathcal{H}_i)$ denotes the density of states, which we may interpret as the number of states with energy U. To verify this, consider the following "intuitive proof." The total number of states with energy $\mathcal{H}_i < U$ is given by the function $\Sigma(U, V, N)$, and should likewise be found by summing 1 over all possible states. Using eq. (3.36), we show that[5]

$$\begin{aligned} \Sigma(U) &= \int_{-\infty}^{U} \mathrm{d}U' \, \omega(U') \\ &= \int_{-\infty}^{U} \mathrm{d}U' \sum_i \delta(U' - \mathcal{H}_i) \\ &= \sum_{\substack{i \\ \mathcal{H}_i < U}} \int_{-\infty}^{\infty} \mathrm{d}U' \, \delta(U' - \mathcal{H}_i) \\ &= \sum_{\substack{i \\ \mathcal{H}_i < U}} 1 \ , \end{aligned} \tag{3.58}$$

as required. In section 3.4, we defined the entropy $S(U)$ as

$$\begin{aligned} & \frac{S(U)}{k_B} := \ln \Gamma(U) = \ln(\omega(U)\Delta) \\ \Rightarrow \quad & \frac{S(U)}{k_B} = \ln \omega(U) + \underbrace{\ln \Delta}_{\to 0} \ . \end{aligned}$$

Here we simply discard the non-extensive constant $\ln \Delta$, which becomes irrelevant for large N. Now, inserting $\ln \omega(U) = S(U)/k_B$ into eq. (3.57), the partition function reads

$$Z = \int_{-\infty}^{\infty} \mathrm{d}U \, e^{-\beta[U - TS(U)]} \ . \tag{3.59}$$

We define $F := U - TS(U)$, yielding

$$\boxed{Z = \int_{-\infty}^{\infty} \mathrm{d}U \, e^{-\beta F(U)} = \int_{-\infty}^{\infty} \mathrm{d}U \, \rho(U) \ , \qquad F = U - TS} \ . \tag{3.60}$$

This expression does not include a sum over all microstates anymore. The partition function is now given in terms of the *free energy* F, which is the relevant potential in the canonical ensemble.

To learn more about the free energy, let us Taylor expand F around the most probable energy

---

[5] Don't get confused with the function Σ and the sum $\sum_i$.

U^*, as we did earlier with the entropy S. For the first derivate evaluated at U^*, we find

$$\begin{aligned} \left.\frac{\mathrm{d}F(U)}{\mathrm{d}U}\right|_{U^*} &= \left.\frac{\mathrm{d}}{\mathrm{d}U}(U - TS(U))\right|_{U^*} \\ &= 1 - T\left.\frac{dS(U)}{\mathrm{d}U}\right|_{U^*} \\ &= 1 - T\,\frac{1}{T} \\ &= 0\ . \end{aligned} \tag{3.61}$$

The fact that the first derivative vanishes implies that the most probable energy U^* also extremizes the free energy F. Note that we deal with two kinds of temperatures here. We must distinguish between the temperature of the reservoir, $T = T(U^*)$, which is independent of U, and the temperature of the canonical system, $T_{can} = T(U)$. In thermal equilibrium (at U^*), both temperatures coincide and we simply denote them by T. Let us also calculate the second derivative,

$$\begin{aligned} \left.\frac{\mathrm{d}^2F(U)}{\mathrm{d}U^2}\right|_{U^*} &= -T\left.\frac{\mathrm{d}^2S(U)}{\mathrm{d}U^2}\right|_{U^*} \\ &= -T\left.\frac{\mathrm{d}}{\mathrm{d}U}\frac{1}{T(U)}\right|_{U^*} \\ &= \left.\frac{T}{T^2(U)}\,\frac{\mathrm{d}T(U)}{\mathrm{d}U}\right|_{U^*} \\ &= \left.\frac{T}{T^2(U)}\left(\frac{\mathrm{d}U}{\mathrm{d}T}\right)^{-1}\right|_{T(U^*)} \\ &= \frac{1}{T\cdot C_V} \sim \frac{1}{N} > 0\ . \end{aligned} \tag{3.62}$$

The inequality in the last line holds for any substance; energy must be added to raise the temperature. We draw the following conclusions from the preceding investigation.

- The most probable energy U^* is determined by the extremum of the free energy F where $\left.\frac{\mathrm{d}F}{\mathrm{d}U}\right|_{U^*} = 0$. At this energy, the temperature of the canonical system coincides with the temperature T of the reservoir: $\left.\frac{\mathrm{d}S}{\mathrm{d}U}\right|_{U^*} = \frac{1}{T}$.
- We know that the most probable energy U^* exists because $\left.\frac{\mathrm{d}^2F}{\mathrm{d}U^2}\right|_{U^*} > 0$. This result also implies, together with $\left.\frac{\mathrm{d}F}{\mathrm{d}U}\right|_{U^*} = 0$, that F has a minimum at U^*.
- The variance of the energy, $\langle(U - U^*)^2\rangle \approx C_V \sim N$ is extensive.
- The relative variance becomes negligible for large N: $\left\langle\left(\frac{U}{U^*} - 1\right)^2\right\rangle \sim \frac{1}{N} \to 0$ $(N \to \infty)$.
- We have shown that $\frac{\mathrm{d}^2S}{\mathrm{d}U^2} < 0$, as used in the derivation of the entropy as an extensive function (see sec. 3.4). Thus, the curvature of the entropy is positive. The most probable state of the system is characterized by a balance between minimizing the energy and maximizing the entropy. At this equilibrium point, the free energy is minimal.

Since the relative variance of the energy U approaches zero for very large N, the distribution of energies is sharply peaked at the dominating energy U^*. In the thermodynamic limit, in fact, energies other than U^* become irrelevant; the distribution approaches a δ-distribution. For the partition function Z this means

$$Z = \int \mathrm{d}U \, e^{-\beta F(U)} \overset{N\to\infty}{=} e^{-\beta F(U^*)} \ . \tag{3.63}$$

Rearranging this equation for F yields

$$\boxed{F = -k_B T \ln Z} \ . \tag{3.64}$$

Though derived in the thermodynamic limit, this result for the free energy F also holds for finite N. The mean energy $\langle U \rangle = U^*$ can by obtained by the following recipe: We combine $\beta F = -\ln Z$ (eq. (3.64)) with the previously established result

$$\begin{aligned}\langle U \rangle &= -\frac{\partial \ln Z}{\partial \beta} \\ &= \frac{\int \mathrm{d}U \, U e^{-\beta F(U)}}{\int \mathrm{d}U \, e^{-\beta F(U)}} \\ &= \frac{\int \mathrm{d}U \, U e^{-\beta U + S(U)/k_B}}{\int \mathrm{d}U \, e^{-\beta U + S(U)/k_B}}\end{aligned}$$

to obtain

$$\langle U \rangle = \frac{\partial (F\beta)}{\partial \beta} = F + \beta \frac{\partial F}{\partial \beta} = F + \frac{1}{k_B T} \frac{\partial T}{\partial \beta} \frac{\partial F}{\partial T} \ . \tag{3.65}$$

The derivative $\partial T / \partial \beta$ is the reciprocal of

$$\frac{\partial \beta}{\partial T} = \frac{\partial (1/(k_B T))}{\partial T} = -\frac{1}{k_B T^2} \ ,$$

such that $\langle U \rangle$, often simply written as U while tacitly referring to the mean value, becomes

$$\begin{aligned} U = F - T\frac{\partial F}{\partial T} \ , \quad F = U - TS \\ \Rightarrow \quad \boxed{\frac{\partial F}{\partial T} = -S} \ . \end{aligned} \tag{3.66}$$

The boxed equation states an important thermodynamic relation between free energy and entropy. Another central relation that we already know is $\partial S / \partial U = 1/T$ (eq. (3.43)).

In the canonical ensemble, the equilibrium energy $U \equiv \langle U \rangle = U^*$ is a function of the temperature of the reservoir, $U(T)$. Therefore, functions like entropy or the free energy depend on temperature T, particle number N, and volume V:

$$S = S(T,V,N) \ , \quad F = F(T,V,N) \ , \quad U(T,V,N) \ , \quad Z(T,V,N) \ , \quad \ldots$$

3.6 Examples of the canonical distribution

3.6.1 Velocity distribution of ideal gas atoms

The Hamiltonian of a gas is given by

$$\mathcal{H}_{\mathrm{tot}} = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} + \sum_{i=1}^{N} V(\vec{q}_i) + \sum_{i=1}^{N} \sum_{j=i+1}^{N} V_2(\vec{q}_i \vec{q}_j) \tag{3.67}$$

The first term is the kinetic energy. The second term describes the potential energy due to an external potential. Finally, the third term represents two-body-interactions between gas particles. In the *ideal gas*, however, particle-particle interactions are neglected; it is assumed that a gas atom moves without being influenced by other gas atoms. Thus, the Hamiltonian for the ideal gas becomes

$$\mathcal{H}_{\text{tot}} = \sum_{i=1}^{N} \left(\frac{\vec{p}_i^2}{2m} + V(\vec{q}_i) \right) \tag{3.68}$$

The partition function is given by

$$Z = \int_{-\infty}^{\infty} d\vec{p}_1 \dots dp_N \int_V d\vec{q}_1 \dots \vec{q}_i q e^{-\beta \mathcal{H}_{\text{tot}}} \tag{3.69}$$

$$Z = \left[\int d\vec{p} e^{-\beta \frac{\vec{p}^2}{2m}} \int d\vec{q} e^{-\beta \vec{V}(\vec{q})} \right]^N = Z_1^N \tag{3.70}$$

Here Z_1 is the single particle partition function. It can be seen that the ideal many-body partition function factorizes into the single-particle partition function. The kinetic energy of one gas atom is $E_{\text{kin}} = \frac{p^2}{2m} = \frac{p_x^2 + p_y^2 + p_z^2}{2m}$. For an isotropic system the expectation values of the momenta are equal: $\langle p_x^2 \rangle = \langle p_y^2 \rangle = \langle p_z^2 \rangle$. Each of them can be written as

$$\begin{aligned} \left\langle \frac{p_x^2}{2m} \right\rangle &= \frac{\int dp_x \frac{p_x^2}{2m} e^{-\beta \frac{p_x^2}{2m}}}{\int dp_x e^{-\beta \frac{px^2}{2m}}} \\ &= \frac{-\partial}{\partial \beta} \ln \left(\int dp_x e^{-\beta \frac{p_x^2}{2m}} \right) \\ &= \frac{-\partial}{\partial \beta} \ln \sqrt{\frac{2\pi m}{\beta}} \\ &= \frac{\partial}{\partial \beta} \frac{1}{2} \ln \beta \\ &= \frac{1}{2\beta} \\ &= \frac{k_B T}{2} \end{aligned} \tag{3.71}$$

Thus, the kinetic energy of one coordinate is $\frac{k_B T}{2}$. This generalizes into the *equipartition theorem.* It states that for a degree of freedom x described by a quadratic energy $\mathcal{H} = \alpha x^2$, the average energy is $\langle \mathcal{H} \rangle = k_B T/2$. For a system with f degrees of freedom this means

$$U = \langle \mathcal{H} \rangle = \frac{f}{2} k_b T \ . \tag{3.72}$$

In the ideal gas, each atom has three degrees of freedom (x,y,z). Therefore, each atom has the kinetic energy of

$$E_{\text{kin}} = 3 \left\langle \frac{p_x^2}{2m} \right\rangle \tag{3.73}$$

$$= \frac{3}{2} k_B T \tag{3.74}$$

For N atoms this becomes

$$E_{\text{kin}} = \frac{3N}{2} k_B T = U \tag{3.75}$$

The heat capacity is constant as a function of temperature:

$$C = \frac{dU}{dT} = \frac{f k_B}{2} \tag{3.76}$$

For a monatomic gas $f = 3N$ and for a diatomic gas $f = (6+1)N$ (6 momenta and 1 distance coordinate). Since $k_B T$ at $T = 300\,\mathrm{K}$ is $4 \cdot 10^{-21}\,\mathrm{J}$, the kinetic energy of a gas atom at room temperature is a very small. The velocity can be calculated by setting the energy equal to the formula for the kinetic energy.

$$\frac{3}{2} k_B T = \left\langle \frac{m v^2}{2} \right\rangle \tag{3.77}$$

$$\left\langle v^2 \right\rangle = 3 \frac{k_B T}{m} \tag{3.78}$$

$$\sqrt{\langle v^2 \rangle} = 460 \frac{\mathrm{m}}{\mathrm{s}} \tag{3.79}$$

So the velocity is very high: of the order of the speed of sound. The expectation value of the velocity $\langle v \rangle = 0$ vanishes, however. Otherwise the gas as a whole would move in a certain direction!

3.6.2 Maxwell-Boltzmann distribution

The un-normalised distribution of one velocity component v_x is, according to the Boltzmann distribution, given by

$$\rho v_x \propto e^{-\frac{p_x^2 \beta}{2m}} \tag{3.80}$$

$$\propto e^{-\frac{\beta m v_x^2}{2}}. \tag{3.81}$$

The expectation value of the second moment of $\vec{v}$ is

$$\left\langle \vec{v}^2 \right\rangle = \frac{\int d\vec{v} \vec{v}^2 \rho(\vec{v})}{\int d\vec{v} \rho(\vec{v})} \tag{3.82}$$

$$= \frac{\int dv_x dv_y dv_z v^2 e^{-\frac{m\beta(v_x^2+v_y^2+v_z^2)}{2}}}{\ldots} \tag{3.83}$$

$$= \frac{\int_0^\infty dv 4\pi v^2 v^2 e^{-\frac{m\beta v^2}{2}}}{\int_0^\infty dv 4\pi v^2 e^{-\frac{m\beta v^2}{2m}}} \tag{3.84}$$

$$= \frac{\int_0^\infty dv v^2 v^2 e^{-\frac{m\beta v^2}{2}}}{\int_0^\infty dv v^2 e^{-\frac{m\beta v^2}{2m}}} \tag{3.85}$$

$$\equiv \int_0^\infty dv v^2 \rho_{MB}(v). \tag{3.86}$$

Here $\rho_{MB}(v)$ denotes the Maxwell-Boltzmann distribution which is given by

$$\rho_{MB}(v) = v^2 e^{-\frac{m\beta v^2}{2}} \cdot 4\pi \left(\frac{m\beta}{2\pi} \right)^{-\frac{3}{2}} \tag{3.87}$$

since

$$\begin{aligned}\int_0^\infty dv v^2 e^{-\frac{m\beta v^2}{2}} &= \frac{1}{2}\int_{-\infty}^\infty dv v^2 e^{-\frac{m\beta v^2}{2}} &(3.88)\\ &= -\frac{1}{2}\frac{d}{d(m\beta/2)}\int_{-\infty}^\infty dv e^{-\frac{m\beta v^2}{2}} &(3.89)\\ &= -\frac{1}{2}\frac{d}{d(m\beta/2)}\left(\frac{2\pi}{m\beta}\right)^{\frac{1}{2}} &(3.90)\\ &= \frac{1}{2}\pi^{\frac{1}{2}}\frac{1}{2}\left(\frac{m\beta}{2}\right)^{-\frac{3}{2}} &(3.91)\\ &= \frac{\pi^{\frac{1}{2}}}{4}\left(\frac{2}{m\beta}\right)^{\frac{3}{2}} &(3.92)\\ &= \frac{1}{4\pi}\left(\frac{2\pi k_B T}{m}\right)^{\frac{3}{2}}. &(3.93)\end{aligned}$$

The maximum of $\rho_{MB}(v)$ is at

$$\begin{aligned}\frac{d\rho_{MB}(v)}{dv^2} &= 4\pi\left(\frac{m\beta}{2\pi}\right)^{\frac{3}{2}}\left(e^{-\frac{m\beta v^2}{2}} - \frac{v^2 m\beta}{2}e^{-\frac{m\beta v^2}{2}}\right) &(3.94)\\ v_{\max}^2 &= \frac{2}{m\beta} &(3.95)\\ &= \frac{2k_B T}{m}. &(3.96)\end{aligned}$$

Compared with equation (3.78) this yields

$$\frac{\langle\vec{v}^2\rangle}{\vec{v}_{\max}^2} = \frac{3}{2}. \tag{3.97}$$

So the mean and the maximum are not the same. In other words, the average velocity does not coincide with the most probable velocity. The law of large numbers does not apply here because the Maxwell-Boltzmann distribution describes a single-particle property and is independent of N.

3.6.3 Barometric Height Formula

Close to the ground of the earth the total Hamiltonian for N gas atoms is

$$\mathcal{H}_{\text{total}} = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} + mgz_i. \tag{3.98}$$

Here $g \approx 9.81\frac{\text{m}}{\text{s}^2}$ is the acceleration due to gravity. The momentum and position distribution factorises, so the height distribution of one atom is

$$\rho(z) = e^{-\beta mgz} \tag{3.99}$$

and with that the expectation value for the height z is

$$\begin{aligned}\langle z\rangle &= \frac{\int_0^\infty dz z e^{-\beta mgz}}{\int_0^\infty dz e^{-mgz}}\\ &= -\frac{d}{d(\beta mg)}\int_0^\infty dz e^{-\beta mgz}\\ &= -\frac{d}{d(\beta mg)}\ln\left(\frac{1}{\beta mg}\right)\\ &= \frac{1}{\beta mg} &(3.100)\end{aligned}$$

For a nitrogen atom at around 0°C this yields an average height of around 15 km above the surface of the earth. If anything, this is merely a rough approximation of the height of the atmosphere, as $\beta = 1/(k_B T)$ is not at all constant across different heights and, furthermore, g decreases with height.

3.7 Ideal Gas in the Canonical Ensemble

The partition function for an ideal gas of N identical mono-atomic gas particles in a volume V is given by

$$Z(N,V,T) = \frac{1}{N!}\Pi_{j=1}^{N}\left[\frac{1}{h^3}\int d^3p_j \int_V d^3q_j\right] e^{-\beta\mathcal{H}(p_{3N},q_{3N})} \tag{3.101}$$

- The pre-factor $\frac{1}{N!}$ avoids over-counting of micro-states due to permutation of indices. This avoids the Gibbs paradox of a non-extensive entropy.
- h is Planck's constant and has units of an action (Js). It makes the integral over phase space and thus Z itself unitless. For all physical observables, the value of h is not important since $F = -k_B T\ln(Z)$.
- The (inverse) pre-factor ($N!h^3$) follows from quantum statistics.
- For an ideal gas $\mathcal{H}(p_{3N},q_{3N}) = \sum_{i=1}^{N}\frac{\vec{p}_i^2}{2m}$ and there is no dependence on $\vec{q}_i$ for an ideal gas, while the finite volume V enters via the integration boundaries.

$$\begin{aligned} \int_V d^3q_i &= \int_0^L dq_i^x \int_0^L dq_i^y \int_0^L d_i^z 1 \qquad (3.102)\\ &= L^3 \qquad (3.103)\\ &= V \qquad (3.104) \end{aligned}$$

- Clearly Z depends on the thermodynamic parameters N, T, V so the partition function has variables $Z(T,V,N)$ as well as the free energy $F(T,V,N)$.

$$\begin{aligned} Z &= \frac{1}{N!}\Pi_{j=1}^{N}\left[\frac{1}{h^3}\int d^3p_j \int_V d^3q_j\right] e^{-\beta\sum_{i=1}^{N}\frac{\vec{p}_i^2}{2m}} \qquad (3.105)\\ &= \frac{1}{N!}\left[\frac{1}{h^3}\int d^3p e^{-\beta\frac{\vec{p}^2}{2m}}\int_V d^3q\right]^N \qquad (3.106)\\ &= \frac{1}{N!}\left(\frac{V}{h^3}\right)^N \left[\int_{-\infty}^{\infty} dp e^{-\beta\frac{p^2}{2m}}\right]^{3N} \qquad (3.107)\\ &= \frac{1}{N!}\left(\frac{V}{h^3}\right)^N (2\pi m k_B T)^{\frac{3N}{2}} \qquad (3.108)\\ &= \frac{1}{N!}\left(\frac{V}{\lambda_t^3}\right)^N \qquad (3.109)\\ \lambda_t &= \frac{h}{\sqrt{2\pi m k_B T}} \qquad (3.110) \end{aligned}$$

Here λ_t is the de Broglie wavelength at temperature T. The free energy follows as

$$\begin{aligned} F &= -k_B T\ln(Z) \qquad (3.111)\\ &= k_B T N\ln\left(\frac{\lambda_t^3}{V}\right) + k_B T\ln(N!) \qquad (3.112) \end{aligned}$$

Using the Stirling approximation this yields

$$\approx k_B T N \left[\ln\left(\frac{N\lambda_t^3}{V}\right) - 1\right] \tag{3.113}$$

The free energy F is extensive and scales like N. Without the factorial in the definition of Z, the free energy would be proportional to $N\ln(N)$. Now other state variables can be calculated from F.

$$\begin{aligned}
F(N,T,V) &= k_B T N \left[\ln\left(\frac{N}{V}h^3\left(2\pi m k_B T\right)^{-\frac{3}{2}}\right) - 1\right] && (3.114)\\
&= k_B T N \left[\ln\left(\frac{N}{V}h^3\right) - \frac{3}{2}\ln\left(2\pi m k_B T\right) - \frac{3}{2}\ln\left(T\right) - 1\right] && (3.115)\\
-S &= \frac{\partial F}{\partial T} && (3.116)\\
&= \frac{F}{T} - \frac{3}{2}k_B N && (3.117)\\
S &= -\frac{F}{T} + \frac{3}{2}k_B N && (3.118)\\
&= -k_B N \left[\ln\left(\frac{N}{V}\lambda_t^3\right) - 1\right] + \frac{3}{2}k_B N && (3.119)\\
&= -k_B N \left[\ln\left(\frac{N}{V}\lambda_t^3\right) - \frac{5}{2}\right] && (3.120)\\
F &= U - TS && (3.121)\\
U(N,T,V) &= F + TS && (3.122)\\
&= \frac{3}{2}N k_B T && (3.123)\\
&= U(T,N) && (3.124)
\end{aligned}$$

So the internal energy of the ideal gas does not depend on the volume V.

$$\begin{aligned}
-\frac{\partial F}{\partial V} &= \frac{k_B T N}{V} && (3.125)\\
&\equiv P(T,V,N) && (3.126)
\end{aligned}$$

This is called the thermal equation of state of the ideal gas. And the total differential of the free Energy dF can be written down.

$$dF = \left.\frac{\partial F}{\partial T}\right|_{V,N} dT + \left.\frac{\partial F}{\partial V}\right|_{T,N} dV + \left.\frac{\partial F}{\partial N}\right|_{V,T} dN \tag{3.127}$$

For $N =$ const. this becomes

$$dF = -S(T,V,N)dT - P(T,V,N)dV \tag{3.128}$$

furthermore with $U = F + TS$ this becomes

$$\begin{aligned}
dU &= -SdT - PdV + TdS + SdT && (3.129)\\
dU &= -PdV + TdS && (3.130)\\
&= -P(V,S)dV + T(V,S)dS. && (3.131)
\end{aligned}$$

To summarise

$$P = -\frac{\partial U}{\partial V} \tag{3.132}$$
$$T = \frac{\partial U}{\partial S} \tag{3.133}$$
$$\frac{1}{T} = \frac{\partial S}{\partial U} \tag{3.134}$$

3.7.1 First law of Thermodynamics in Differential Form

The total differential of U corresponds to the first law of thermodynamics.

$$dU(S,V) = TdS - PdV \tag{3.135}$$
$$dU(S,V) = \Delta Q - \Delta W \tag{3.136}$$

Here $\Delta W = PdV$ is the mechanical work done by the gas and $\Delta Q = TdS$ is the heat transferred to the gas. As discussed later in more detail, ΔQ and ΔW are not total differentials; That means that Q and W are not state functions.

3.8 Ideal Gas in the Microcanonical Ensemble

$$\Gamma = \int_{U<\mathcal{H}<U+\Delta} d^{3N}\vec{q}d^{3N}\vec{p} \tag{3.137}$$
$$= T \tag{3.138}$$

Here Γ is the number of micro-states in the energy range $U < \mathcal{H} < U + \Delta$. So the number N' of micro-states with the energy $\mathbb{H} < U$ is given by

$$\Sigma(U,V,N) = \int_{\mathcal{H}<U+\Delta} d^{3N}\vec{q}d^{3N}\vec{p} \tag{3.139}$$

and

$$\Sigma'(U) = \frac{d\Sigma}{dU} \tag{3.140}$$
$$= \omega(U) \tag{3.141}$$
$$= \frac{\Gamma}{\Delta} \tag{3.142}$$

For a mono-atomic ideal gas of N atoms in a volume V the shape of the container is not relevant. To calculate Σ a (quantum mechanical) correction factor has to be introduced.

$$\Sigma = \frac{1}{N!h^{3N}} \int_{\mathcal{H}<U} d\vec{q}_1 \dots d\vec{q}_N d\vec{p}_1 \dots d\vec{p}_N \tag{3.143}$$

$$\mathbb{H} = \sum_{i=1}^{N} \left[\frac{\vec{p}_1^{\,2}}{2m} + V_{\text{Wall}}(\vec{q}_i) \right] \tag{3.144}$$

$$V_{\text{Wall}} = \begin{cases} 0 & |\vec{q}_i \in V \\ \infty & |\vec{q}_i \notin V \end{cases} \tag{3.145}$$

$$\sum(U) = \frac{1}{N!h^3} \int_V d\vec{q}_1 \dots \int_V d\vec{q}_N \int_{\sum_{i=1}^N \frac{\vec{p}_i^{\,2}}{2m}<U} d\vec{p}_1 \dots d\vec{p}_N \tag{3.146}$$

$$= \frac{V^N}{N!h^{3N}} (2m)^{\frac{3N}{2}} \int_{\sum_{i=1}^{3N} x_i^2<U} dx_1 \dots dx_{3N} \tag{3.147}$$

$$x_1 = \frac{p_i}{\sqrt{2m}} \tag{3.148}$$

$$dp_1 = (2m)^{\frac{1}{2} dx_1} \tag{3.149}$$

$$\sum(U) = \frac{V^N}{N!h^{3N}} (2m)^{\frac{3N}{2}} \nu_{3N} \left(U^{\frac{1}{2}} \right) \tag{3.150}$$

$$\tag{3.151}$$

Here ν_{3N} is the volume of a $3N$-dimensional hypersphere which is $\nu_n(R) = c_n R^n$. The coefficient c_n turns out to be $c_n = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)}$. With this the volume becomes

$$\nu_n(R) = \frac{\pi^{\frac{n}{2}}}{\Gamma\left(\frac{n}{2}+1\right)} R^n \tag{3.152}$$

$$\Gamma(U) = \Delta \frac{d\Sigma}{dU} \tag{3.153}$$

$$= \left(V \left(\frac{2\pi m U}{h^2} \right)^{\frac{3}{2}} \right)^N \frac{1}{N!} \frac{1}{\Gamma\left(\frac{3N}{2}+1\right)} \frac{3N}{2} \frac{\Delta}{U} \tag{3.154}$$

$$S = k_B \ln\left(\Gamma(U)\right) \tag{3.155}$$

$$= k_B \left(N \ln \left(V \left(\frac{2\pi m U}{h^2} \right)^{\frac{3}{2}} \right) - \ln(N!) - \ln \left(\Gamma \left(\frac{3N}{2} \right) \right) + \ln \left(\frac{\Delta}{U} \right) \right) \tag{3.156}$$

For very large systems ($N >> 1, V >> 1$) $N >> \ln(N)$ and the Stirling approximation ($\ln(N!) = N\ln(N) - N$) the expression simplifies a lot.

$$S = k_B N \left[\ln \left(\frac{V}{N} \left(\frac{4\pi m U}{3h^2 N} \right)^{\frac{3}{2}} \right) + \frac{5}{2} \right] \tag{3.157}$$

So the entropy S is extensive, but only because the $\frac{1}{N!}$ factor was introduced in the beginning. With the thermal wavelength λ_t, introduced earlier and using that
$U = \frac{3}{2} N k_B T$, it can be written that

$$S = k_B N \left(\ln \left(\frac{V}{N} \lambda_t^{-3} \right) + \frac{5}{2} \right) \tag{3.158}$$

$$= -k_B N \left(\ln \left(\frac{N}{V} \lambda_t^3 \right) - \frac{5}{2} \right), \tag{3.159}$$

which is the same result derived in in the canonical ensemble (see equation (3.120)).

3.8.1 Heat capacity of a solid

A very simple model for a solid is just a cube lattice with localised atoms that can be modelled as $3N$ harmonic oscillators with the same frequency ω. The Hamiltonian $\mathbb{H}$ is given by

$$\mathbb{H} = \sum_{i=1}^{N} \left(\frac{\vec{p}_i^{\,2}}{2m} + \frac{1}{2} m\omega^2 \vec{q}_i^{\,2} \right) \tag{3.160}$$

As the particles are distinguishable since they are localised on lattice sites, no $\frac{1}{N!}$ is needed.

$$\begin{aligned}
\Sigma(U) &= \frac{1}{h^{3N}} \int_{\mathcal{H}} d\vec{q}_1 \dots d\vec{q}_N d\vec{p}_1 \dots d\vec{p}_N &(3.161)\\
&= \frac{1}{h^{3N}} \frac{2}{\omega}^{3N} \int_{\mathcal{H}} dx_1 \dots dx_{6N} &(3.162)\\
&= \frac{1}{h^{3N}} \frac{2}{\omega}^{3N} \nu_{6N} \left(U^{\frac{1}{2}} \right) &(3.163)\\
U &= 3Nk_BT &(3.164)\\
S &= 3Nk_B \left(\ln \left(\frac{2\pi U}{3h\omega N} \right) \right) &(3.165)
\end{aligned}$$

3.9 Grand Canonical Ensemble

For an open system, where particles can exchange with the reservoir, also the particle number can fluctuate. This is important for example in cells, chemical reactions, and others. Again considering two coupled systems, and $\mathcal{H}(p_{3N}, q_{3N}, N) \approx \mathcal{H}(p_{3N}, q_{3N}, N_1) + \mathcal{H}(p_{3N}, q_{3N}, N_2) +$ interactions.

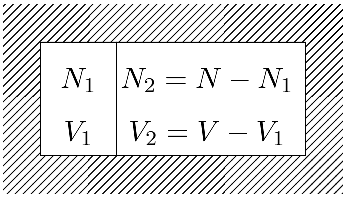


Figure 3.2: A small volume V_1 in contact with a reservoir of volume V_2. The total number of particles stays constant $N = N_1 + N_2$, but the volume V_1 can exchange particles with the reservoir; Further $V_1 << V_2$ and $N_1 << N_2$. (The stripy pattern represents thermal insulation).

The canonical partition function is

$$\begin{aligned}
Z(N, V, T) &= \frac{1}{h^{3N} N!} \sum_{N_1=1}^{N} \frac{N!}{N_1! N_2!} \\
&\times \int d^{3N_1} p_1 d^{3N_1} q_1 d^{3N_2} p_2 d^{3N_2} q_2 e^{-\beta\mathcal{H}(p_1,q_1,N_1) - \beta\mathcal{H}(p_2,q_2,N_2)} &(3.166)\\
&= \sum_{N_1=0}^{N} Z_{N_1}(V_1, T) Z_{N_2}(V_2, T) &(3.167)
\end{aligned}$$

Introduce the probability distribution ρ

$$\rho(p_{3N_1}, q_{3N_2,N_1}) \equiv \frac{Z_{N_2}(V_2, T)}{Z_N} \frac{e^{-\beta\mathcal{H}(p_{3N_1}, q_{3N_1}, N_1)}}{h^{3N_1} N_1!} \tag{3.168}$$

which is normalized

$$\sum_{N_1=0}^{N} \int d^{3N_1} p_1 d^{3N_1} q_1 \rho(p_1, q_1, N_1) = 1. \tag{3.169}$$

$$\tag{3.170}$$

With the free energy $F(N, V, T) = -k_B T \ln(Z_n(V, T))$ it can be obtained that

$$\frac{Z(N_2, V_2, T)}{Z(N, V, T)} = e^{-\beta(F(N-N_1, V-V_1, T) - F(N,V,T))} \tag{3.171}$$

Now $N_1 << N$, which means that system 2 is a particle and volume reservoir. Now $F(N - N_1, V - V_1, T)$ can be expanded around $N_1 = 0$ and $V_1 = 0$.

$$\frac{\partial F(N, V, T)}{\partial V}|_{N,T} = -P(N, V, T) \tag{3.172}$$

Here $\frac{\partial F(N,V,T)}{\partial N}|_{V,T} \equiv \mu(N, V, T)$ is the chemical potential. Furthermore μ is the free energy needed to add a particle to the system.

$$F(N - N_1, V - V_1, T) \approx F(N, V, T) - N_1 \mu + V_1 P \tag{3.173}$$

$$\frac{Z_{N_2}(V_2, T)}{Z_N(V, T)} = e^{\beta N_1 \mu - \beta P V_1} \tag{3.174}$$

$$\rho(p, q, N) = \frac{e^{-\beta(PV - \mu N + \mathcal{H}(p,q))}}{N! h^{3N}} \tag{3.175}$$

With this the grand canonical partition function can be written

$$\begin{aligned}
\mathbb{Z}_\mu(V, T) &\equiv \sum_{N=0}^{\infty} e^{\beta N \mu} Z_N(V, T) && (3.176)\\
&= \sum_{N=0}^{\infty} e^{\beta N \mu} \int \frac{d^{3N} q d^{3N} p}{N! h^{3N}} e^{-\beta \mathcal{H}(p_{3N}, q_{3N}, N)} && (3.177)\\
&= e^{\beta P V} \sum_{N=0}^{\infty} d^{3N} q d^{3N} p \rho(p, q, N) && (3.178)\\
&= e^{\beta P V}, && (3.179)
\end{aligned}$$

which is the grand-canonical equation of state.

$$\begin{aligned}
k_B T \ln(Z_\mu(V, T)) &= -\Omega(\mu, V, T) && (3.180)\\
&= PV && (3.181)\\
&= P(\mu, V, T) V(\mu, T) && (3.182)
\end{aligned}$$

Here $\Omega(\mu, V, T)$ is the grand (canonical) potential.

3.9.1 Properties of Particles in the Grand (Canonical) Ensemble

Mean Number of Particles

From the definition of the grand-canonical partition function it can be seen that

$$\langle N \rangle = \frac{\partial \ln \mathbb{Z}_\mu}{\beta \partial \mu}. \tag{3.183}$$

Next, we calculate the particle number fluctuation $\langle N^2\rangle - \langle N\rangle^2$

$$\begin{aligned}
\frac{\partial}{\beta\partial\mu}\frac{\partial}{\beta\partial\mu}\ln \mathbb{Z}_\mu(V,T) &= \frac{\partial}{\beta\partial\mu}\frac{\partial}{\beta\partial\mu}\ln\left(\sum_{N=0}^{\infty} e^{\beta N_\mu} Z_N\right) && (3.184)\\
&= \frac{\partial}{\beta\partial\mu}\frac{\sum_N N e^{\beta N\mu} Z_N}{\sum_N e^{\beta N\mu} Z_N} && (3.185)\\
&= \frac{\sum_N N^2 e^{\beta N\mu} Z_N}{\sum_N e^{\beta N\mu} Z_N} - \left(\frac{\sum_N N e^{\beta N\mu} Z_N}{\sum_N e^{\beta N\mu} Z_N}\right)^2 && (3.186)\\
&= \langle N^2\rangle - \langle N\rangle^2 && (3.187)\\
&= \frac{\partial}{\beta\partial\mu}\frac{\partial}{\beta\partial\mu}\frac{PV}{k_BT} && (3.188)\\
&= \frac{1}{\beta^2}\frac{\partial^2}{\partial\mu^2}\bigg|_{V,T}\frac{PV}{k_BT} && (3.189)\\
&= k_BTV\frac{\partial^2 P}{\partial\mu^2}\bigg|_{T,V} && (3.190)\\
&= k_BTV\frac{\partial^2 P(\mu,T,V)}{\partial\mu^2} && (3.191)\\
&\propto N && (3.192)
\end{aligned}$$

So (as for energy fluctuations in the canonical ensemble) the relative particle number deviations from the mean number are $\frac{\sqrt{\langle N^2\rangle - \langle N\rangle^2}}{\langle N\rangle} \propto N^{-\frac{1}{2}}$. This tends to 0 as $N \to \infty$. This means that there are no particle fluctuations as N becomes large and thus $\langle N\rangle = N' = N$ for large systems, where N' denotes the most probable particle number

Conclusions

The important conclusion from this is that the grand canonical ensemble, the canonical ensemble and the micro canonical ensembles are equivalent. Therefore the grand canonical partition function $\mathbb{Z}_\mu(V,T)$ becomes

$$\begin{aligned}
\mathbb{Z}_\mu(V,T) &= \sum_{N=0}^{\infty} e^{\beta N\mu} Z_N(V,T) && (3.193)\\
&\approx e^{\beta N'\mu} Z_{N'}(V,T) && (3.194)
\end{aligned}$$

Here N' is the most probable particle number which is of course equal to the mean. Thus the logarithm of the partition function becomes

$$\begin{aligned}
\ln \mathbb{Z}_\mu(V,T) &= \beta N'\mu + \ln Z_{N'} && (3.195)\\
& && (3.196)
\end{aligned}$$

The grand potential results from the free energy via a Legendre transformation $\Omega = F - \mu N$, using that $F = -k_BT\ln Z_{N'}$ and $\Omega = -k_BT\ln\mathbb{Z}_\mu$.

- With $\Omega = -PV$ it follows $F = \mu N - PV$. And from $F = U - TS$ it follows that $U = TS + \mu N - PV$. This is the fundamental equation. All thermodynamic potentials $(U, F, \Omega, (G))$ can be expresses in a bilinear form as products of extensive(S, N, V) and intensive (T, μ, P) state variables.

3.9.2 Total Differential of Ω

$$dF(T,V,N) = \mu dN - PdV - SdT \tag{3.197}$$

$$\Omega = F - \mu N \tag{3.198}$$

$$d\Omega = dF - d(\mu N) \tag{3.199}$$

$$= \mu dN - PdV - SdT - \mu dN - Nd\mu \tag{3.200}$$

$$= -PdV - SdT - Nd\mu \tag{3.201}$$

$$= d\Omega(\mu, T, V) \tag{3.202}$$

3.9.3 Gibbs-Duhem Equation

Since $\Omega = -PV$

$$d\Omega = -PdV - VdP \tag{3.203}$$

$$= -PdV - SdT - Nd\mu \tag{3.204}$$

it can be concluded that

$$0 = VdP - SdT - Nd\mu. \tag{3.205}$$

This is the Gibbs-Duheme equation which is the relation between all intensive differentials.

3.9.4 Derivation of $\frac{\partial^2 P}{\partial \mu^2}|_{T,V}$

The free energy is $F(N,V,T) = Nf(N,V,T)$. By construction f is intensive, it can only depend on $\frac{V}{N} = v$, so $F(N,V,T) = Nf\left(\frac{V}{N}, T\right) = Nf(v,t)$, where v is the volume per particle. The chemical potential μ again is defined as

$$\mu = \left.\frac{\partial F}{\partial N}\right|_{T,V} \tag{3.206}$$

$$= \frac{\partial}{\partial N}[Nf(v,T)] \tag{3.207}$$

$$= f(v,T) - v\frac{\partial f(v,T)}{\partial v} \tag{3.208}$$

$$= \mu(v,T) \tag{3.209}$$

The pressure P is

$$P = -\left.\frac{\partial F}{\partial V}\right|_{T,N} \tag{3.210}$$

$$= -\frac{\partial}{\partial V}(Nf(v,T)) \tag{3.211}$$

$$= -\frac{\partial}{\partial V}f(v,T) \tag{3.212}$$

$$= P(v,T) \tag{3.213}$$

From this

$$P(v,T) = P\left(v(\mu,T), T\right) \tag{3.214}$$

$$\hat{=} P(\mu,T) \tag{3.215}$$

$$\left.\frac{\partial P}{\partial \mu}\right|_{T} = \left.\frac{\partial P}{\partial v}\right|_{T} \left.\frac{\partial v}{\partial \mu}\right|_{T} \tag{3.216}$$

So with 3.209 $\frac{\partial \mu}{\partial v}\Big|_T$ can be calculated.

$$\frac{\partial \mu}{\partial v}\bigg|_T = \frac{\partial f}{\partial v}\bigg|_T - \frac{\partial f}{\partial v}\bigg|_T - v\,\frac{\partial^2 f}{\partial v^2}\bigg|_T \tag{3.217}$$

$$= -v^2\,\frac{\partial^2 f}{\partial v^2}\bigg|_T \tag{3.218}$$

$$\frac{\partial P}{\partial v}\bigg|_T = -\,\frac{\partial^2 f}{\partial v^2}\bigg|_T \tag{3.219}$$

and therefore

$$\frac{\partial \mu}{\partial v}\bigg|_T = v\,\frac{\partial P}{\partial v}\bigg|_T \tag{3.220}$$

With all of this 3.216 becomes

$$\frac{\partial P}{\partial \mu}\bigg|_{T,V} = \frac{1}{v} \tag{3.221}$$

and

$$\frac{\partial^2 P}{\partial \mu^2}\bigg|_{V,T} = \frac{\partial \frac{1}{v}}{\partial \mu}\bigg|_T \tag{3.222}$$

$$= -\,\frac{1}{v^2}\frac{\partial v}{\partial \mu}\bigg|_T \tag{3.223}$$

$$= -\,\frac{1}{v^3}\frac{\partial v}{\partial P}\bigg|_T \tag{3.224}$$

Going back to variables T, V, N and using the isothermal compressibility $\kappa_T(P,T)$ this becomes

$$\frac{\partial^2 P}{\partial \mu^2}\bigg|_{T_V} = -\,\frac{N^2}{V^3}\frac{\partial V(P,T,N)}{\partial P}\bigg|_{T,N} \tag{3.225}$$

$$= \frac{N^2}{V^2}\kappa_T(P,T) \tag{3.226}$$

3.10 Ideal Gas in the Grand Canonical Ensemble

The grand-canonical partition function was defined as

$$\mathbb{Z}(\mu,V,T) = \sum_{N=0}^{\infty} e^{\beta N \mu} Z(N,V,T) \tag{3.227}$$

$$\tag{3.228}$$

The free energy was derived in section 3.7 as $F(N,V,T) = k_B T N \ln\!\left(\frac{N\lambda_t^3}{V}\right) - k_B T N = N f(v,T)$ and the chemical potential is $\mu = \frac{\partial F}{\partial N} = k_B T \ln\!\left(\frac{N\lambda_t^3}{V}\right)$. So the partition function becomes

$$\mathbb{Z} = \sum_{N=0}^{\infty} \frac{1}{N!}\left(\frac{V}{\lambda_t^3} e^{\beta\mu}\right)^N \tag{3.229}$$

$$= e^{\frac{V}{\lambda_t^3} e^{\beta\mu}}. \tag{3.230}$$

And the grand potential Ω is

$$\Omega(\mu, V, T) = -k_B T \ln(\mathbb{Z}(\mu, V, T)) \tag{3.231}$$

$$= -k_B T \frac{V}{\lambda_t^3} e^{\beta\mu} \tag{3.232}$$

$$= -PV. \tag{3.233}$$

and from that follows

$$-k_B T N = -PV \tag{3.234}$$

$$PV = N k_B T. \tag{3.235}$$

The ideal equation of state is equation (3.235).

3.10.1 Chemical Reactions

A bimolecular reaction is an example of a simple chemical reaction. Here two atoms react to one molecule which can split back into the two atoms again: $A + A \rightleftharpoons B$. Examples are gases like hydrogen $H + H \rightleftharpoons H_2$. Here μ_A and μ_B are the chemical potentials, V is the volume and T is the temperature. The partition function is

$$\mathbb{Z}(\mu_A, \mu_B, V, T) = \sum_{N_A=0}^{\infty} \sum_{N_B=0}^{\infty} e^{\beta N_A \mu_A} e^{\beta N_B \mu_B} Z_A(N_A, V, T) Z_B(N_B, V, T). \tag{3.236}$$

Here Z_A (and similar for Z_B) are ideal gas partition functions,

$$Z_A(N_A, V, T) = \frac{1}{N_A!} \left(\frac{V}{\lambda_{t_A}^3} \right)^{N_B}. \tag{3.237}$$

$$\mathbb{Z} = \sum_{N_A=0}^{\infty} \frac{1}{N_A!} \left(e^{\beta\mu_A} \frac{V}{\lambda_{t_A}^3} \right)^{N_A} \sum_{N_B=0}^{\infty} \frac{1}{N_B!} \left(e^{\beta\mu_B} \frac{V}{\lambda_{t_B}^3} \right)^{N_B} \tag{3.238}$$

$$\mathbb{Z} = \exp\left[e^{\beta\mu_A} \frac{V}{\lambda_{t_A}^3} + e^{\beta\mu_B} \frac{V}{\lambda_{t_B}^3} \right]. \tag{3.239}$$

The expectation values of N_A and N_B are

$$\langle N_A \rangle = N_A \tag{3.240}$$

$$= \frac{1}{\beta} \frac{\partial \ln(\mathbb{Z})}{\partial \mu_A} \tag{3.241}$$

$$= e^{\beta\mu_A} \frac{V}{\lambda_{t_A}^3} \tag{3.242}$$

$$\langle N_B \rangle = N_B \tag{3.243}$$

$$= e^{\beta\mu_B} \frac{V}{\lambda_{t_B}^3}. \tag{3.244}$$

The chemical energy of the reaction $2A \to B$ is given by $\Delta\mu = \mu_B - 2\mu_A$ (one B created, two A destroyed).

$$N_B = e^{\beta(\Delta\mu + 2\mu_A)} \frac{V}{\lambda_{t_B}^3} \tag{3.245}$$

$$= e^{\beta\Delta\mu} \frac{V}{\lambda_{t_B}^3} N_A^2 \frac{\lambda_{t_A}^6}{V^6} \tag{3.246}$$

$$\frac{N_B}{V} \left(\frac{V}{N_A} \right)^2 = \frac{C_B}{C_A^2} \tag{3.247}$$

Here $C_A = \frac{N_A}{V}$ and $C_B = \frac{N_B}{V}$.

$$\begin{aligned} \frac{C_B}{C_A^2} &= e^{\beta\Delta\mu}\frac{\lambda_{t_A}^6}{\lambda_{t_B}^3} && (3.248)\\ &\equiv K && (3.249) \end{aligned}$$

This is the law of mass action, K is the equilibrium reaction constant. The concentrations of components in a chemical reaction are related by power laws; The powers are given by the multiplicity in the reaction. The total or maximal concentration of A is given by $C_A + 2\ C_B = C_A^{\mathrm{TOT}}$ (if $C_B = 0$ then $C_A = C_A^{\mathrm{TOT}}$ and if $C_A = 0$ then $C_B = {C_A^{\mathrm{TOT}}}/{2}$).

$$\begin{aligned} \frac{C_B}{(C_A^{\mathrm{TOT}} - 2C_B)^2} &= K && (3.250)\\ \frac{C_B}{K} &= \left(C_A^{\mathrm{TOT}}\right)^2 + 4(C_B)^2 - 4C_BC_A^{\mathrm{TOT}} && (3.251)\\ 0 &= (C_B)^2 - C_B\left(C_A^{\mathrm{TOT}} + \frac{1}{4K}\right) + \frac{(C_A^{\mathrm{TOT}})^2}{4} && (3.252)\\ C_B &= \frac{C_A^{\mathrm{TOT}} + \frac{1}{4K}}{2} \pm \sqrt{\frac{(C_A^{\mathrm{TOT}} + \frac{1}{4K})^2}{4} - \frac{\left(C_A^{\mathrm{TOT}}\right)^2}{4}} && (3.253) \end{aligned}$$

Since C_B cannot be larger than $\frac{C_A^{\mathrm{TOT}}}{2}$ the negative root is the correct one.

$$C_B = \frac{C_A^{\mathrm{TOT}} + \frac{1}{4K}}{2} - \sqrt{\frac{1}{64K^2} + \frac{C_A^{\mathrm{TOT}}}{8K}} \tag{3.254}$$

So this is the concentration of the product.

4 Thermodynamics

As already mentioned in Chapter **??**, Thermodynamics (Thermostatics would be a better name) describes macroscopical systems with the help of a few phenomenological rules which are called the laws of thermodynamics. These rules are not mathematically derived, but rather are generalisations or idealisations of experimental results. With these laws a large number of predictions becomes possible. The advantage of Thermodynamics is the generality of the predictions, the disadvantage is that material specific properties like the heat capacity of gases are not deducible.

4.1 Axiomatic Thermodynamic

Starting with the fundamental equation

$$U = TS - PV + \mu N \tag{4.1}$$

and taking the differential form

$$dU = TdS - PdV + \mu dN \tag{4.2}$$

$T(S, V, N)$, $P(S, V, N)$ and $\mu(S, V, N)$ can be derived, so $U(S, V, N)$ contains the complete information and everything can be derived from it. $T(S, V, N)$ does not contain all the informations since one would need the three functions $T(S, V, N)$, $P(S, V, N)$ and $\mu(S, V, N)$ to reconstruct $U(S, V, N)$ - they are the slopes of the 3-dimensional function U along the 3 directions. Now one can do the Legendre transform.

$$\begin{aligned} F &= U - TS && (4.3) \\ dF &= dU - TdS - SdT && (4.4) \\ &= -SdT - PdV + \mu dN && (4.5) \\ dF(T, V, N) &= -S(T, V, N)dT - P(T, V, N)dV + \mu(T, V, N)dN && (4.6) \end{aligned}$$

So the above Legendre transform is performed by

- obtaining $S(T, V, N)$ by inversion of $T(S, V, N)$
- replacing S by $S(T, V, N)$ so that

$$\begin{aligned} F(T, V, N) &= U(S, V, N) - TS && (4.7) \\ &= U(S(T, V, N), V, N) - TS(T, V, N) && (4.8) \end{aligned}$$

In fact $F(T, V, N)$ does also contain the complete thermodynamic information, so do the inversions $V(T, F, N)$, $T(F, V, N)$ and N(F,T,V): For example, from the expression for dF the expression for dV can be derived:

$$dV = -\frac{S}{P}dT - \frac{1}{P}dF + \frac{\mu}{P}dN. \tag{4.9}$$

From this differential the functions $\frac{S(T,F,N)}{P(T,F,N)}$, $P(T, F, N)$ and $\mu(T, F, N)$ can be obtained. There are many ways of formulating thermodynamics, but here the focus will lie only on the most important ones.

4.2 Equation of State and Response Functions

We again consider the differential of F

$$dF(T,V,N) = -S(T,V,N)dT - P(T,V,N)dV + \mu(T,V,N)dN. \tag{4.10}$$

$F(T,V,N)$ contains the complete information, the functions $S(T,V,N)$, $P(T,V,N)$, $\mu(T,V,N)$ however do not. They contain other useful information though and so they are called equations of state and they are first derivatives of thermodynamic potentials. The following one is called the thermal equation of state:

$$P(T,V,N) = -\frac{\partial F(T,V,N)}{\partial V}. \tag{4.11}$$

The functions $P(T,V,N), V(P,T,N), N(P,T,V)$ are equivalent; in fact $P(T,V,N) = P(T,v)$, where $v = \frac{V}{N}$. The so-called caloric equation of state can be derived from

$$\begin{aligned} dU &= TdS - PdV + \mu dN && (4.12)\\ dS &= \frac{1}{T}dU + \frac{P}{T}dV - \frac{\mu}{T}dN && (4.13)\\ \frac{\partial S(U,V,N)}{\partial U} &= \frac{1}{T(U,V,N)}. && (4.14)\\ & && (4.15) \end{aligned}$$

$T(U,V,N)$ connects the variables T, U, V and N and is called the caloric equation of state. Typically one writes $U(T,V,N)$, which is derived by inversion. So in conclusion, first derivatives of thermodynamic potentials are called equations of state. Now one can also take the second derivatives. These give the so-called response functions. We again consider the thermal equation of state

$$V(P,T,N) = \frac{\partial G(P,T,N)}{\partial P}. \tag{4.16}$$

Here $G = U - TS + PV$ is the Gibbs free energy. Considering the total differential of $V(P,T,N)$, we obtain

$$dV = \left.\frac{\partial V}{\partial P}\right|_{T,N} dP + \left.\frac{\partial V}{\partial T}\right|_{P,N} dT + \left.\frac{\partial V}{\partial N}\right|_{P,T} dN. \tag{4.17}$$

So the second derivatives of thermodynamic potentials describe how state variables change, when other state variables are varied. These response functions are extremely important to characterise systems, so they are very useful in applications and hence they are tabulated in books. Now with this the isothermal compressibility κ_T can be written down as

$$\begin{aligned} \kappa_T(P,T) &= -\frac{1}{V}\frac{\partial V(P,T,N)}{\partial P} && (4.18)\\ -V\kappa_T(P,T) &= \frac{\partial V(P,T,N)}{\partial P} && (4.19)\\ &= \frac{\partial^2 G(P,T,N)}{\partial P^2}. && (4.20) \end{aligned}$$

Now the expansion coefficient α is

$$\begin{aligned} \alpha(P,T) &= \frac{1}{V}\frac{\partial V(P,T,N)}{\partial T}. && (4.21)\\ & && (4.22) \end{aligned}$$

And the volume per particle is

$$\frac{\partial V(N,P,T)}{\partial N} = \frac{\partial N v(P,T)}{\partial N} \tag{4.23}$$

$$= v(P,T) \tag{4.24}$$

With that equation (4.17) can be written as

$$dV = -V\kappa_T dP + V\alpha dT + v dN. \tag{4.25}$$

Hence differentials of equations of state define response functions.

4.3 Maxwell Relations

Again starting from the fundamental differential form $dU = TdS - PdV + \mu dN$:

$$T = \frac{\partial U(S,V,N)}{\partial S} \tag{4.26}$$

$$= T(S,V,N) \tag{4.27}$$

$$-P(S,V,N) = \frac{\partial U(S,V,N)}{\partial V} \tag{4.28}$$

So since

$$\left.\frac{\partial}{\partial V}\right|_{S,N} \left.\frac{\partial U}{\partial S}\right|_{V,N} = \left.\frac{\partial}{\partial S}\right|_{V,N} \left.\frac{\partial U}{\partial V}\right|_{S,N} \tag{4.29}$$

it follows that

$$\frac{\partial T(S,V,N)}{\partial V} = -\frac{\partial P(S,V,N)}{\partial S}. \tag{4.30}$$

This is called a Maxwell relation (not to be confused with Maxwell's equation from electrodynamics), which relates derivatives of state variables. This concept is very powerful, but it can be confusing, because there are many Maxwell relations. Just from $U(S,V,N)$ many pairs can be formed and there are many more:

$$F(T,V,N), G(T,P,N), H(S,P,N), \Omega(T,V,N)\ldots \tag{4.31}$$

4.4 Adiabatic Processes and the Application of Thermodynamic Calculus

The differential form of the first law of thermodynamics is (again)

$$TdS = dU + PdV - \mu dN \tag{4.32}$$

$$= \Delta Q \tag{4.33}$$

$$\neq dQ. \tag{4.34}$$

The last equation stresses that heat is not a potential! Adiabatic processes are processes where no heat is exchanged with the environment, for example by doing the process as quickly as possible, so nearly no heat can be conducted away from the system. To describe typical adiabatic processes, such as compression of a gas, it is useful to switch the independent variables of S from (U,V,N), to for example (V,P,N) or (V,T,N). To make calculations easier, the μN in equation (4.32) is dropped, as N is constant most of the time anyway.

4.4.1 From $S(U,V)$ to $S(P,T)$ and $S(V,T)$

From the caloric equation of state $U(P,T)$, dU is written as

$$dU = \frac{\partial U}{\partial P}|_T dP + \frac{\partial U}{\partial T}|_P dT. \tag{4.35}$$

From the thermal equation of State $V(P,T)$, dV becomes

$$dV = \left.\frac{\partial V}{\partial P}\right|_T dP + \left.\frac{\partial V}{\partial T}\right|_P dT. \tag{4.36}$$

With this equation (4.32) becomes

$$TdS = \left[\left.\frac{\partial U}{\partial P}\right|_T + P\left.\frac{\partial V}{\partial P}\right|_T\right] dP + \left[\left.\frac{\partial U}{\partial T}\right|_P + P\left.\frac{\partial V}{\partial T}\right|_P\right] dT \tag{4.37}$$

Alternatively, TdS can be written in terms of dT and dV.

$$TdS = \left.\frac{\partial U}{\partial T}\right|_V dT + \left[P + \left.\frac{\partial U}{\partial V}\right|_T\right] dV \tag{4.38}$$

The expansions (equations 4.37 and 4.38) are the total differential of dS in terms of the non-canonical variables P, V and V, T. In contrast to dS in terms of U and V they do not contain the complete thermodynamical information.

Obtaining the Heat Capacity

The heat capacity at constant volume is defined as

$$C_V = \left.\frac{\Delta Q}{\Delta T}\right|_V. \tag{4.39}$$

Using equation (4.38) and that $\Delta Q = TdS$, we find

$$C_V = \left.\frac{\partial U}{\partial T}\right|_V. \tag{4.40}$$

The heat capacity at constant pressure is

$$C_P = \left.\frac{\Delta Q}{\Delta T}\right|_P. \tag{4.41}$$

Using equation (4.37) and again that $\Delta Q = TdS$ we thus obtain

$$C_P = \left.\frac{\partial U}{\partial T}\right|_P + P\left.\frac{\partial V}{\partial T}\right|_P \tag{4.42}$$

$$= \left.\frac{\partial (U + PV)}{\partial T}\right|_P \tag{4.43}$$

$$= \left.\frac{\partial H}{\partial T}\right|_P. \tag{4.44}$$

Here $H = U + PV$ is the enthalpy. For an ideal gas $U = \frac{3}{2}k_BTN$ and hence

$$c_V = \left.\frac{\partial U}{\partial T}\right|_V \tag{4.45}$$

$$= \frac{3}{2}k_BN \tag{4.46}$$

$$c_P = \left.\frac{\partial U}{\partial T}\right|_P + P\left.\frac{\partial V}{\partial T}\right|_P \tag{4.47}$$

$$= \frac{3}{2}k_BN + Nk_B \tag{4.48}$$

$$= \frac{5}{2}Nk_B > c_V. \tag{4.49}$$

$$\tag{4.50}$$

In deriving equation (4.48) the equation of state

$$PV = Nk_BT \tag{4.51}$$

$$\Rightarrow \left.\frac{\partial V}{\partial T}\right|_P = \frac{Nk_B}{P} \tag{4.52}$$

has been used. So in general $c_P \geq c_V$, because part of the absorbed heat is converted to mechanical work.

TdS **equations**

Now the task is to relate the prefactors in equations (4.38) and (4.37) to response functions.

$$dS = \frac{1}{T}\left.\frac{\partial U}{\partial T}\right|_V dT + \left(\frac{P}{T} + \frac{1}{T}\left.\frac{\partial U}{\partial V}\right|_T\right) dV \tag{4.53}$$

Using

$$\frac{\partial^2 S}{\partial T \partial V} = \frac{\partial^2 S}{\partial V \partial T} \tag{4.54}$$

this yields

$$\left.\frac{\partial}{\partial V}\right|_T \frac{1}{T}\left.\frac{\partial U}{\partial T}\right|_V = \left.\frac{\partial}{\partial T}\right|_V \left(\frac{P}{T} + \frac{1}{T}\left.\frac{\partial U}{\partial V}\right|_T\right) \tag{4.55}$$

$$\frac{1}{T}\frac{\partial^2 U}{\partial V \partial T} = -\frac{1}{T^2}\left(P + \left.\frac{\partial U}{\partial V}\right|_T\right) + \frac{1}{T}\left.\frac{\partial P}{\partial T}\right|_V + \frac{1}{T}\frac{\partial^2 U}{\partial T \partial V} \tag{4.56}$$

$$P + \left.\frac{\partial U}{\partial V}\right|_T = T\left.\frac{\partial P}{\partial T}\right|_V. \tag{4.57}$$

The second step now is to relate $\left.\frac{\partial P}{\partial T}\right|_V$ to something already known.

$$dV = \left.\frac{\partial V}{\partial T}\right|_P dT + \left.\frac{\partial V}{\partial P}\right|_T dP \tag{4.58}$$

$$dP = \frac{dV}{\frac{\partial V}{\partial P}|_T} - \frac{\frac{\partial V}{\partial T}|_P}{\frac{\partial V}{\partial P}|_T} dT \tag{4.59}$$

With that the derivative of P with respect to T can be rewritten as

$$\frac{\partial P}{\partial T}|_V = -\frac{\frac{\partial V}{\partial T}|_P}{\frac{\partial V}{\partial P}|_T} \tag{4.60}$$

$$= \frac{\alpha}{k_T} \tag{4.61}$$

With all of this

$$P+\frac{\partial U}{\partial V}|_T = T\frac{\partial P}{\partial T}|_V \tag{4.62}$$

$$= \frac{T\alpha}{k_T} \tag{4.63}$$

Therefore

$$TdS = c_V dT + \frac{T\alpha}{\kappa_T}dV \tag{4.64}$$

$$TdS = c_P dT - \alpha TV dP. \tag{4.65}$$

Where the first one was derived above and the second one is left as an exercise. These so-called TdS equations are important since for adiabatic processes $TdS = 0$.

4.5 Different Ways of Expanding a Gas

There are many different paths of expanding a gas from an initial Volume V_i to a final volume $V_f > V_i$. The mechanical work W performed by the gas is

$$W = \int_{V_i}^{V_f} P(V)dV. \tag{4.66}$$

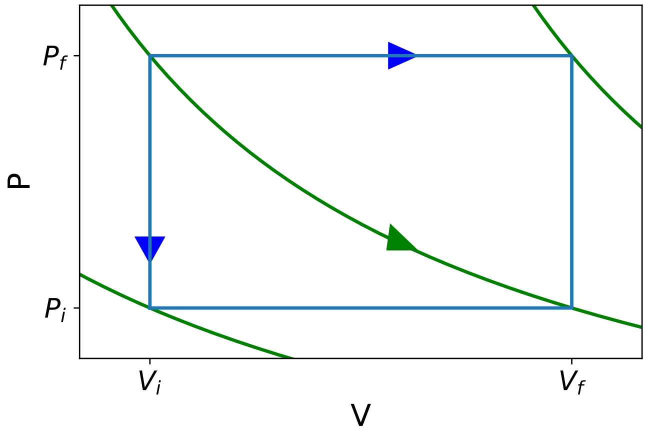


Figure 4.1: $P - V$ diagram of an expanding gas from initial volume V_i to final volume V_f. Green lines are isotherms and blue lines are isochors (vertical) and isobars (horizontal).

Here it can be assumed that the state changes are slow, so that the equilibrium holds and $T(P,V)$ is given by the equation of state. So the path 1 can be assumed to be an isothermal expansion ($P = \frac{Nk_BT}{V}$). The work W^1 along this path is

$$W^1 = Nk_BT\int_{V_i}^{V_f}\frac{dV}{V} \tag{4.67}$$

$$= Nk_BT\ln\left(\frac{V_F}{V_i}\right) \tag{4.68}$$

Obviously path 2 and 3 are not isothermal

$$W^2 = P_i(V_f - V_i) = \frac{Nk_BT}{V_i} = Nk_BT\left(\frac{V_f}{V_i} - 1\right) \tag{4.69}$$

$$W^3 = P_f(V_f - V_i) = \frac{Nk_BT}{V_f} = Nk_BT\left(1 - \frac{V_i}{V_f}\right) \tag{4.70}$$

Clearly $W^2 > W^1 > W^3$, since $\frac{V_f}{V_i} - 1 > \ln\left(\frac{V_f}{V_i}\right) > 1 - \frac{V_i}{V_f}$ or $x - 1 > \ln(x) > 1 - \frac{1}{x}$ for $x > 1$. It can be clearly seen that the work is not an exact differential, since $W_{i\to f}$ is path dependent. So the total work W_{tot} done in a cyclic process is

$$W_{\text{tot}} = \oint \Delta W \tag{4.71}$$

$$= \oint PdV \tag{4.72}$$

$$\neq 0. \tag{4.73}$$

This is the area of a cyclic process in the PV diagram, so the work done is not a state function. Because of that a work function $W(P, V)$ has no meaning.

4.5.1 Conclusions

- For a closed contour integral over the first law of thermodynamics

  $$dU = \Delta Q - \Delta W \tag{4.74}$$

  we have

  $$\oint dU = 0. \tag{4.75}$$

  Thus, it becomes clear that heat is also not a state function, since in general

  $$\oint \Delta Q = \oint \Delta W \tag{4.76}$$

  $$\neq 0. \tag{4.77}$$

- In this example, it also becomes clear that $U_f - U_i = 0$ and with that the absorbed heat Q is equal to the work W performed along the isotherm for an ideal gas, since $U = \frac{3}{2}Nk_BT$ (for an ideal mono-atomic gas).

- For a thermally insulated system, so that $\Delta Q \overset{!}{=} 0$, it can be seen that

  $$dU = -\Delta W \tag{4.78}$$

  $$= -dW \tag{4.79}$$

  and the differential work becomes an exact differential.

4.5.2 Adiabatic Expansion of an Ideal Gas

Now a third TdS equation can be derived. Inserting

$$dT = \frac{1}{\alpha V}dV + \frac{\kappa_T}{\alpha}dP \tag{4.80}$$

into the first (equation (4.65)) and second (equation (4.65)) TdS equation yields

$$TdS = \frac{c_V \kappa_T}{\alpha} dP + \left(\frac{c_V}{\alpha V} + \frac{T\alpha}{\kappa_T} \right) dV \tag{4.81}$$

and

$$TdS = \frac{c_P}{\alpha V} dV + \left(\frac{c_P \kappa_T}{\alpha} - \alpha T V \right) dP. \tag{4.82}$$

Comparing the coefficients of dP yields

$$\frac{c_V \kappa_T}{\alpha} = \frac{c_P \kappa_T}{\alpha} - \alpha T V \tag{4.83}$$

$$c_P - c_V = \alpha^2 \frac{TV}{\kappa_T}. \tag{4.84}$$

Comparing the coefficients of dV also yields

$$\frac{c_V}{\alpha V} + \frac{T\alpha}{\kappa_T} = \frac{c_P}{\alpha V} \tag{4.85}$$

$$c_P - c_V = \alpha^2 \frac{TV}{\kappa_T}. \tag{4.86}$$

Inserting either of equations (4.84) or (4.86) into either of the modified TdS equations (4.82) or (4.82) yields

$$TdS = \frac{c_V \kappa_T}{\alpha} dP + \frac{c_P}{\alpha V} dV. \tag{4.87}$$

This now is the third TdS equation. It can be used to describe an adiabatic process ($TdS = 0$).

$$TdS = \frac{c_V \kappa_T}{\alpha} dP + \frac{c_P}{\alpha V} dV \tag{4.88}$$

$$= 0 \tag{4.89}$$

For the ideal gas we have

$$\kappa_T = - \frac{1}{V} \frac{\partial V}{\partial P} \bigg|_T \tag{4.90}$$

$$= - \frac{1}{V} \frac{\partial N k_B T}{\partial P} \bigg|_T \tag{4.91}$$

$$= + \frac{N k_B T}{V P^2} \tag{4.92}$$

$$= \frac{1}{P}. \tag{4.93}$$

For a monoatomic, ideal gas

$$c_V = \frac{3}{2} N k_B \tag{4.94}$$

$$c_P = \frac{5}{2} N k_B \tag{4.95}$$

which are independent of V and P. Inserting these expressions into equation (4.89) this yields

$$\frac{dP}{P} = -\gamma \frac{dV}{V} \tag{4.96}$$

$$\ln(P)\big|_i^f = -\gamma \ln(V)\big|_i^f \tag{4.97}$$

$$\ln\left(\frac{P_f}{P_i}\right) = -\gamma \ln\left(\frac{V_i}{V_f}\right) \tag{4.98}$$

$$\frac{P_f}{P_i} = \left(\frac{V_i}{V_f}\right)^\gamma. \tag{4.99}$$

This is the adiabatic equation and $\gamma = \frac{c_P}{c_V} \geq 1$ is the adiabatic exponent. Since $\gamma \geq 1$ the adiabatic $P(V)$ curves fall off faster than the isotherms. The temperature decreases during this adiabatic transformation. So for the ideal gas we find

$$\gamma = \frac{c_P}{c_V} \tag{4.100}$$

$$= \frac{5}{3}, \tag{4.101}$$

which can be used to describe the Carnot process.

4.6 Carnot Process

The Carnot process is the idealisation of real heat engines which convert heat into work. The $P-V$-diagram can be seen in figure 4.2.

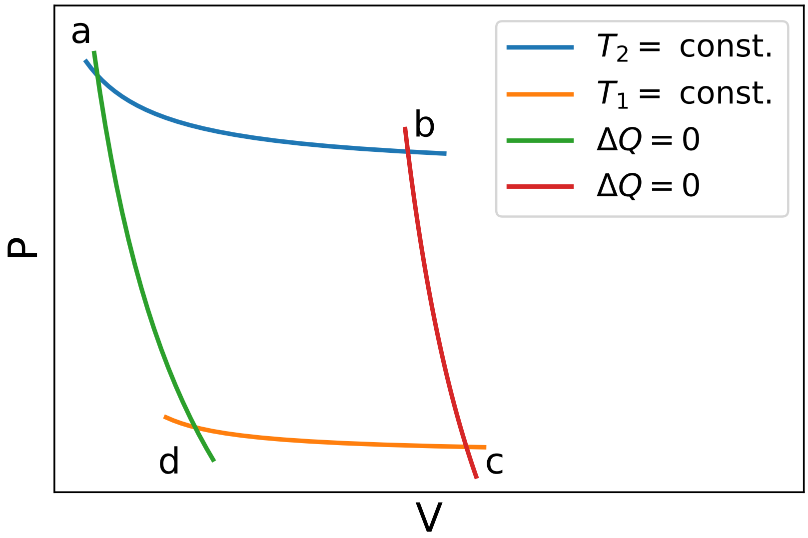


Figure 4.2: P-V-diagram of a typical heat engine / Carnot cycle

The cycle consist of four steps (colours refer to figure 4.2).

- $a \to b$: isothermal expansion at T_1 (green)
- $b \to c$: adiabatic expansion (light blue)
- $c \to d$: isothermal compression at $T_2 < T_1$ (blue)
- $d \to a$: adiabatic compression (red)

Figures 4.3 and 4.4 show these four steps again in a semi-realistic realisation with a cylinder that is periodically connected to two different heat reservoirs. As a result heat is transferred from the hot reservoir (T_1) to the cold reservoir (T_2).

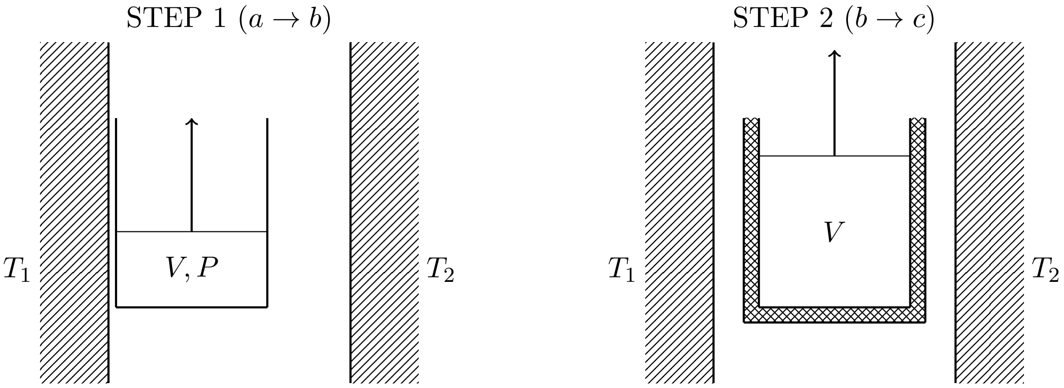


Figure 4.3: Carnot process steps 1 and 2: isothermal and adiabatic expansion. (In step 2 the cylinder is thermally insulated - shown with the cross-hatching pattern).

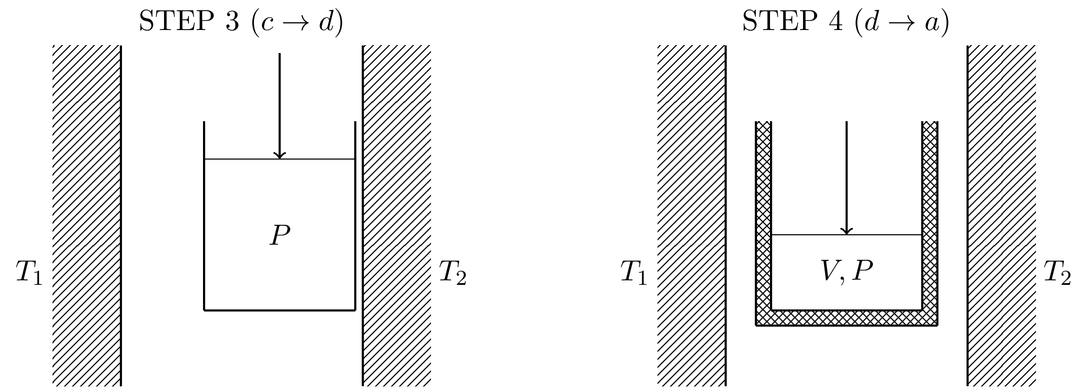


Figure 4.4: Carnot process steps 3 and 4: isothermal and adiabatic compression. (In step 4 the cylinder is thermally insulated - shown with the cross-hatching pattern)

As in every cycle the integral over U is zero

$$\oint dU = 0, \tag{4.102}$$

it follows that

$$\begin{align} \oint \Delta W &\equiv W_{\text{tot}} \tag{4.103} \\ &= Q_1 + Q_2. \tag{4.104} \end{align}$$

In the adiabatic processes no heat is absorbed.

Here W_{tot} is the total work done by the system. This is equivalent to the area within one cycle in the $P - V$ diagram. Q_1 is the heat absorbed from reservoir 1 (which has the temperature T_1). Q_2 is the heat absorbed from reservoir 2 (which has the temperature T_2). Further

$$\begin{align} \oint dS &= 0 \tag{4.105} \\ &= {}^{Q_1}/_{T_1} + {}^{Q_2}/_{T_2} \tag{4.106} \end{align}$$

and thus

$$Q_2 = -\frac{T_2}{T_1} Q_1 \tag{4.107}$$

So heat is injected into the reservoir 2 and the total work W_{tot} is

$$\begin{align} W_{\text{tot}} &= Q_1 - \frac{T_2}{T_1} Q_1 \tag{4.108} \\ &= Q_1 \left(1 - \frac{T_2}{T_1}\right) \tag{4.109} \\ &= Q_1 \eta. \tag{4.110} \end{align}$$

Here $\eta = \frac{W_{\text{tot}}}{Q_1} = 1 - \frac{T_2}{T_1}$ is defined as the efficiency of the heat engine; it is the net work done per absorbed heat from the hotter reservoir. Since $0 < \frac{T_2}{T_1} < 1$, it can be seen that $0 < \eta < 1$. W_{tot} is larger, the smaller T_2 is. For $T_2 = T_1$ the work and the efficiency are zero.

4.6.1 Principle of Cogeneration (of Heat and Electricity)

The idea is to heat a house as efficiently as possible. In the house is a tank with natural gas to burn and below the house ground water can be found. By just burning the gas per one Joule of burned gas the house is fed with (obviously) one Joule of heat. By using a Carnot engine this can be done much better: The first idea would be to just burn gas in a heat engine and use the electricity to power other appliances. The room is at $T_1 = 313K$ and the gas burns at $T_2 = 873K$. The efficiency of the heat engine is

$$\eta = 1 - \frac{313}{873} \tag{4.111}$$
$$= 0.64 \tag{4.112}$$

So for every Joule of chemical energy $0.64J$ go into mechanical (quasi electrical) work and $0.36J$ go into the house. But this can be improved further: The $0.64J$ can be fed into the heat pump to transfer heat from the ground water at $283K$ to the house as well. The coefficient of performance can be calculated via

$$\frac{1}{\eta} = \frac{T_1}{T_1 - T_2} \tag{4.113}$$
$$= \frac{313}{313 - 283} \tag{4.114}$$
$$= 10.4. \tag{4.115}$$

This means that for every one Joule of electric power 10.4 Joule of heat can be transferred. Using the $0.64J$ of electricity from above to power this ground-water-based heat engine

$$E = 0.36J + 10.4 \cdot 0.64J \tag{4.116}$$
$$\approx 6.7J \tag{4.117}$$

of heat can be transferred to the house per one Joule of burned gas. This seems like a violation of energy conservation, but it is not to forget that the groundwater actually became a lot colder by doing this. This is where the energy comes from. There is just a huge amount of ground water, so the effect is not really noticeable. These systems are used in many private and public buildings. The process is called co-generation (of heat and electricity)

4.7 Transfer and Creation of Entropy

Now we discuss why real engines are less efficient than the idealised Carnot process. ΔQ is the heat transfer from reservoir 1 at T_1 to reservoir 2 at T_2. The two reservoirs are connected via a wire that conducts heat (see Fig. 4.5).

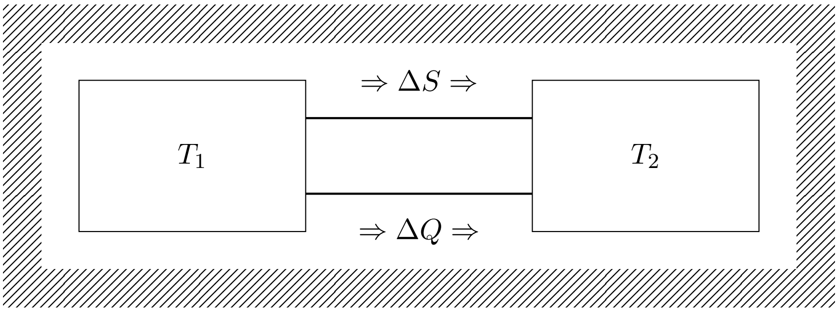


Figure 4.5: Two reservoirs with temperatures T_1 and T_2 can exchange heat and entropy via wires. (The stripy pattern represents thermal insulation).

4.7.1 Reversible Case

The entropy changes in the reservoirs are $\Delta S_1 = -\frac{\Delta Q}{T_1} < 0$ and $\Delta S_2 = -\frac{\Delta Q}{T_2} > 0$. For $T_1 = T_2$ the total entropy change $\Delta S_{\text{tot}} = \Delta S_1 + \Delta S_2 = 0$ is zero. This is an example of a reversible heat transfer. Heat is transferred, but the total entropy stays constant. The second law of thermodynamics allows these kinds of reversible processes to happen.

4.7.2 Irreversible Case

Opposite to the reversible case, here the assumption is that $T_1 > T_2$. The heat will flow from the hot to the cold reservoir. ΔQ is so small that T_1 and T_2 stay quasi constant. The entropy however increases:

$$\begin{aligned}
\Delta S_{\text{tot}} &= \Delta S_1 + \Delta S_2 && (4.118)\\
&= \Delta Q \left(\frac{1}{T_2} - \frac{1}{T_1} \right) && (4.119)\\
&= \Delta Q \frac{T_1 - T_2}{T_1 T_2} && (4.120)\\
&> 0, && (4.121)
\end{aligned}$$

and thus this process is irreversible. Note that the opposite process, i.e. heat flowing from the cold to the hot reservoir, would increase the entropy and is not allowed.

4.7.3 Comments

- The Carnot process is reversible if no heat leaks from the hot to the cold reservoir.
- The heat flow from the hot to the cold reservoir without the conversion to mechanical work is the main reason for the low efficiency of real heat engines.
- There is no heat engine with a higher efficiency than the Carnot engine.

4.8 Extremal Properties of Thermodynamic Potentials

- In an insulated system the entropy $S(U, V, N)$ is maximised, as shown in section **??**.
- If energy transfer is allowed between a small system and a reservoir (see Fig. 4.6),

$$\begin{aligned}
dU &= TdS - PdV + \mu dN && (4.122)\\
dS &= \frac{1}{T}dU + \frac{P}{T}dV - \frac{\mu}{T}dN && (4.123)\\
S_{\text{tot}} &= S(U_1, V_1, N_1) + S(U - U_1, V - V_1, N - N_1) && (4.124)\\
&= S(U_1, V_1, N_1) + S(U, V, N) - \frac{U_1}{T} && (4.125)\\
&= S(U_1, V_1, N_1) - \frac{U_1}{T}. && (4.126)\\
&= -\frac{1}{T}(U_1 - S_1 T) && (4.127)
\end{aligned}$$

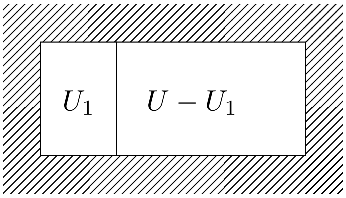


Figure 4.6: A small reservoir in heat contact with a reservoir (The stripy pattern represents thermal insulation).

So from the second law of thermodynamics follows that $S_1 - \frac{U_1}{T}$ is maximised. And thus $U_1 - TS_1$ is minimised. The free energy $F = U - TS$ is minimised when T, V, N are fixed.

- Now entropy transfer and volume transfer are allowed: V_1 and U_1 can change (see Fig. 4.6):

$$S_{\text{tot}} = S(U_1, V_1, N_1) + S(U - U_1, V - V_1, N - N_1) \tag{4.128}$$

$$= S_1 + S(U, V, N) - \frac{U_1}{T} - \frac{V_1 P}{T} \tag{4.129}$$

The total entropy S_{tot} is maximised, and thus $TS_1 - U_1 - V_1 P$ is maximised. With that the Gibbs free energy (free enthalpy) $G(T, P, N) = U - TS + PV$ is minimised for a system at fixed T, P, N.

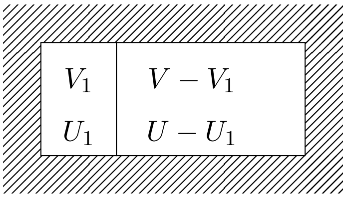


Figure 4.7: A small reservoir in contact with a reservoir, but now V_1 and U_1 can change (The stripy pattern represents thermal insulation).

4.9 Thermodynamic Description of Phase Transitions

Simple substances can exist in one of three phases, solid, liquid, vapour (gas).

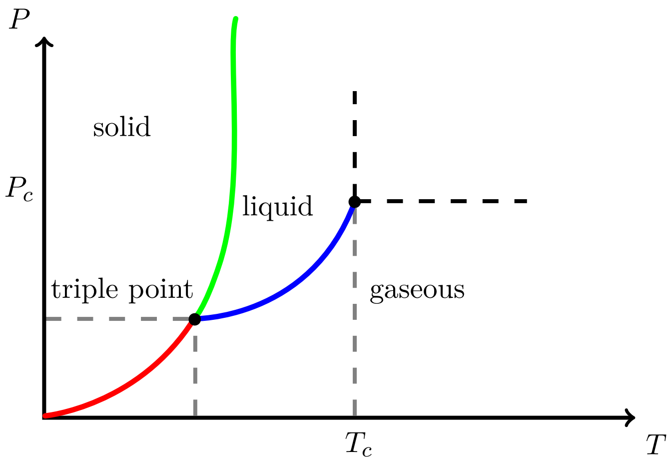


Figure 4.8: Example of a phase diagram.

Along the lines in the diagram, two phases coexist, here are two solutions to the equation of state with different volumes $v = \frac{V}{N}$ per particle. On the two-phase coexistence lines, the system splits into two coexisting phases (phase separation / phase equilibrium). At the triple point all three phases coexist. At the critical point the difference between two phases (liquid/vapour) vanishes.

4.9.1 Thermodynamic Stability at the Phase Transitions

Consider a two phase system like liquid-vapour at constant T and P. As shown in section 4.8, at constant T,P, equilibrium is characterised by a minimum of the Gibbs free energy (free enthalpy)

$$\begin{aligned} G &= U - TS + PV && (4.130)\\ &= \sum_i \mu_i N_i. && (4.131)\end{aligned}$$

Now let $N = N_1 + N_2$, where N_1 is the number of particles in phase 1 and N_2 the number of particles in phase 2.

$$\begin{aligned} G &= N_1\mu_1 + N_2\mu_2 && (4.132)\\ &= N_1\mu_1 + (N - N_1)\mu_2 && (4.133)\end{aligned}$$

This can be minimised with respect to N_1

$$\begin{aligned} \frac{\partial G}{\partial N_1} &= \frac{\partial}{\partial N_1}(N_1\mu_1 + (N - N_1)\mu_2) && (4.134)\\ &= \mu_1 - \mu_2 && (4.135)\\ &\overset{!}{=} 0. && (4.136)\end{aligned}$$

We see that in equilibrium the chemical potential has to be equal for the two phases.

4.9.2 Clausius-Clapeyron Equation

Consider two coexisting phases $i = 1, 2$. For each phase we can write

$$\begin{aligned} dG_i &= -S_i dT + V_i dP + \mu_i dN_i && (4.137)\\ &= d(\mu_i N_i) && (4.138)\\ &= \mu_i dN_i + N_i d\mu_i && (4.139)\\ d\mu_i &= -\frac{S_i}{N_i}dT + \frac{V_i}{N_i}dP. && (4.140)\end{aligned}$$

With the definition of the total differential this yields

$$\begin{aligned} \left.\frac{\partial \mu_i}{\partial T}\right|_P &= -\frac{S_i}{N_i} && (4.141)\\ \left.\frac{\partial \mu_i}{\partial P}\right|_T &= \frac{V_i}{N_i} && (4.142)\\ \left.\frac{\partial(\mu_2 - \mu_1)}{\partial T}\right|_P &= -\left(\frac{S_2}{N_2} - \frac{S_1}{N_1}\right) && (4.143)\\ &= -\Delta s. && (4.144)\end{aligned}$$

Here $s = \frac{S}{N}$ is the entropy per particle.

$$\begin{aligned} \left.\frac{\partial(\mu_2 - \mu_1)}{\partial P}\right|_T &= \frac{V_2}{N_2} - \frac{V_1}{N_1} && (4.145)\\ &= \Delta v && (4.146)\end{aligned}$$

Here $v = \frac{V}{N}$ is the volume per particle. With $\Delta\mu = \mu_2 - \mu_1$ and using the differential chain rule this becomes

$$\begin{aligned} \frac{\left.\frac{\partial \Delta\mu}{\partial T}\right|_P}{\left.\frac{\partial \Delta\mu}{P}\right|_T} &= -\frac{\Delta s}{\Delta v} \qquad (4.147) \\ \frac{dP_{co}(T)}{dT} &= \left.\frac{\partial P}{\partial T}\right|_{\Delta\mu=0} \qquad (4.148) \\ &= \frac{\Delta s}{\Delta v}. \qquad (4.149) \end{aligned}$$

With $\Delta q_i = T\Delta s$, which is the latent heat required for the phase change, this becomes

$$\frac{dP_{co}}{dT}(T) = \frac{\Delta q_i}{T\Delta v} \qquad (4.150)$$

This is the Clausius-Clapeyron equation.

5 Statistical Mechanics of Real Systems

5.1 Virial Theorem

Now considering a real gas of interacting atoms, again N denotes the number of gas particles. The Hamiltonian is

$$\mathbb{H} = \mathbb{H}_{\text{kin}} + \mathbb{H}_{\text{pot}} \tag{5.1}$$

$$\mathbb{H}_{\text{kin}} = \sum_{j=1}^{N} \frac{p_j^2}{2m_j} \tag{5.2}$$

$$\mathbb{H}_{\text{pot}} = \sum_{j=1}^{N} \sum_{k=j+1}^{N} \left[(\vec{q}_j - \vec{q}_k)^2 \right]^{\frac{n}{2}} a_{jk} \tag{5.3}$$

Here the parameters a_{jk} give the strength of the potential and n denotes the exponent of the power law interaction. With this, different scenarios can be considered.

- Gravitation, where $n = -1$ and $a_{jk} \propto -m_j m_k$
- One-component plasma, where $n = -1$ and $a_{jk} \propto q^2$; q is the particle charge.
- Two-component plasma, where $n = -1$ and $a_{jk} \propto q_j q_k$; q_j is the charge of particle j.

This model is very general, as this is a power law interaction model and n can be any arbitrary number. The partition function Z is given by

$$Z = \frac{1}{N!} \prod_{j=1}^{N} \left[\int_V d^3 q_j \int \frac{d^3 p_j}{h^3} \right] e^{-\beta(H_{\text{kin}} + H_{\text{pot}})}. \tag{5.4}$$

An important note is that Z cannot be calculated explicitly. However some interaction information can be obtained analytically. For this it is useful to rescale T,p and q, using a rescaling factor of λ, according to

$$T = \lambda^n \tilde{T} \tag{5.5}$$

$$q = \lambda \tilde{q} \tag{5.6}$$

$$p = \lambda^{\frac{n}{2}} . \tilde{p} \tag{5.7}$$

This rescaling is constructed so that the Hamiltonian stays invariant under this rescaling.

$$\frac{\frac{p_j^2}{2m_j} + (q_j - q_k)^n a_j k}{k_B T} = \frac{\frac{\tilde{p}_j^2}{2m_j} + (\tilde{q}_j - \tilde{q}_k)^n a_j k}{k_B \tilde{T}} \tag{5.8}$$

The integration boundaries however do change as

$$\int_0^L dq = \int_0^{\frac{L}{\lambda}} \lambda d\tilde{q} \tag{5.9}$$

$$= \int_0^{\tilde{L}} \lambda d\tilde{q} \tag{5.10}$$

$$L = \lambda \tilde{L} \tag{5.11}$$

$$V = \lambda^3 \tilde{V}. \tag{5.12}$$

Here L is the length of a cube with volume V. Now the partition function in rescaled coordinates becomes

$$Z(V,T,N) = \frac{1}{N!}\prod_{j=1}^{N}\left[\int_{\tilde{V}} d^3\tilde{q}_j\lambda^3 \int \frac{d^3\tilde{p}_j}{h^3}\lambda^{\frac{3n}{2}}\right] e^{-\frac{\mathbb{H}_{\text{kin}}(\tilde{p})+\mathbb{H}_{\text{pot}}(\tilde{q})}{k_B\tilde{T}}} \tag{5.13}$$

$$= \lambda^{3N(1+\frac{n}{2})} Z(\tilde{V},\tilde{T},N) \tag{5.14}$$

This λ factor has no physical meaning, it is just a scaling factor. Now from the three variables V, T, N three new variables $VT^{-\frac{3}{n}}$, T and N are defined; with those the partition function can be written as

$$Z(V,T,N) \equiv (VT^{-\frac{3}{n}}, T, N) \tag{5.15}$$

$$= (\lambda^3\tilde{V}\tilde{T}^{-\frac{3}{n}}\lambda^{-3}, \lambda^n\tilde{T}, N) \tag{5.16}$$

$$= h(\tilde{V}\tilde{T}^{-\frac{3}{n}}\lambda^n\tilde{T}, N). \tag{5.17}$$

Here h is an unknown function and it can be seen that λ only appears as a pre-factor to $\tilde{T}$. Since Z is proportional to $\lambda^{3N(1+\frac{n}{2})}$, the partition function must have the functional form

$$Z(N,V,T) = (\lambda^n\tilde{T})^{\frac{3N(1+\frac{n}{2})}{n}}\phi(\tilde{V}\tilde{T}^{-\frac{3}{n}}, N), \tag{5.18}$$

where ϕ is another unknown function. This formula for Z is exact, as no approximation were made. The scaling factor λ was used to reveal certain symmetry relations but can well be set to $\lambda = 1$ now. With that Z becomes

$$Z(N,V,T) = T^{\frac{3N(1+\frac{n}{2})}{n}}\phi(VT^{-\frac{3}{n}}, N) \tag{5.19}$$

And from that it can be concluded that the free energy reads

$$F = -k_BT\left[\frac{3N}{n}\left(1+\frac{n}{2}\right)\ln(T) + \ln\left(\phi(VT^{-\frac{3}{N}})\right)\right] \tag{5.20}$$

$$P = -\left.\frac{\partial F}{\partial V}\right|_T \tag{5.21}$$

$$= k_BTT^{-\frac{3}{n}}\frac{\phi'\left(VT^{-\frac{3}{n}}\right)}{\phi\left(VT^{-\frac{3}{n}}\right)} \tag{5.22}$$

$$S = -\left.\frac{\partial F}{\partial T}\right|_V \tag{5.23}$$

$$= -\frac{F}{T} + k_BT\frac{3N}{n}(1+\frac{n}{2})\frac{1}{T} - k_BT\frac{3}{n}T^{-\frac{3}{n}-1}V\frac{\phi'\left(VT^{-\frac{3}{n}}\right)}{\phi\left(VT^{-\frac{3}{n}}\right)} \tag{5.24}$$

by comparison with equation (5.22) this yields

$$= -\frac{F}{T} + k_B\frac{3N}{n}\left(1+\frac{n}{2}\right) - \frac{3}{n}\frac{V}{T}P \tag{5.25}$$

and with

$$F = U - TS \tag{5.26}$$

this becomes

$$U = k_B T \left[\frac{3N}{n}\left(1+\frac{n}{2}\right) - \frac{3}{n} V T^{-\frac{3}{n}} \frac{\phi'(VT^{-\frac{3}{n}})}{\phi(VT^{-\frac{3}{n}})}\right] \tag{5.27}$$

$$= k_B T \frac{3N}{n}\left(1+\frac{n}{2}\right) - \frac{3}{n} VP \tag{5.28}$$

This is an exact relation between U, T, N, V and P for an interacting system. It is not an equation of state, as it depends on 5, state variables. This is called the virial theorem. For $n = -1$ (gravitational or Coulombic systems this becomes)

$$U = -\frac{3k_B T N}{2} + 3VP \tag{5.29}$$

$$\tag{5.30}$$

For a vanishing interaction $n = 0$ it becomes

$$U = \frac{3}{n}\left[k_B T N\right] + \frac{3N}{2} k_B T N \tag{5.31}$$

$$= \frac{3N}{2} k_B T N \tag{5.32}$$

$$k_B T N = VP, \tag{5.33}$$

which is the equation for the ideal gas, which is what was expected.

5.2 Virial Expansion

This virial expansion is a perturbative analysis of a real gas. This chapter is also about perturbation theory. The partition function for an interacting gas is

$$Z\left(N,T,V\right) = \frac{1}{N!}\prod_{j=1}^{N}\left[\int \frac{d^3 p_j d^3 q_j}{h^3}\right] e^{-\beta\left(\mathbb{H}_{\text{kin}}+\mathbb{H}_{\text{pot}}\right)} \tag{5.34}$$

$$\mathbb{H}_{\text{kin}} = \sum_{l=1}^{N} \frac{\vec{p}_l^{\,2}}{2m} \tag{5.35}$$

$$\mathbb{H}_{\text{pot}} = \sum_{l=1}^{N}\sum_{k=l+1}^{N} w\left(q_l - q_k\right) \tag{5.36}$$

Here $w(q)$ is an arbitrary pair potential. The momenta can be integrated out exactly.

$$Z\left(N,T,V\right) = \frac{1}{N!}\prod_{j}\int \frac{d^3 q_j}{\lambda_t^3} e^{-\beta\sum_{k>l}^{N} w(q_l-q_k)} \tag{5.37}$$

Here λ_t is again the thermal wavelength. Now switch to the grand-canonical ensemble.

$$\mathbb{Z}\left(\mu, V, T\right) = \sum_{N=0}^{\infty} e^{\beta N\mu} Z_N\left(V,T\right) \tag{5.38}$$

The grand canonical equation of state was previously derived as

$$PV = k_B T \ln\left(\mathbb{Z}\left(\mu, V, T\right)\right) \tag{5.39}$$

and also

$$N = \frac{1}{\beta}\frac{\partial \ln\left(\mathbb{Z}\left(\mu, V, T\right)\right)}{\partial \mu}. \tag{5.40}$$

$Z = e^{\beta\mu}$ is called the fugacity and with this the grand-canonical partition function from equation (5.38) can be written as a power series in z.

$$\mathbb{Z}\left(\mu, V, T\right) = \sum_{N=0}^{\infty} z^N Z_N\left(V, T\right) \tag{5.41}$$

With that the equation of state can also be written as a power series.

$$\begin{aligned} \frac{PV}{k_B T} &= \ln\left(\mathbb{Z}\right) && (5.42)\\ &\equiv V \sum_{l=1}^{\infty}\left(\frac{z}{\lambda_t^3}\right)^l b_l && (5.43)\\ &= \ln\left[\sum_{N=0}^{\infty} z^N Z_N\right] && (5.44)\end{aligned}$$

With $Z_0 \equiv 1$ the first few terms can be written as

$$= \ln[1 + zZ_1 + z^2 Z_2 + z^3 Z_3 + \ldots] \tag{5.45}$$

using the Taylor expansion for $\ln\left(1+\epsilon\right)$ this yields

$$= zZ_1 + z^2\left(Z_2 - \frac{1}{2}Z_1^2\right) + z^3\left(Z_3 - Z_1 Z_2 + \frac{1}{3}Z_1^3\right) + O\left(z^4\right) \tag{5.46}$$

By comparison with the definition from equation(5.43) the coefficients of b_l can be read of as

$$\begin{aligned} b_1 &= \frac{\lambda_t^3}{V} Z_1 && (5.47)\\ b_2 &= \frac{\lambda_t^6}{V}\left(Z_2 - \frac{1}{2}Z_1^2\right) && (5.48)\\ b_3 &= \frac{\lambda_t^9}{V}\left(Z_3 - Z_1 Z_2 + \frac{1}{3}Z_1^3\right) && (5.49)\\ b_4 &= \ldots . && (5.50)\end{aligned}$$

Using the Mayer function $f_{12} = e^{-\beta w(q_1 - q_2)} - 1$ the Z_n become

$$\begin{aligned} Z_1 &= \int \frac{d^3 q}{\lambda_t^3} e^0 && (5.51)\\ &= \frac{V}{\lambda_t^3} && (5.52)\\ Z_2 &= \frac{1}{2!\lambda_t^6}\int d^3 q_1 d^3 q_2 e^{-\beta w(q_1 - q_2)} && (5.53)\\ &= \frac{1}{2!\lambda_t^6}\int d^3 q_1 d^3 q_2 (1 + f_{12}) && (5.54)\\ Z_3 &= \frac{1}{3!\lambda_t^6}\int d^3 q_1 d^3 q_2 d^3 q_3 \left(1 + f_{12}\right)\left(1 + f_{23}\right)\left(1 + f_{13}\right) && (5.55)\end{aligned}$$

and finally the coefficients b_l follows as

$$b_1 = 1 \tag{5.56}$$

$$b_2 = \frac{1}{2}\int \frac{d^3q_1 d^3q_2}{V} f_{12} \tag{5.57}$$

$$b_3 = \frac{1}{6}\int \frac{d^3q_1 d^3q_2 d^3q_3}{V}\left[(1+f_{12})(1+f_{23})(1+f_{13}) - 3(1+f_{12}) + 2\right] \tag{5.58}$$

$$= \frac{1}{6}\int \frac{d^3q_1 d^3q_2 d^3q_3}{V}\left[f_{12}f_{23}f_{13} + f_{12}f_{23} + f_{12}f_{13} + f_{23}f_{13}\right] \tag{5.59}$$

$$= \frac{1}{6}\int \frac{d^3q_1 d^3q_2 d^3q_3}{V}\left[f_{12}f_{23}f_{13} + 3f_{12}f_{23}\right]. \tag{5.60}$$

With this N becomes

$$N = \frac{1}{\beta}\frac{\partial \ln(\mathbb{Z}(\mu))}{\partial \mu} \tag{5.61}$$

$$= \frac{1}{\beta}\frac{\partial z}{\partial \mu}\frac{\partial \ln(\mathbb{Z}(\mu))}{\partial z} \tag{5.62}$$

$$= z\frac{\partial \ln(\mathbb{Z}(z))}{\partial z} \tag{5.63}$$

With the concentration being defined as

$$c = \frac{N}{V} \tag{5.64}$$

this becomes

$$= \sum_{l=1}^{\infty} l\left(\frac{z}{\lambda_t^3}\right)^l b_l \tag{5.65}$$

$$c \approx b_1\frac{z}{\lambda_t^3} + b_2 2\left(\frac{z}{\lambda_t^3}\right)^2 + \ldots \tag{5.66}$$

$$\frac{z}{\lambda_t^3} \approx c - 2b_2c^2 + O\left(c^3\right) \tag{5.67}$$

$$\frac{P}{k_BT} = \frac{z}{\lambda_t^3} + \left(\frac{z}{\lambda_t^3}\right)^2 b_2 + O\left(\left(\frac{z}{\lambda_t^3}\right)^3\right) \tag{5.68}$$

$$= c - b_2c^2 + O\left(c^3\right). \tag{5.69}$$

The leading term is the ideal gas law and the correction term is proportional to $-b_2c^2$. With the model of hard spheres with diameter d this becomes

$$b_1 = 1 \tag{5.70}$$

$$b_2 = \frac{1}{2}\int \frac{d^3q_1 d^3q_2}{v}\left(e^{-\beta w(q_1-q_2)} - 1\right) \tag{5.71}$$

$$= -\frac{1}{2}\frac{4\pi}{3}d^3 \tag{5.72}$$

$$= -\frac{1}{2}v_{ex} \tag{5.73}$$

Here $v_{ex} = \frac{4\pi}{3}d^3$ is the excluded volume. And with all of that $\frac{P}{k_BT}$ becomes

$$\frac{P}{k_BT} = c + \frac{1}{2}v_{ex}c^2 + O\left(c^3\right) \tag{5.74}$$

So the first correction to the ideal gas law is proportional to the excluded volume of the hard-sphered particles.

5.2.1 Side Note on the Inversion of Power Series

$$c = \sum_{l=1}^{\infty} d_l z^l \tag{5.75}$$

Starting from the Ansatz

$$z = \sum_{l=1}^{\infty} n_l c^l \tag{5.76}$$

and inserting it c becomes

$$\begin{aligned} c &= d_1 z + d_2 z^2 + d_3 z^3 && (5.77) \\ &= d_1 \left(n_1 c + n_2 c^2 + \ldots\right) + d_2 \left(n_1 c + n_2 c^2 + \ldots\right)^2 + \ldots && (5.78) \\ &= d_1 n_1 c + \left(d_1 n_2 + d_2 n_1^2\right) c^2 + \ldots && (5.79) \\ & && (5.80) \end{aligned}$$

By equating the coefficients this yields

$$\begin{aligned} d_1 n_1 &\overset{!}{=} 1 && (5.81) \\ n_1 &= 1/d_1 && (5.82) \\ d_1 n_2 + d_2 n_1^2 &\overset{!}{=} 0 && (5.83) \\ n_2 &= -\frac{d_2 n_1^2}{d_1} && (5.84) \\ &= -\frac{d_2}{d_1^3} && (5.85) \end{aligned}$$

5.3 Van-der-Waals Equation of State.

The typical pair potential $w(r)$ (figure 5.1) of a real gas is

$$w\left(r\right) = \begin{cases} \infty & |r < D \\ -w_0 \left(\frac{D}{r}\right)^s & |r > D \end{cases} \tag{5.86}$$

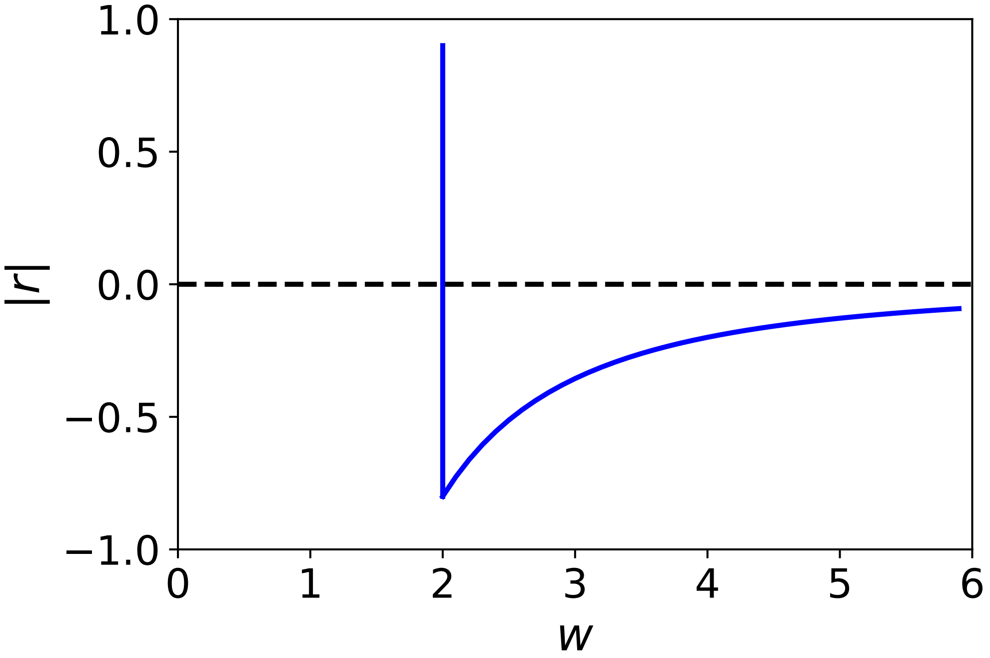


Figure 5.1: Typical pair potential (Here for example: $D = 2$, $s = 2$)

The first non-trivial virial coefficient is the second one, which becomes

$$a_2 = \frac{1}{2}\int_0^\infty dr 4\pi r^2 \left(1 - e^{-\beta w(r)}\right) \tag{5.87}$$

$$= 2\pi\int_0^D dr r^2 + 2\pi\int_D^\infty dr r^2 \left(1 - e^{\frac{w_0}{k_B T}\left(\frac{D}{r}\right)^S}\right) \tag{5.88}$$

The second integral can not be solved in closed form as it will result in a gamma function. So rather weak attraction can be assumed and

$$\frac{w_0}{k_B T} < 1 \tag{5.89}$$

$$a_2 = \frac{2\pi}{3}D^3 - 2\pi D^S \frac{w_0}{k_B T}\int_D^\infty dr r^{2-s}. \tag{5.90}$$

For $s > 3$ the integral converges and it can be written as

$$a_2 = \frac{2\pi}{3}D^3 - \frac{2\pi}{s-3}D^3\frac{w_0}{k_B T} \tag{5.91}$$

with

$$a \equiv \frac{2\pi}{s-3}D^3 w_0 \tag{5.92}$$

$$b \equiv \frac{2\pi}{3}D^3 \tag{5.93}$$

this becomes

$$a_2 \equiv b' - \frac{a'}{k_B T} \tag{5.94}$$

$$\tag{5.95}$$

Which is now the second virial coefficient. It can be plugged into the virial expansion.

$$\frac{P}{k_B T} = c + a_2 c^2 \tag{5.96}$$

$$= c + \left(b - \frac{a}{k_B T}\right) c^2 \tag{5.97}$$

$$P = c k_B T + \left(b k_B T - a\right) c^2 \tag{5.98}$$

$$P + a' c^2 = c k_B T \left(1 + bc\right) \tag{5.99}$$

with the expansion for $\frac{1}{1-\epsilon} \approx 1 + \epsilon + O\left(\epsilon^2\right)$ this becomes

$$\approx \frac{c k_B T}{1 - bc} \tag{5.100}$$

$$P + a' c^2 = \frac{c k_B T}{1 - bc} \tag{5.101}$$

$$\left(P + a c^2\right)\left(1 - bc\right) = c k_B T \tag{5.102}$$

$$\tag{5.103}$$

With $c = \frac{N}{V}$ this becomes the Van-der-Waals equation of state.

$$\boxed{\left(P + \frac{aN^2}{V^2}\right)\left(V - b'N\right) = N K_B T} \tag{5.104}$$

5.3.1 Conclusions

- In the ideal limit ($w_0 \to 0$ and $V \to 0$ and from that results that $a = b = 0$ and the Van-der-Waals equation of state becomes the ideal equation of state.
- The real volume $V_{\text{real}} = V - bN$ is reduced by the volume of particles.
- The real pressure $P_{\text{real}} = P + \frac{aN^2}{V^2}$ is increased by the attraction between the particles, as $a_2 \propto w_0 D^3$.

Equation (5.104) can be rewritten as

$$\boxed{P = \frac{N k_B T}{-bN} - \frac{qN^2}{V^2}} \tag{5.105}$$

which is plotted in figure 5.2.

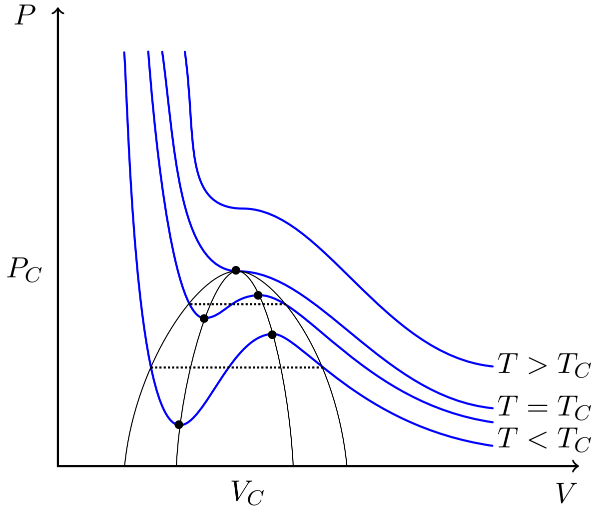


Figure 5.2: P-V diagram of the Van-der-Waals equation

So there is a finite range of volume for $T < T_c$ within which the volume goes up when the pressure increases. As a negative compressibility is unphysical, the Maxwell construction is used.

$$dF = -SdT - PdV \tag{5.106}$$

With a constant T this becomes

$$= -PdV. \tag{5.107}$$

Hence the free energy difference between two volumes V_1 and V_2 along an isotherm is

$$\Delta F = -\int_{V_1}^{V_2} P\left(V\right) dV \tag{5.108}$$

The ΔG for two system parts with the same P_0 is

$$\begin{aligned} \Delta G &= \Delta F + P_0 \Delta V && (5.109) \\ &= \int_{V_1}^{V_2} \left(P_0 - P\left(V\right)\right) dV && (5.110) \\ &= 0 && (5.111) \end{aligned}$$

This is the equilibrium condition at phase coexistence.

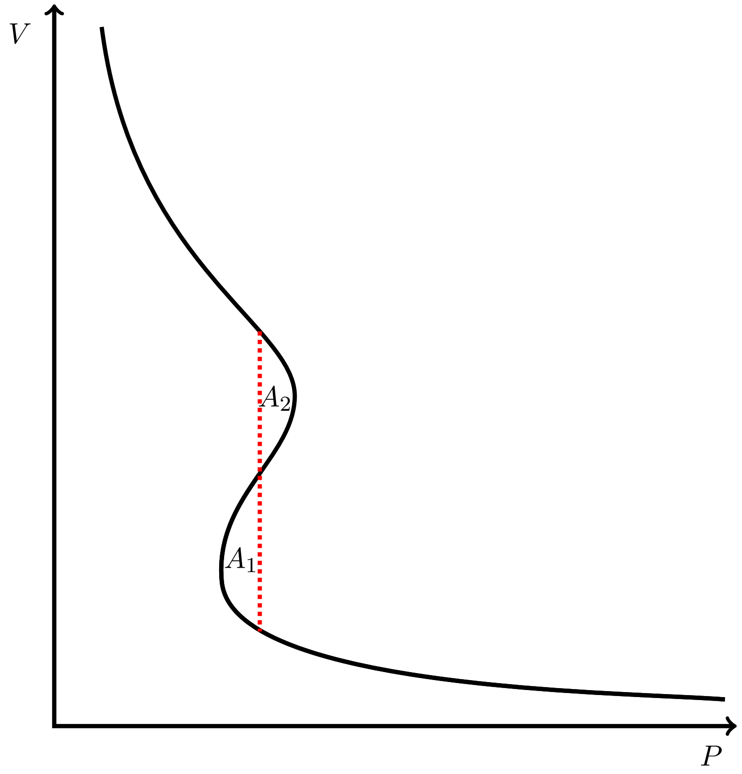


Figure 5.3: V-P-diagram. The two areas A_1 and A_2 must be equal: only then the integral becomes zero.

The region with $\frac{dP}{dV} > 0$ is never observed, because the system splits into two volumes V_1 and V_2 at pressure P_0. These so called binodals limit the two-plane coexistence region. Below the critical temperature $T < T_c$ there are two planes and there are volumes, where the slope is zero, there is a volume where the curvature is zero. At the critical point the slope and the curvature are zero. This is how the critical point can be found. So from equation (5.105) the volume, temperature (up to a factor of k_B) and the pressure at the critical (c) point Q_c are $V_{\text{critical}} = 3b'N$, $k_B T_c = \frac{8}{27}\frac{a'}{b'}$ and $P_{\text{critical}} = \frac{a'}{27b'^2}$.

6 Quantum Statistics

6.1 Quantum Theory

There are a number of experimental discoveries that led to the development of quantum theory. One of them is the photo electric effect, where an electron is emitted from a metal surface when hit by radiation, which suggested that the energy is quantised by h with $E = h\nu$. The second one was the electron diffraction experiment which resulted in $\lambda = \frac{h}{p}$. In classical mechanics Newtons equation of motion describes the energy E of a system $E = \frac{p^2}{2m}$. This can also be applied to waves. The simplest possible wave is a planar wave

$$\Psi(x,t) \propto e^{i(kx-\omega t)} \tag{6.1}$$

with the momentum

$$p = \frac{h}{\lambda} = \hbar k. \tag{6.2}$$

$$\tag{6.3}$$

Here $k = \frac{2\pi}{\lambda}$ and $\hbar = \frac{h}{2\pi}$. With $\omega = 2\pi\nu$ the energy can be written as

$$E = h\nu = \hbar\omega. \tag{6.4}$$

And thus the Schrödinger equation emerges:

$$i\hbar\frac{\partial\Psi}{\partial t} = -\frac{\hbar^2}{2m}\frac{\partial^2\Psi}{\partial x^2}. \tag{6.5}$$

With an external potential V the energy becomes

$$E = \frac{p^2}{2m} + V(x). \tag{6.6}$$

And the stationary Schrödiger equation is

$$E_n\Psi_n(x) = \hat{\mathbb{H}}\Psi_n(x), \tag{6.7}$$

with the Hamiltonian

$$\hat{\mathbb{H}} = -\frac{h^2}{2m}\frac{\partial^2}{\partial x^2} + V(x). \tag{6.8}$$

6.1.1 N-Particle Wave Function

The N-particle Hamiltonian reads

$$\hat{\mathbb{H}} = \frac{\hbar^2}{2m}\sum_{i=1}^{N}\vec{\nabla}_i^2 + \underbrace{\sum_{i=1}^{N}u(\vec{r}_i)}_{\text{external potential}} + \underbrace{\sum_{i=1}^{N}v(\vec{r}_i-\vec{r}_j)}_{\text{interaction potential}} . \tag{6.9}$$

The Schrödinger equation becomes

$$\hat{\mathbb{H}}\Psi_n(\vec{r}_1,\dots,\vec{r}_N) = E_n\Psi_n(\vec{r}_1,\dots,\vec{r}_N) \tag{6.10}$$

Pauli 1940 in his famous paper (Physical Reviews 58, 716) introduced a permutation operator P_{ij}, exchanging the ith and the jth particle.

$$\hat{P}_{ij}\Psi_n(\vec{r}_1,\vec{r}_i,\dots,\vec{r}_j,\dots,\vec{r}_N) = \Psi_n(\vec{r}_1,\vec{r}_j,\dots,\vec{r}_i,\dots,\vec{r}_N) \tag{6.11}$$

Clearly double application of P results in the original order.

$$\hat{P}_{ij}^2\Psi_n(\vec{r}_1,\vec{r}_i,\dots,\vec{r}_j,\dots,\vec{r}_N) = \Psi_n(\vec{r}_1,\vec{r}_i,\dots,\vec{r}_j,\dots,\vec{r}_N) \tag{6.12}$$

But how is this interesting? Assuming that Ψ_n is an eigenfunction of $\hat{P}$, with the eigenvalue of P^2 being 1, in principle the sign in front of Ψ after application of P can be either plus or minus one, i.e.

$$\hat{P}_{ij}\Psi_n(\vec{r}_1,\vec{r}_i,\dots,\vec{r}_j,\dots,\vec{r}_N) = \pm\Psi_n(\vec{r}_1,\vec{r}_i,\dots,\vec{r}_j,\dots,\vec{r}_N). \tag{6.13}$$

The sign tells what kind of particles one is dealing with.

Sign	Spin	Particles
$+$	integer	bosons (photons, phonons,He^4 atoms, Li, Na, ...)
$-$	half integer	fermions(electrons, He^3 atoms, ...)

Table 6.1: Particle overview, sorted by spin and sign of Ψ after application of the permutation operator

Consequences for fermions

A product wave function can be expressed as

$$\Psi_{(\alpha_1,\alpha_2,\dots)}(\vec{r}_1,\dots,\vec{r}_N) = \phi_{(\alpha_1)}(\vec{r}_1)\phi_{(\alpha_2)}(\vec{r}_2)\dots\phi_{(\alpha_N)}(\vec{r}_N) \tag{6.14}$$

Where α are quantum numbers characterising the N particle state. Assuming the fermionic case, $\hat{P}$ can be applied.

$$\begin{aligned} &\hat{P}_{ij}\phi_{(\alpha_1)}(\vec{r}_1)\dots\phi_{(\alpha_i)}(\vec{r}_i)\dots\phi_{(\alpha_j)}(\vec{r}_j)\dots\phi_{(\alpha_N)}(\vec{r}_N) && (6.15)\\ &= -\phi_{(\alpha_1)}(\vec{r}_1)\dots\phi_{(\alpha_i)}(\vec{r}_j)\dots\phi_{(\alpha_j)}(\vec{r}_i)\dots\phi_{(\alpha_N)}(\vec{r}_N) && (6.16)\end{aligned}$$

If the particles i and j are in the same quantum state $\alpha_i = \alpha_j = \alpha$ then

$$\begin{aligned} \hat{P}_{ij}\phi_{\alpha_i}(\vec{r}_i)\phi_{\alpha_j}(\vec{r}_j) &\overset{!}{=} -\phi_{\alpha_i}(\vec{r}_j)\phi_{\alpha_j}(\vec{r}_i) && (6.17)\\ &= \phi_{\alpha_j}(\vec{r}_j)\phi_{\alpha_i}(\vec{r}_i) && (6.18)\end{aligned}$$

Which would result in $1 = -1$ which means that the assumption was wrong. So what this shows is that all particles, that have a negative sign when the permutation operator is applied, can never be in the same state. This is called Pauli's exclusion principle, which states that all fermions must be in different states.

In the occupation number picture n_α is the occupation number of state α with the normalisation

$$N = \sum_{\alpha=1}^{\infty} n_\alpha \tag{6.19}$$

For fermions there can be either one or no particle in any state, $n_\alpha = 0, 1$. For bosons however there can be any number of particles $n_\alpha = 0, 1, 2, 3, \dots$ in any state at the same time.

6.2 Boson and Fermion Statistics: General Results

A quantum system of N particles is characterised by the occupation numbers n_i of state i and single particle energy levels ϵ_i. The total number of particles N is

$$N = \sum_{i=1}^{\infty} n_i. \tag{6.20}$$

And the Hamiltonian $\mathbb{H}$ is given by

$$\mathbb{H}(\{n_i\}) = \sum_{i=1}^{\infty} n_i \epsilon_i. \tag{6.21}$$

The partition function Z_N is the sum over all distributions n_i:

$$Z_N = \sum_{\{n_i\}} e^{-\beta \mathbb{H}(\{n_i\})}. \tag{6.22}$$

And the grand canonical partition function $\mathbb{Z}(z)$ becomes

$$\begin{aligned}
\mathbb{Z}(z) &= \sum_{N=0}^{\infty} z^N \sum_{\{n_i\}} e^{-\beta \sum_i \epsilon_i n_i} && (6.23)\\
&= \sum_{\{n_i\}} z^{\sum_i n_i} e^{-\beta \sum_i n_i \epsilon_i} && (6.24)\\
&= \sum_{n_1} \sum_{n_2} \ldots z^{\sum_i n_i} e^{-\beta \sum_i n_i \epsilon_i} && (6.25)\\
&= \sum_{n_1} z^{n_1} e^{-\beta \epsilon_1 n_1} \sum_{n_2} z^{n_2} e^{-\beta \epsilon_2 n_2} \ldots && (6.26)
\end{aligned}$$

so this factorises completely. For fermions this becomes

$$\begin{aligned}
&= \sum_{n_1=0}^{1} z^{n_1} e^{-\beta \epsilon_1 n_1} \sum_{n_2=0}^{1} z^{n_2} e^{-\beta \epsilon_2 n_2} \ldots && (6.27)\\
&= (1 + z e^{-\beta \epsilon_1})(1 + z e^{-\beta \epsilon_2}) \ldots && (6.28)
\end{aligned}$$

And the grand canonical partition function for fermions becomes

$$\mathbb{Z}(z) = \prod_{i=1}^{\infty} (1 + z e^{-\beta \epsilon_i}) \tag{6.29}$$

For bosons the grand canonical partition function is

$$\begin{aligned}
\mathbb{Z}(z) &= \sum_{n_1=0}^{\infty} z^{n_1} e^{-\beta \epsilon_1 n_1} \sum_{n_2=0}^{\infty} z^{n_2} e^{-\beta \epsilon_2 n_2} \ldots && (6.30)\\
\sum_{n_1=0}^{\infty} z^{n_1} e^{-\beta \epsilon_1 n_1} &= 1 + z e^{-\beta \epsilon_1} + (z e^{-\beta \epsilon_1})^2 + (z e^{-\beta \epsilon_1})^3 + \ldots && (6.31)\\
&= \frac{1}{1 - z e^{-\beta \epsilon_1}} && (6.32)\\
\mathbb{Z}(z) &= \frac{1}{1 - z e^{-\beta \epsilon_1}} \cdot \frac{1}{1 - z e^{-\beta \epsilon_1}} \ldots \cdot && (6.33)
\end{aligned}$$

And the grand canonical partition function for bosons becomes

$$\mathbb{Z}(z) = \prod_{i=1}^{\infty} \frac{1}{1 - ze^{-\beta\epsilon_i}} \tag{6.34}$$

There is a problem, when $ze^{-\beta\epsilon_i} \to 1$ as $\mathbb{Z}$ diverges. This results in a phase transition, the Bose-Einstein condensation. So the grand canonical partition functions $\mathbb{Z}(z, T, V)$ are calculated exactly for bosons and fermions. And from these all thermodynamic relations can be calculated. The mean occupation of energy level m is

$$\langle n_m \rangle = \frac{\sum_{\{n_j\}} n_m z^{\sum_k n_k} e^{-\sum_j \epsilon_j n_j}}{\sum_{n_j} z^{\sum_k n_k} e^{-\beta \sum_j \epsilon_j n_j}} \tag{6.35}$$

$$= -\frac{1}{\beta} \frac{\partial}{\partial \epsilon_m} \ln\left(\sum_{n_j} z^{\sum_k n_k} e^{-\beta \sum_j \epsilon_j n_j} \right), \tag{6.36}$$

$$\langle n_m \rangle = -\frac{1}{\beta} \frac{\partial \ln(\mathbb{Z})}{\partial \epsilon_m}, \tag{6.37}$$

$$\ln(\mathbb{Z}) = \pm \sum_{i=1}^{\infty} \ln\Big(1 \pm ze^{-\beta\epsilon_j}\Big) \tag{6.38}$$

Here the equation with the plus sign is true for fermions and the one with the minus sign corresponds to the bosons.

$$\langle n_m \rangle = \frac{ze^{-\beta\epsilon_m}}{1 \pm ze^{-\beta\epsilon_m}} \tag{6.39}$$

$$= \frac{1}{e^{\beta\epsilon_m} z^{-1} \pm 1} \tag{6.40}$$

This is the distribution function for fermions and bosons. The total particle number is

$$N = \sum_m \langle n_m \rangle \tag{6.41}$$

$$= \sum_m \frac{1}{e^{+\beta\epsilon_m} z^{-1} \pm 1}. \tag{6.42}$$

And the internal energy U is

$$U = \sum_m \langle n_m \rangle \epsilon_m \tag{6.43}$$

$$= \sum_m \frac{\epsilon_m}{e^{+\beta\epsilon_m} z^{-1} \pm 1}. \tag{6.44}$$

$$\tag{6.45}$$

6.3 Free Ideal Fermions

How to correctly count quantum states? A normalised plane wave can be written as

$$\phi_p(\vec{r}) = \frac{1}{\sqrt{V}} e^{\frac{i\vec{p}\vec{r}}{\hbar}} \tag{6.46}$$

with

$$\int_V d^3r |\phi_p(\vec{r})|^2 = 1. \tag{6.47}$$

$$\tag{6.48}$$

So the kinetic energy operator operates on the plane wave such that

$$-\frac{\hbar^2}{2m}\vec{\nabla}^2\phi_p(\vec{r}) = \frac{\vec{p}^2}{2m}\phi_p(\vec{r}) \tag{6.49}$$

So $\phi_p(\vec{r})$ are eigenfunctions to the kinetic energy (and many more) operator(s). With periodic boundary conditions in a box of length L with Volume $V = L^3$ the wave function is invariant under translation.

$$\phi_p(\vec{r} + L\hat{e}_x) = \phi_p(\vec{r}) \tag{6.50}$$

This means that

$$\vec{p} = \frac{2\pi\hbar}{L}\vec{n} \tag{6.51}$$

$$\vec{n} = \begin{pmatrix} 0, 1, 2, \dots \\ 0, 1, 2, \dots \\ 0, 1, 2, \dots \end{pmatrix} \tag{6.52}$$

So the momentum quantum numbers form a cubic lattice with lattice constant

$$\frac{2\pi\hbar}{L} = \frac{h}{L}. \tag{6.53}$$

In the continuum limit $L \to \infty$ this becomes.

$$\int d^3p \to (\frac{h}{L})^3 \sum_{\vec{p}} \tag{6.54}$$

$$\sum_{\vec{p}} = \frac{V}{h^3}\int d\vec{p}. \tag{6.55}$$

So for Fermions such as electrons $\langle n_m \rangle$ is

$$\langle n_m \rangle = \frac{1}{e^{\beta\epsilon_m}z^{-1} + 1} \tag{6.56}$$

with $z \equiv e^{\beta\epsilon_F}$ $n(\epsilon)$ is

$$n(\epsilon) = \frac{1}{1 + e^{\beta(\epsilon - \epsilon_F)}} \tag{6.57}$$

So the Fermi energy is the chemical potential.

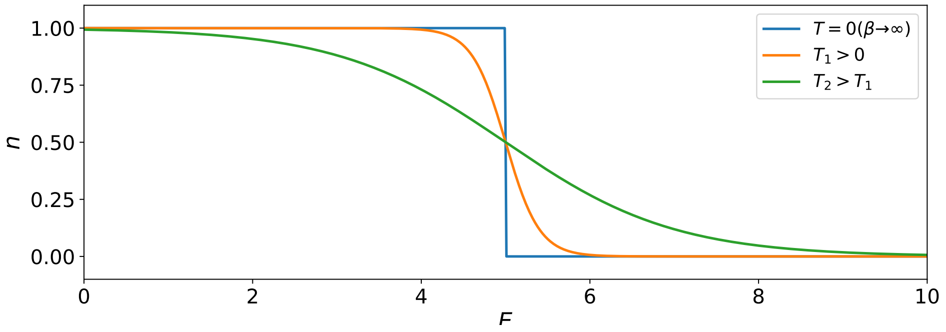


Figure 6.1: Fermi distribution for three different temperatures

The Fermi energy ϵ_F is determined by the total number of particles N. At $T = 0$ N is

$$N = g \sum_{p_i < p_F} 1, \tag{6.58}$$

where g is the multiplicity of a state; further

$$\begin{aligned} N &= \frac{gV}{h^3} \int_{|p|<p_F} d^3p 1 && (6.59) \\ &= \frac{gV}{h^3} \frac{4\pi}{3} p_F^3. && (6.60) \end{aligned}$$

Using $\epsilon_F = \frac{p_F^2}{2m}$ and the multiplicity which is $g = 2$ for spin $s = \frac{1}{2}$ particles, p_F^3 is

$$\begin{aligned} p_F^3 &= \frac{N}{V} \frac{h^3}{g} \frac{3}{4\pi} && (6.61) \\ &= (2m\epsilon_F)^{\frac{3}{2}}. && (6.62) \end{aligned}$$

and the Fermi energy is

$$\epsilon_F = \frac{h^2}{2m} \left(\frac{N}{V} \frac{3}{4\pi g} \right)^{\frac{2}{3}}. \tag{6.63}$$

With the electron mass $m_e = 9 \cdot 10^{-31} kg$, Planck's constant $h = 7 \cdot 10^{-34} Js$ and $n = \frac{N}{V} = \frac{1}{(0.3 \cdot 10^{-9} m)^3}$ the Fermi energy at room temperature is

$$\begin{aligned} \epsilon_F &\approx 2 \cdot 10^{-19} J && (6.64) \\ &= 50 k_B T. && (6.65) \end{aligned}$$

So typical electrons in metals are in fact very hot.

6.4 Free Ideal Bosons and Black Body Radiation

From $E = mc^2$ follows

$$\begin{aligned} E &= mc^2 && (6.66) \\ &= \frac{m_0 c^2}{(1 - \frac{v^2}{c^2})^{\frac{1}{2}}}, && (6.67) \\ E^2 \left(1 - \frac{v^2}{c^2} \right) &= m_0^2 c^4 && (6.68) \\ E^2 &= m_0^2 c^4 + E^2 \frac{v^2}{c^2} && (6.69) \\ &= m_0^2 c^4 + m^2 v^2 c^2, && (6.70) \\ E &= \sqrt{m_0^2 c^4 + p^2 c^2} && (6.71) \end{aligned}$$

So in the non-relativistic limit this becomes

$$\begin{aligned} E &= m_0 c^2 \sqrt{1 + \frac{p^2}{m_0^2 c^2}} && (6.72) \\ &\approx m_0 c^2 + \frac{p^2}{2m_0}, && (6.73) \end{aligned}$$

where the first term is a constant and the second term is the normal energy-momentum relation. In the ultra-relativistic limit this becomes

$$E = pc\sqrt{1+\frac{m_0^2c^2}{p^2}} \tag{6.74}$$

$$\approx pc, \tag{6.75}$$

which is the correct limit for photons, which do not have a resting mass. From the ultra-relativistic dispersion relations the internal energy becomes

$$U = g\sum_p \epsilon_p\langle n_p\rangle \tag{6.76}$$

$$= g\sum_p \frac{pcze^{-\beta pc}}{1-ze^{-\beta pc}}. \tag{6.77}$$

Here g is the multiplicity which is $g=2$ for photons and U becomes

$$U = 2\frac{V}{h^3}\int_{-\infty}^{\infty} d^3p\frac{|p|ce^{-\beta|p|c}}{1-ze^{-\beta|p|c}} \tag{6.78}$$

Here the fugacity $z=1$ as $\mu=0$ and the energy is

$$U = \frac{2V4\pi c}{h^3}\int_0^{\infty} dpp^3\frac{e^{-\beta pc}}{1-e^{-\beta pc}}. \tag{6.79}$$

With $p=\frac{\hbar}{c}\omega$ this becomes

$$\frac{U}{V} = \frac{8\pi\hbar^4}{c^3h^3}\int_0^{\infty} d\omega\frac{\omega^3}{e^{\beta\hbar\omega}-1} \tag{6.80}$$

$$\equiv \int_0^{\infty} d\omega u(\omega) \tag{6.81}$$

$$u(\omega) = \frac{\hbar}{\pi^2c^3}\frac{\omega^3}{e^{\beta\hbar\omega}-1}. \tag{6.82}$$

Here $u(\omega)$ is the spectral energy density. For small ω the denominator $e^{\beta\hbar\omega}-1$ becomes $\beta\hbar\omega$ and from that the famous Rayleigh-Jeans law

$$u(\omega) \approx \frac{k_BT\omega^2}{c^3\pi^2} \tag{6.83}$$

is obtained. For large ω this becomes Wien's law

$$u(\omega) \approx \frac{\hbar}{\pi^2c^3}\omega^3e^{-\beta\hbar\omega}. \tag{6.84}$$

The maximum of $u(\omega)$ is located at

$$\omega_{\text{max}} = 2.8\frac{k_BT}{\hbar} \tag{6.85}$$

and shifts to higher frequency with rising temperature. This can be seen in figure 6.2.

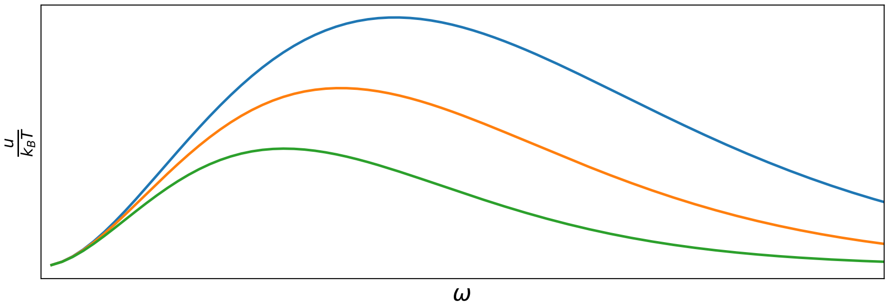


Figure 6.2: Wien's law for three different temperatures

7 Appendix

7.1 About this Script

This lecture was given by Prof. Roland Netz in the summer of 2016. This document started off as private notes that were turned into a full script.

Most of the script was proof read multiple times, but if you find any errors, please mark them on a printed or PDF version of the script (or in a plain list with page numbers) and send them to me. I will try to correct them directly. Most importantly we want to correct the contextual and mathematical errors, but if you find layout errors, wrong labels on figures, simple typos or wrong commata or anything else, please also send them to me. I hope this script helps you with your studies :)

Martin Borchert (`martin.b@robothek.de`)

Update (June 2019): This version includes fully revised chapters 1-3.

7.1.1 Distribution

For the rights about distribution to people outside the physics department or for other uses, please contact Professor Netz.

7.1.2 Graphs and Figures

All graphs and figures were done in latex via the Tikz package or in python 3 with Matplotlib (of course via Jupyter Notebook!) or with Inkscape. Note: svg2tikz is an awesome extension for Inkscape (awesome vector graphics drawing tool) to create Tikz much more easily!