

Introduction to Statistical Physics

Johar M. Ashfaque

1. Introduction

All matter we see around us consists of vast number of microscopic particles that are in constant motion, and in interaction with one another. At any given instant of time, it is possible in principle to find the state of this kind of *many-body system*. The laws of physics are such that the future evolution of the state can then be solved from the governing equations of motion. However in practise, the number of the constituent particles is typically of the order of the Avogadro's number

$$N_A = 6 \cdot 10^{23},$$

leading to stupendously large set of coupled equations. Even the storing of the initial state of these equations would exceed in multitude the combined memory of all of the computers in the world. This is the so-called *many-body problem* of physics.

When the number of the constituent particles is immense, it has turned out useful to use the tools of probability calculus to describe their emergent behaviour. In this course, we will not be interested in the details of individual particles, but in the relations between the macroscopic properties that result. In fact, we lack most of the information required to describe the internal state of the system, and therefore our approach has to be probabilistic in nature. This is the starting point of *statistical physics*.

The equations governing the macroscopic properties (e.g. energy, pressure, temperature) of a many-body system are statistical averages of those of the constituent particles. Therefore, it is implicitly assumed that the averaged quantities experience statistical *deviations*.

For example, the familiar equation of state of the ideal gas

$$PV = nRT$$

relates the average pressure P and average volume V to the average temperature T . Quite generally, the deviations from the average in a many-body system with N particles go as $1/\sqrt{N}$. The astounding number of particles, which makes the calculations of the macroscopic properties from first principles practically impossible, leads to such an accuracy of the statistical results that they can be taken as exact physical laws.

Motivation

The systematic studies of macroscopic systems started from the phenomenological studies in the 18th and 19th centuries. These were done by the likes of James Watt, William Rankine, Rudolf Clausius, Sadi Carnot, and William Thomson. The resulted theory of thermodynamics is concerned on heat and its relation to other forms of energy and work. The strength of thermodynamics is in its generality, emerging from the fact that the theory does not depend on any detailed assumptions on the microscopic properties of the system. This is also the main weakness of thermodynamics, since only relatively few statements can be made on such general grounds. This restricts the

scope of the theory and leaves many interesting properties of nature unexplained.

The idea that the macroscopic properties of matter are caused by the microscopic constituents was first introduced in the second half of the 19th century by James Clerk Maxwell, Ludwig Boltzmann, Max Planck, Rudolf Clausius, and J. Willard Gibbs. The theory was completed in the first half of 20th century by the advent of quantum mechanics that gives the correct description of atomic size particles. By applying the probabilistic ideas to systems in equilibrium, one can obtain all results of the thermodynamics together with the ability to relate the macroscopic parameters of the system to its microscopic constituents. This powerful method will be subject in the first half of this course. The description of systems not in equilibrium is much more difficult task. One can still make some general statements of such systems which leads to statistical physics of irreversible processes. This will be discussed briefly in the latter part of the course.

Statistical physics can, in principle, be applied to any state of matter: solids, liquids, gases, matter composed of several phases/or several components, matter under extreme conditions of temperature and pressure, matter in equilibrium with radiation, and so on. In addition, the concepts and methods of statistical physics have turned out useful in many other fields of science: chemistry, study of dynamical systems, communications, bioinformatics, complexity, and even stock markets.

"If a million monkeys typed ten hours a day, it is extremely unlikely that their output would exactly equal all the books of the richest libraries of the world; and yet, in comparison, it is even more unlikely that the laws of statistical mechanics would ever be violated, even briefly."

- Emile Borel, 1913

2. Background

2.1 Ensemble and Probability

The laws of Physics are deterministic in a sense that given the state of a system at one time one can calculate the state at all later times, by using the time-dependent Schrödinger equation or classical mechanics. Then, the outcome of an experiment on the system should be determined completely by this final state. Even though this can be done in principle, it is practically impossible to determine an initial state of a system with N_A particles, let alone to calculate its time evolution.

Statistical physics circumvents this problem by considering a (infinitely) large set of similar experiments, instead of a single one. This kind of set of mental copies of similar systems is called an *ensemble*¹.

Let then S denote a general system and consider an experiment on the system with a general outcome X . We denote with Σ the ensemble of mental copies of S . Now, the *probability* of observing the outcome X is given by

$$P(X) = \lim_{\Omega(\Sigma) \rightarrow \infty} \frac{\Omega(X)}{\Omega(\Sigma)}, \quad (1)$$

where $\Omega(\Sigma)$ is the number of systems in the ensemble and $\Omega(X)$ is the number of systems exhibiting the outcome X . This is the scientific definition of probability.

Example: Throwing a dice

One could in principle determine the initial position and orientation of a dice. Also, one could figure out exactly how the dice is thrown out of hand. By using these as initial conditions, one could use the classical equations of motion to determine exactly which of the six numbers turns up. Practically, this is impossible.

Instead, we usually consider a large group of (nearly) similarly thrown dices. The result of a single throw is not anymore deterministic, since we do not have the exact knowledge of the initial state. Nevertheless, we can find out the probability of a particular outcome if we can assume that each outcome is equally likely². In such a case, we obtain

$$P(X) = \frac{1}{6}.$$

This can, of course, be compared with the actual dice throws. Indeed, if we notice a deviation from this assumption of equally probable outcomes, we immediately suspect that the dice is crooked!

2.2 Random Walk

Next, we will consider *one-dimensional random walk* as a simple example, grasping the relevant concepts in probability calculus needed in this course. Assume, that a particle

performs successive steps, each of which has a length l and is taken on a random direction. The successive steps are statistically independent, so we can denote

p = probability of a right step

$q = 1 - p$ = probability of a left step

One obtains a great insight by considering the number p as the fraction of right steps in the ensemble of mental copies of single steps.

The random walk of N individual steps is formed by taking a subset, i.e. a *sample*, from the ensemble and calculating the sum of its members. After N steps, the particle lies at

$$x = ml,$$

where

$$m = n_r - n_l \quad (2)$$

and n_r (n_l) denotes the number of steps to right (left). Naturally,

$$-N \leq m \leq N$$

and

$$N = n_r + n_l. \quad (3)$$

By taking into account the number of different ways taking n_r steps out of N , we arrive at the probability

$$W_N(n_r) = \frac{N!}{n_r!n_l!} p^{n_r} q^{n_l} \quad (4)$$

of taking n_r steps to the right. Probability function (4) is referred to as the *binomial distribution* since it resembles the terms in the binomial expansion.

By using Eqs. (2) and (3), one can show (exercise) that the probability of finding a particle at position m after N steps is

$$\begin{aligned} P_N(m) &= W_N(n_r) \\ &= \frac{N!}{[(N+m)/2]![(N-m)/2]!} p^{(N+m)/2} q^{(N-m)/2}. \end{aligned} \quad (5)$$

Moments of a discrete distribution

Before continuing with the random walk example, let us recall a couple of concepts that characterize a given discrete distribution (later, we will generalize these for continuous distributions).

The *mean* of an arbitrary (normalized) distribution $P(u)$ is given by

$$\mu \equiv \langle u \rangle = \sum_{i=1}^M u_i P(u_i) = \lim_{\Omega(\Sigma) \rightarrow \infty} \sum_{i=1}^M \frac{u_i \Omega(u_i)}{\Omega(\Sigma)}, \quad (6)$$

where u_i are the M possible values the variable u can have. The latter equality allows the interpretation of the mean as the *ensemble average*. The mean has the familiar interpretation due to the *law of large numbers*:

¹Ensemble is the French word for group. We will discuss in a later chapter the reasons why this kind of treatment can be done.

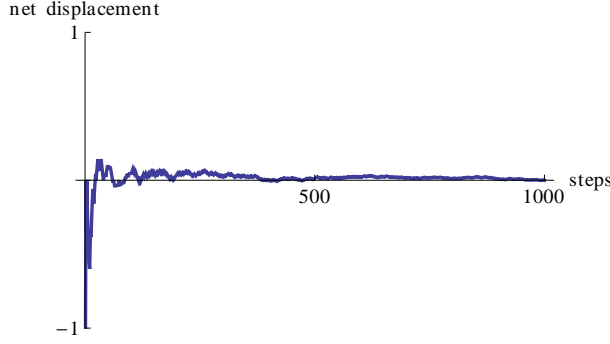
²This is in fact a postulate of so-called *a priori* probabilities that lies in the foundation of statistical physics, and will be discussed more thoroughly in the following chapter!

IF ONE REPEATS ENDLESSLY THE SAME RANDOM EXPERIMENT, THE AVERAGE OF THE RESULTS APPROACHES THE MEAN OF THE ENSEMBLE FORMED BY MENTAL COPIES OF THE SINGLE EXPERIMENT.

This law is, of course, a straightforward consequence of our definition of probability.

Let us consider a series of random steps ($l = 1$) and, after each step, calculate the average displacement of a single random walk

$$\bar{x} = l \frac{n_r - n_l}{N} \rightarrow l(p - q) = \langle x \rangle.$$



We have assumed $p = 0.5$ and $l = 1$ in the figure. As one increases the number of steps, the size of the sample grows and the average of the steps approaches that of the ensemble.

Assume that $f(u)$ and $g(u)$ are any functions of the discrete variable u and c is a constant. Then,

$$\langle f(u) \rangle = \sum_{i=1}^M f(u_i) P(u_i)$$

$$\langle f(u) + g(u) \rangle = \langle f(u) \rangle + \langle g(u) \rangle$$

$$\langle cf(u) \rangle = c \langle f(u) \rangle.$$

The mean is called the *first moment* of the distribution. In general, the n^{th} *moment* is defined as

$$\langle u^k \rangle = \sum_{i=1}^M u_i^k P(u_i). \quad (7)$$

In principle, the whole distribution can be resolved by calculating all of its moments.

A useful mean value is called the *variance*, or the *second moment about the mean*

$$\sigma^2 \equiv \langle (u - \langle u \rangle)^2 \rangle = \sum_{i=1}^M (u_i - \langle u \rangle)^2 P(u_i). \quad (8)$$

The variance has a great importance since its square root σ , the *standard deviation*, can be used to characterize the width of the range over which the random variable is distributed around its mean. The following relation is often useful

$$\sigma^2 = \langle u^2 \rangle - \langle u \rangle^2.$$

Mean values of the random walk

First, let us calculate the mean and variance of the whole ensemble

$$\mu = p - q, \quad \sigma^2 = p(1 - (p - q))^2 + q(-1 - (p - q))^2 = 4pq.$$

The random walk with length N is formed by taking a sample from the ensemble. It can be characterized either with distribution (4) or (5). According to Eq. (4), we can show that the binomial distribution is normalized, i.e.

$$\sum_{n_r=0}^N W_N(n_r) = 1.$$

Thus, we can calculate the mean number of steps to the right as

$$\langle n_r \rangle = Np,$$

where we have used the result

$$n_r p^{n_r} = p \frac{\partial}{\partial p} (p^{n_r}).$$

The mean displacement is then given by

$$\langle m \rangle = N(p - q) = N\mu,$$

which is (naturally) zero when $p = q$. We see that the average displacement is formed by N independent single displacements with the mean $\mu = p - q$.

Correspondingly, the variance of the right steps in the random walk is

$$\sigma_r^2 = \langle (n_r - \langle n_r \rangle)^2 \rangle = Npq.$$

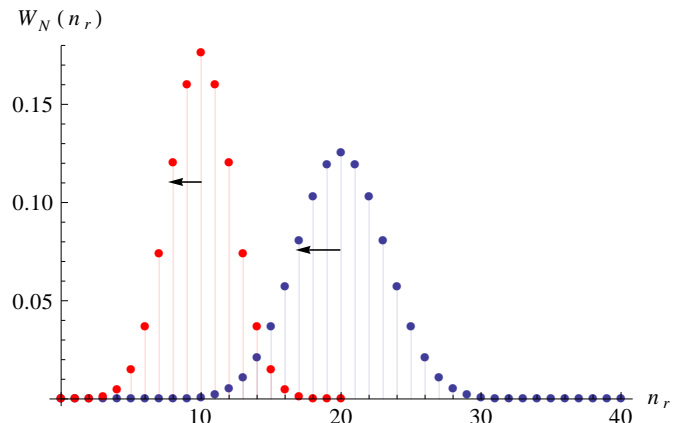
Now, the variance of the net displacement is

$$\sigma_m^2 = \langle (m - \langle m \rangle)^2 \rangle = 4\sigma_r^2 = N\sigma^2,$$

where we have used the relations (2) and (3). For the number of right steps, we obtain the standard deviation

$$\frac{\sigma_r}{\langle n_r \rangle} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}}. \quad (9)$$

We see that the relative width decreases as $1/\sqrt{N}$ with increasing sample size N .³



³This kind of $1/\sqrt{N}$ -law is common in statistics, and is the reason why statistical mechanics gives (nearly) exact results.

Binomial distributions of the number of right steps with $N = 20$ (red) and $N = 40$ (blue). The arrows indicate the standard deviations.

Large N limit

Let us then assume that N is large, meaning that we are in the limit of a long random walk, or, of a large sample. Due to simpler calculations, we do the calculation for the number of right steps n_r , but the same reasoning applies also for the net displacement m . First, we note that since

$$\langle n_r \rangle = Np \quad \text{and} \quad \sigma_r = \sqrt{Npq} \gg 1.$$

we have that near the mean $\langle n_r \rangle$ and in the limit of large N

$$|W_N(n_r + 1) - W_N(n_r)| \ll W_N(n_r).$$

This allows us to consider W_N as a continuous function of continuous n_r .

Now, the location of the maximum $n_r = \tilde{n}_r$ can be found approximately from the equation

$$\frac{d \ln W_N}{dn_r} = 0.$$

We make a Taylor expansion of $\ln W_N$ in the vicinity of \tilde{n}_r , and obtain⁴

$$\ln W_N(n_r) = \ln W_N(\tilde{n}_r) + B_1\eta + \frac{1}{2}B_2\eta^2 + \frac{1}{6}B_3\eta^3 + \dots, \quad (10)$$

where $\eta = n_r - \tilde{n}_r$ and

$$B_k = \left. \frac{d^k \ln W_N}{dn_r^k} \right|_{n_r=\tilde{n}_r}.$$

By using Stirling's formula (exercise)

$$\ln n! \approx n \ln n - n,$$

we obtain

$$B_1 = -\ln n_r + \ln(N - n_r) + \ln p - \ln q.$$

At the maximum ($n_r = \tilde{n}_r$), $B_1 = 0$ which implies

$$\tilde{n}_r = Np = \langle n \rangle_r.$$

Similarly,

$$B_2 = -\frac{1}{Npq} = -\frac{1}{\sigma_r^2} < 0$$

as it should for a maximum. The higher order derivatives $B_k \propto N^{-(k-1)}$, and can be neglected in the limit of large N in the expansion (10).

We have thus obtained

$$W_N(n_r) = \tilde{W}_N e^{-\frac{1}{2}|B_2|\eta^2},$$

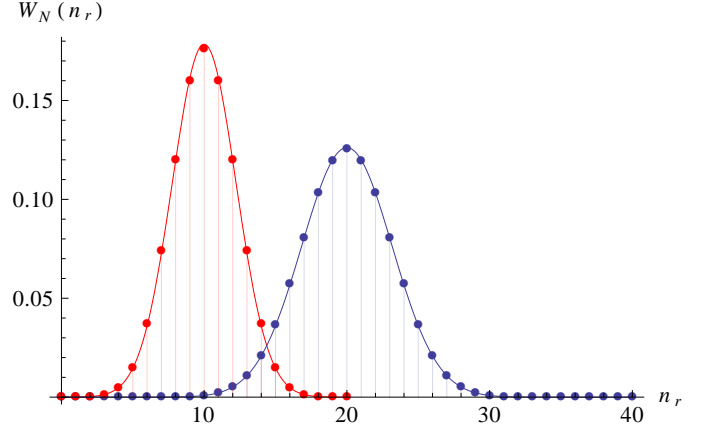
⁴We use $\ln W_N$ instead of W_N since its expansion converges more rapidly.

where $\tilde{W}_N = W_N(\langle n \rangle_r)$ can be determined from the normalization condition⁵

$$\sum_{n_r=0}^N W_N(n_r) \approx \int_{-\infty}^{\infty} W_N(\langle n \rangle_r + \eta) d\eta = 1 \Rightarrow \tilde{W}_N = \sqrt{\frac{|B_2|}{2\pi}}.$$

Thus, we have obtained that, in the limit of large N and n_r , the binomial distribution converges to the Gaussian distribution

$$W_N(n_r) = \frac{1}{\sqrt{2\pi\sigma_r^2}} e^{-\frac{(n_r - \langle n \rangle_r)^2}{2\sigma_r^2}}. \quad (11)$$



Binomial distribution (dots) with $p = 0.5$, $N = 20$ and $N = 40$, together with the corresponding Gaussian distributions (solid lines).

2.3 Central Limit Theorem

The preceding discussion is one special case of the so-called *central limit theorem*⁶:

IF WE SUM N ($N \rightarrow \infty$) SYSTEMS THAT ARE STATISTICALLY INDEPENDENT AND IDENTICAL, THE RESULTANT DISTRIBUTION IS UNDER VERY GENERAL CONDITIONS A GAUSSIAN.

In terms of the random walk, we started with a single step with a probability distribution

$$P(X) = \begin{cases} p & X = \text{right} \\ q & X = \text{left}. \end{cases}$$

Then, we combine a large number N of such random steps and calculate the net displacement to the right $m = n_r - n_l$, where n_r (n_l) is the number of steps to the right (left). In the limit of $N \rightarrow \infty$, we obtain

$$\begin{aligned} \frac{n_r}{N} &\rightarrow p \\ \frac{n_l}{N} &\rightarrow q \\ \frac{m}{N} &\rightarrow p - q = \mu. \end{aligned}$$

⁵We have seen that the distribution has a pronounced maximum, so the continuation of the integral from $-\infty$ to ∞ is an excellent approximation.

⁶Central limit theorem is notoriously difficult to prove, which is therefore omitted here.

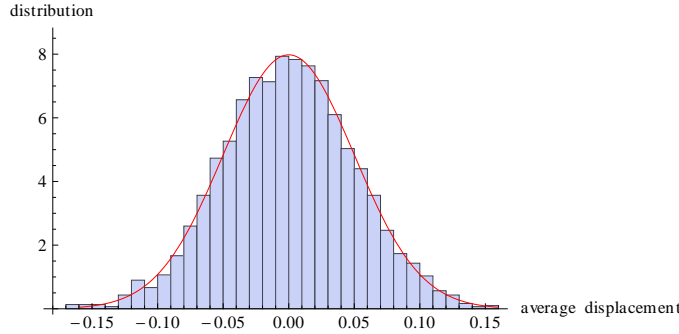
This is the law of large numbers. Additionally for a fixed N , the central limit theorem tells us that the measured values for individual random walks are spread around the mean. This spread is characterized by the relevant standard deviation⁷.

We showed that for a random walk, the number of right steps obeys a Gaussian distribution (11) when N is large. We can use the relations (2) and (3) to show that also other resultant random variables are distributed normally. For example, let us study the average displacement $\bar{m} = m/N$. Now,

$$m = 2n_r - N, \quad \langle \bar{m} \rangle = \frac{\langle m \rangle}{N} = \mu, \quad \sigma_{\bar{m}}^2 = \frac{\sigma^2}{N}.$$

Because the change of variables is linear, we have that

$$P_N(\bar{m}) = W_N(n_r) \left| \frac{dn_r}{d\bar{m}} \right| = \frac{1}{\sqrt{2\pi\sigma_{\bar{m}}^2}} e^{-\frac{(\bar{m} - \langle \bar{m} \rangle)^2}{2\sigma_{\bar{m}}^2}}.$$



In the figure, $N = 400$ and $p = 0.5$. We have calculated \bar{m} for 3000 random walks. The height of the bars denote the number of random walks with \bar{m} within the interval $m + \delta m$. The interval δm is the width of the bars. The height of the bars is normalized so, that the total area under them is 1. The red curve is the Gaussian distribution with

$$\langle \bar{m} \rangle = 0 \text{ and } \sigma_{\bar{m}} = \frac{1}{\sqrt{N}}.$$

2.4 Several Variables and Continuous Distributions

Several variables

In statistical physics, the typical problems deal with systems with 10^{24} , or so, particles. Therefore, it is evident that we have to deal with probability distributions with more than one variable. We consider here the case of two variables, but the results are straightforwardly generalizable to cases with any number of variables.

Assume that we have two random variables u and v with possible values u_i and v_j , where $i = 1, \dots, N$ and $j = 1, \dots, M$. We denote with

$$P(u_i, v_j)$$

the probability that u has the value u_i and v the value v_j . We require the normalization

$$\sum_{i=1}^N \sum_{j=1}^M P(u_i, v_j) = 1.$$

Naturally,

$$P(u_i) = \sum_{j=1}^M P(u_i, v_j) \text{ and } P(v_j) = \sum_{i=1}^N P(u_i, v_j).$$

When the outcome of one variable is independent of the outcome of the other, we say that the two variables are *statistically independent*, or *uncorrelated*. In this case,

$$P(u_i, v_j) = P(u_i)P(v_j).$$

Let $F(u, v)$, $G(u, v)$, $f(u)$ and $g(v)$ be arbitrary functions. Now,

$$\langle F \rangle = \sum_{i=1}^N \sum_{j=1}^M F(u_i, v_j) P(u_i, v_j)$$

$$\langle F + G \rangle = \langle F \rangle + \langle G \rangle$$

$$\langle f \rangle = \sum_{i=1}^N \sum_{j=1}^M f(u_i) P(u_i, v_j) = \sum_{i=1}^N f(u_i) P(u_i)$$

$$\langle fg \rangle = \langle f \rangle \langle g \rangle,$$

where the last equality holds only for uncorrelated u and v . The proofs are left as an exercise.

Continuous distributions

We have already had one example in the previous sections on the probability distribution of a continuous random variable in the form of the Gaussian. Let us consider such cases more generally.

Assume that u can obtain any value in the continuous range $a_1 < u < a_2$. The probability of finding the variable between u and $u + du$ is

$$P(u) = \mathcal{P}(u)du, \quad (12)$$

where \mathcal{P} is the *probability density*. The normalization condition and the definition of the expectation value are straightforwardly generalized from the discrete case. Thus,

$$\int_{a_1}^{a_2} \mathcal{P}(u)du = 1 \quad (13)$$

and

$$\langle f(u) \rangle = \int_{a_1}^{a_2} f(u) \mathcal{P}(u)du, \quad (14)$$

where f is an arbitrary function of u . These results are easily generalized to the cases where one has more than one continuous random variable (exercise).

⁷The mean is only the most probable outcome. For example, if you toss a coin four times, the mean is two heads and two tails. But, sometimes you might get all four tails!

3. Statistical Formulation

In this chapter, we will present the formal description of the statistical approach to the analysis of the internal motions of a many-body system. This approach holds for systems in *equilibrium*⁸. The non-equilibrium statistics will be studied in the latter part of the course.

Characterization of the state

We want to describe a system that is *macroscopic*, by which we mean that the number N of its constituent particles obeys

$$\frac{1}{\sqrt{N}} \ll 1. \quad (15)$$

Typically, $N \approx N_A \approx 10^{24}$. This means that the statistical arguments discussed in the previous chapter can be applied.

Macroscopic state can be characterized with the so-called state variables. *Extensive variables* are proportional to the size of the system, i.e. to the particle number N or to the volume V . For example, the total energy E of N non-interacting particles is

$$E = \sum_i n_i \varepsilon_i, \quad (16)$$

where n_i denotes the number of particles with energy ε_i . Naturally,

$$N = \sum_i n_i. \quad (17)$$

Intensive variables are independent on the size, and can be determined from any sufficiently large volume element of the system. Examples include temperature T , pressure p , chemical potential μ , ratios of extensive variables, such as density $\rho = N/V$.

The macroscopic state, or *macrostate*, is determined by the values of the extensive variables, e.g. N , V and E .

Statistical ensemble

Example: Coin flips

Assume that we flip three coins. There are altogether four different possible outcomes: 0,1,2 or 3 heads.

State index	Outcome	Number of heads
1	hhh	3
2	hht	2
3	hth	2
4	thh	2
5	htt	1
6	tht	1
7	tth	1
8	ttt	0

⁸We will define the equilibrium in the course of the discussion. At the moment, it is sufficient to consider it as the state of the system which does not show any temporal macroscopic changes.

Assume that we have tossed two heads. This is the "macrostate" of the three coin system. Nevertheless from the Table, we see that we can obtain two heads in three different ways. If we do not possess any additional knowledge on the system (e.g. labels on the coins), we can only say that the system is after the toss in one of the corresponding three "microstates" 2,3 or 4.

Assume then that the number of coins is of the order of N_A . One can easily imagine that there exists, in general, an extremely large number of possible states for a given value of heads.

We can generalize the above coin flip example to the system consisting of macroscopic number of microscopic particles. The extensive macroscopic parameters, that determine the macroscopic state, give us only partial information about the system. Nevertheless, they impose constraints to the possible states the system is allowed to occupy. The macroscopic state, i.e. *macrostate*, of the system is a function of a set of f coordinates, where f denotes the number of the *degrees of freedom* in the system (i.e. the position coordinates and the possible spins of the constituent particles) and is proportional to the number N of the constituent particles. Any given macrostate can be achieved by several different sets of values for the coordinates. Each such set is called a *microstate*.

Postulate of equal *a priori* probabilities

From the simple coin flip example, we can see that for a given macrostate the number of possible microstates can increase very rapidly as the number of random variables increases. What can we then say about the probability of finding the system in any of its allowed microstates? As an example, consider an isolated system. By definition, the energy is conserved which imposes a constraint on the possible microstates the system can occupy. Usually, there are however many such states.

General statements can be made when the system is in *equilibrium*. The equilibrium is characterized by the fact that the probability of finding the system in *any* of its accessible microstates is independent on time. There is nothing in the laws of physics that would make one assume that some of these microstates would be occupied more frequently than the others. So, in the case of any other prior knowledge, we assume that:

ISOLATED SYSTEM IN EQUILIBRIUM IS EQUALLY LIKELY TO BE FOUND IN ANY OF ITS ACCESSIBLE STATES.

In statistical physics, the above assumption is called the postulate of equal *a priori* probabilities. Therefore, it cannot be proven but one can only derive its consequences and compare them with measurements. This kind of measurements have been done in multitude, and the agreement has so far been very good⁹.

⁹We naturally assume this postulate when we play a game of chance, such as coin toss, dice toss, or cards. In the absence of any additional information, we always assume that the different outcomes occur with equal probabilities, as we discussed in the first dice toss example!

Approach to equilibrium

The postulate of a priori probabilities indicates that the equilibrium of the isolated system is characterized by a uniform distribution of its accessible microstates. The accessible states are those that do not violate the constraints set by the known values of systems macroscopic parameters y_1, y_2, \dots, y_n (such as E, V and N). Thus, the number of the accessible states can be written in functional form

$$\Omega = \Omega(y_1, y_2, \dots, y_n),$$

meaning that the parameter α lies in the range y_α and $y_\alpha + \delta y_\alpha$.

Assume that initially the isolated system is in equilibrium, meaning that it can be found equally likely in any of its Ω_i states. A removal of a constraints leads to an increase of accessible states

$$\Omega_f \geq \Omega_i.$$

When $\Omega_f > \Omega_i$, this means that *immediately* after the removal of the constraint the system is not distributed equally between its allowed microstates, it is not in equilibrium, and its macroscopic parameters change in time.

Consider a system in such a state of *non-equilibrium*. Let P_r be the probability of the system to be in the microstate r . The so-called *H-theorem*¹⁰ says that (exercise)

$$\frac{dH}{dt} \leq 0, \quad (18)$$

where $H = \langle \ln P_r \rangle$ and the equality holds when $P_r = P_s$ for all microstates r and s . This means that the non-equilibrium state changes so, that the quantity H decreases until the uniform equilibrium distribution of the microstates is reached.

The H-theorem guarantees that isolated systems approach the equilibrium, eventually. The time scale that characterizes the process is called the *relaxation time* τ . The relaxation time depends on the particular system, especially on the inter-particle interactions. The postulate of equal a priori probabilities holds for the isolated systems in equilibrium. In order for the postulate to hold, it is therefore important to wait a number of relaxation times after the decoupling of the system from the environment. By comparing with the experimental time scale t_{exp} we can separate three distinctive cases:

1. $\tau \ll t_{\text{exp}}$: System reaches equilibrium very quickly and equilibrium statistics are applicable.
2. $\tau \gg t_{\text{exp}}$: Equilibrium is reached very slowly. In fact, it is so slow in the experimental time scale that it can be considered to be in equilibrium (with additional constraints that prevent the reaching to the equilibrium). Hence, the statistical arguments apply.

¹⁰The theorem was first proposed and proven by L. Boltzmann in 1872.

3. $\tau \approx t_{\text{exp}}$: Distribution over accessible states is not uniform and changes in the experimental time scales. Problem cannot be reduced to a discussion of equilibrium situations.

According to the H-theorem, after the relaxation the quantity H is a static constant, and can be written for an isolated system as

$$H_{\text{eq}} = \sum_r P_r \ln P_r = -\ln \Omega_f.$$

Thus, we see that

IF SOME CONSTRAINTS OF AN ISOLATED SYSTEM ARE REMOVED THEN THE PARAMETERS OF THE SYSTEM TEND TO SPONTANEOUSLY READJUST THEMSELVES IN SUCH A WAY THAT

$$\Omega(y_1, \dots, y_n) \rightarrow \text{maximum}.$$

We call *process* the transition from the initial to the final macrostate. When $\Omega_f = \Omega_i$, the process is called *reversible*. The initial macrostate can be recovered by reimposing the constraint since the system is throughout the process in equilibrium. If $\Omega_f > \Omega_i$ the corresponding process is *irreversible*. The imposition or removal of additional constraint cannot *spontaneously* recover the initial macrostate because of the uniform distribution over Ω_f states.

We see that the quantity $\ln \Omega$ measures the degree of irreversibility of the macrostate of an isolated system in equilibrium. Especially for any physical process operating on an isolated system,

$$\ln \Omega_f - \ln \Omega_i \geq 0.$$

For historical reasons, this measure of irreversibility is written in the units joules/kelvin. Thus, it is multiplied with the Boltzmann constant ($k_B = 1.381 \cdot 10^{-23}$ J/K)

$$S = k_B \ln \Omega, \quad (19)$$

and termed the (equilibrium) *entropy*.

Implications of the statistical definition of entropy

The H-theorem gives straightforwardly the *second law of thermodynamics*¹¹:

THE ENTROPY OF AN ISOLATED SYSTEM TENDS TO INCREASE WITH TIME AND CAN NEVER DECREASE.

As one goes to lower energies, one has to turn into quantum mechanical treatment¹². This is because in classical physics there is no lower bound on the energy of the system. Accordingly, we can write the *third law of thermodynamics*:

¹¹In thermodynamics, the second law results after rather cumbersome considerations of Carnot cycles. This has been extensively studied in in the Thermodynamics course and will be omitted here.

¹²In order to have unique value for the entropy, one actually has to treat it quantum mechanically. This is because the volume elements in classical phase-space (discussed later) are not bounded by the uncertainty principle.

ENTROPY OF A SYSTEM APPROACHES TO A CONSTANT VALUE AS THE TEMPERATURE APPROACHES ZERO.

The constant is determined solely by the *degeneracy* of the quantum mechanical ground state of the system. In the case of non-degenerate ground state (one accessible microstate, $\Omega = 1$), the entropy is zero.

Probability calculations

Consider an isolated system in equilibrium whose (constant) energy is known to be somewhere between E and $E + \delta E$. We denote with $\Omega(E)$ the total number of states with energy in this range. Let us assume that we measure the property y , and let y_k be the possible outcomes of that measurement and $\Omega(E; y_k)$ the number of the state with such outcome. Based on our discussions in the previous chapter, we form an ensemble of the $\Omega(E)$ states. Then, our fundamental postulate of a priori probabilities tells us that each of the $\Omega(E)$ states in the ensemble is equally likely to occur. Thus, the probability of obtaining y_k from the measurement is

$$P(y_k) = \frac{\Omega(E; y_k)}{\Omega(E)}.$$

We assume in the above that δE is sufficiently large, so that we can take $\Omega(E) \rightarrow \infty$. This is generally true in macroscopic systems.

We see that the statistical physics calculations reduce simply to *counting states*, subject to relevant constraints (such as for E in the above discussion).

4. Statistical Thermodynamics

The previous chapter indicates that the statistical physics of a macrostate follows from the number $\Omega(E, V, N)$ of allowed microstates. In general, this number is a function of the macroscopic energy of the system, in addition to the *external variables* such as the particle number and volume. We leave the detailed computations of Ω to the next chapter, and concentrate, in this chapter, in its relation with thermodynamic quantities. It is worthwhile to remember that thermodynamics is expressed in terms of macroscopic parameters, not relying on any microscopic knowledge on the matter under study.

We have learned that the macroscopic quantities are statistical in nature. This means that in equilibrium their values are statistical averages of N_A , or so, particles. The largeness of N_A means that the distributions of the averages are Gaussians whose deviations are small compared with the most probable value, e.g.

$$\frac{\sigma_E}{\langle E \rangle} \sim \frac{1}{\sqrt{N_A}}.$$

Thus, we can replace macroscopic quantities, such as energy, temperature and pressure, by their mean values. This is the essence of classical thermodynamics.

4.1 Contact Between Statistics and Thermodynamics

Thermal interaction

Consider two systems A_1 and A_2 separately in equilibrium. The number of microstates in system A_1 in a macrostate with energy in range E_1 to $E_1 + \delta E_1$ is denoted with $\Omega_1(E_1)$. Correspondingly, we denote the number of microstates of the system A_2 with $\Omega_2(E_2)$. Assume that the systems are brought in a *thermal* contact, meaning that they can exchange energy, but that their particle numbers and volumes stay fixed. We thus assume that the energies E_1 and E_2 can vary, but the variations are restricted by the energy conservation condition

$$E^{(0)} = E_1 + E_2 = \text{constant}.$$

Each microstate in A_1 can be combined with a state in A_2 , so the total number of microstates in the composite system is

$$\Omega^{(0)} = \Omega_1(E_1)\Omega_2(E_2) = \Omega_1(E_1)\Omega_2(E^{(0)} - E_1).$$

Based on the discussion in the previous chapter, after we have brought the systems together and after waiting several relaxation times, the combined system can be assumed to have reached the equilibrium. We have shown with the H-theorem that the equilibrium is such that $\Omega^{(0)}$ is maximized.

Thus, (we use partial derivative since we assume purely thermal interaction in which the other parameters are kept

fixed)

$$\left(\frac{\partial \Omega_1(E_1)}{\partial E_1}\right)_{E_1=\langle E_1 \rangle} \Omega_2(\langle E_2 \rangle) + \Omega_1(\langle E_1 \rangle) \left(\frac{\partial \Omega_2(E_2)}{\partial E_2}\right)_{E_2=\langle E_2 \rangle} \frac{\partial E_2}{\partial E_1} = 0.$$

This implies ($\partial E_2 / \partial E_1 = -1$) that the equilibrium condition of the composite system can be written as

$$\left(\frac{\partial \ln \Omega_1(E_1)}{\partial E_1}\right)_{E_1=\langle E_1 \rangle} = \left(\frac{\partial \ln \Omega_2(E_2)}{\partial E_2}\right)_{E_2=\langle E_2 \rangle} \quad \text{or} \quad \beta_1 = \beta_2,$$

where

$$\beta = \left(\frac{\partial \ln \Omega(N, V, E)}{\partial E}\right)_{E=\langle E \rangle}. \quad (20)$$

The parameter β has the following properties:

1. If two systems in equilibrium with the same value of β are brought into thermal contact, they will remain in equilibrium.
2. If two systems in equilibrium with different values of β are brought into thermal contact, the system with higher value of β will absorb heat from the other until the two β values are the same (exercise).

In addition, we notice that β has the units of inverse energy. We define the *thermodynamic* or *absolute temperature* with

$$\frac{1}{T} = k_B \beta = \left(\frac{\partial S(E, V, N)}{\partial E}\right)_{V, N, E=\langle E \rangle}, \quad (21)$$

where the subscripts contain the assumption of constant volume and particle number. We thus see that the thermodynamic temperature of a macroscopic system depends only of the rate of changes of the number of accessible microstates with the total energy.

According to the previous discussion, it is easy to arrive at the *zeroth law of thermodynamics*:

IF TWO SYSTEMS ARE SEPARATELY IN THERMAL EQUILIBRIUM WITH A THIRD SYSTEM THEN THEY MUST ALSO BE IN THERMAL EQUILIBRIUM WITH ONE ANOTHER.

General interaction

In the most general case, the external parameters do not remain fixed and the system is not thermally insulated. Additionally, the (essentially quantum mechanical) energies E_r of the accessible microstates are dependent on the external parameters $E_r = E_r(x_1, \dots, x_n)$, where x_i are the relevant external macroscopic parameters. For simplicity, we consider here the case where only one external parameter x is allowed to vary. In addition to the constraint due to x , we assume that the energy of the system is found from the interval E to $E + \delta E$. When we change x by dx the energy E_r of a microstate r changes by $(\partial E_r / \partial x) dx$.

We want to calculate how the total number of accessible states $\Omega(E, x)$ changes as we increase x with dx . First, we define the *generalized force* conjugate to the parameter x as

$$X = -\frac{\partial E_r}{\partial x}. \quad (22)$$

For a fixed value of X the number of such states that are located within energy $-X dx$ below E , and change to a value greater than E is

$$\sigma_X(E) = -\frac{\Omega_X(E, x)}{\delta E} X dx,$$

where $\Omega_X(E, x)$ is the number of such microstates that have their energy E_r within the interval above, and the generalized force within X and $X + \delta X$.

Thus, we obtain the total number of states $\sigma(E)$ that change from an energy value lower than E to one that is larger

$$\sigma(E) = \sum_X \sigma_X(E) = -\frac{\Omega(E, x)}{\delta E} \bar{X} dx,$$

where

$$\bar{X} = \frac{1}{\Omega(E, x)} \sum_X \Omega_X(E, x) X$$

is the average of the generalized forces over all accessible states obeying uniform distribution. We have also denoted the total number of states with energy within E to $E + \delta E$ as

$$\Omega(E, x) = \sum_X \Omega_X(E, x).$$

The total number of states within E and $E + \delta E$ changes as

$$\frac{\partial \Omega(E, x)}{\partial x} dx = \sigma(E) - \sigma(E + \delta E) = -\frac{\partial \sigma}{\partial E} \delta E.$$

We obtain

$$\frac{\partial \Omega}{\partial x} = \frac{\partial}{\partial E} (\Omega \bar{X}).$$

This implies that

$$\frac{\partial \ln \Omega}{\partial x} = \frac{\partial \ln \Omega}{\partial E} \bar{X} + \frac{\partial \bar{X}}{\partial E}.$$

Since X is intensive but Ω and E are extensive variables, one sees that the second term is negligible compared with the first in the macroscopic limit.

Thus, we have obtained that a change in an external parameter results in a mean generalized force given by the relation

$$\frac{\partial S}{\partial x} = \frac{1}{T} \bar{X}. \quad (23)$$

In general, if we calculate the infinitesimal change in entropy $S = S(E, x_1, \dots, x_n)$, we obtain

$$dS = \left(\frac{\partial S}{\partial E}\right)_{x_1, \dots, x_n} dE + \sum_{i=1}^n \left(\frac{\partial S}{\partial x_i}\right)_{E, x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n} dx_i, \quad (24)$$

where the subscripts denote the variables that are kept constant during partial derivation.

In order to apply the results (20) and (23), one should take care that the process leading to the change in S is reversible. This guarantees that the average values have throughout the process well defined values. The reversibility is achieved by requiring that the changes in E and the external parameters x_i are carried out so slowly that the system remains arbitrarily close to equilibrium at all stages of the process. This is called a *quasi-static process*. The slowness is, naturally, a characteristic of the system and determined by the length of the relaxation time.

For a quasi-static process, one can write

$$dS = \frac{1}{T}dE + \frac{1}{T} \sum_{i=1}^n \bar{X}_i dx_i. \quad (25)$$

We see that the latter term is a sum of components of the form generalized force times the a change in the corresponding coordinate. Thus, it can be related to the *decrease* in the internal energy dE during the process. Accordingly, we call it the *work done by the system* and denote it with

$$dW = \sum_{i=1}^n \bar{X}_i dx_i.$$

The change in the internal energy caused by the thermal interaction is called *heat*

$$dQ = dE + dW, \quad (26)$$

One should pay attention to the notation. Special symbols dQ and dW denote the fact that the heat and work are infinitesimal quantities that do *not* correspond to any differences between two heats or two works of the initial and final states of the process. Such infinitesimal quantities are called *inexact differentials*. In contrast, the state variables (such as E , p , V , T , ...) that have well defined values for the initial and final states are referred to as *exact differentials*.

Generally speaking, consider a function $F(x, y)$ of two independent variables x and y . The differential

$$dF = \frac{\partial F}{\partial x} dx + \frac{\partial F}{\partial y} dy$$

is exact if

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x}.$$

Thus, $dF = dF$ is *integrable* and

$$\int_1^2 dF = F(2) - F(1), \quad (27)$$

meaning that the integral of an exact differential does not depend on the path of integration.

In the case of inexact differentials, the integrals depend on the taken path, but there always exists an *integrating factor* $\lambda(x, y)$ so that

$$\lambda dF = df$$

is an exact differential. For example, the integrating factor of heat is $1/T$ ¹³

$$dS = \frac{dQ}{T}.$$

Now, the total change in the mean energy changes during the process is

$$\Delta E = E_f - E_i = Q - W, \quad (28)$$

where W is the macroscopic work done by the system as a result of the change in external parameters. The quantity Q is the *heat absorbed by the system*, i.e. the change in the macroscopic energy due to the thermal interaction. For a quasi-static process, this is a restatement of the *first law of thermodynamics*, which in its most general form it says:

ENERGY OF AN ISOLATED SYSTEM IS CONSERVED.

Equilibrium conditions

Let us consider the case with an exchange of the external parameters V and N between the subsystems A_1 and A_2 , in addition to energy. Again, we consider a process leading from initial to final equilibrium. The change in entropy $S_1 = S_1(E_1, V_1, N_1)$ due to *infinitesimal* changes in the arguments can be written as

$$dS_1 = \left. \frac{\partial S_1}{\partial E_1} \right|_{V_1, N_1} dE_1 + \left. \frac{\partial S_1}{\partial V_1} \right|_{E_1, N_1} dV_1 + \left. \frac{\partial S_1}{\partial N_1} \right|_{E_1, V_1} dN_1.$$

One can write a similar formula for the change in entropy of the system A_2 by replacing $1 \rightarrow 2$ in the above formula. Due to the fact that the system is isolated, we have that $dE_2 = -dE_1$, $dV_2 = -dV_1$ and $dN_2 = -dN_1$. Thus, the change in entropy of the total system can be written as

$$\begin{aligned} dS = dS_1 + dS_2 = & \left(\left. \frac{\partial S_1}{\partial E_1} \right|_{V_1, N_1} - \left. \frac{\partial S_2}{\partial E_2} \right|_{V_2, N_2} \right) dE_1 \\ & + \left(\left. \frac{\partial S_1}{\partial V_1} \right|_{E_1, N_1} - \left. \frac{\partial S_2}{\partial V_2} \right|_{E_2, N_2} \right) dV_1 \\ & + \left(\left. \frac{\partial S_1}{\partial N_1} \right|_{E_1, V_1} - \left. \frac{\partial S_2}{\partial N_2} \right|_{E_2, V_2} \right) dN_1 \end{aligned} \quad (29)$$

In the equilibrium, small deviations in the total entropy have to vanish, leading to new equilibrium conditions

$$\left. \frac{\partial S_1}{\partial V_1} \right|_{E_1, N_1} = \left. \frac{\partial S_2}{\partial V_2} \right|_{E_2, N_2} \quad (30)$$

$$\left. \frac{\partial S_1}{\partial N_1} \right|_{E_1, V_1} = \left. \frac{\partial S_2}{\partial N_2} \right|_{E_2, V_2}. \quad (31)$$

By recalling from thermodynamics, the definitions of *pressure* and *chemical potential* are respectively

$$\left. \frac{\partial E}{\partial V} \right|_{S, N} = -p, \quad \left. \frac{\partial E}{\partial N} \right|_{S, V} = \mu.$$

¹³Note that $dS \geq dQ/T$ for an arbitrary process, the equality holds only for reversible ones.

Using the mathematical identities, valid for any function f of variables x and y ,

$$\left. \frac{\partial f}{\partial x} \right|_y \left. \frac{\partial x}{\partial y} \right|_f \left. \frac{\partial y}{\partial f} \right|_x = -1, \quad \left. \frac{\partial f}{\partial x} \right|_y = \left(\left. \frac{\partial x}{\partial f} \right|_y \right)^{-1}$$

one arrives at statistical definition for pressure

$$\frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E,N} \quad (32)$$

and that for chemical potential

$$-\frac{\mu}{T} = \left. \frac{\partial S}{\partial N} \right|_{E,V}. \quad (33)$$

What we have obtained is the *condition of equilibrium* in the case where all the state variables E , V and N are allowed to vary

$$T_1 = T_2, \quad p_1 = p_2, \quad \mu_1 = \mu_2. \quad (34)$$

Also, we have obtained the first law of thermodynamics for infinitesimal reversible processes

$$dE = dQ - dW = TdS - pdV + \mu dN. \quad (35)$$

This is often called the *fundamental thermodynamic relation*. It should be noted that the differential of the internal energy consists terms AdB , of which one is always an extensive variable and the other intensive. This observation holds generally.

Note also that for *thermally isolated systems*

$$dS = \frac{dQ}{T} = 0$$

which is the definition of an *adiabatic* process.

4.2 Equations of State

Relations (20) and (23) are of particular interest

$$\frac{1}{T} = \frac{\partial S}{\partial E}, \quad \frac{\partial S}{\partial x} = \frac{1}{T} \bar{X}. \quad (36)$$

Namely, they constitute a way to find relations between temperature, generalized forces (like p and μ) and external parameters (like N and V). The relations are called *equations of state* and are important since they connect quantities that are readily measurable. We will list here some examples of well known equations of state.

Ideal gas

Let us consider a gas of N monatomic particles. We assume that the gas is confined in a container of volume V , and that the interactions and quantum effects between the particles are negligible. This is called *classical ideal gas*. Ideal gas is a good approximation for dilute gases that have

$\rho = N/V \rightarrow 0$. For such a system, one can readily write the dependence of $\Omega(E, V, N)$ on V . We will settle for now in order-of-magnitude calculation, and improve it once we start to discuss the calculation of Ω more generally.

Let us consider a single particle in the gas. Since the system has a fixed number of particles, the number of its accessible microstates depends only on E and V , as $\Omega(E, V)$. We have to find the number of states within the interval E to $E + \delta E$ and in the volume V . Thus, Ω has to be proportional to the total volume V and that of the energy space. Since we assume that the particles are classical and that the interactions are negligible, they can occupy the same space. For the N particle system, we end up with

$$\Omega(E, V) \propto V^N \chi(E), \quad (37)$$

where $\chi(E)$ does not depend on volume.

Now, the calculation of the equation of state is straightforward. We have

$$S = k_B \ln \Omega = k_B N \ln V + k_B \ln \chi(E) + \text{constant}$$

which gives the pressure

$$p = k_B T \frac{N}{V} \text{ or } pV = Nk_B T. \quad (38)$$

Chemists prefer the form

$$pV = nRT,$$

where $n = N/N_A$ is the number of moles of the gas particles and $R = k_B N_A$ is the *gas constant*. Note that equation of state (38) (nor R) does not depend on the kind of the molecules constituting the gas.

In order to acquire complete knowledge on the macrostate of the ideal gas, one should also calculate the equation of state for the internal energy E from equation

$$\frac{1}{T} = \frac{\partial S}{\partial E} = k_B \frac{\partial \ln \chi(E)}{\partial E}.$$

This means that the mean energy $E = E(T)$ is function of only temperature, not volume V . The explicit dependence requires the knowledge on $\chi(E)$. We will return to its determination later.

If the gas is composed of m different species of molecules, the equation of state is still (38) where

$$N = \sum_{i=1}^m N_i \text{ and } p = \sum_{i=1}^m p_i.$$

Here

$$p_i = N_i k_B T / V$$

is the *partial pressure* of the i^{th} gas.

Virial expansion of real gases

The interactions between the constituent molecules become important in real gases. Nevertheless, the equation

of state of any gas can be written in a general series of the particle density $\rho = N/V$ as

$$p = k_B T \left[\rho + B_2(T)\rho^2 + B_3(T)\rho^3 + \dots \right], \quad (39)$$

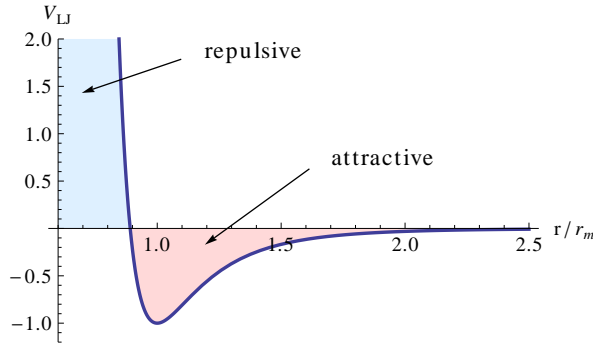
where B_k is the k^{th} virial coefficient.

van der Waals equation

The interactions between the molecules of real gases are

- repulsive at short distances (due to the Pauli principle); every particle needs at least the volume b which implies $V \gtrsim Nb$.
- attractive at large distances due to the induced dipole momenta. The pressure decreases when two particles are separated by the attraction distance. The probability of this is $\propto \rho^2$.

This can be modelled with the *Lennard-Jones potential*



$$V_{LJ}(r) = \varepsilon \left[\left(\frac{r_m}{r} \right)^{12} - 2 \left(\frac{r_m}{r} \right)^6 \right],$$

where ε is the depth of the potential and r_m is the distance where the potential is at minimum.

We improve the ideal gas equation of state

$$p'V' = Nk_B T$$

so that

$$V' = V - Nb \text{ and } p = p' - a\rho^2 = \text{true pressure.}$$

Thus,

$$(p + a\rho^2)(V - Nb) = Nk_B T$$

Stretched wire

Consider a wire at rest with length L_0 . The stretching of the wire to length L causes the *tension*

$$\sigma = E(T)(L - L_0)/L_0,$$

where $E(T)$ is the temperature dependent constant of elasticity.

Surface tension

Surface tension is a property of a surface of a liquid that allows it to resist external forces. It has temperature dependence

$$\sigma = \sigma_0 \left(1 - \frac{T}{T_C} \right)^n$$

where the temperature is measured in $^{\circ}\text{C}$, T_C and n are dependent on the liquid, and σ_0 is the tension at $T = 0^{\circ}\text{C}$.

Electric polarization

When a piece of material is in an external electric field \mathbf{E} we define

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P},$$

where

\mathbf{D} = electric flux density

\mathbf{P} = electric polarization

= atomic total dipole moment density.

If the material is homogeneous and dielectric it has

$$\mathbf{P} = \left(a + \frac{b}{T} \right) \mathbf{E},$$

where a and b are almost constant and $a, b \geq 0$.

Curie's law

When a piece of paramagnetic material is in magnetic field \mathbf{H} we write

$$\mathbf{B} = \mu_0(\mathbf{H} + \mathbf{m}),$$

where

\mathbf{B} = magnetic flux density

\mathbf{m} = magnetic polarization

= atomic total magnetic moment density.

Polarization obeys roughly *Curie's law*

$$\mathbf{m} = \frac{\rho C}{T} \mathbf{H},$$

where C is a constant related to the individual atom.

4.3 Thermodynamic Potentials

Consider a conservative mechanical system, such as a spring, where work can be stored in the form of potential energy and subsequently retrieved. Under *certain* circumstances the same is true for thermodynamic systems. Energy can be stored in a thermodynamic system by doing work on it through a reversible process. This work can later be retrieved in the form of work. This kind of energy that can be stored and retrieved in the form of work is called *free energy*. We will show in the following that there are as many free energies as there are combinations of constraints.

The free energies are analogous to the potential energy of a spring and are, thus, called *thermodynamic potentials*.

Internal energy

The internal energy of a thermodynamic system is determined as a sum of all forms of energy intrinsic to the system. According to the first law (35),

$$dE = TdS - pdV + \mu dN,$$

which means that S , V and N are the *natural variables* of the internal energy¹⁴

$$E = E(S, V, N).$$

The first law implies the relations

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

Because E , S , V and N are extensive, we have that

$$E(\lambda S, \lambda V, \lambda N) = \lambda E(S, V, N),$$

where λ is an arbitrary constant characterizing the size of the system.

Assume then that we make infinitesimal changes to the internal energy, i.e. $S \rightarrow S + \epsilon S$, $V \rightarrow V + \epsilon V$ and $N \rightarrow N + \epsilon N$. Now

$$\begin{aligned} E(S + \epsilon S, V + \epsilon V, N + \epsilon N) &= E(S, V, N) \\ &+ \left(\frac{\partial E}{\partial S}\right)_{V,N} \epsilon S + \left(\frac{\partial E}{\partial V}\right)_{S,N} \epsilon V + \left(\frac{\partial E}{\partial N}\right)_{S,V} \epsilon N. \end{aligned}$$

On the other hand, we have that

$$E(S + \epsilon S, V + \epsilon V, N + \epsilon N) = E(S, V, N) + \epsilon E(S, V, N).$$

We end up with the *Euler equation for homogeneous functions*

$$E = S \left(\frac{\partial E}{\partial S}\right)_{V,N} + V \left(\frac{\partial E}{\partial V}\right)_{S,N} + N \left(\frac{\partial E}{\partial N}\right)_{S,V},$$

or,

$$E = TS - pV + \mu N \quad (40)$$

which is the *fundamental equation*.

The combinations of parameters in the above equation are such that they lead to an exact differential for E . Thus, the second derivatives of E have to be independent on the order of differentiation, e.g.

$$\frac{\partial^2 E}{\partial V \partial S} = \frac{\partial^2 E}{\partial S \partial V}.$$

This leads to a useful set of so called *Maxwell's relations*

$$\begin{aligned} \left(\frac{\partial T}{\partial V}\right)_{S,N} &= -\left(\frac{\partial p}{\partial S}\right)_{V,N} \\ \left(\frac{\partial T}{\partial N}\right)_{S,V} &= \left(\frac{\partial \mu}{\partial S}\right)_{V,N} \\ \left(\frac{\partial p}{\partial N}\right)_{S,V} &= -\left(\frac{\partial \mu}{\partial V}\right)_{S,N} \end{aligned}$$

In an irreversible process,

$$T\Delta S > \Delta Q = \Delta E + \Delta W,$$

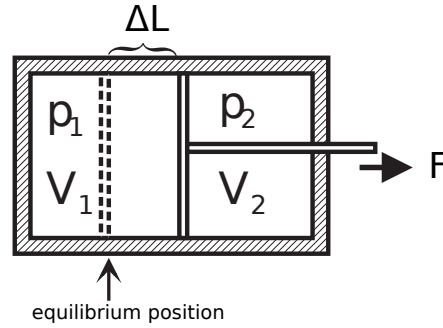
which means that

$$\Delta E < T\Delta S - p\Delta V + \mu\Delta N.$$

If S , V and N stay constant in the process then the internal energy decreases. Thus we can deduce that

In an equilibrium with given S , V and N the internal energy is at the minimum.

We consider a reversible process in a thermally isolated system



We partition ΔW into the components

$$\begin{aligned} \int p dV &= \left[\begin{array}{l} \text{work due to the} \\ \text{change of the total} \\ \text{volume} \end{array} \right] \\ \Delta W_{\text{free}} &= \left[\begin{array}{l} \text{work done by the} \\ \text{gas against the} \\ \text{force } \mathbf{F} \end{array} \right]. \end{aligned}$$

Now

$$\begin{aligned} \Delta W_{\text{free}} &= \Delta W_1 + \Delta W_2 = p_1 \Delta V_1 + p_2 \Delta V_2 \\ &= (p_1 - p_2) \Delta V_1 = (p_1 - p_2) A \Delta L \\ &= -F \Delta L. \end{aligned}$$

According to the first law we have

$$\begin{aligned} \Delta E &= \Delta Q - \Delta W = \Delta Q - \int p dV - \Delta W_{\text{free}} \\ &= \Delta Q - \Delta W_{\text{free}}. \end{aligned}$$

Because now $\Delta Q = 0$, we have

$$\Delta E = -\Delta W_{\text{free}} = F \Delta L,$$

¹⁴Often internal energy is also denoted with U .

i.e. when the variables S , V and N are kept constant the change of the internal energy is completely exchangeable with the work. ΔE is then called *free energy* and E *thermodynamic potential*.

Note If there are irreversible processes in an isolated system (V and N constants) then

$$\Delta W_{\text{free}} \leq -\Delta E.$$

Helmholtz free energy

Internal energy is the appropriate thermodynamic potential in processes where the S , V and N are kept constant. Additionally, the internal energy can be determined as a function of these natural variables, and all thermodynamic properties of the system can be calculated by taking partial derivatives of the internal energy with respect to its natural variables. Often, the experimental circumstances are such that the macrostate of the studied system is more conveniently represented with different set of independent state variables. The change in the variable set made with the *Legendre transformation*.

For example, let us look at the function $f(x, y)$ of two variables. We denote

$$z = f_y = \frac{\partial f(x, y)}{\partial y}$$

and define the function

$$g = f - y f_y = f - y z.$$

Now,

$$\begin{aligned} dg &= df - y dz - z dy = f_x dx + f_y dy - y dz - z dy \\ &= f_x dx - y dz. \end{aligned}$$

Thus we can take x and z as independent variables of the function g , i.e. $g = g(x, z)$. Obviously

$$y = -\frac{\partial g(x, z)}{\partial z}.$$

Corresponding to the Legendre transformation $f \rightarrow g$, there is the inverse transformation $g \rightarrow f$

$$f = g - z g_z = g + y z.$$

In a process where T , V and N are constant, it is worthwhile to study the transformed potential

$$E \rightarrow A = E - S \left(\frac{\partial E}{\partial S} \right)_{V, N}$$

or

$$A = E - TS$$

which is the (*Helmholtz*) *free energy*.

Now

$$dA = -S dT - p dV + \mu dN,$$

so the natural variables of A are T , V and N . We immediately obtain the partial derivatives

$$\begin{aligned} S &= - \left(\frac{\partial A}{\partial T} \right)_{V, N} \\ p &= - \left(\frac{\partial A}{\partial V} \right)_{T, N} \\ \mu &= \left(\frac{\partial A}{\partial N} \right)_{T, V}. \end{aligned}$$

Due to the fact that dA is an exact differential, we obtain the Maxwell relations

$$\begin{aligned} \left(\frac{\partial S}{\partial V} \right)_{T, N} &= \left(\frac{\partial p}{\partial T} \right)_{V, N} \\ \left(\frac{\partial S}{\partial N} \right)_{T, V} &= - \left(\frac{\partial \mu}{\partial T} \right)_{V, N} \\ \left(\frac{\partial p}{\partial N} \right)_{T, V} &= - \left(\frac{\partial \mu}{\partial V} \right)_{T, N}. \end{aligned}$$

Similar to the case with internal energy, we have for an irreversible change

$$\Delta A < -S \Delta T - p \Delta V + \mu \Delta N,$$

i.e. when the variables T , V and N are constant the system drifts to the minimum of the free energy. Correspondingly

$$\Delta W_{\text{free}} \leq -\Delta A,$$

when (T, V, N) is constant.

Helmholtz free energy is suitable for processes that are done at constant volume and temperature. The constant temperature can be achieved by embedding the studied system into *heat bath*. Generally, a bath is an equilibrium system, much larger than the system under consideration, which can exchange given extensive property with the system¹⁵.

Enthalpy

Using the Legendre transform

$$E \rightarrow H = E - V \left(\frac{\partial E}{\partial V} \right)_{S, N} = E + pV$$

We move from the variables (S, V, N) to the variables (S, p, N) . The quantity

$$H = E + pV$$

is called *enthalpy*.

Now,

$$dH = T dS + V dp + \mu dN.$$

¹⁵In the case of heat bath, the exchanged property is entropy, resulting in the constant temperature of the system determined by the bath.

From this we can read the partial derivatives

$$\begin{aligned} T &= \left(\frac{\partial H}{\partial S} \right)_{p,N} \\ V &= \left(\frac{\partial H}{\partial p} \right)_{S,N} \\ \mu &= \left(\frac{\partial H}{\partial N} \right)_{S,V} \end{aligned}$$

Corresponding Maxwell relations are

$$\begin{aligned} \left(\frac{\partial T}{\partial p} \right)_{S,N} &= \left(\frac{\partial V}{\partial S} \right)_{p,N} \\ \left(\frac{\partial T}{\partial N} \right)_{S,p} &= \left(\frac{\partial \mu}{\partial S} \right)_{p,N} \\ \left(\frac{\partial V}{\partial N} \right)_{S,p} &= \left(\frac{\partial \mu}{\partial p} \right)_{S,N} \end{aligned}$$

In an irreversible process one has

$$\Delta Q = \Delta E + \Delta W - \mu \Delta N < T \Delta S.$$

Now $\Delta E = \Delta(H - pV)$, so that

$$\Delta H < T \Delta S + V \Delta p + \mu \Delta N.$$

We see that

In a process where S , p and N are constant spontaneous changes lead to the minimum of H , i.e. in an equilibrium of a (S, p, N) -system the enthalpy is at the minimum.

The enthalpy is a suitable potential for an isolated system in a *pressure bath* (p is constant).

Let us look at an isolated system in a pressure bath. Now

$$dH = dE + d(pV)$$

and

$$dE = dQ - dW + \mu dN.$$

Again we partition the work into two components:

$$dW = p dV + dW_{\text{free}}.$$

Now

$$dH = dQ + V dp - dW_{\text{free}} + \mu dN$$

and for a finite process

$$\Delta H \leq \int T dS + \int V dp - \Delta W_{\text{free}} + \int \mu dN.$$

When (S, p, N) is constant one has

$$\Delta H \leq -\Delta W_{\text{free}}$$

i.e. ΔW_{free} is the minimum work required for the change ΔH .

Note Another name of enthalpy is *heat function* (in constant pressure).

Gibbs' function

The Legendre transformation

$$E \rightarrow G = E - S \left(\frac{\partial E}{\partial S} \right)_{V,N} - V \left(\frac{\partial E}{\partial V} \right)_{S,N}$$

defines the *Gibbs function* or the *Gibbs free energy*

$$G = E - TS + pV.$$

Its differential is

$$dG = -S dT + V dp + \mu dN,$$

so the natural variables are T , p and N . For the partial derivatives we can read the expressions

$$\begin{aligned} S &= - \left(\frac{\partial G}{\partial T} \right)_{p,N} \\ V &= \left(\frac{\partial G}{\partial p} \right)_{T,N} \\ \mu &= \left(\frac{\partial G}{\partial N} \right)_{T,p} \end{aligned}$$

From these we obtain the Maxwell relations

$$\begin{aligned} \left(\frac{\partial S}{\partial p} \right)_{T,N} &= - \left(\frac{\partial V}{\partial T} \right)_{p,N} \\ \left(\frac{\partial S}{\partial N} \right)_{T,p} &= - \left(\frac{\partial \mu}{\partial T} \right)_{p,N} \\ \left(\frac{\partial V}{\partial N} \right)_{T,p} &= \left(\frac{\partial \mu}{\partial p} \right)_{T,N} \end{aligned}$$

In an irreversible process

$$\Delta G < -S \Delta T + V \Delta p + \mu \Delta N,$$

holds, i.e. when the variables T , p and N stay constant the system drifts to the minimum of G .

Correspondingly

$$\Delta W_{\text{free}} \leq -\Delta G,$$

when (T, p, N) is constant.

The Gibbs function is suitable for systems which are allowed to exchange mechanical energy and heat.

Grand potential

The Legendre transform

$$E \rightarrow \Omega = E - S \left(\frac{\partial E}{\partial S} \right)_{V,N} - N \left(\frac{\partial E}{\partial N} \right)_{S,V}$$

defines the *grand potential*

$$\Omega = E - TS - \mu N.$$

Its differential is

$$d\Omega = -S dT - p dV - N d\mu,$$

so the natural variables are T , V and μ .

The partial derivatives are now

$$\begin{aligned} S &= -\left(\frac{\partial \Omega}{\partial T}\right)_{p,\mu} \\ p &= -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} \\ N &= -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V}. \end{aligned}$$

We get the Maxwell relations

$$\begin{aligned} \left(\frac{\partial S}{\partial V}\right)_{T,\mu} &= \left(\frac{\partial p}{\partial T}\right)_{V,\mu} \\ \left(\frac{\partial S}{\partial \mu}\right)_{T,V} &= \left(\frac{\partial N}{\partial T}\right)_{V,\mu} \\ \left(\frac{\partial p}{\partial \mu}\right)_{T,V} &= \left(\frac{\partial N}{\partial V}\right)_{T,\mu}. \end{aligned}$$

In an irreversible process

$$\Delta \Omega < -S \Delta T - p \Delta V - N \Delta \mu,$$

holds, i.e. when the variables T , V and μ are kept constant the system moves to the minimum of Ω .

Correspondingly

$$\Delta W_{\text{free}} \leq -\Delta \Omega,$$

when (T, V, μ) is constant.

The grand potential is suitable for systems that are allowed to exchange heat and particles.

4.4 Application of Thermodynamics

Specific heat

Consider an increase in temperature by ΔT due to absorption of heat ΔQ by the system. We define the *specific heat* or *heat capacity* as

$$C = \frac{\Delta Q}{\Delta T}.$$

If the increase in temperature is done without the change in volume, the process is called *isochoric*. According to the 1st law, we have for infinitesimal reversible processes

$$dE = TdS = dQ$$

which results in *specific heat at constant volume*

$$C_V = T \left(\frac{\partial S}{\partial T}\right)_{V,N} = \left(\frac{\partial E}{\partial T}\right)_{V,N}$$

For (*isobaric*) processes in constant pressure, we have

$$dH = TdS$$

and

$$C_p = T \left(\frac{\partial S}{\partial T}\right)_{p,N} = \left(\frac{\partial H}{\partial T}\right)_{p,N} \quad (41)$$

Compressibility and expansion

Typically, one can control T and p in the experiments (instead of S and V). If we consider the isobaric change in volume as a function of temperature, we can define the *volume heat expansion coefficient*

$$\alpha_p = \frac{1}{V} \left(\frac{\partial V}{\partial T}\right)_{p,N}.$$

Similarly, in *isothermic* volume change we define the *compressibility*

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_{T,N}.$$

In the case of adiabatic compression, we define the *adiabatic compressibility*

$$\kappa_S = -\frac{1}{V} \left(\frac{\partial V}{\partial p}\right)_{S,N}.$$

The sound propagation is related to adiabatic compression of the gas, rather than isothermic, resulting in the relation

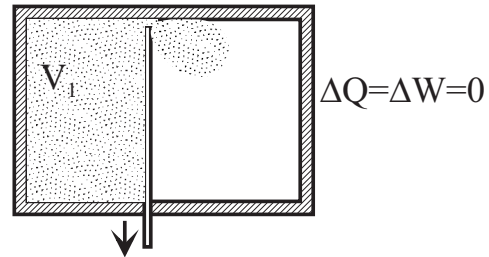
$$c_S = \sqrt{\frac{1}{\rho_m \kappa_S}}$$

between the velocity of sound and the adiabatic compressibility. In the above formula, we have denoted the mass density of the gas with ρ_m . One can show (exercise) that the two compressibilities are related by

$$\kappa_T = \kappa_S + VT \frac{\alpha_p^2}{C_p}.$$

Free expansion

Consider thermally insulated rigid container with two compartments. Assume that the volume V_1 is filled with gas at temperature T_1 . Open the valve separating the compartments so that the gas is free to expand and to fill the whole container with volume V_2 .



We are interested in the temperature T_2 after the equilibrium is reached. It is worthwhile to remember, that right after the valve is opened, the system is in non-equilibrium state and that the expansion of the gas is irreversible. Now, the system is adiabatically insulated and thus

$$\Delta Q = 0.$$

Also, the gas does not do work in the process, i.e.

$$\Delta W = 0.$$

According to the first law,

$$\Delta E = 0.$$

This means that due to the conservation of energy we have that

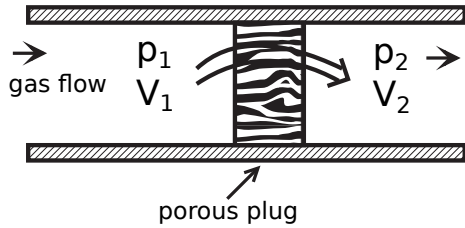
$$E(T_2, V_2) = E(T_1, V_1).$$

Remember, that for ideal gas we had that $E(T)$ does not depend on volume. Thus, the temperature does not change in the free expansion of the ideal gas, i.e. $T_1 = T_2$. One can show that for a real gas (like a van der Waals gas) with interactions, the free expansion results in *cooling*, i.e.

$$T_2 < T_1.$$

Joule-Thomson (throttling) process

Consider then a gas flow in a pipe with thermally insulated walls.



Introduction of a constriction to the flow, such as a porous plug, results in a constant pressure difference $p_1 > p_2$ across the constriction. In a steady state p_1 and p_2 are temporal constants, and the process is again irreversible. Given the temperature T_1 before (on the left hand side of the plug), we are interested in the temperature T_2 after the plug. When an infinitesimal amount of matter passes through the choke the work done by the system is

$$dW = p_2 dV_2 + p_1 dV_1.$$

	V_1	V_2
Initial state	V_{init}	0
Final state	0	V_{final}

The work done by the system is

$$\Delta W = \int dW = p_2 V_{\text{final}} - p_1 V_{\text{init}}.$$

According to the first law we have

$$\Delta E = E_{\text{final}} - E_{\text{init}} = \Delta Q - \Delta W = -\Delta W,$$

so that

$$E_{\text{init}} + p_1 V_{\text{init}} = E_{\text{final}} + p_2 V_{\text{final}}.$$

Thus in this process, the enthalpy $H = E + pV$ is constant, i.e. the process is *isenthalpic*,

$$\Delta H = H_{\text{final}} - H_{\text{init}} = 0.$$

We consider now a *reversible* isenthalpic (and $dN = 0$) process initial \rightarrow final. Here

$$dH = T dS + V dp = 0,$$

so

$$dS = -\frac{V}{T} dp. \quad (*)$$

Now $T = T(S, p)$, so that

$$dT = \left(\frac{\partial T}{\partial S} \right)_p dS + \left(\frac{\partial T}{\partial p} \right)_S dp.$$

On the other hand

$$\left(\frac{\partial T}{\partial S} \right)_p = \frac{T}{C_p},$$

where C_p is the isobaric heat capacity (see Eq. (41)).

Using the Maxwell relation

$$\left(\frac{\partial T}{\partial p} \right)_S = \left(\frac{\partial V}{\partial S} \right)_p$$

and the partial derivative relation

$$\left(\frac{\partial V}{\partial S} \right)_p = \frac{\left(\frac{\partial T}{\partial S} \right)_p}{\left(\frac{\partial T}{\partial V} \right)_p}$$

we can write

$$dT = \frac{T}{C_p} dS + \frac{T}{C_p} \left(\frac{\partial V}{\partial T} \right)_p dp.$$

Substituting into this the differential dS in constant enthalpy (*) we get so called *Joule-Thomson coefficients*

$$\left(\frac{\partial T}{\partial p} \right)_H = \frac{T}{C_p} \left[\left(\frac{\partial V}{\partial T} \right)_p - \frac{V}{T} \right].$$

Defining the *heat expansion coefficient* α_p so that

$$\alpha_p = \frac{1}{V} \left(\frac{\partial V}{\partial T} \right)_p,$$

we can rewrite the Joule-Thomson coefficient as

$$\mu_{\text{JT}} = \left(\frac{\partial T}{\partial p} \right)_H = \frac{V}{C_p} (T\alpha_p - 1).$$

We see that when the pressure decreases the gas

- cools down, if $T\alpha_p > 1$.
- warms up, if $T\alpha_p < 1$.

For ideal gases $\left(\frac{\partial T}{\partial p} \right)_H = 0$ holds. For real gases $\left(\frac{\partial T}{\partial p} \right)_H$ is positive below the *inversion temperature*, so the gas cools down. In the case of van der Waals gas, one can show that

$$\mu_{\text{JT}} = \frac{1}{C_p} \left[\frac{2a}{RT} - b \right],$$

where one assumes that the gas has low density implying $RTv \gg a$ and $v \gg b$ ($v = V/n$ is the molar volume and $n = N/N_A$).

Joule-Thomson effect is often used in the process of liquefying gases. In order to achieve cooling, one has to work in such temperature and pressure range that $\mu_{JT} > 0$. At room temperature, all gases except hydrogen, helium and neon cool upon expansion by the Joule-Thomson process. For example, for helium one has to pre-cool it first (e.g. with liquid hydrogen) below 34 K. Also, the throttling process is applied in many thermal machines, such as refrigerators, air conditioners, and heat pumps.

Chemical reaction

Chemical reactions involve large numbers of molecules breaking and reforming their chemical bonds. The *chemical bonds* characterize the manner by which the atoms are joined to form molecules¹⁶. Examples include:

- *Covalent bond*, caused by the mutual sharing of electron pairs, such as in H_2 .
- *Ionic bond*, molecule held together by electrostatic attraction of two oppositely charged ions, e.g. $NaCl$.
- *Polar bond*, covalent bond but unequal sharing of electrons which causes permanent dipole moment of the molecule, e.g. H_2O .

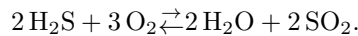
Regardless of the type of bonds, the macroscopic properties of all chemical reactions can be described by thermodynamics. In the following, we will discuss only reactions involving electrically neutral species of molecules.

Chemical equilibrium is reached when the rate of production equals the rate of depletion for each species of molecules involved. Reactions themselves do not stop, even at equilibrium. In general, the *chemical reaction formula* is written as

$$0 = \sum_j \nu_j M_j.$$

Here $\nu_j \in \mathcal{I}$ are the *stoichiometric coefficients* and M_j stand for the molecular species.

Example Consider for example the chemical reaction (burning of hydrogen sulphide)



j	A	B	C	D
M_j	H_2S	O_2	H_2O	SO_2
ν_j	-2	-3	2	2

Typically, the chemical reactions take place in constant pressure and temperature. Thus, the relevant thermodynamic potential is the Gibbs' potential. It turns out that the Gibbs' free energy can be generalized to such situations where the system is out of equilibrium. This can be

¹⁶More info in the courses Solid State Physics and Condensed Matter Physics.

done with a single additional variable called the *degree of reaction* ξ so that

$$dN_j = \nu_j d\xi.$$

When ξ increments by one, one reaction of the reaction formula from left to right takes place.

Convention: When $\xi = 0$ the reaction is as far left as it can be. Then we always have

$$\xi \geq 0.$$

The differential for the Gibbs potential can now be written as

$$dG = \sum_j \mu_j dN_j = d\xi \sum_j \nu_j \mu_j.$$

We define

$$\Delta_r G \equiv \Delta_r \equiv \left(\frac{\partial G}{\partial \xi} \right)_{p,T} = \sum_j \nu_j \mu_j.$$

Δ_r is thus the change in the Gibbs function per one reaction.

Since (p, T) is constant G has a minimum at an equilibrium. The equilibrium condition is thus

$$\Delta_r G^{\text{eq}} = \sum_j \nu_j \mu_j^{\text{eq}} = 0.$$

In a non-equilibrium $dG/dt < 0$, so if $\Delta_r > 0$ we must have $d\xi/dt < 0$, i.e. the reaction proceeds to left and vice versa.

We assume that the components obey the ideal gas equation of state. Then one can show¹⁷

$$\mu_j = k_B T [\phi_j(T) + \ln p + \ln x_j],$$

where $x_j = N_j/N$ is the concentration of j^{th} component, $N = \sum N_j$,

$$\phi_j(T) = \frac{\mu_j^0}{k_B T} - \eta_j - \frac{5}{2} \ln T,$$

and μ_j^0 and η_j are constants. So,

$$\Delta_r G = k_B T \sum_j \nu_j \phi_j(T) + k_B T \ln \left(p^{\sum \nu_j} \prod x_j^{\nu_j} \right).$$

The equilibrium condition can now be written as

$$\prod_j x_j^{\nu_j} = p^{-\sum_j \nu_j} K(T),$$

where

$$K(T) = e^{-\sum_j \nu_j \phi_j(T)}$$

is the *equilibrium constant* of the reaction. The equilibrium condition is called the *law of mass action*.

The *reaction heat* is the change of heat $\Delta_r Q$ per one reaction to right. A reaction is

¹⁷For the rigorous derivation, one needs the concepts of quantum statistics, which will be introduced later. For now the derivation of this formula is omitted.

- *Endothermic*, if $\Delta_r Q > 0$ i.e. the reaction takes heat.
- *Exothermic*, if $\Delta_r Q < 0$ i.e. the reaction releases heat.

We write $\Delta_r G$ as

$$\Delta_r G = -k_B T \ln K(T) + k_B T \sum_j \nu_j \ln p x_j.$$

Now

$$\begin{aligned} \Delta Q &= \Delta E + \Delta W = \Delta E + p \Delta V = \Delta(E + pV) \\ &= \Delta H, \end{aligned}$$

since $\Delta p = 0$.

When the total amount matter is constant

$$dG = -S dT + V dp$$

holds in a reversible process and

$$\begin{aligned} d\left(\frac{G}{T}\right) &= \frac{1}{T} dG - \frac{G}{T^2} dT = -\left(\frac{G}{T^2} + \frac{S}{T}\right) dT + \frac{V}{T} dp \\ &= -\frac{H}{T^2} dT + \frac{V}{T} dp, \end{aligned}$$

because $G = H - TS$. We see that

$$H = -T^2 \left[\frac{\partial}{\partial T} \left(\frac{G}{T} \right) \right]_{p,N}.$$

Now

$$\frac{\partial}{\partial T} \left(\frac{\Delta_r G}{T} \right) = -k_B \frac{d}{dT} \ln K(T),$$

so that

$$\Delta_r H = k_B T^2 \frac{d}{dT} \ln K(T).$$

This expression is known as the *reaction heat*.

5. Ensemble Theory

The preceding chapter was devoted in studying the macroscopic consequences of statistical physics in the thermodynamic limit. We argued that the whole of the classical thermodynamics can be derived from the knowledge of the functional dependence of the number of microstates on the macroscopic parameters, e.g. $\Omega = \Omega(S, V, N)$. In this chapter, we will consider the way how the actual explicit dependence can be resolved.

5.1 Phase Space of a Classical System

The *microstate* of a *classical* system can be determined up to arbitrary accuracy, at least in principle. This is achieved by measuring the *instantaneous* positions and momenta of all particles constituting the system. This requires $3N$ position coordinates $(q_1, q_2, \dots, q_{3N})$ and $3N$ momentum coordinates $(p_1, p_2, \dots, p_{3N})$. The space spanned by the position coordinates is often referred to as the *configuration space*, and that spanned by the momentum coordinates as the *momentum space*. Together, the $6N$ coordinates form the so-called *phase space*.

The phase point (q_i, p_i) , where $i = 1, 2, \dots, 3N$ is a *representative point*. Every representative point corresponds to a possible state of the system. As it evolves in time, a representative point carves a *trajectory* into the phase space. The direction of the trajectory is determined by the velocity vector $\mathbf{v} = (\dot{q}_i, \dot{p}_i)$, whose components are solved from Hamilton's equations of motion

$$\dot{q}_i = \frac{\partial H(q_i, p_i)}{\partial p_i} \quad (42)$$

$$\dot{p}_i = -\frac{\partial H(q_i, p_i)}{\partial q_i}, \quad (43)$$

where $H(q_i, p_i) \equiv H(q, p)$ is the Hamiltonian of the system.

Macroscopic parameters limit the accessible regions in phase space. For example, the knowledge of the volume of the system V immediately tells us which are the possible values for the coordinates q_i in the configuration space. Moreover, if the system has the energy E , the trajectory is restricted to the *hypersurface*

$$H(q, p) = E.$$

More realistically, the energy is determined in the interval E to $E + \delta E$, and the trajectories are within the *hypershell* defined by these limits.

Given the values of the relevant macroscopic state variables (such as E , V and N), the system can be in any of the possible microstates within the corresponding hypershell. We define the *ensemble in phase space* as a group of mental copies of the given system, one for each possible representative point.

At any given time t , the *number of representative points* within the volume element $d^{3N}q d^{3N}p$ around the point (q, p) is given by

$$\rho(q, p; t) d^{3N}q d^{3N}p, \quad (44)$$

where $\rho(q, p; t)$ is the *density function*. Accordingly, we define the *ensemble average* of a given physical quantity $f(q, p)$ as

$$\langle f \rangle = \frac{\int f(q, p) \rho(q, p; t) d^{3N} q d^{3N} p}{\int \rho(q, p; t) d^{3N} q d^{3N} p}. \quad (45)$$

We call the ensemble *stationary* if

$$\frac{\partial \rho}{\partial t} = 0.$$

This implies time-independent averages $\langle f \rangle$ and, thus, the stationary ensemble describes a system in *equilibrium*.

Ergodicity

We can imagine two kinds of averages in the ensemble. If we consider a single representative point, we can follow it along its trajectory for a long time T and calculate the *time average*

$$\bar{f} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T f(q(t), p(t)) dt.$$

On the other hand, we can look at the whole swarm of representative points and take the ensemble average (45) *at any given time*. If the ensemble is stationary, the $\langle f \rangle$ is time-independent and

$$\langle f \rangle = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \langle f \rangle dt.$$

The two averages in the previous equation are completely independent. Thus,

$$\langle f \rangle = \langle \bar{f} \rangle.$$

According to *ergodic hypothesis*¹⁸, the trajectory of a representative point travels through every point in the ensemble in a sufficiently long period of time. Thus, the time averages of physical quantities have to be the same for *every* member of the ensemble. Thus,

$$\langle f \rangle = \bar{f},$$

where the latter is calculated for a trajectory of *any* member of the ensemble.

In the experiments, what we measure are long-time averages. In other word,

THE ENSEMBLE AVERAGE OF ANY PHYSICAL QUANTITY f IS IDENTICAL TO THE VALUE ONE EXPECTS TO OBTAIN BY MAKING AN APPROPRIATE MEASUREMENT ON THE GIVEN SYSTEM.

¹⁸The proof of the ergodic hypothesis is not at all trivial and, thus, omitted here. It was first proposed by L. Boltzmann (1871). Notice also the similarity with the *law of large numbers*.

5.2 Liouville's Theorem and Microcanonical Ensemble

We consider the "flow" of representative points in the phase space. Let ω be an arbitrary volume in the phase space and σ the corresponding surface. The number of points within the volume increases as

$$\frac{\partial}{\partial t} \int_{\omega} \rho d\omega,$$

where $d\omega = d^{3N} q d^{3N} p$. On the other hand, the flow out of the volume can be written as

$$\int_{\sigma} \rho \mathbf{v} \cdot \hat{\mathbf{n}} d\sigma = \int_{\omega} \nabla \cdot (\rho \mathbf{v}) d\omega,$$

where \mathbf{v} is the velocity of the representative points, $d\sigma$ is the surface element, and $\hat{\mathbf{n}}$ the unit vector normal to the element. We have also applied the divergence theorem, where the divergence in the phase space is naturally given by

$$\nabla \cdot (\rho \mathbf{v}) = \sum_i^{3N} \left[\frac{\partial}{\partial q_i} (\rho \dot{q}_i) + \frac{\partial}{\partial p_i} (\rho \dot{p}_i) \right].$$

The number of representative points is conserved, and thus

$$\frac{\partial}{\partial t} \int_{\omega} \rho d\omega = - \int_{\omega} \nabla \cdot (\rho \mathbf{v}) d\omega.$$

This has to hold for an arbitrary volume ω , so we end up with the *continuity equation* of the ensemble

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0. \quad (46)$$

By applying the definition of divergence, the continuity equation can be rewritten as

$$\frac{d\rho}{dt} = \frac{\partial \rho}{\partial t} + \{\rho, H\} = 0, \quad (47)$$

which is called *Liouville's theorem*. In the above, we have denoted the classical Poisson bracket with

$$\{A, B\} = \sum_{i=1}^{3N} \left(\frac{\partial A}{\partial q_i} \frac{\partial B}{\partial p_i} - \frac{\partial A}{\partial p_i} \frac{\partial B}{\partial q_i} \right)$$

Liouville's theorem means that the representative points in the ensemble move in the phase space in the same manner as an incompressible fluid in the physical space.

Microcanonical ensemble

Liouville's theorem holds quite generally for classical systems, whereas

$$\frac{\partial \rho}{\partial t} = 0$$

is merely a requirement for the equilibrium. Thus, in equilibrium we get

$$\{\rho, H(q, p)\} = \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0.$$

Consider then a system of N molecules confined to a volume V with the total energy within the range $E \leq H(q, p) \leq E + \delta E$. One possible way to satisfy both of the above conditions is to assume that the density is completely independent on the coordinates, i.e.

$$\rho(q, p) = \begin{cases} \text{constant} & \text{if } E \leq H(q, p) \leq E + \delta E \\ 0 & \text{otherwise} \end{cases}$$

This is the definition of *microcanonical* ensemble. Physically, this implies that the ensemble is *uniformly* distributed among all possible microstates. Thus, there is equal a priori probability for a given member of the microcanonical ensemble of being in any of the accessible microstates. Microcanonical ensemble is, therefore, describing isolated systems in equilibrium. It is of great importance, since whenever we have two interacting systems A and A' the situation can be reduced to the case of an isolated system by studying the combined system $A + A'$.

It should be kept in mind that even though classical physics allows in principle the determination of the coordinates q_i and p_i up to an arbitrary accuracy, in practice there always exists uncertainties Δq_i and Δp_i (e.g. due to measurement precision). We denote the product of uncertainties of conjugated variables with

$$h_0 = \Delta q_i \Delta p_i.$$

This means that the phase space is divided into volume elements of size h_0^f ($f = 3N$ is the number of degrees of freedom), each of which can be seen to contain one distinguishable microstate. Thus, the number of microstates within the relevant hypershell of the phase space is, asymptotically,

$$\Omega = \frac{\omega}{h_0^f},$$

where

$$\omega = \int d\omega = \int d^f q d^f p$$

and the integration is done over the hypershell. From the knowledge of Ω , thermodynamics follows via entropy

$$S = k_B \ln \Omega, \quad (48)$$

as in the previous chapters.

Fundamentally, the *simultaneous* measurement of canonically conjugated variables is restricted by the Heisenberg uncertainty principle

$$\Delta q_i \Delta p_i \geq \hbar. \quad (49)$$

This implies that the ultimate limit of accuracy is $h_0 \rightarrow \hbar$.

Example: Ideal gas

Consider the equilibrium of monatomic ideal gas consisting of N particles and confined in volume V . Previously, we argued in Eq. (37) that the number of accessible microstates in ideal gas is

$$\Omega \propto V^N \chi(E).$$

Let us try to justify this more rigorously in terms of the phase space arguments presented above. The Hamiltonian of the ideal gas is independent on the configuration space, due to the lack of interactions, i.e.

$$H(q, p) = \sum_{i=1}^{3N} \frac{p_i^2}{2m},$$

where m is the mass of the atom. The volume of the hypershell of the accessible states is given by

$$\omega = \int d^{3N} q \int d^{3N} p = V^N \int d^{3N} p.$$

The condition $E \leq \sum_{i=1}^{3N} \frac{p_i^2}{2m} \leq E + \delta E$ defines a shell in the momentum space, lying in between a sphere with radius $R(E) = \sqrt{2mE}$ and that of slightly larger radius $R(E + \delta E)$. In the limit $\delta E \ll E$, the thickness of the shell is approximately

$$R(E + \delta E) - R(E) \approx \delta E \sqrt{\frac{m}{2E}}.$$

The volume integral in momentum space is then given by the product of this thickness and the surface area of the (momentum) hypersphere with radius $R(E)$. In general, the surface area $S_n(R)$ of n -dimensional sphere with radius R is given by

$$S_n(R) = \frac{2\pi^{n/2} R^{n-1}}{\Gamma(n/2)},$$

where Γ is the gamma-function. Thus, we obtain

$$\omega = \frac{\delta E}{E} V^N \frac{(2\pi m E)^{3N/2}}{[(3N/2) - 1]!}.$$

Now, the number of accessible microstates is

$$\Omega(E, V, N) = \frac{\omega}{h^{3N}} = \frac{\delta E}{E h^{3N}} V^N \chi(E),$$

where $\chi(E) = (2\pi m E)^{3N/2} / (E[(3N/2) - 1]!)$.

5.3 Canonical Ensemble

Often the system under study is not isolated, but interacting with a much larger system (i.e. a bath or a reservoir) fixing one or more of the systems macroscopic properties. One example of such instances is when the system is immersed in a heat bath.

We consider a small system A in thermal interaction with a heat bath A' . In equilibrium, the two systems have the same temperature T , but the energy E of the studied system is a variable. This kind of treatment is often beneficial, since in experimental situations it is typically easier to control temperature, instead of energy.

We assume that the bath contains an infinitely large number of mental copies of the system A . These copies form the so-called *canonical* ensemble in which the macroscopic state of the systems is determined by the parameters

T , V and N . The system A can in principle be any relatively small macroscopic system, or even a distinguishable *microscopic* system.

Then, we would like to know what is the probability P_r that the system is in any one particular microstate r of energy E_r . In other words, we are not interested in measuring any properties that depend upon the heat bath, but want a statistical ensemble for the system that averages over the relevant states of the bath.

Consider that the common temperature of the system in equilibrium with the heat bath is T . This determines the total energy $E^{(0)}$ of the combined system $A + A'$. Now, assuming that the system A is in a state with energy E_r , then the bath has such energy E'_r that

$$E_r + E'_r = E^{(0)} = \text{constant}.$$

The bath is much larger than the system, and thus

$$\frac{E_r}{E^{(0)}} = \left(1 - \frac{E'_r}{E^{(0)}}\right) \ll 1. \quad (50)$$

We denote with $\Omega'(E'_r)$ the number of states of accessible states of the bath with energy E'_r . Because the combined system $A + A'$ is isolated and in equilibrium, all states should be found with equal probabilities. Thus,

$$P_r \propto \Omega'(E'_r) = \Omega'(E^{(0)} - E_r).$$

Because of Eq. (50), we have that¹⁹

$$\begin{aligned} \ln \Omega'(E'_r) &= \ln \Omega'(E^{(0)}) + \left(\frac{\partial \ln \Omega'}{\partial E'}\right)_{E'=E^{(0)}} (E'_r - E^{(0)}) + \dots \\ &\simeq \text{constant} - \beta' E_r. \end{aligned}$$

In equilibrium $\beta' = \beta = 1/k_B T$, and we obtain

$$P_r \propto \exp(-\beta E_r).$$

So, we see that the probability P_r is proportional to the *Boltzmann constant*. The probabilities should be normalized, so by summing over all possible states r of A we should have $\sum_r P_r = 1$, and

$$P_r = \frac{\exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)}. \quad (51)$$

This is called the *canonical distribution*. The sum over states normalizing the distribution,

$$Z(\beta) = \sum_r \exp(-\beta E_r), \quad (52)$$

is often referred to as the *partition function*. Canonical ensemble consists of mental copies of A , all of which obey the distribution (51).

Gibbs entropy

The entropy of the whole system (system+bath) is again given by Eq. (48). But, we are interested merely on the properties of the small system A , and thus we have to generalize our definition of entropy to²⁰

$$S = -k_B \int d^{3N} q d^{3N} p \rho(q, p) \ln \rho(q, p) \quad (53)$$

$$= -k_B \sum_r P_r \ln P_r. \quad (54)$$

This is the *Gibbs entropy* which is valid even for non-equilibrium systems, as well as system that interact with surroundings. The latter form contrasts the fact that there is a limit in how small "compartments" one can divide the phase space (ultimately set by the Heisenberg uncertainty principle), given by h^{3N} . Thus,

$$P_r = \rho(q_r, p_r) h^{3N}.$$

We see immediately that in the case of an isolated system in equilibrium, $P_r = 1/\Omega$ for all r and

$$S = k_B \ln \Omega.$$

Now, we can show alternatively²¹ that the canonical distribution results from the maximization of the entropy with the constraints (exercise)

$$\begin{aligned} \langle E \rangle &= \sum_r E_r P_r = \text{constant} \\ \sum_r P_r &= 1. \end{aligned}$$

Physical significance of canonical distribution

The internal energy for the system interacting with a heat bath is given by the expectation value

$$\langle E \rangle = \frac{\sum_r E_r \exp(-\beta E_r)}{\sum_r \exp(-\beta E_r)} = -\frac{\partial}{\partial \beta} \ln Z(\beta).$$

We recall from the chapter discussing the thermodynamic potentials, that the relevant potential for processes in constant temperatures is the Helmholtz free energy A . Now,

$$\langle E \rangle = A + TS = \left(\frac{\partial(A/T)}{\partial(1/T)} \right)_{N,V},$$

which implies

$$A(T, V, N) = -k_B T \ln Z(T, V). \quad (55)$$

This is the most fundamental result of canonical ensemble theory, since now the all thermodynamic quantities are given in terms of the partition function.

²⁰Compare this with the proof of the H-theorem.

²¹By using Lagrange's multipliers, check e.g. the course of Analytical Mechanics.

¹⁹Again, we study the logarithm of Ω due to reasons of convergence.

Example: Classical ideal gas

Again, we consider classical ideal gas as an example. We recall that $\rho(q, p)$ is the measure of the probability of finding a representative point in the vicinity of the phase space point (q, p) . In canonical ensemble,

$$\rho(q, p) \propto \exp(-\beta H(q, p)),$$

The number of distinct microstates in a volume element $d\omega$ was found to be²²

$$\frac{d\omega}{N!h^{3N}}. \quad (56)$$

Thus, the partition function is

$$Z(T, V) = \frac{1}{N!h^{3N}} \int \exp(-\beta H(q, p)) d\omega.$$

In the case of ideal gas,

$$H(q, p) = \sum_{i=1}^N \frac{p_i^2}{2m}$$

and the partition function can be written in the form

$$Z(T, V) = \frac{1}{N!} \left[\frac{V}{h^3} (2\pi m k_B T)^{3/2} \right]^N = \frac{1}{N!} [Z_1(T, V)]^N. \quad (57)$$

This means that because of the *absence* of the interparticle interactions, the partition function can be written as a product of single molecule partition functions.

Now, we can write

$$A(T, V, N) = N k_B T \left[\ln \left\{ \frac{N}{V} \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \right\} - 1 \right],$$

$$\mu = \left(\frac{\partial A}{\partial N} \right)_{T, V} = k_B T \ln \left\{ \frac{N}{V} \left(\frac{h^2}{2\pi m k_B T} \right)^{3/2} \right\}$$

and

$$P = - \left(\frac{\partial A}{\partial V} \right)_{V, N} = \frac{N k_B T}{V}.$$

Also, we obtain the expression for the internal energy of the classical ideal gas

$$\langle E \rangle = A + TS = - \left[\frac{\partial}{\partial \beta} \ln Z \right]_{E_r} = \frac{3}{2} N k_B T.$$

Energy fluctuations in canonical ensemble

In many physical cases, the accessible energy levels can be *degenerate*, meaning that there exists in general g_i states belonging to the same energy value E_i . Thus,

$$Z(T, V) = \sum_i g_i \exp(-\beta E_i),$$

and

$$P_i = \frac{g_i \exp(-\beta E_i)}{\sum_i g_i \exp(-\beta E_i)}.$$

Often one can assume that the accessible energy levels are very close to one another compared with the average energy E . Accordingly, one can assume that the energy E is *continuous* within any reasonable interval $(E, E + \delta E)$. Correspondingly, the probability that the given system from the canonical ensemble has the energy in this energy range is

$$P(E) = \mathcal{P}(E) dE \propto \exp(-\beta E) g(E) dE, \quad (58)$$

where $\mathcal{P}(E)$ is the probability density and $g(E)$ is the *density of states*. We obtain that

$$Z(T, V) = \int_0^\infty \exp(-\beta E) g(E) dE.$$

Naturally, if δE is sufficiently small energy interval, we have that the number of states within range $(E, E + \delta E)$ is

$$\Omega \approx g(E) \delta E.$$

Since the Boltzmann factor is a monotonically decreasing function of E , and the density of states monotonically increasing one, we have that

$$\left. \frac{\partial}{\partial E} \left(e^{-\beta E} g(E) \right) \right|_{E=E^*} = 0$$

determines a maximum of $P(E)$. This can be rewritten as

$$\left. \frac{\partial \ln g(E)}{\partial E} \right|_{E=E^*} = \beta.$$

Now, we can write

$$S = k_B \ln g \quad \text{and} \quad \left(\frac{\partial S(E)}{\partial E} \right)_{E=\langle E \rangle} = \frac{1}{T}.$$

Thus, we see that the most probable value is equal to the mean, i.e. $E^* = \langle E \rangle$.

Consider then the mean

$$\langle E \rangle = \frac{\int E g(E) \exp(-\beta E) dE}{\int g(E) \exp(-\beta E) dE}.$$

we are interested in the variance

$$\sigma^2 = \langle E^2 \rangle - \langle E \rangle^2 = - \frac{\partial \langle E \rangle}{\partial \beta} = k_B T^2 C_V.$$

We see that because C_V and $\langle E \rangle$ are extensive variables, one has

$$\frac{\sigma}{\langle E \rangle} \propto \frac{1}{\sqrt{N}}.$$

Thus, in the limit $N \rightarrow \infty$ we have that the energy of a system in the canonical ensemble is for all practical purposes equal to the mean energy. This means that the situation is more or less the same as in the microcanonical ensemble!

²²The additional factor $N!$ comes from the quantum mechanical fact that the constituting particles in the system are *indistinguishable*. We will return to this later when we talk about quantum statistics more thoroughly.

Moreover, if we make the Taylor expansion around $E^* = \langle E \rangle$

$$\ln \left[e^{-\beta E} g(E) \right] \approx \left(-\beta \langle E \rangle + \frac{S}{k_B} \right) - \frac{1}{2\sigma^2} (E - \langle E \rangle)^2 + \dots$$

Thus, we see that

$$\mathcal{P}(E) \propto e^{-\beta E} g(E) \approx e^{-\beta(\langle E \rangle - TS)} \exp \left(-\frac{(E - \langle E \rangle)^2}{2\sigma^2} \right),$$

meaning that the probability distribution is *Gaussian* with the mean $\langle E \rangle$ and the standard deviation $\sigma = \sqrt{k_B T^2 C_V}$.

Equipartition theorem

Consider the expectation value

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = \frac{\int \left(x_i \frac{\partial H}{\partial x_j} \right) e^{-\beta H} d\omega}{\int e^{-\beta H} d\omega},$$

where x_i and x_j are any of the $6N$ generalized coordinates (q, p) , and H is the (classical) Hamiltonian of the system. By using partial integration to the numerator, we obtain

$$\langle x_i \frac{\partial H}{\partial x_j} \rangle = \delta_{ij} k_B T.$$

We thus obtain

$$\langle p_i \frac{\partial H}{\partial p_i} \rangle = \langle p_i \dot{q}_i \rangle = k_B T = -\langle q_i \dot{p}_i \rangle = \langle q_i \frac{\partial H}{\partial q_i} \rangle.$$

By summing over i , we obtain

$$\sum_{i=1}^{3N} \langle p_i \dot{q}_i \rangle = -\sum_{i=1}^{3N} \langle q_i \dot{p}_i \rangle = 3N k_B T. \quad (59)$$

Hamiltonian of a system that is *quadratic* in its coordinates can be written as

$$H = \sum_j A_j P_j^2 + \sum_j B_j Q_j^2,$$

where Q_j and P_j are canonically conjugate and A_j and B_j constants of the problem. Clearly,

$$\sum_j \left(P_j \frac{\partial H}{\partial Q_j} + Q_j \frac{\partial H}{\partial P_j} \right) = 2H. \quad (60)$$

This implies

$$\langle H \rangle = \frac{1}{2} f k_B T, \quad (61)$$

where f is the number of non-zero coefficients A_j and B_j . So, we see that each harmonic term in the Hamiltonian makes a contribution of $\frac{1}{2} k_B T$ to the internal energy of the system. This also results to a contribution of $\frac{1}{2} k_B$ to the specific heat C_V . This the theorem of *equipartition of energy* among the degrees of freedom of the system²³.

²³There will corrections to this theorem in the quantum limit where some of the degrees of freedom can be "frozen" (cannot be excited) due to the insufficient amount of available (thermal) energy.

The derivation of the equipartition theorem embodies also the so-called *viral theorem* by Clausius (1870). The *virial* of the system

$$\mathcal{V} = \sum_i \langle q_i \dot{p}_i \rangle = -3N k_B T. \quad (62)$$

Immediately from Eq. (60), one obtains

$$\mathcal{V} = -2\mathcal{K}, \quad (63)$$

where \mathcal{K} is the average kinetic energy of the system.

5.4 Applications of Canonical Ensemble

A system of harmonic oscillators

Harmonic oscillator is one of the most practical models in physics. We will now examine a system of N non-interacting oscillators, mainly to illustrate the canonical ensemble formulation, but also for the later development of the theory of black-body radiation and the theory of lattice vibrations.

First, consider the classical Hamiltonian of a single oscillator

$$H(q_i, p_i) = \frac{1}{2} m \omega^2 q_i^2 + \frac{1}{2m} p_i^2.$$

The single oscillator partition function is then

$$Z_1(\beta) = \int \int e^{-\beta H(q,p)} \frac{dq dp}{h} = \frac{k_B T}{\hbar \omega},$$

where $\hbar = h/2\pi$. This is the classical counting of the number of accessible microstates. Similar to the non-interacting gas, the partition function of N oscillators is a product of N single oscillator partition functions²⁴

$$Z(\beta) = [Z_1(\beta)]^N = \left(\frac{k_B T}{\hbar \omega} \right)^N.$$

Again, from the knowledge of the partition function, the thermodynamics follows. For example,

$$A = N k_B T \ln \left(\frac{\hbar \omega}{k_B T} \right),$$

$$\mu = k_B T \ln \left(\frac{\hbar \omega}{k_B T} \right),$$

$$p = 0,$$

$$S = N k_B \left[\ln \left(\frac{k_B T}{\hbar \omega} \right) + 1 \right],$$

$$\langle E \rangle = N k_B T.$$

The internal energy agrees with the equipartition of energy, because there are two independent quadratic terms for each single-oscillator Hamiltonian.

It should be noted at this point that when we derived the canonical distribution, we did not make assumptions on the

²⁴Note, that the oscillators themselves are *distinguishable*, but the excitations (i.e. photons/phonons) are *distinguishable*. We will return to this.

nature of the energy levels E_r of the system. So, instead of classical energies, we can calculate the thermodynamics starting from the quantum mechanical eigenvalues for the single oscillators, namely

$$E_n = \left(n + \frac{1}{2}\right)\hbar\omega.$$

Accordingly, the single-oscillator partition function is

$$Z_1(\beta) = \sum_{n=0}^{\infty} e^{-\beta(n+1/2)\hbar\omega} = \left[2 \sinh\left(\frac{\hbar\omega}{2k_B T}\right)\right]^{-1}.$$

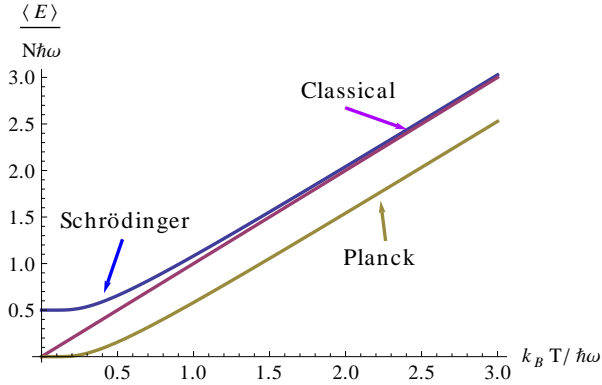
This results in the N -oscillator partition function

$$Z(\beta) = [Z_1(\beta)]^N = e^{-(N/2)\beta\hbar\omega} (1 - e^{-\beta\hbar\omega})^{-N}.$$

Again, we obtain

$$\begin{aligned} A &= N \left[\frac{1}{2}\hbar\omega + k_B T \ln(1 - e^{-\hbar\omega/k_B T}) \right], \\ \mu &= A/N, \\ p &= 0, \\ S &= N k_B \left[\frac{\beta\hbar\omega}{e^{\beta\hbar\omega} - 1} - \ln(1 - e^{-\beta\hbar\omega}) \right], \\ \langle E \rangle &= N \left[\frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \right]. \end{aligned} \quad (64)$$

It is interesting to notice, that the energy per oscillator approaches the equipartition energy only when $k_B T \gg \hbar\omega$. In low temperatures, the internal energy is given by the zero-point energy of the oscillators $\hbar\omega/2$.



Average energy (64) per oscillator, labelled as Schrödinger oscillator. Note that if there were no zero-point oscillations, the classical limit would be off by $-\frac{1}{2}N\hbar\omega$ (Planck oscillator).

Statistics of paramagnetism

Paramagnetism is a form of magnetism caused by *permanent* dipole moments of the constituent atoms, generated by the electron spin and angular momentum. Magnetization of paramagnetic materials is created only in the presence of an external magnetic field \mathbf{H} . The magnetization disappears when the field is removed.

Consider N such magnetic dipoles, each with magnetic moment $\boldsymbol{\mu}$. We assume that the dipoles are not interacting with one another, and that they are localized and classical in the sense that they are distinguishable and freely orientable. The internal energy of the N dipole system is given by its potential energy due to the external field \mathbf{H}

$$E = \sum_{i=1}^N E_i = - \sum_{i=1}^N \boldsymbol{\mu}_i \cdot \mathbf{H} = - \sum_{i=1}^N \mu H \cos \theta_i.$$

Again, due to the fact that the dipoles do not interact, we can write

$$\begin{aligned} Z(\beta) &= [Z_1(\beta)]^N, \\ Z_1(\beta) &= \sum_{\theta} e^{\beta\mu H \cos \theta}. \end{aligned}$$

The dipoles will obviously align themselves along the field. The magnitude of the magnetization will be

$$M_z = N \langle \mu \cos \theta \rangle = \frac{N}{\beta} \frac{\partial}{\partial H} \ln Z_1(\beta) = - \left(\frac{\partial A}{\partial H} \right)_T. \quad (65)$$

By denoting with $d\Omega = \sin \theta d\theta d\phi$ the elemental solid angle representing the orientation of the dipole, we obtain

$$Z_1(\beta) = \int_0^{2\pi} \int_0^{\pi} e^{\beta\mu H \cos \theta} \sin \theta d\theta d\phi = 4\pi \frac{\sinh(\beta\mu H)}{\beta\mu H}.$$

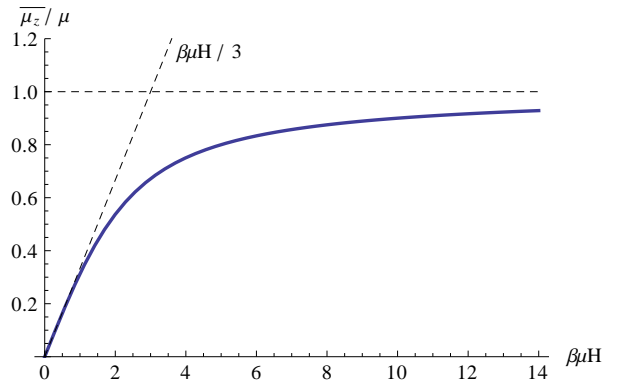
We have, thus, obtained the magnetization per dipole

$$\bar{\mu}_z = \frac{M_z}{N} = \mu L(\beta\mu H),$$

where

$$L(x) = \coth x - \frac{1}{x}$$

is the *Langevin function*.



We note that the parameter $\beta\mu H$ gives the strength of the field in terms of thermal energy. When $\mu H \gg k_B T$, we have magnetic saturation

$$M_z \simeq N\mu.$$

In the large temperature limit, $\mu H \ll k_B T$ and

$$M_z \simeq \frac{N\mu^2}{3k_B T} H \equiv \frac{C}{T} H,$$

which is the *Curie law* of paramagnetism and C is the Curie constant.

Quantum mechanics limits the number of possible orientations of the dipoles. Generally, the magnetic moment and the total angular momentum \mathbf{l} of the dipole are related by

$$\boldsymbol{\mu} = \left(g \frac{e}{2m}\right) \mathbf{l},$$

where $g(e/2mc)$ is the *gyromagnetic ratio*, g is *Lande's g-factor*,

$$l^2 = J(J+1)\hbar^2, \quad J = \frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots \text{ or } 0, 1, 2, \dots$$

and

$$g = \frac{3}{2} + \frac{S(S+1) - L(L+1)}{2J(J+1)}$$

where S and L are the spin and orbital quantum numbers of the dipole. We obtain

$$\mu^2 = g^2 \mu_B^2 J(J+1)$$

and

$$\mu_z = g\mu_B m, \quad m = -J, -J+1, \dots, J-1, J$$

where $\mu_B = e\hbar/2m$ is the *Bohr magneton*.

Now, the single-dipole partition function is

$$\begin{aligned} Z_1(\beta) &= \sum_{m=-J}^J e^{\beta g\mu_B m H} \\ &= \sinh \left[\left(1 + \frac{1}{2J}\right) \beta g\mu_B m H \right] / \sinh \left[\frac{1}{2J} \beta g\mu_B m H \right]. \end{aligned}$$

The mean magnetic moment per dipole is

$$\bar{\mu}_z = \frac{M_z}{N} = \frac{1}{\beta} \frac{\partial}{\partial H} \ln Z_1(\beta) = (g\mu_B J) B_J(\beta g\mu_B m H),$$

where

$$B_J(x) = \left(1 + \frac{1}{2J}\right) \coth \left[\left(1 + \frac{1}{2J}\right) x \right] - \frac{1}{2J} \coth \left[\frac{1}{2J} x \right]$$

is the *Brillouin function* of order J . Similar to classical case, we have saturation for strong fields. In the weak field limit, we also obtain the Curie law

$$\bar{\mu}_z \simeq \frac{g^2 \mu_B^2 J(J+1)}{3k_B T} H \equiv \frac{C_J}{T} H, \quad (66)$$

where C_J is the Curie constant.

It should be noted that when $J \rightarrow \infty$ the number of possible orientations of the dipoles become infinitely large and the problem reduces to its classical counterpart.

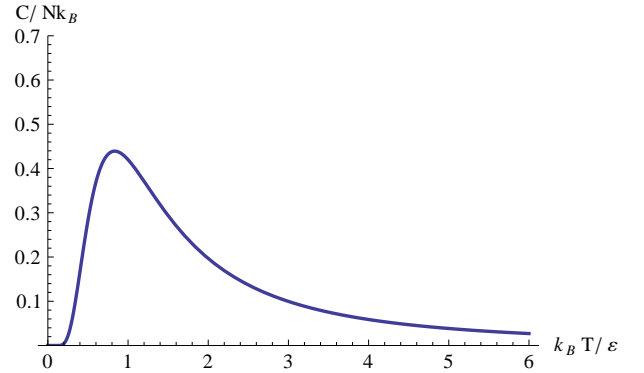
Peculiarities of spin systems

Consider system with N non-interacting spin- $\frac{1}{2}$ ($J = \frac{1}{2}$) dipoles. Such a dipole has only two possible orientations,

and the corresponding energies are $\pm\varepsilon = \pm\mu_B H$. Thus,

$$\begin{aligned} Z(\beta) &= [2 \cosh(\beta\varepsilon)]^N, \\ A &= -Nk_B T \ln[2 \cosh(\beta\varepsilon)], \\ S &= Nk_B \left[\ln \left\{ 2 \cosh \left(\frac{\varepsilon}{k_B T} \right) \right\} - \frac{\varepsilon}{k_B T} \tanh \left(\frac{\varepsilon}{k_B T} \right) \right], \\ E &= -N\varepsilon \tanh \left(\frac{\varepsilon}{k_B T} \right), \\ M &= N\mu_B \tanh \left(\frac{\varepsilon}{k_B T} \right), \\ C &= \left(\frac{\partial E}{\partial T} \right)_H = Nk_B \left(\frac{\Delta}{k_B T} \right)^2 e^{\Delta/k_B T} (1 + e^{\Delta/k_B T})^{-2}, \end{aligned}$$

where $\Delta = \varepsilon - (-\varepsilon) = 2\mu_B H$ is the difference between the two allowed dipole energies. Let us analyse these a bit. First, we found the natural result $E = -MH$. In the limit $T \rightarrow 0$, $S \rightarrow 0$ and $E \rightarrow -N\varepsilon$. This is a result of complete ordering of the dipoles along the field. At high temperatures, the dipoles are randomly oriented and the energy and magnetization vanish.



The specific heat displays the *Schottky anomaly*, i.e. a peak as function of temperature. Generally, the Schottky anomaly occurs in systems with limited number of energy levels.

Interestingly,

$$\frac{1}{T} = -\frac{k_B}{\varepsilon} \tanh^{-1} \left(\frac{E}{N\varepsilon} \right),$$

which implies that $T < 0$ when $E > 0$ ²⁵. Positive values of energy are abnormal, since they correspond to a situation where the magnetization is in *opposite* direction to the applied field. Yet, it has been realized in experiments with nuclear spins in a crystal (with LiF-crystal by Purcell and Pound in 1951). A sudden reversal of a strong magnetic field will result in negative magnetization and thus in negative temperature, assumed that the relaxation time for the mutual interaction among the spins is small compared with that for the spins and the lattice.

We showed before that the heat will flow from the system with lower value of β to the one with a higher value. Interestingly, one can show that this holds also for negative temperatures. In a sense then, systems at negative temperatures are *hotter* than those at positive temperatures!

²⁵This should be compared with the exercise 3.1!

5.5 Grand Canonical Ensemble

In addition to energy, also the number N of particles in a statistical system is hardly ever measured. Thus, the next natural step is to regard both N and E as variables and allow them to fluctuate.

Now, one can proceed in the same way as in the case of canonical ensemble. One can either consider a small system A immersed in the large reservoir A' with which it can exchange energy and particles, or, maximize the entropy with the constraints

$$\begin{aligned}\langle E \rangle &= \text{constant} \\ \langle N \rangle &= \text{constant} \\ \sum_r P_r &= 1.\end{aligned}$$

In both cases, one ends up with the probability

$$P_{r,N} = \frac{\exp(-\beta(E_r - \mu N))}{\sum_{r,N} \exp(-\beta(E_r - \mu N))} \quad (67)$$

of observing the system with energy E_r and N particles. This is the *grand canonical distribution*. Similar to the canonical distribution, the studied system can be any relatively small macroscopic system or a distinguishable microscopic system. Moreover, the *grand canonical ensemble* consists of mental copies of system A obeying the grand canonical distribution in the particle-energy bath.

We can write the *grand canonical partition function*.

$$Z_G = \sum_{r,N} e^{-\beta(E_r - \mu N)}$$

Number of particles and energy

Now

$$\begin{aligned}\langle N \rangle &= \frac{\sum_{r,N} N e^{-\beta(E_r - \mu N)}}{Z_G} \\ &= \frac{1}{\beta} \frac{\partial \ln Z_G}{\partial \mu}\end{aligned} \quad (68)$$

and

$$\begin{aligned}\langle E \rangle &= \frac{\sum_{r,N} E_r e^{-\beta(E_r - \mu N)}}{Z_G} \\ &= k_B T^2 \frac{\partial \ln Z_G}{\partial T} + k_B T \mu \frac{\partial \ln Z_G}{\partial \mu}.\end{aligned}$$

Entropy

According to the definition, the Gibbs entropy is

$$S = -k_B \sum_{r,N} P_{r,N} \ln P_{r,N}.$$

Now

$$\ln P_{r,N} = -\beta E_r + \beta \mu N - \ln Z_G,$$

so

$$S = \frac{\langle E \rangle}{T} - \mu \frac{\langle N \rangle}{T} + k_B \ln Z_G.$$

Grand potential

In thermodynamics we defined the

$$\Omega = E - TS - \mu N,$$

so in the grand canonical ensemble the *grand potential* is

$$\Omega = -k_B T \ln Z_G.$$

With the help of this the grand canonical distribution can be written as

$$P_{r,N} = e^{\beta(\Omega - E_r + \mu N)}.$$

Notify The grand canonical state sum depends on the variables T , V and μ , i.e.

$$Z_G = Z_G(T, V, \mu).$$

Density and energy fluctuations

Let us evaluate the fluctuations in the number of particles around the mean

$$\sigma_N^2 = \langle N^2 \rangle - \langle N \rangle^2 = \frac{1}{\beta} \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V}.$$

We obtain for the particle density $n = \langle N \rangle / V$

$$\begin{aligned}\frac{\sigma_n^2}{\langle n \rangle^2} &= \frac{\sigma_N^2}{\langle N \rangle^2} = \frac{k_B T}{\langle N \rangle^2} \left(\frac{\partial \langle N \rangle}{\partial \mu} \right)_{T,V} = -\frac{k_B T}{V} \left(\frac{\partial v}{\partial \mu} \right)_T \\ &\propto \frac{1}{\sqrt{\langle N \rangle}}\end{aligned}$$

where $v = 1/n$.

It is straightforward to show that $G = \mu N$, implying that the chemical potential is just another name for the amount free Gibbs energy per particle. Thus, one obtains the *Gibbs-Duhem equation*

$$d\mu = v dP - \frac{S}{\langle N \rangle} dT. \quad (69)$$

By assuming that T is constant, we obtain

$$\frac{\sigma_n^2}{\langle n \rangle^2} = \frac{k_B T}{V} \kappa_T,$$

where $\kappa_T = -(1/V)(\partial V / \partial P)_T$ is the isothermal compressibility.

Typically, the relative fluctuations of the particle go as $1/\sqrt{\langle N \rangle}$. There are, however, exceptions. For example, in *phase transitions*²⁶ the compressibility can become excessively large. Especially, at critical points²⁷ the compressibility diverges leading to large fluctuations in the particle

²⁶We will discuss more on phase transitions in the following section, and also later in the course.

²⁷See definition later.

density. In this kind of situations the grand canonical ensemble can lead to different result than the corresponding canonical ensemble, and should be preferred due to the correct physical picture it gives.

We can repeat the calculation for the variance of the energy. It gives

$$\sigma_E^2 = k_B T^2 C_V + \left[\left(\frac{\partial \langle E \rangle}{\partial N} \right)_{T,V} \right]^2 \sigma_N^2.$$

This means that the fluctuations in energy of the grand canonical ensemble is equal to that of the canonical ensemble plus a contribution arising from the fluctuating N . These fluctuations are typically negligible, but can also become large in phase transitions due to the second term in the formula.

5.6 Phase Equilibrium and Phase Transitions

A big part of the success of statistical physics and thermodynamics has been gained in the study of phase transitions. Thermodynamic phases are such regions in the *phase diagram* where the thermodynamic functions are analytic. *Phase transitions* on the other hand arise at point, lines or surfaces of the diagram where such properties are non-analytic. The typical phases of matter are

- *vapor phase* is a low density gas obeying ideal-gas equation of state, and possible corrections described by the virial expansion (such as van der Waals equation),
- *liquid phase* is dense fluid with strong interactions between the atoms,
- *solid phase* is static structure where the constituent atoms are bound together either in regularly or irregularly.
- *plasma phase* is similar to gas in which certain portion of the atoms is ionized. No definite shape or volume, but unlike gas, can form structures in magnetic fields.

In addition to these, there exists exotic phases in low temperatures that arise from the corrections in statistics introduced by quantum mechanics. Such phases include the *superfluid* and *supersolid* phases in helium, and the *superconducting* phase that is observed in numerous compounds.

Phase equilibrium

Consider two phases A and B contained in a cylinder at constant pressure P and temperature T . Assume that the total number of particles is $N = N_A + N_B$. If the two phases do not *coexist* at these P and T , the numbers N_A and N_B will change in order to approach equilibrium. This leads to the change in Gibbs energy $G = G_A(N_A, P, T) +$

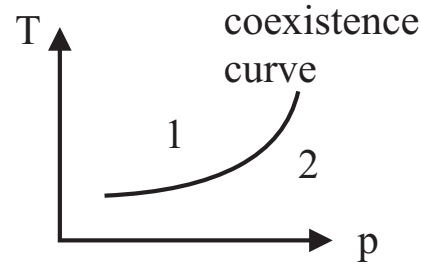
$$G_B(N_B, P, T)$$

$$dG = \left(\frac{\partial G_A}{\partial N_A} \right)_{T,P} dN_A + \left(\frac{\partial G_B}{\partial N_B} \right)_{T,P} dN_B = (\mu_A - \mu_B) dN_A.$$

Because Gibbs energy is minimized at the equilibrium, we have that if $\mu_A > \mu_B$ the number of molecules in phase B will increase and that of A decrease. Similarly, if $\mu_A < \mu_B$ the N_A will increase and N_B decrease. When

$$\mu_A(P, T) = \mu_B(P, T)$$

the Gibbs free energy is independent on N_A and N_B and the phases are said to *coexist*. This equilibrium condition draws a *coexistence curve* in the (p, T) -plane. If three phases are simultaneously in equilibrium, the crossing of three coexistence curves is called the *triple point*. The point where the coexistence curve ends is called *critical point*.



Now, let us define $p_\sigma(T)$ on the coexistence curve in the (p, T) -plane, determining the phase boundary between the phases A and B . Accordingly,

$$\left(\frac{\partial \mu_A}{\partial T} \right)_p + \left(\frac{\partial \mu_A}{\partial p} \right)_T \frac{dp_\sigma}{dT} = \left(\frac{\partial \mu_B}{\partial T} \right)_p + \left(\frac{\partial \mu_B}{\partial p} \right)_T \frac{dp_\sigma}{dT}.$$

According to the Gibbs-Duhem equation (69), we have

$$\frac{S}{N} = - \left(\frac{\partial \mu}{\partial T} \right)_p, \\ \frac{V}{N} = \left(\frac{\partial \mu}{\partial p} \right)_T.$$

We obtain the *Clausius-Clapeyron equation*

$$\frac{dp_\sigma}{dT} = \frac{S_B - S_A}{V_B - V_A} = \frac{\Delta S}{\Delta V} = \frac{\Delta H}{T \Delta V}, \quad (70)$$

where

$$\Delta Q = T \Delta S = \Delta E + p \Delta V = \Delta H$$

is the *phase transition heat* or *latent heat* released in a process where the coexistence curve is crossed at constant T and p .

Phase transitions

Generally, when the free energy is an analytic function only single phase exists. This is denoted in the phase diagrams with the open spaces between the coexistence curves. When the parameters (in our case T or p) are changed so,

that the system reaches the coexistence curve, the free energy becomes non-analytic and a phase transition occurs. In a phase transition the chemical potential

$$\mu = \left(\frac{\partial G}{\partial N} \right)_{p,T}$$

is continuous. Instead

$$S = - \left(\frac{\partial G}{\partial T} \right)_p$$

and

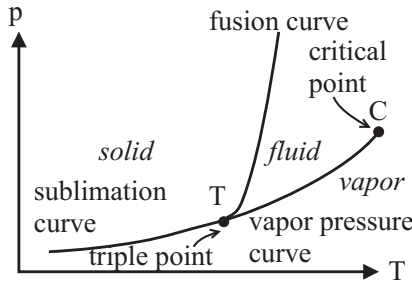
$$V = \left(\frac{\partial G}{\partial p} \right)_T$$

are not necessarily continuous.

A transition is of *first order*, if the first order derivatives (S, V) of G are discontinuous and of *second order*, if the second order derivatives are discontinuous.

Examples

a) Vapor pressure curve



We consider the transition

liquid \rightarrow vapor.

Supposing that we are dealing with ideal gas we have

$$\Delta V = V_v = \frac{Nk_B T}{p},$$

since

$$V_{l(\text{iquid})} \ll V_{v(\text{apor})}.$$

Because the vaporization heat (the phase transition heat) ΔH_{lv} is roughly constant on the vapor pressure curve we have

$$\frac{dp}{dT} = \frac{\Delta H_{lv} p}{Nk_B T^2}.$$

Integration gives us

$$p = p_0 e^{-\Delta H_{lv}/Nk_B T}.$$

b) Fusion curve

Now

$$\Delta V_{ls} = V_{l(\text{iquid})} - V_{s(\text{olid})}$$

can be positive or negative (for example H_2O).

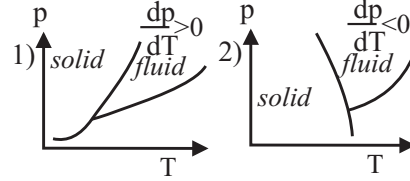
According to the Clausius-Clapeyron equation

$$\frac{dp}{dT} = \frac{\Delta H_{ls}}{T \Delta V_{ls}}$$

we have

$$\frac{dp}{dT} > 0, \quad \text{if } \Delta V_{ls} > 0 \quad 1)$$

$$\frac{dp}{dT} < 0, \quad \text{if } \Delta V_{ls} < 0 \quad 2) \quad .$$



We see that when the pressure is increased in constant temperature the system

- 1) drifts "deeper" into the solid phase,
- 2) can go from the solid phase to the liquid phase.

c) Sublimation curve

Now

$$dH = T dS + V dp = C_p dT + V(1 - T\alpha_p) dp,$$

since $S = S(p, T)$ and using Maxwell relations and definitions of thermodynamic response functions

$$dS = \left(\frac{\partial S}{\partial p} \right)_T dp + \left(\frac{\partial S}{\partial T} \right)_p dT = - \left(\frac{\partial V}{\partial T} \right)_p dp + \frac{C_p}{T} dT.$$

The vapor pressure is small so $dp \approx 0$, and

$$\begin{aligned} H_s &= H_s^0 + \int_0^T C_p^s dT && \text{solid phase} \\ H_v &= H_v^0 + \int_0^T C_p^v dT && \text{vapor (gas).} \end{aligned}$$

Let us suppose that the vapor satisfies the ideal gas state equation. Then

$$\Delta V_{vs} = \frac{Nk_B T}{p} - V_s \approx \frac{Nk_B T}{p},$$

so

$$\frac{dp}{dT} = \frac{\Delta H_{vs}}{T \Delta V_{vs}} \approx \frac{p \Delta H_{vs}}{Nk_B T^2},$$

where $\Delta H_{vs} = H_s - H_v$.

For a monatomic ideal gas $C_p = \frac{5}{2} k_B N$, so that

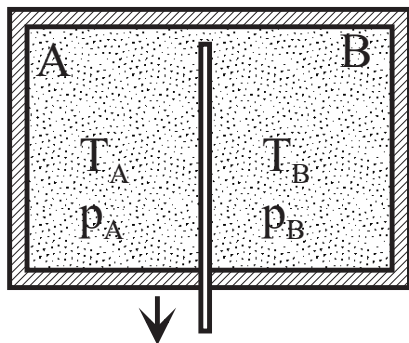
$$\ln p = -\frac{\Delta H_{vs}^0}{Nk_B T} + \frac{5}{2} \ln T - \frac{1}{k_B N} \int_0^T \frac{C_p^s dT'}{T'^2} + \text{constant}.$$

Here ΔH_{vs}^0 is the sublimation heat at 0 temperature and pressure.

5.7 Gibbs Paradox and Correct Enumeration of Microstates

Up to this point, our discussions on enumeration of the accessible microstates has been classical. According to classical physics, each microstate corresponds to a point in a phase space, spanned by the coordinates and momenta of the constituent particles of the studied system. Now, classical physics consider elementary particles as *distinguishable*, meaning that an pairwise exchange of phase-space coordinates of two particles leads to a new microstate. However, this line of thinking results in overcounting of the accessible states, which is traditionally demonstrated by the so-called *Gibbs paradox* related to the entropy of mixing of two ideal gases.

Entropy of mixing



Consider two ideal gases (A and B) confined compartments separated by a valve in volume V . Suppose that initially $p_A = p_B = p$ and $T_A = T_B = T$. If we open the valve, both gas expand adiabatically ($\Delta Q = 0$) to the whole volume.

In a mixture of ideal gases every component satisfies the state equation

$$p_j V = N_j k_B T.$$

The concentration of the component j is

$$x_j = \frac{N_j}{N} = \frac{p_j}{p},$$

where the total pressure p is

$$p = \sum_j p_j.$$

Each constituent gas expands in turn into the volume V . Since $p_A = p_B$ and $T_A = T_B$, we have $V_j = V x_j$. The change in the entropy is (exercise)

$$\Delta S = \sum_j N_j k_B \ln \frac{V}{V_j}$$

or

$$\Delta S_{\text{mix}} = -N k_B \sum_j x_j \ln x_j. \quad (71)$$

Now $\Delta S_{\text{mix}} \geq 0$, since $0 \leq x_j \leq 1$, and the equality holds only when either $N_A = 0$ or $N_B = 0$.

The problem arises when the two gases are the same. Equation for the entropy of mixing (71) still holds, implying that the entropy will increase after the opening of the valve. However, since the two gases are the same, the process should be reversible! This is the *Gibbs paradox*.

When the numbers of particles are large, we can use Stirling's formula and write

$$\begin{aligned} \Delta S_{\text{mix}} &= S_{A+B} - S_A - S_B \\ &= -k_B (N_A \ln N_A - N_A \ln N + N_B \ln N_B - N_B \ln N) \\ &\approx k_B (\ln N! - \ln N_A! - \ln N_B!). \end{aligned}$$

Thus, we see that the problem would be solved by reducing the entropies by a factor $k_B \ln N_{(A,B)}!$.

This kind of a factor appears to the entropy when the particles constituting the system are considered *indistinguishable*²⁸. By correcting the counting of the states in this way, one obtains results that agree with the reality, at least at the *classical limit*. We will later show, independently, that this kind of reduction has to be made to the number of accessible microstates in order to have classical statistical physics as the true limit of quantum statistical physics.

²⁸In fact, this was done already in Eq. (56).

6. Quantum Statistics

The ensemble theory developed in the preceding chapter is extremely general. However when applied to quantum-systems composed of *indistinguishable* entities, one has to proceed with great care. One benefits in writing the ensemble theory in a manner that is more suitable to quantum mechanical treatment, i.e. in the language of *operators*.

6.1 Density Operator

Let us consider a quantum system characterized by the *Hamiltonian operator* \hat{H} . Recall from quantum mechanics, that the (normalized) *state vector* $|\psi; t\rangle$ gives the maximal description of the system. Quantum states that can be represented in the above manner as vectors of a *Hilbert space*, are called *pure states*. The time evolution of the state vector is given by the Schrödinger equation

$$\hat{H}|\psi; t\rangle = i\hbar \frac{\partial}{\partial t} |\psi; t\rangle.$$

Then, let $\{|n\rangle\}$ be a complete set of orthonormal vectors, i.e. the *basis*, in the Hilbert space under consideration. Every vector can be written as a linear superposition of the basis vectors, especially

$$|\psi; t\rangle = \sum_n a_n(t) |n\rangle,$$

where

$$a_n(t) = \langle n | \psi; t \rangle.$$

Thus, the state of the system can be described equally well with the coefficients $a_n(t)$

$$i\hbar \dot{a}_n(t) = \sum_m a_m(t) \langle n | \hat{H} | m \rangle = \sum_m a_m(t) H_{nm},$$

where $H_{nm} = \langle n | \hat{H} | m \rangle$ are the *matrix elements* of \hat{H} . The coefficients $a_n(t)$ have the standard interpretation as the *probability amplitudes* of the system to be in states $|n\rangle$. Since we assume that the state is normalized, we have that

$$\sum_n |a_n(t)|^2 = 1.$$

Each *measurable* quantity α of the system has a corresponding *Hermitian operator* \hat{A} in the Hilbert space \mathcal{H} of the system. The possible outcomes of a measurement of the property α are the (real) eigenvalues of \hat{A} . The *mean* outcome is given by the expectation value

$$\langle \hat{A} \rangle \equiv \langle \psi; t | \hat{A} | \psi; t \rangle.$$

A description of the system, equivalent to the one given by the state vector, can be obtained by defining the so-called *density operator*

$$\begin{aligned} \hat{\rho}(t) &= |\psi; t\rangle \langle \psi; t| = \sum_{m,n} a_m(t) a_n^*(t) |m\rangle \langle n| \\ &= \sum_{mn} \rho_{mn}(t) |m\rangle \langle n|, \end{aligned} \quad (72)$$

where the matrix elements $\langle m | \hat{\rho} | n \rangle \equiv \rho_{mn}(t) = a_m(t) a_n^*(t)$. One can show that all of the *measurable* information in the state vector is contained in the density operator. For example, the expectation value can be written as²⁹

$$\langle \hat{A} \rangle = \text{Tr}(\hat{A} \hat{\rho}) \quad (73)$$

and the Schrödinger equation as

$$\dot{\hat{\rho}} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}(t)]. \quad (74)$$

The latter equation describing the time-evolution of the density operator is called *von Neumann equation*. It should be noted that von Neumann equation follows from the Liouville equation by the application of Dirac's correspondence principle, i.e.

$$\{H, \rho\} \rightarrow -\frac{i}{\hbar} [\hat{H}, \hat{\rho}].$$

The above discussion holds an implicit assumption that the system is isolated, and that we have a complete knowledge of the preparation of the quantum state $|\psi; t\rangle$. In practice however, we have to settle for less. We can only restrict the set of accessible states into a set, *ensemble*, $\{|\psi^k; t\rangle | k = 1, \dots, \mathcal{N}\}$, whose members do not have to be mutually orthogonal. Then, what we can say is that the state k is *prepared* with probability $P(k)$. Accordingly, we generalize the definition of the density operator into

$$\begin{aligned} \hat{\rho}(t) &= \sum_k P(k) |\psi^k; t\rangle \langle \psi^k; t| = \sum_k P(k) \left[\sum_{mn} \rho_{mn}^k(t) |m\rangle \langle n| \right] \\ &\equiv \sum_{mn} \rho_{mn}(t) |m\rangle \langle n|. \end{aligned}$$

We see that the matrix elements of the density operator are ensemble averages of the corresponding elements of the states in the ensemble. Especially, the diagonal elements

$$\rho_{nn}(t) = \sum_k P(k) \rho_{nn}^k(t) = \sum_k P(k) |a_n(t)|^2$$

give the probabilities of finding a system in state n in the ensemble. Notice the two type of averaging, one quantum mechanical due to the probability amplitudes, and one statistical.

This describes a *mixed state* where we have an incomplete knowledge on the state of the system. The density operator presented in this form contains both the statistical and quantum mechanical information about the system. It should be noted that Eqs. (73) and (74) hold also for the general definition.

Entanglement with the environment

One important aspect, that will be discussed more rigorously later when we study non-equilibrium systems, is

²⁹It should be noted that trace operation is independent on the choice of the basis. Also, $\text{Tr} \hat{\rho} = 1$ which follows from the normalization of the state vectors.

that the studied systems in general interact with their surroundings, one example being a system immersed to a heat bath. In such cases, the Hamiltonian of the whole system (system+environment+interactions) is generally of form³⁰

$$\hat{H}_T = \hat{H} + \hat{H}_B + \hat{H}_{\text{int}}.$$

Only in the cases where $\hat{H}_{\text{int}} = 0$, it is possible to write down the state vector of the bare system. In the presence of interactions, the system and its environment become *entangled*, meaning that the state vector cannot be written as a tensor product of a system part and a bath part, i.e. $|\Psi; t\rangle \neq |\psi; t\rangle \otimes |\psi_B; t\rangle$. Therefore, it is reasonable to speak only about the state vector of the combined system³¹

$$|\Psi; t\rangle = \sum_{nB} A_{nB}(t) |n\rangle |B\rangle.$$

The Schrödinger equation gives, in principle, the time-evolution of the state vector of the combined system. But even if could produce a solution³², it would give an unnecessarily large amount of information about the bath. This excess information is naturally interesting in its own right, but at the moment we are only interested in the behaviour of the system. Therefore, it is practical and convenient to "trace out" the bath and hide its contributions into the probabilities $P(k)$. This can be seen by considering the expectation value of the operator \hat{A} in systems Hilbert space

$$\begin{aligned} \langle \hat{A} \rangle &= \langle \Psi; t | \hat{A} | \Psi; t \rangle \\ &= \sum_{nn'} \sum_{BB'} \langle \Psi; t | nB \rangle \langle nB | \hat{A} | n'B' \rangle \langle n'B' | \Psi; t \rangle \\ &\equiv \text{Tr}_{\text{sys}}(\hat{A} \hat{\rho}_{\text{sys}}), \end{aligned}$$

where

$$\hat{\rho}_{\text{sys}} = \sum_B \sum_{nn'} |n'\rangle \langle n'B | \Psi; t \rangle \langle \Psi; t | nB \rangle \langle n|.$$

is the *reduced density operator* of the system. Clearly,

$$\hat{\rho}_{\text{sys}} = \text{Tr}_B(|\Psi; t\rangle \langle \Psi; t|). \quad (75)$$

We thus see that the density operator of the system is written as a *partial trace* of that of the combined system, taken over the environment variables only.

6.2 Equilibrium Density Operator

The density operator is an extremely powerful tool in the study of both equilibrium and non-equilibrium quantum systems, both closed and open. The non-equilibrium cases

³⁰It should be noted that the environment can consists of a large set of quantum systems, or even of a single quantum system.

³¹Formally, we have to combine the Hilbert spaces \mathcal{H}_A and \mathcal{H}_B of the interacting quantum systems with the tensor product $\mathcal{H}_A \otimes \mathcal{H}_B$. Accordingly in our case, $\hat{H} \rightarrow \hat{H} \otimes \hat{I}_B$, $\hat{H}_B \rightarrow \hat{I} \otimes \hat{H}_B$, $|n\rangle |B\rangle \equiv |n\rangle \otimes |B\rangle$, where $|B\rangle$ are the basis vectors of \mathcal{H}_B and \hat{I} (\hat{I}_B) is the identity operator of \mathcal{H} (\mathcal{H}_B).

³²This is a formidable task, since the bath contains in general a macroscopic amount of degrees of freedom!

will be studied later. In the case of an equilibrium system, the density operator should be static. This implies that

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\hat{H}, \hat{\rho}] = 0.$$

Thus, the density operator commutes with the Hamiltonian, which means that they can be diagonalized simultaneously. Let $\{|n\rangle\}$ be the eigenbasis of the Hamiltonian. In this basis, the density operator is diagonal, i.e.

$$\hat{\rho} = \sum_n \rho_n |n\rangle \langle n|.$$

Entropy

Recall that the Gibbs entropy was defined as

$$S = -k_B \sum_n P_n \ln P_n,$$

where P_r is the probability of finding a system in the state r in the ensemble. Based on the above discussion, $P_n = \rho_n$ and we can write

$$\begin{aligned} S &= -k_B \sum_n \rho_n \ln \rho_n \\ &= -k_B \sum_n \langle n | \hat{\rho} \ln \hat{\rho} | n \rangle \\ &= -k_B \text{Tr}(\hat{\rho} \ln \hat{\rho}). \end{aligned}$$

The last equality is the general definition of entropy in quantum mechanics. It should be noted that it is independent on the choice of basis (since the trace is). Also, the definition bears great generality as it holds also in non-equilibrium systems.

Microcanonical ensemble

Recall, that in microcanonical ensemble the constituting mental copies of a system are characterized by fixed N in a fixed volume V . The energy of the system is lying in the interval $(E, E + \delta E)$, where $\delta E \ll E$. The number of accessible microstates is denoted with $\Omega(E, V, N; \delta E)$, and each microstate occurs with equal probability (in accordance of the postulate of equal a priori probabilities). Thus, the matrix elements of the density operator in the energy eigenbasis are of form

$$\rho_n = \begin{cases} 1/\Omega & \text{for each accessible state} \\ 0 & \text{for all other states} \end{cases}$$

We see immediately that $\text{Tr} \hat{\rho} = 1$, as it should for normalized states. We see that the entropy in the microcanonical ensemble is given by the familiar formula

$$S = k_B \ln \Omega.$$

We first emphasize that Ω is calculated quantum mechanically, so that the indistinguishability of the particles has been taken into account right from the start. This means that we should avoid paradoxes, such as the Gibbs'. Also,

if the number of quantum states within δE is $\Omega = 1$ we obtain $S = 0$, i.e. the Nernst theorem (third law of thermodynamics).

Moreover, the case $\Omega = 1$ corresponds to the pure state, in which case the construction of the ensemble is quite superfluous because every system in the ensemble would be in the same state.

Canonical ensemble

Consider then the canonical ensemble where the macrostate is characterized by constant parameters N , V and T , and the energy E is a variable quantity. The density matrix in energy basis is again diagonal, and

$$\rho_n = \frac{e^{-\beta E_n}}{Z}, \quad n = 0, 1, 2, \dots$$

where

$$Z = \sum_n e^{-\beta E_n}$$

is the partition function. This means that the density operator can be written as

$$\hat{\rho} = \frac{e^{-\beta \hat{H}}}{\text{Tr}(e^{-\beta \hat{H}})}, \quad (76)$$

where

$$e^{-\beta \hat{H}} = \sum_{j=0}^{\infty} (-1)^j \frac{(\beta \hat{H})^j}{j!}.$$

The expectation values are thus of form

$$\langle \hat{A} \rangle = \frac{\text{Tr}(\hat{A} e^{-\beta \hat{H}})}{\text{Tr}(e^{-\beta \hat{H}})}.$$

Grand canonical ensemble

Similarly, in the case of grand canonical ensemble, where both E and particle number N are variables, we can write

$$\hat{\rho} = \frac{e^{-\beta(\hat{H} - \mu \hat{N})}}{Z_G}, \quad (77)$$

where \hat{N} is the *particle number operator* and

$$Z_G = \sum_{r, N} = e^{-\beta(E_r - \mu N)} = \text{Tr}(e^{-\beta(\hat{H} - \mu \hat{N})})$$

is the grand partition function. Again, the ensemble average

$$\langle \hat{A} \rangle = \frac{\text{Tr}(\hat{A} e^{-\beta(\hat{H} - \mu \hat{N})})}{Z_G}.$$

6.3 Systems of Indistinguishable Particles

Let us then turn into quantum mechanical description of a system of N identical particles. For simplicity, we will deal first systems where the particles are *non-interacting*. The Hilbert space of the system is a tensor product of single-particle Hilbert spaces \mathcal{H}_1

$$\mathcal{H} = \underbrace{\mathcal{H}_1 \otimes \mathcal{H}_1 \otimes \dots \otimes \mathcal{H}_1}_{N \text{ copies}}.$$

If $|\mathbf{q}_i\rangle$ are the position eigenstates in the single-particle Hilbert spaces, one can write a general state of the N -particle system as

$$|\Psi\rangle = \int \dots \int d\mathbf{q}_1 \dots d\mathbf{q}_N |\mathbf{q}_1 \dots \mathbf{q}_N\rangle \Psi(\mathbf{q}_1, \dots, \mathbf{q}_N),$$

where $|\mathbf{q}_1 \dots \mathbf{q}_N\rangle = |\mathbf{q}_1\rangle \otimes \dots \otimes |\mathbf{q}_N\rangle$, and $\Psi(\mathbf{q}_1, \dots, \mathbf{q}_N) = \langle \mathbf{q}_1 \dots \mathbf{q}_N | \Psi \rangle$ is the *wave function*. Now, because the particles do not interact, we can write the solution of the time-independent Schrödinger equation (we denote $\mathbf{q} = (\mathbf{q}_1, \dots, \mathbf{q}_N)$)

$$\hat{H} \Psi_E(\mathbf{q}) = E \Psi_E(\mathbf{q})$$

as

$$\Psi_E(\mathbf{q}) = \underbrace{\psi_{\varepsilon_1}(\mathbf{q}_1) \otimes \dots \otimes \psi_{\varepsilon_N}(\mathbf{q}_N)}_{N \text{ times}}$$

with

$$E = \sum_{i=1}^N \varepsilon_i.$$

The pairs $(\varepsilon_i, \psi_i(i))$ are the solutions of the single-particle Schrödinger equation

$$\hat{H}_i \psi_i(i) = \varepsilon_i \psi_i(i), \quad (78)$$

where we have denoted $\psi_i(i) = \psi_{\varepsilon_i}(\mathbf{q}_i)$.

We denote with n_i the number of particles in the state with energy ε_i . Thus,

$$N = \sum_i n_i$$

and

$$E = \sum_i n_i \varepsilon_i.$$

Now it should be noted that, according to quantum mechanics, even if two particles are in different single-particle states, an interchange between them should not lead into a new microstate. This means that as long the numbers n_i remain the same, any *permutation* P among the N particles should result into same microstate. In this way, the *indistinguishability* of the particles is guaranteed. We define the permutation P as the reordering of the coordinates as

$$P \Psi(1, 2, \dots, N) = \Psi(P1, P2, \dots, PN).$$

For arbitrary P , we should have

$$|P \Psi|^2 = |\Psi|^2. \quad (79)$$

The simplest way to form a wave function Ψ that obeys the above condition is make a linear combination of all possible $N!$ permutations of the wave function Ψ_E .

There are two possibilities with Eq. (79) fulfilled:

$$P\Psi = \Psi \text{ for all } P$$

meaning that the wave function is *symmetric*, or

$$P\Psi = \begin{cases} +\Psi & \text{if } P \text{ is an even permutation,} \\ -\Psi & \text{if } P \text{ is an odd permutation} \end{cases}$$

which means that the wave function is *antisymmetric*. One can show that every permutation of coordinates can be decomposed into successive pairwise exchanges called *transpositions*. If the number of transpositions forming the permutation is even (odd) the permutation is called even (odd). The *sign* of the permutation

$$\sigma(P) = \begin{cases} +1 & \text{if } P \text{ is an even permutation,} \\ -1 & \text{if } P \text{ is an odd permutation} \end{cases}$$

The symmetric wave functions are of form

$$\Psi_S(\mathbf{q}) = \frac{1}{\sqrt{N!}} \sum_P P\Psi_E(\mathbf{q}). \quad (80)$$

The antisymmetric wave functions can be written as

$$\begin{aligned} \Psi_A(\mathbf{q}) &= \frac{1}{\sqrt{N!}} \sum_P \sigma(P) P\Psi_E(\mathbf{q}) \\ &= \frac{1}{\sqrt{N!}} \begin{vmatrix} \psi_1(1) & \psi_1(2) & \cdots & \psi_1(N) \\ \psi_2(1) & \psi_2(2) & \cdots & \psi_2(N) \\ \vdots & \vdots & \cdots & \vdots \\ \psi_N(1) & \psi_N(2) & \cdots & \psi_N(N) \end{vmatrix} \end{aligned} \quad (81)$$

which is the so-called *Slater determinant*. One should notice that in the antisymmetric state, if we change two particles in the same single-particle state, we should obtain the same state. This means that $P\Psi_A = \Psi_A$. But, since exchange of two particles is an odd permutation, we should have $P\Psi_A = -\Psi_A$. Thus, it is impossible to have an antisymmetric wave function with two particles in the same single-particle state. This is a result equivalent to *Pauli's exclusion principle* for electrons.

The systems composed of particles obeying the exclusion principle are described by an antisymmetric wave function. The statistics governing the particles is called *Fermi-Dirac* statistics, and the constituent particles *fermions*. For symmetric wave functions we have no such problems. The particle numbers n_i are not restricted. The statistics governing such systems is *Bose-Einstein* statistics, and the constituent particles are *bosons*. It should be noted that the Hilbert space of a many particle system not the whole space spanned by the function Ψ_E but a subspace spanned by either symmetric (bosons) or antisymmetric (fermions) state vectors.

Even though the presentation has been done for non-interacting particles, one can show that the wave functions

have to be either symmetric or antisymmetric also in the systems of interacting particles.

Spin-Statistics Theorem

One can show by using relativistic quantum field theory, that there exists an intrinsic link between the spin of the particle and the statistics it obeys. Namely, particles with integral spin must have symmetric wave functions and are, therefore, bosons. Correspondingly, particles with half-odd integer spin have antisymmetric wave functions and are fermions. All particles in nature are either fermions or bosons. The fermions constitute the matter around us: quarks (up, down, strange, charm, bottom, top), leptons (electron, muon, tau, corresponding neutrinos and antiparticles), and their compositions with odd number of constituents. Bosons are responsible for interactions: photons, gluons, W and Z bosons, Higgs boson; but include also compositions of fermions with even number of constituents.

In fact, since the Hilbert spaces of quantum systems are in general complex vector spaces, one could in principle have for every transposition P_T

$$P_T\Psi = e^{i\theta}\Psi$$

where θ is a real variable, which is connected to the spin s of the particle by $\theta = 2\pi s$. One immediately realizes that

$$e^{2i\pi s} = (-1)^{2s} = \begin{cases} +1 & \text{for bosons} \\ -1 & \text{for fermions} \\ e^{2i\pi s} & \text{for anyons} \end{cases}$$

Thus, *anyons* are such particles that can acquire *any* phase when two of them are interchanged. For topological reasons, this can occur only in dimensions smaller than 3. When the spin of an anyon is a rational fraction, other than integer or half-integer, it is said to obey *fractional statistics*. Naturally, elementary particles cannot be anyons. Nevertheless, the *collective excitations*, i.e. *quasiparticles*, of interacting elementary particles restricted to two dimensions can display fractional statistics, and are essential ingredients e.g. in the theory of *fractional Hall effect*.

One interesting "hot topic" at the moment is the search of anyons obeying *non-abelian* statistics. This means that instead of a phase factor, the many particle state is changed by a matrix operation in a transposition, implying that two successive transpositions do not in general commute. One candidate for this are the so-called *Majorana fermions*, fermions that are their own anti-particles. Again, elementary particles are not found to be Majorana fermions, but they have been realized as (anyonic) quasiparticles in superconductors.

6.4 Non-Interacting Bosons and Fermions

We want to calculate the statistics for N non-interacting bosons or fermions, i.e. for *quantum ideal gas*. Recall, that the quantum state has n_i particles in the single-particle

energy state with energy ε_i . The eigenstate of the Hamiltonian of N non-interacting particles is then denoted with

$$|n_1, n_2, \dots, n_i, \dots\rangle \quad (82)$$

resulting in

$$\hat{H}|n_1, n_2, \dots, n_i, \dots\rangle = \left(\sum_i n_i \varepsilon_i \right) |n_1, n_2, \dots, n_i, \dots\rangle$$

with

$$N = \sum_i n_i \text{ and } E = \sum_i n_i \varepsilon_i.$$

It is straightforward to show that the basis $\{|n_1, n_2, \dots, n_i, \dots\rangle\}$ is complete

$$\sum_{\{n_i\}} |n_1, n_2, \dots, n_i, \dots\rangle \langle n_1, n_2, \dots, n_i, \dots| = \hat{I},$$

where \hat{I} is the identity operator, and orthonormal

$$\langle n'_1, n'_2, \dots | n_1, n_2, \dots \rangle = \delta_{n_1 n'_1} \delta_{n_2 n'_2} \dots$$

For such energy eigenstate, we can always define the *creation operator* \hat{a}_i^\dagger so that

$$\hat{a}_i^\dagger |n_1, n_2, \dots, n_i, \dots\rangle = C |n_1, n_2, \dots, n_i + 1, \dots\rangle,$$

i.e. \hat{a}_i^\dagger creates one particle into the state i . Correspondingly, the *annihilation operator* obeys

$$\hat{a}_i |n_1, n_2, \dots, n_i, \dots\rangle = C' |n_1, n_2, \dots, n_i - 1, \dots\rangle$$

meaning that \hat{a}_i removes one particle from the state i .

It is beneficial to use grand canonical ensemble, since we are dealing with both particle and energy exchange with the bath. Again, because the particles do not interact we can write the grand partition function Z_G as a product of single-particle partition functions Z_G^k as

$$Z_G = \prod_k Z_G^k.$$

Bosons

Creation and annihilation operators for bosons

For bosons, the creation and annihilation operators obey the commutation relations

$$\begin{aligned} [\hat{a}_i, \hat{a}_j^\dagger] &= \delta_{ij} \\ [\hat{a}_i, \hat{a}_j] &= [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0. \end{aligned}$$

One can show that (this has been done for harmonic oscillator creation and annihilation operators in quantum mechanics courses)

$$\begin{aligned} \hat{a}_i |n_1, n_2, \dots, n_i, \dots\rangle &= \sqrt{n_i} |n_1, n_2, \dots, n_i - 1, \dots\rangle \\ \hat{a}_i^\dagger |n_1, n_2, \dots, n_i, \dots\rangle &= \sqrt{n_i + 1} |n_1, n_2, \dots, n_i + 1, \dots\rangle \end{aligned}$$

Also, the (*occupation*) *number operator* is the familiar

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$$

and obeys the relation

$$\hat{n}_i |n_1, n_2, \dots, n_i, \dots\rangle = n_i |n_1, n_2, \dots, n_i, \dots\rangle,$$

with $n_i = 0, 1, \dots$

Bose-Einstein distribution

For bosons, all fillings n_k of the eigenstate ψ_k are allowed. Each particle in the state contributes with energy ε_k and chemical potential μ , so the single-particle grand partition function is

$$\begin{aligned} Z_G^k &= \text{Tre}^{-\beta(\hat{H}_k - \mu \hat{n}_k)} = \sum_{n_k=0}^{\infty} e^{-\beta(\varepsilon_k - \mu)n_k} \\ &= \frac{1}{1 - e^{-\beta(\varepsilon_k - \mu)}}, \end{aligned}$$

where the trace is taken over the complete set of symmetric many-body states. Thus, the grand partition function for bosons is

$$Z_G^{\text{boson}} = \prod_k \frac{1}{1 - e^{-\beta(\varepsilon_k - \mu)}}.$$

The grand potential is

$$\Omega_G = -k_B T \ln Z_G^{\text{boson}} = k_B T \sum_k \ln \left(1 - e^{-\beta(\varepsilon_k - \mu)} \right).$$

We recall that the average number of particles is given by Eq. (68)

$$\langle N \rangle = -\frac{\partial \Omega_G}{\partial \mu}.$$

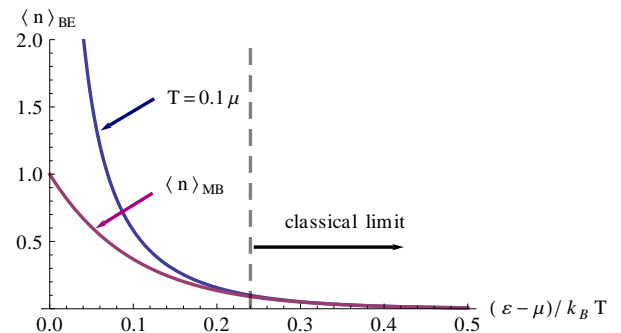
This means that since the grand potential of the whole N particle system is composed of a sum of grand potentials of the individual states, we can write the expected number of particles in state ψ_k as

$$\langle \hat{n}_k \rangle = -\frac{\partial \Omega_G^k}{\partial \mu} = \frac{1}{e^{\beta(\varepsilon_k - \mu)} - 1}.$$

We have thus seen that the number of particles in a single-particle bosonic state obeys the *Bose-Einstein distribution*

$$\langle \hat{n} \rangle_{\text{BE}} = \frac{1}{e^{\beta(\varepsilon - \mu)} - 1}. \quad (83)$$

This describes the filling of single-particle eigenstates with non-interacting bosons.



When the occupancy of a state is low, $\langle n \rangle \ll 1$, we obtain the Boltzmann distribution

$$\langle \hat{n} \rangle_{\text{BE}} \approx e^{-\beta(\varepsilon - \mu)}.$$

The condition for this is $\varepsilon - \mu \gg k_B T$ ³³. As $\varepsilon \rightarrow \mu$ we obtain $\langle \hat{n} \rangle_{\text{BE}} \rightarrow \infty$. In systems of non-interacting bosons μ is always less than the lowest single-particle eigenenergy.

As a final remark, Eq. (64) corresponds to an average excitation level of a quantum harmonic oscillator

$$\langle \hat{n} \rangle_{\text{qho}} = \frac{1}{e^{\beta \hbar \omega} - 1}.$$

One can use this to show that the excitations inside a harmonic oscillator can be treated as particles obeying Bose-Einstein statistics.

Fermions

Creation and annihilation operators for bosons

For fermions, the creation and annihilation operators obey the *anti-commutation* relations

$$\begin{aligned} \{\hat{a}_i, \hat{a}_j^\dagger\} &= \hat{a}_i \hat{a}_j^\dagger + \hat{a}_j^\dagger \hat{a}_i = \delta_{ij} \\ \{\hat{a}_i, \hat{a}_j\} &= \{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0. \end{aligned}$$

One can show that

$$\begin{aligned} \hat{a}_i |n_1, n_2, \dots, n_i, \dots\rangle &= \begin{cases} (-1)^{S_i} \sqrt{n_i} |n_1, n_2, \dots, n_i - 1, \dots\rangle & \text{if } n_i = 1 \\ 0 & \text{otherwise} \end{cases} \\ \hat{a}_i^\dagger |n_1, n_2, \dots, n_i, \dots\rangle &= \begin{cases} (-1)^{S_i} \sqrt{n_i + 1} |n_1, n_2, \dots, n_i + 1, \dots\rangle & \text{if } n_i = 0 \\ 0 & \text{otherwise} \end{cases} \end{aligned}$$

In the above, $S_i = n_1 + n_2 + \dots + n_{i-1}$. The (*occupation*) *number operator*

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$$

and obeys again the relation

$$\hat{n}_i |n_1, n_2, \dots, n_i, \dots\rangle = n_i |n_1, n_2, \dots, n_i, \dots\rangle,$$

with $n_i = 0, 1$

Example: Spin- $\frac{1}{2}$ system

Spin- $\frac{1}{2}$ particle is characterized by two quantum states, spin-“up” and spin-“down”, denoted with

$$|\uparrow\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, |\downarrow\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

One can show that every Hermitian operator \hat{O} in the two-dimensional Hilbert space spanned by $\{|\downarrow\rangle, |\uparrow\rangle\}$ can be written as

$$\hat{O} = c\hat{I} + a_1\hat{\sigma}_x + a_2\hat{\sigma}_y + a_3\hat{\sigma}_z,$$

³³Rather interestingly, this is fulfilled usually at *high* temperatures, because then there are more available states for occupation, leading to large and negative μ .

where c and a_i are complex coefficients and

$$\hat{\sigma}_x = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \hat{\sigma}_y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \hat{\sigma}_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

are the *Pauli spin matrices*.

The creation and annihilation operators are now

$$\hat{\sigma}_+ = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \hat{\sigma}_- = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

Fermi-Dirac distribution

Due to the antisymmetry of the wave function of fermions, only $n_k = 0$ and $n_k = 1$ are allowed. This means that the single-particle partition function is

$$Z_G^k = \sum_{n_k=0}^1 e^{-\beta(\varepsilon_k - \mu)n_k} = 1 + e^{-\beta(\varepsilon_k - \mu)}.$$

The grand partition function for the N fermion system is thus

$$Z_G^{\text{fermion}} = \prod_k \left(1 + e^{-\beta(\varepsilon_k - \mu)} \right)$$

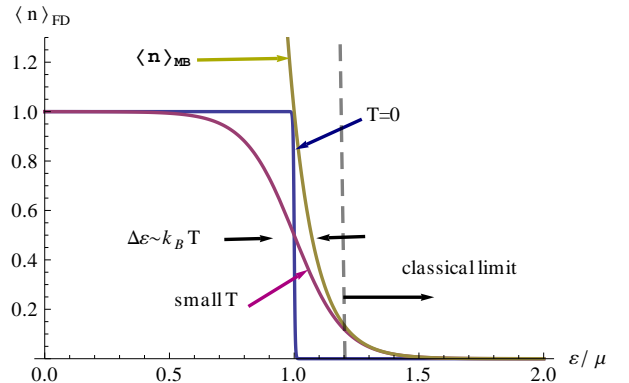
We obtain the average occupation in the state ψ_k to be

$$\langle \hat{n}_k \rangle = \frac{\sum_{n_k=0}^1 n_k e^{-\beta(\varepsilon_k - \mu)n_k}}{\sum_{n_k=0}^1 e^{-\beta(\varepsilon_k - \mu)n_k}} = \frac{1}{e^{\beta(\varepsilon_k - \mu)} + 1}.$$

We have thus shown that the average occupation of a single-particle state in a fermionic system obeys the *Fermi-Dirac distribution*

$$f(\varepsilon) = \langle \hat{n} \rangle_{\text{FD}} = \frac{1}{e^{\beta(\varepsilon - \mu)} + 1}, \quad (84)$$

where $f(\varepsilon)$ is often called the *Fermi function*. Again, in the limit $\varepsilon_k - \mu \gg k_B T$, the occupation of the state ψ_k is low and the distribution is given by the Boltzmann distribution. In fermionic systems, the chemical potential can be greater or less than any eigenenergy; at low temperatures the chemical potential separates the filled states from the empty states, and is often called as the *Fermi energy* $\varepsilon_F = \mu(T=0)$.



Note that only states within $|\varepsilon - \mu| \lesssim k_B T$ are partially filled.

Maxwell-Boltzmann gas

We calculate the trace in the grand partition function over the *whole* Hilbert space, i.e. over states Ψ_E . For each value of N , we eliminate the overcounting by dividing with $N!$

$$\begin{aligned} Z_G &= \sum_N \frac{1}{N!} \left(\sum_k e^{-\beta \varepsilon_k} \right)^N e^{N\beta\mu} \\ &= \prod_k e^{e^{-\beta(\varepsilon_k - \mu)}}. \end{aligned} \quad (85)$$

The grand potential is

$$\Omega_G = -k_B T \sum_k e^{-\beta(\varepsilon_k - \mu)}.$$

Thus, we end up with the average number of particles in a single-particle eigenstate with energy ε

$$\langle \hat{n} \rangle_{\text{MB}} = e^{-\beta(\varepsilon - \mu)}, \quad (86)$$

which is called the *Maxwell-Boltzmann distribution*. One sees that this indeed is the limiting case of both the Bose-Einstein and Fermi-Dirac distributions in the limit $\varepsilon - \mu \gg k_B T$. It is worthwhile to notice, that the correct result was obtained by tracing over the whole Hilbert space, and by requiring the indistinguishability of the particles with the factor $1/N!$.

Interestingly, the non-interacting quantum particles fill the single-particle states according to the same law

$$\langle \hat{n} \rangle = \frac{1}{e^{\beta(\varepsilon - \mu)} + c},$$

where $c = -1$ for bosons, $c = 1$ for fermions and $c = 0$ for Maxwell-Boltzmann statistics.

7. Ideal Bose Gas

We see that the distributions for both bosons and fermions imply that in certain parameter range the statistics deviate from the classical Maxwell-Boltzmann statistics. For bosons, the standard textbook examples include Bose-Einstein condensation, black body radiation and vibrations of the crystal lattice. The last example is vital for understanding, e.g., the specific heats of metals at high temperatures. Nevertheless, we will omit the discussion here³⁴.

Degeneracy discriminant

Before we continue the exploration of the resulting phenomena, we would like to have some kind of criterion to separate between the classical and quantum realms.

We have seen that the Maxwell-Boltzmann distribution is valid when

$$\langle \hat{n}_k \rangle \ll 1$$

for all single-particle states k . This can be rewritten as

$$e^{\beta\mu} \ll e^{\beta\varepsilon_k} \leq 1,$$

since $\min \varepsilon_k = 0$. In the above, $e^{\beta\mu}$ is often called the *fugacity*, which describes the ease of adding particles to the system. Now, recall that average value of particles can be written as

$$N = \langle \hat{N} \rangle = -\frac{\partial \Omega_G}{\partial \mu}$$

where the grand potential Ω_G for the Maxwell-Boltzmann gas is

$$\Omega_G = -k_B T e^{\beta\mu} \sum_k e^{-\beta\varepsilon_k} = -k_B T e^{\beta\mu} Z_1$$

where $Z_1(\beta)$ is the single-particle canonical partition function. In the classical limit, we obtained the result (57)

$$\begin{aligned} Z_1(\beta) &= \sum_k e^{-\beta\varepsilon_k} \\ &\approx \frac{V}{h^3} (2\pi m k_B T)^{3/2} \\ &\equiv V \lambda_T^{-3}, \end{aligned}$$

where $\lambda_T = \sqrt{h^2/(2\pi m k_B T)}$ is the *thermal de Broglie wave length*.

Thus, we can write the condition for classical statistics as

$$\rho \lambda_T^3 \ll 1, \quad (87)$$

where $\rho = N/V$ is the density of the gas.

One gets an intuitive picture of the above condition by considering the volume per particle as a sphere of radius r

$$\frac{1}{\rho} = \frac{4\pi}{3} r^3.$$

³⁴Interested reader can refer to the material on the course of condensed matter physics.

In addition, it is beneficial to consider the particles composing the gas as quantum mechanical *wave packets*³⁵. Next, we make a rough estimate of the size of the wave packet. We assume that the average energy of a particle is given by the thermal energy (equipartition theorem):

$$\frac{p^2}{2m} \approx k_B T \rightarrow p \approx \frac{h}{\sqrt{h^2/(2\pi m k_B T)}} = \frac{h}{\lambda_T}.$$

If we assume that the quantum mechanical uncertainty in the momentum Δp is less than the momentum p itself, we obtain according to Heisenberg's uncertainty principle

$$\Delta x \geq \lambda_T.$$

Now, the thermal wave length λ_T can be considered as the minimum size of the wave packet.

Thus, we see that the classical limit occurs when the wave packets do not overlap. In other words, this means that the quantum correlations between the particles are suppressed when the particles are small compared to the volume they are occupying.

One should note that the quantum effects reveal themselves when the density of the particles is high and/or the temperature of the system is low, together with small particle mass³⁶.

Density of free particles in a box

Before we start the actual demonstration of the power of quantum statistics, we note that one often has to study statistical quantities of type

$$\sum_{\varepsilon} F_{\varepsilon}$$

where F_{ε} is a function defined by the single-particle energy eigenstates. If the studied volume is large, the eigenstates lie close to one another and F_{ε} can be taken as continuous functions of energy. Thus, the summations are easily evaluated as integrals.

In the case of free particles, we assume that they are confined to a cube with $V = L^3$ and subjected to *periodic boundary conditions*³⁷, resulting in a solution of the Schrödinger equation in the form of plane waves

$$\psi_{\mathbf{k}} = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}, \quad (88)$$

with the *wave vector* $\mathbf{k} = (2\pi/L)(n_1, n_2, n_3)$ where n_i are integers. Each state occupies the volume $(2\pi/L)^3$ in the \mathbf{k} -space, so the density of plane waves in the \mathbf{k} -space is

$$D(\mathbf{k}) = \frac{V}{8\pi^3}.$$

Now, the energy of a plane wave is given by the *dispersion relation*³⁸

$$\varepsilon_k = \frac{\hbar^2 k^2}{2m}.$$

Thus, the number of plane waves in a small range of energy $d\varepsilon$ is

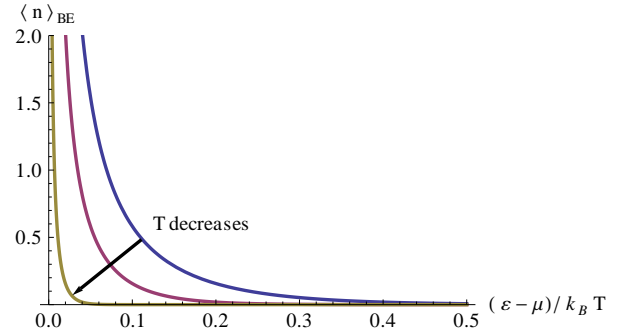
$$D(\varepsilon)d\varepsilon = (4\pi k^2) \left(\frac{V}{8\pi^3} \right) \left(\frac{dk}{d\varepsilon} d\varepsilon \right),$$

where the first term is the surface area of a sphere in \mathbf{k} -space, the second term is the density of plane waves per unit volume in \mathbf{k} -space, and the third is the thickness of the spherical shell. The above equation defines the *energy density of states*

$$D(\varepsilon) = \frac{2\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2}.$$

7.1 Bose-Einstein Condensation

We saw in the previous section that the number of particles in a single particle state of a many-boson system can take any non-negative integer value.



Especially interesting is the case in extremely low temperatures where all of the particles tend to go to the lowest accessible quantum state. This results into a new form of matter, *Bose-Einstein condensate*. Let us try to estimate the temperature region where this *macroscopic quantum phenomenon* occurs.

Now, the number of particles in the ground state ($\varepsilon = 0$) is

$$n_0 = \langle \hat{n}_0 \rangle = \frac{1}{z^{-1} - 1} = \frac{z}{1 - z} \rightarrow \infty,$$

because $z \rightarrow 1$, when $T \rightarrow 0$ (exercise). Thus, we obtain for the total number of particles

$$N = \sum_{\varepsilon} \frac{1}{z^{-1} e^{\beta \varepsilon} - 1} = \frac{z}{1 - z} + \sum_{\varepsilon \neq 0} \frac{1}{z^{-1} e^{\beta \varepsilon} - 1}.$$

Now, when the volume V is large, the spectrum of single particle energy states is almost continuous. Thus, one can

³⁵See Quantum mechanics.

³⁶It is vital to acknowledge that the temperature and density should be compared with each other. For example, in white dwarfs one has a degenerate quantum gas even in temperatures of the order 10^7 K.

³⁷Check the course of Solid state physics or Condensed matter physics for more details on this calculation.

³⁸Dispersion relation simply means the dependence of the energy on the wave vector.

approximate the summation with an integral³⁹

$$\begin{aligned} N &= \sum_{\varepsilon} \frac{1}{z^{-1}e^{\beta\varepsilon} - 1} = n_0 + \int_0^{\infty} \frac{D(\varepsilon)}{z^{-1}e^{\beta\varepsilon} - 1} d\varepsilon \\ &= n_0 + \frac{V}{\lambda_T^3} g_{3/2}(z). \end{aligned}$$

We have also identified the *Bose-Einstein functions*

$$g_{\nu}(z) = \frac{1}{\Gamma(\nu)} \int_0^{\infty} \frac{x^{\nu-1} dx}{z^{-1}e^x - 1} = \sum_{n=1}^{\infty} \frac{z^n}{n^{\nu}}.$$

Similarly, one can show that the grand potential is

$$\Omega_{\text{BE}} = k_B T \ln(1 - z) - \frac{V k_B T}{\lambda_T^3} g_{5/2}(z)$$

which implies

$$p = - \left(\frac{\partial \Omega_{\text{BE}}}{\partial V} \right)_{T,N} = \frac{k_B T}{\lambda_T^3} g_{5/2}(z). \quad (89)$$

Similarly, we can calculate the internal energy of the system

$$E = \int_0^{\infty} \frac{\varepsilon D(\varepsilon)}{z^{-1}e^{\beta\varepsilon} - 1} d\varepsilon = \frac{3}{2} k_B T \frac{V}{\lambda_T^3} g_{5/2}(z)$$

Thus, we obtain for ideal Bose gas

$$p = \frac{2}{3} \frac{E}{V}$$

In the classical limit we have $n_l \ll 1$ for all states l . Thus, n_0 is negligible compared with $V g_{3/2}(z)/\lambda_T^3$ and

$$\rho \lambda_T^3 = g_{3/2}(z).$$

Since $\mu < 0$, we have that $z = 0$ when $T \rightarrow \infty$, which is in accordance with our condition (87) for classical statistics.

Also, when $T \rightarrow 0$ the fugacity $z \rightarrow 1$. So, the fugacity obtains values $0 < z < 1$, and within this interval the Bose-Einstein function is monotonically increasing function with limits

$$\begin{aligned} g_{3/2}(0) &= 0 \\ g_{3/2}(1) &= 2.612. \end{aligned}$$

When the fugacity $z = 1$ (i.e. $\mu(T, N) = 0$, the density ρ and temperature T obey

$$\rho \lambda_T^3 = g_{3/2}(1) = 2.612,$$

This implies that $n_0 = 0$. If we still increase the density, or decrease the temperature, the increase in the quantity $\rho \lambda_T^3$ must originate from the particles occupying the ground state.

³⁹Notice, that we removed the grand potential of the ground state since it would be taken into account incorrectly by the density of states, giving a zero weight.

We thus see, that

$$\begin{aligned} \rho \lambda_T^3 &= g_{3/2}(z), & \text{when } \rho \lambda_T^3 < 2.612 \\ \rho \lambda_T^3 &= \frac{n_0}{V} \lambda_T^3 + g_{3/2}(1), & \text{when } \rho \lambda_T^3 \geq 2.612. \end{aligned}$$

In other words, when

$$\lambda_T^3 \geq 2.612 \frac{1}{\rho}$$

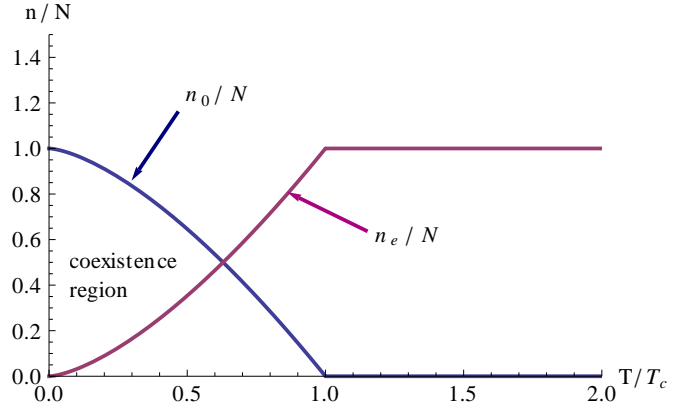
there will be a macroscopic occupation in the ground state, forming a *Bose-Einstein condensate*. Again, the condensate starts to arise when the thermal wave packets describing the bosons start to overlap.

Typically, the density of the gas is kept constant and the temperature is lowered. The temperature below which the chemical potential μ is zero, i.e. $z = 1$, is

$$T_c = \left(\frac{2\pi\hbar^2}{mk_B} \right) \left(\frac{\rho}{2.612} \right)^{2/3}.$$

This is called the *critical temperature*, below which the condensate starts to form. Below the T_c , the relative fraction of particles in the condensate is

$$\frac{n_0}{N} = 1 - \frac{2.612}{\lambda_T^3 \rho} = 1 - \left(\frac{T}{T_c} \right)^{3/2}. \quad (90)$$



We see that at T_c the gas goes through a phase transition, and starts to coexist in two phases. The particles in the ground state form the *condensed phase*, characterised by the *order parameter* n_0/N ⁴⁰. The particles in the excited states form a *normal phase* with relative fraction of particles given by

$$\frac{n_e}{N} = 1 - \frac{n_0}{N}.$$

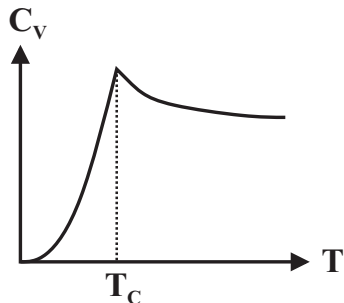
One can also imagine that the density of the gas is increased at constant temperature, resulting in the *critical density*

$$\rho_c = 2.612 \times \lambda_T^{-3}.$$

Now, the critical temperature leads together with Eq. (89) in *critical pressure*

$$p_c = \frac{k_B T_c}{\lambda_T^3} g_{5/2}(1).$$

⁴⁰We will return to the concept of order parameter when we study more about phase transitions.



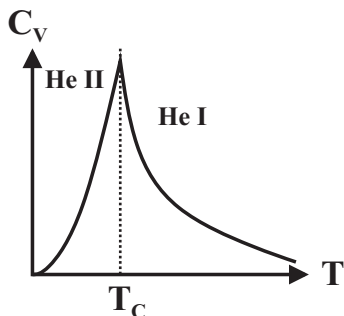
The phase transition is of first order, meaning that the first derivative of the Gibbs' energy is discontinuous.

Experimental realisation

Helium-4

Helium is a peculiar element, since it resists to go into solid phase at standard pressure, even at the lowest temperatures. This is because of the small mass of the helium atom that causes so large zero-point motion that it is enough to prevail over the weak inter-atomic attractive (van der Waals) forces. Consequently, the boiling point of Helium is extremely low, 4.2 kelvins. Helium has two stable isotopes ^3He and ^4He . ^3He is a fermion and, thus, we concentrate here on the ^4He -atoms, which are bosons.

Helium-4 goes through a phase transition at 2.19 K. This is visible as the discontinuity in the specific heat C_V .



Due to the shape of the C_V -curve, the transition was named λ -transition and the corresponding T_c the λ -temperature. Interestingly, the shape of the C_V curve resembles that of a Bose gas. Moreover, the critical temperature for Bose-Einstein condensation, calculated with the ^4He parameters, gives $T_c = 3.13$ K, which is not that far from the observed transition temperature. This lead people to believe that the transition in ^4He was a manifestation of Bose-Einstein condensation.

However, the ^4He is not an ideal liquid. There exists strong repulsive interactions between ^4He -atoms at short distances, and attractive (although weak) ones at longer distances. Nevertheless, the theoretical basis of the description of the fluid below the critical temperature is based in the so-called *two-fluid model*. Below the critical temperature, the particles occupying the ground state are thought to form a *superfluid*, a state of matter which behaves like fluid with zero viscosity and zero entropy. The phase with a superfluid component is commonly called helium-II, whereas the completely normal ^4He is helium-I. On the

other hand, the particles occupying the excited states are identified as the *normal component*. This can be denoted as

$$\rho = \rho_s + \rho_n$$

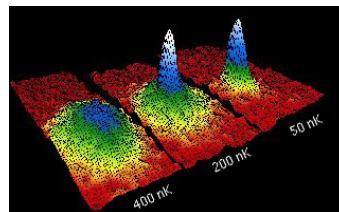
$$\mathbf{j} = \mathbf{j}_s + \mathbf{j}_n$$

$$\vdots$$

It should, nevertheless, be stressed that the quantitative agreement with the experiments requires the inclusion of interactions to the model.

BEC in ultra-cold atomic gases

The first demonstration of *gaseous* Bose-Einstein condensation was made in dilute gases of alkali atoms in 1995⁴¹. In these experiments, the gases were trapped and cooled with magneto-optical traps, which effectively confine the gas in a three dimensional harmonic potential. The experiment is performed many times with different final temperatures, and once the temperature is reached the trapping magnetic field is turned off and the subsequent ballistic motion of the atoms is observed. Atoms in a high temperature gas will obey thermal distribution and expand in thermal manner. The atoms below the condensation temperature are in the ground state of the harmonic trap, and will hardly move after the field is turned off.



The onset of Bose-Einstein condensation in rubidium, taken from the experiments on rubidium atoms at Wiemann and Cornell's group at JILA-NIST.

7.2 Blackbody Radiation (Photon Gas)

Matter emits and absorbs radiation. In equilibrium, the rates of these processes are equal, which is called the *detailed balance*. Now, consider a perfect absorber, a *blackbody*, that can absorb all radiation at all frequencies. What will its emission spectrum look like? First, recall that the electromagnetic radiation consists of plane waves, similar to Eq (88). Since photons are massless, the dispersion relation is $\omega_k = ck$, where c is the speed of light. There are two *modes* for each wave vector \mathbf{k} , one for each transverse polarization. Experimentally, it is impossible to have a perfect blackbody. Closed cavity with a small opening, and in a constant temperature produces nevertheless a good approximation.

The radiation spectrum of a blackbody was one of the key ignitions for the theory of quantum mechanics.

⁴¹Wiemann and Cornell with rubidium atoms, and Ketterle with sodium atoms. They shared the physics Nobel prize for this in 2001.

Namely, the classical physics predicts the so-called *ultra-violet catastrophe*, radiation at infinite powers. This results from the fact that each mode of the radiation field can have the equipartition energy $k_B T$. Moreover, in a 3-dimensional cavity, the number of modes is proportional to the square of the frequency. So, the radiated power should follow the *Rayleigh-Jeans* law, and be proportional to the squared frequency, leading to unlimited radiated power at high frequencies. This is clearly in contradiction with the real world⁴², except for the limit of low frequencies.

The correct description to this problem was given by Max Planck in the beginning of the 20th century, when he proposed that the emitted energies are not distributed continuously, but come in lumps proportional to the frequency, i.e.

$$\varepsilon_k = \hbar \omega_k.$$

Let us apply Planck's ideas and consider a cavity of volume V at temperature T . There exists two conceptually different, but practically identical viewpoints, as the field can be seen as

- an assembly of quantum harmonic oscillators with energies $\varepsilon_s = \hbar \omega_s (n_s + \frac{1}{2})$, where $n_s = 0, 1, 2, \dots$ and ω_s is the angular frequency of the oscillator.
- a gas of identical and indistinguishable quanta, *photons*, with energies $\hbar \omega_s$. Each photon is a boson, and each mode can support any number $n_s = 0, 1, 2, \dots$ of photons.

In both cases, one the same expectation value

$$\langle n_s \rangle = \frac{1}{e^{\beta \hbar \omega_s} - 1}.$$

Notice that this is the Bose-Einstein distribution with $\mu = 0$. Since we are dealing with an equilibrium gas, the free energy should be minimized in order to find the equilibrium number of photons. There are no conservation laws for photons, like those of mass, charge, lepton number, angular momentum and such. Thus, there is no constraints imposed in the minimization, giving just $(\partial A / \partial N)_{T,V} = 0$. This implies the disappearance of the chemical potential. In other words, the total number of particles in the present case is *indefinite*.

We see that there exists a complete correspondence between "an oscillator in eigenstate n_s with eigenenergy $\hbar \omega_s (n_s + \frac{1}{2})$ " and "the occupation of the energy level $\hbar \omega_s$ by n_s photons". From this point on, regardless of the used viewpoint, the treatments are the same.

The number of single-particle plane waves in a small range $d\omega$ is given by

$$D(\omega) d\omega = (4\pi k^2) \left(\frac{2V}{8\pi^3} \right) \left(\frac{dk}{d\omega} d\omega \right),$$

where the factor two is due to the number of modes per k . We thus obtain the density of states per unit frequency

$$D(\omega) = \frac{V \omega^2}{\pi^2 c^3}.$$

⁴²You do not get a sun-tan by opening the door of a hot oven!

Thus, the average total energy in the field is⁴³

$$\langle E \rangle = \int_0^\infty d\omega D(\omega) \frac{\hbar \omega}{e^{\beta \hbar \omega} - 1}.$$

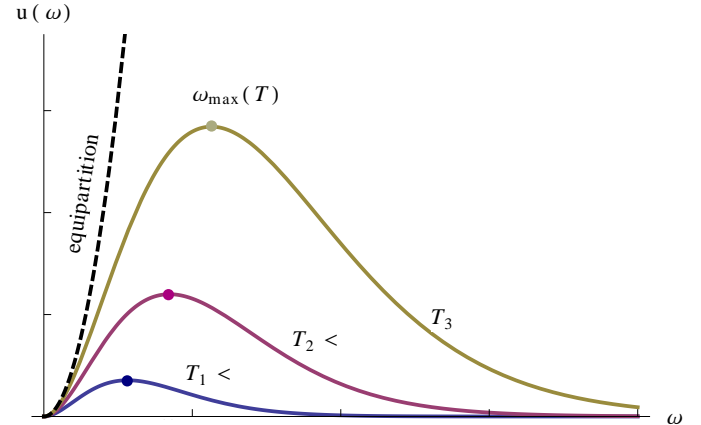
We thus see that the energy density can be written as

$$\frac{E}{V} = \frac{\hbar}{\pi^2 c^3} \int_0^\infty d\omega \frac{\omega^3}{e^{\beta \hbar \omega} - 1}.$$

This implies the energy density per unit frequency

$$u(\omega) = \frac{\hbar \omega^3}{\pi^2 c^3 (e^{\beta \hbar \omega} - 1)}, \quad (91)$$

which is *Planck's law of radiation*. The spectrum of a perfect absorber is determined solely by its temperature, not by its shape, material nor structure. At the room temperature, most of the energy radiated by the blackbody is infra-red, not perceived by the human eye. Only after about 800 K, the black body starts to radiate visible light.



We see that the maximum of the energy density follows the *Wien displacement law*

$$\omega_{\max} = \text{constant} \times T$$

In the limit of small frequencies, one obtains

$$u(\omega) \approx \frac{k_B T}{\pi^2 c^3} \omega^2 = k_B T \frac{D(\omega)}{V},$$

which is the Rayleigh-Jeans law, equivalent to the equipartition energy density: $k_B T$ per classical oscillator.

In general, the total energy density obeys the *Stefan-Boltzmann law*

$$\frac{E}{V} = \frac{4}{c} \sigma T^4,$$

⁴³Notice that here we neglect the (infinite) zero-point energy. This is the so-called *renormalization* of the energy, justified by stating that only the differences in the energy can be measured. Those readers that have been observant, noticed that in the "photon picture" there seems to be no such term as the zero-point energy. Nevertheless, this is just a sloppy notation, since according to quantum field theory, the zero-point fluctuations in the oscillator model are present as *vacuum fluctuations* in the photon model. The vacuum fluctuations arise as constant creation of photon-anti-photon pairs, which can interact before annihilation, giving rise to the *vacuum energy* equivalent to the zero-point energy of the oscillators.

where

$$\sigma = \frac{\pi^2 k_B^4}{60 \hbar^3 c^2} = 5.67 \cdot 10^{-8} \frac{\text{W}}{\text{m}^2 \text{K}^4}$$

We can calculate also other properties of the photon gas. The free energy is given by the grand potential (because $\mu = 0$)

$$A = \Omega_G + \mu N = \Omega_G = k_B T \int_0^\infty D(\omega) \ln(1 - e^{-\beta \hbar \omega}).$$

We obtain by the change of variable $\beta \hbar \omega = x$, and by integration by parts

$$A = -\frac{4\sigma}{3c} V T^4 = -\frac{1}{3} E.$$

The entropy of the gas is given by

$$S = -\frac{\partial A}{\partial T} = \frac{16\sigma}{3c} V T^3,$$

and the pressure

$$p = -\frac{\partial A}{\partial V} = \frac{4\sigma}{3c} T^4.$$

We thus see that the pressure of the photon gas is one third of its energy density, i.e.

$$p = \frac{1}{3} \frac{E}{V}.$$

This is a well known result of the radiation theory.

One of the most important examples of the blackbody radiation is the 2.7 K cosmic microwave background, which is thought to be the remnant from the Big Bang.

8. Ideal Fermi Gas

In this chapter, we will study the gas of indistinguishable, non-interacting fermions. According to the previous chapters, we have obtained the grand potential

$$\Omega_G^{\text{fermion}} = -k_B T \sum_{\varepsilon} \ln(1 + z e^{-\beta \varepsilon})$$

and the average number of particles

$$N = \sum_{\varepsilon} \langle \hat{n}_{\varepsilon} \rangle = \sum_{\varepsilon} \frac{1}{z^{-1} e^{\beta \varepsilon} + 1}.$$

Unlike bosons, the fugacity of fermion gas is unrestricted, i.e. $0 \leq z < \infty$. Because of the Pauli principle no phenomena, like the Bose-Einstein condensation, where a macroscopic amount of particles go into the same single-particle state can occur. Nevertheless, the fermion gas has its own peculiarities that arise from the characteristic quantum nature of its constituents, and those will be studied here.

Similar to boson gases, one can change the summations of statistical variables into integrations for large enough volumes V . This leads to energy density of states

$$D(\varepsilon) = g_s \frac{2\pi V}{h^3} (2m)^{3/2} \varepsilon^{1/2},$$

where g_s is the number of different spin states (for spin- $\frac{1}{2}$ particle $g_s = 2$). Accordingly, we obtain

$$\frac{p}{k_B T} = \frac{g_s}{\lambda_T^3} f_{5/2}(z)$$

and

$$\rho = \frac{g_s}{\lambda_T^3} f_{3/2}(z).$$

In the above, the *Fermi-Dirac functions* are defined as

$$f_{\nu}(z) = \frac{1}{\Gamma(\nu)} \int_0^\infty \frac{x^{\nu-1} dx}{z^{-1} e^{\beta \varepsilon} + 1} = \sum_{n=1}^{\infty} (-1)^{(n-1)} \frac{z^n}{n^{\nu}}.$$

The internal energy is given by

$$E = \int_0^\infty \frac{\varepsilon D(\varepsilon)}{z^{-1} e^{\beta \varepsilon} + 1} d\varepsilon = \frac{3}{2} k_B T \frac{g_s V}{\lambda_T^3} f_{5/2}(z),$$

giving also for the ideal Fermi gas the relation

$$p = \frac{2}{3} \frac{E}{V}.$$

We see that in the classical limit

$$\frac{\rho \lambda_T^3}{g_s} = f_{3/2}(z) \ll 1.$$

This means that $f_{\nu}(z) \approx z$. Thus, we obtain the familiar classical results

$$pV = N k_B T, \quad E = \frac{3}{2} N k_B T$$

and so on. For the values $\rho \lambda_T^3 / g_s \sim 1$, one has to rely on numerical calculations in expanding the series in the Fermi-Dirac functions. In the quantum limit $\rho \lambda_T^3 / g_s \rightarrow \infty$, one

can write the thermodynamics in a closed form, and the gas is *completely degenerate* meaning that

$$n_{\text{FD}} = \begin{cases} 1 & \text{for } \varepsilon < \mu_0 \\ 0 & \text{for } \varepsilon > \mu_0 \end{cases}$$

Here $\mu_0 = \varepsilon_F$ is the chemical potential at $T = 0$, often referred to as the *Fermi energy*. All single particle states up to Fermi energy are completely filled (according to the Pauli principle), and those above the Fermi energy are completely empty.

The value of the Fermi energy can be calculated from the density, because

$$N = \int_0^{\varepsilon_F} D(\varepsilon) d\varepsilon = \frac{g_s(2\varepsilon_F m)^{3/2}}{6\pi^2 \hbar^3} V$$

$$\Rightarrow \varepsilon_F = \left(\frac{6\pi^2 \rho}{g_s} \right)^{2/3} \frac{\hbar^2}{2m}.$$

Similarly, one can calculate the average energy at $T = 0$, or the so-called *ground-state energy*, as

$$E_0 = \int_0^{\varepsilon_F} \varepsilon D(\varepsilon) d\varepsilon = \frac{\sqrt{2} g_s V m^{3/2}}{5\pi^2 \hbar^3} \varepsilon_F^{5/2},$$

which implies the average energy per particle

$$\frac{E_0}{N} = \frac{3}{5} \varepsilon_F.$$

Often needed quantities are *Fermi temperature*, *momentum* and *wave vector*, which are defined respectively as

$$T_F = \frac{\varepsilon_F}{k_B}$$

$$p_F = \sqrt{2m\varepsilon_F}$$

$$k_F = \frac{p_F}{\hbar}.$$

If we want to study properties like specific heat and entropy, we have to have a *finite* temperature. If we still can assume that the temperature is low, meaning that $T \ll \mu/k_B$, we expect that the deviations from the ground state are not that large and that we need only the first terms from the so-called asymptotic *Sommerfeld expansion*

$$f_\nu = \frac{(\beta\mu)^\nu}{\Gamma(\nu+1)} \left[1 + \nu(\nu-1) \frac{\pi^2}{6} \frac{1}{(\beta\mu)^2} \right. \quad (92)$$

$$\left. + \nu(\nu-1)(\nu-2)(\nu-3) \frac{7\pi^4}{360} \frac{1}{(\beta\mu)^4} + \dots \right].$$

The Sommerfeld expansion gives us the low temperature corrections to the thermodynamic quantities, such as

$$\frac{E}{N} = \frac{3}{5} \varepsilon_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right]$$

$$p = \frac{2}{3} \frac{E}{V} = \frac{2}{5} \rho \varepsilon_F \left[1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right].$$

On immediately sees, that the leading terms are the identical to the ground-state results. From the average energy, we can calculate the specific heat

$$\frac{C_V}{N k_B} = \frac{\pi^2}{2} \frac{k_B T}{\varepsilon_F}.$$

This implies that for $T \ll T_F$ the specific heat is much less than the classical value $\frac{3}{2} N k_B$.

This last result is a consequence of the fundamental property of the Fermi gas. Due to the Pauli principle, the single-particle states far below the Fermi energy are completely filled. It means that the particles occupying such states cannot be (easily) excited because there are no empty states nearby, and therefore do not contribute to the specific heat, or for example to electric conductivity. For metals, the Fermi temperature is typically high (of the order of 10^4 K, meaning that the low temperature approximation is justified at the room temperature. Thus, the low temperature ideal Fermi gas serves as an excellent starting point of the study of metallic properties⁴⁴.

8.1 Magnetism in an Ideal Fermi Gas

Let us study the behaviour of the ideal Fermi gas under the influence of external magnetic field \mathbf{B} . Previously on page 26, we showed that if orientations the spins of the fermions are distributed according to the classical Boltzmann distribution, the net magnetic moment M_z becomes saturated at low temperatures (compared with the magnitude of the field). This implies that the *magnetic susceptibility* $\chi(T)$ at constant temperature becomes zero

$$\chi(T) = \mu_0 \lim_{B \rightarrow \infty} \left(\frac{\partial M_z}{\partial B} \right)_T = 0.$$

If one takes into account the Fermi statistics, one obtains a very different behaviour, especially at low temperatures.

Pauli paramagnetism

For simplicity, we assume that we are dealing with spin- $\frac{1}{2}$ particles at low temperatures ($T \ll T_F$). This kind of situation occurs, for example, with the electron gas in metals. The energy of such a particle in external field is

$$\varepsilon = \begin{cases} \frac{p^2}{2m} - \mu_B B & \text{when } \boldsymbol{\mu} \text{ is parallel to } \mathbf{B} \\ \frac{p^2}{2m} + \mu_B B & \text{when } \boldsymbol{\mu} \text{ is antiparallel to } \mathbf{B}, \end{cases}$$

where $\boldsymbol{\mu}$ is the intrinsic magnetic moment of the particle and $\mu_B = e\hbar/2m$ is the Bohr magneton. Thus, we see that the introduction of the magnetic field divides the spins into two populations,

Now, since the studied Fermi gas is strongly degenerated ($T \ll T_F$), the occupations in the two different populations are given by the Fermi distributions

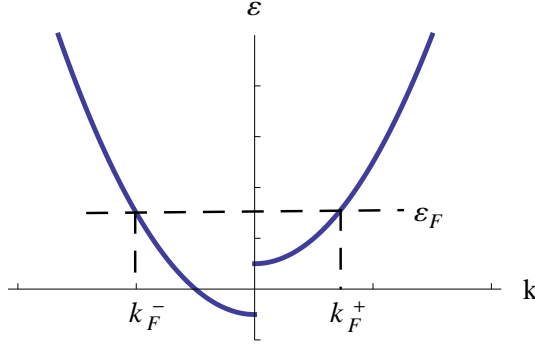
$$n_{p\pm} = \frac{1}{e^{\beta(\varepsilon_p \pm \mu_B B - \varepsilon_F)} + 1},$$

⁴⁴This fact is exhaustively discussed in the condensed matter physics course, and will not be studied further here.

where $\varepsilon_p = p^2/2m$ and $\varepsilon_F = \mu(T = 0)$.

All the states up to the Fermi energy ε_F are filled, which implies that the wave numbers k_F^\pm for the two populations are different and can be determined from the condition

$$\begin{aligned}\frac{\hbar^2 k_F^{+2}}{2m} + \mu_B B &= \varepsilon_F \\ \frac{\hbar^2 k_F^{-2}}{2m} - \mu_B B &= \varepsilon_F.\end{aligned}$$



Clearly, the number of states in the population with the parallel alignment is larger. In terms of the Fermi wave number, we obtain the population occupations

$$\begin{aligned}N_+ &= \int_0^{k_F^+} 4\pi k^2 D(k) dk = V \frac{k_F^{+3}}{6\pi^2} \\ N_- &= \int_0^{k_F^-} 4\pi k^2 D(k) dk = V \frac{k_F^{-3}}{6\pi^2}.\end{aligned}$$

Thus, we obtain a net magnetic moment

$$\begin{aligned}M &= \mu_B(N_- - N_+) \\ &= \frac{\mu_B V}{6\pi^2} (k_F^{-3} - k_F^{+3}) \\ &= \frac{\mu_B V}{6\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} [(\varepsilon_F + \mu_B B)^{3/2} - (\varepsilon_F - \mu_B B)^{3/2}].\end{aligned}$$

Realistic values for fields produced in laboratories are of the order $B \lesssim 10$ T, which corresponds to $\mu_B B/k_B \lesssim 1$ K. This is much smaller than the Fermi temperatures in metals, and thus we can assume that $\mu_B B \ll \varepsilon_F$. This way we have

$$M \approx \frac{V}{2\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \mu_B^2 B \varepsilon_F^{1/2}.$$

We end up with the weak field *Pauli susceptibility* (per unit volume)

$$\chi_0 = \lim_{B \rightarrow 0} \frac{\mu_0}{V} \frac{\partial M}{\partial B} = \frac{3}{2} \mu_0 \rho \frac{\mu_B^2}{\varepsilon_F},$$

which holds for $T \ll T_F$ and $\mu_B B \ll \varepsilon_F$. Thus, we see that instead of being zero, the susceptibility is a constant determined by the density. This should be compared to the classical (high-temperature) Curie law, Eq. (66) (with $g = 2$ and $J = \frac{1}{2}$), which implies

$$\chi_\infty = \rho \frac{\mu_0 \mu_B^2}{k_B T}.$$

At finite temperatures, one should apply the Sommerfeld expansions to resolve the susceptibility. We omit the calculation here, and just state that the lowest order corrections to the low and high temperature susceptibilities are

$$\chi = \begin{cases} \chi_0 \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 \right] & \text{when } T \text{ is low} \\ \chi_\infty \left[1 - \frac{\rho \lambda_T^3}{2^{5/2}} \right] & \text{when } T \text{ is high,} \end{cases}$$

respectively. The high temperature correction is proportional to $(T_F/T)^{3/2}$ and tends to zero as $T \rightarrow \infty$.

Landau diamagnetism

In addition to the (para)magnetism associated to the orientations of the spins, there exists additional form of magnetism which arises from the quantization of the kinetic energy of *charged* particles associated with their motion perpendicular to the direction of the field. This is *diamagnetic* in character, meaning that it creates a field opposite to the external field \mathbf{B} , implying negative susceptibility $\chi(T)$. According to the *Bohr-van Leeuwen theorem*, the phenomenon of diamagnetism cannot be explained in terms of classical physics (exercise).

Consider an ideal Fermi gas formed by particles with charge q . In the presence of an external magnetic field, the single-charge Hamiltonian is written as

$$\hat{H} = \frac{1}{2m} \left(-i\hbar \nabla - q\mathbf{A} \right)^2,$$

where

$$\mathbf{B} = \nabla \times \mathbf{A}.$$

We assume that the magnetic field points in \mathbf{z} -direction, i.e. $\mathbf{B} = B\mathbf{z}$, which allows us to choose the *Landau gauge*⁴⁵ for the vector potential

$$\mathbf{A} = (0, B\hat{x}, 0).$$

Thus, we obtain the Hamiltonian

$$\hat{H} = \frac{\hat{p}_x^2}{2m} + \frac{1}{2m} \left(\hat{p}_y - qB\hat{x} \right)^2 + \frac{\hat{p}_z^2}{2m}.$$

As a result, a charged particle has a *helical* trajectory whose axis is parallel to the external field, and whose projection to the (x, y) -plane is a circle.

We see that the Hamiltonian does not depend on the operator \hat{y} , meaning that \hat{H} and \hat{p}_y can be diagonalized in the same basis. This means that in the y -direction the Hamiltonian describes a free particle and altogether in the (x, y) -plane we have

$$\hat{H}_{xy} = \frac{\hat{p}_x^2}{2m} + \frac{1}{2} m \omega_c^2 \left(\hat{x} - \frac{\hbar k_y}{m \omega_c} \right)^2,$$

⁴⁵Recall that a given field can be created with different choices for potentials. The behaviour of a particle in the field is independent on the choice for the potential. This is called *gauge invariance*.

i.e. a harmonic oscillator Hamiltonian with a *displacement* in the x -coordinate by $X = k_y l_0^2$ where $l_0 = \sqrt{\hbar/qB}$ is the *magnetic length*. We have also defined the *cyclotron frequency* $\omega_c = qB/m$. The eigenvalues of the harmonic oscillator Hamiltonian are not dependent on the translations in the coordinates. Thus, the eigenenergies are of form

$$E_n = \left(n + \frac{1}{2}\right) \hbar \omega_c, \quad n = 0, 1, 2, \dots$$

These are the so-called *Landau levels* of a charged particle in a magnetic field.

By using the periodic boundary conditions, we obtain in the (free) y -direction

$$k_y = \frac{2\pi n_y}{L_y}, \quad n_y = 0, \pm 1, \pm 2, \dots$$

and in the harmonic coordinate x

$$0 \leq X < L_x.$$

These restrict the possible values for n_y :

$$\Rightarrow N_s \equiv n_y^{\max} = \frac{L_x L_y}{2\pi l_0^2} = Z \frac{e}{h} L_x L_y B = Z \frac{\Phi}{\Phi_0},$$

where $q = Ze$, $\Phi_0 = h/e$ is the *flux quantum*, and Φ is the magnetic flux through the system. We thus see that each Landau level is degenerate, with exactly one state for each flux quantum and for each elementary charge.

The total energy of the ideal charged Fermi gas is

$$\varepsilon_n = \frac{p_z^2}{2m} + E_n,$$

where the kinetic energy associated with the linear motion along the field is assumed to be a continuous variable. Now, we can calculate the grand partition function

$$\begin{aligned} \ln Z_G &= \sum_{\varepsilon} \ln(1 + ze^{-\beta\varepsilon}) \\ &= \int_{-\infty}^{\infty} dp_z \frac{L_z}{h} \\ &\quad \times \left[\sum_{n=0}^{\infty} \left(L_x L_y \frac{qB}{h} \ln [1 + ze^{-\beta(\hbar\omega_c(n+1/2) + p_z^2/2m)}] \right) \right]. \end{aligned}$$

Similar to Eq. (65), we can calculate the magnetization as

$$M = \frac{1}{\beta} \left(\frac{\partial}{\partial B} \ln Z_G \right)_{z, V, T}.$$

After somewhat lengthy calculation (which we skip here), we obtain the low-field ($\mu_{\text{eff}} B \ll k_B T$) susceptibility

$$\chi(T) = -\mu_0 \frac{(2\pi m)^{3/2} \mu_{\text{eff}}^2}{3\hbar^3 \beta^{1/2}} f_{1/2}(z),$$

where $\mu_{\text{eff}} = q\hbar/2m$ (this is the Bohr magneton if $q = e$). Notice that the susceptibility is diamagnetic in nature, irrespective of the sign of the charge of the particle.

In the classical limit, we have that $f_{1/2}(z) \approx z \approx \rho \lambda_T^3$. This implies

$$\chi_{\infty} = -\frac{\rho \mu_0 \mu_{\text{eff}}^2}{3k_B T}.$$

We thus see that at high temperatures the diamagnetism obeys the Curie law, but with the opposite sign. In the quantum limit ($T \ll T_F$ but still $\mu_{\text{eff}} B \ll k_B T$), we obtain

$$\chi_0(T) \approx -\frac{\rho \mu_0 \mu_{\text{eff}}^2}{2\varepsilon_F}.$$

We have applied the first term of the asymptotic Sommerfeld expansion (92). Note that the asymptotic results χ_0 and χ_{∞} are precisely one third in magnitude of the paramagnetic results, given that the particles are electrons. This gives us the intuition that the diamagnetic contribution to the materials response to a magnetic field is negligible compared with the paramagnetic contribution.

8.2 Relativistic Electron Gas in White Dwarfs

The first application of Fermi statistics was related to the equilibrium studies of *white dwarf* stars. White dwarfs have burned most of their hydrogen in the nuclear fusion ($2\text{H} \rightarrow \text{He}$). When this energy release becomes absent, the gravitational force is unbalanced and the star collapses into a white dwarf.

Consider then a typical idealization of white dwarf, consisting of a ball of helium with a mass $M \approx 10^{30}$ kg and mass density $\rho_m \approx 10^{10} \text{ kg m}^{-3}$. The temperature within the white dwarf is of the order $T = 10^7$ K. This means that the thermal energy per helium atom is

$$\frac{3}{2} k_B T \approx 1 \text{ keV} \gg E_{\text{ion}} \approx 10 \text{ eV},$$

where E_{ion} is the *ionization energy* of a helium atom. Thus, the helium atoms in white dwarfs are completely ionized.

Now, the mass of the helium consists mostly of that of the nucleus, and can be approximated as (${}^4\text{He}$ is more abundant)

$$M \approx N(m_e + 2m_p) \approx 2Nm_p,$$

where N is the number of electrons, m_e is the mass of the electron, and $m_p = 1.673 \cdot 10^{-27}$ kg is the mass of the proton. This gives us the approximate electron density

$$\rho = \frac{N}{V} = \frac{\rho_m}{2m_p} \approx 10^{36} \text{ electrons/m}^3.$$

This gives us the Fermi energy of the electron gas

$$\varepsilon_F \approx 10^6 \text{ eV} \approx mc^2 = 0.512 \cdot 10^6 \text{ eV}.$$

Thus, the kinetic energies of the electrons are of the order of their relativistic rest energy, meaning that in white dwarfs we are dealing with *relativistic Fermi gas*. Additionally, the Fermi energy gives $T_F \approx 10^{10}$ K, which means that the electrons are (almost) completely degenerate.

In the equilibrium of the white dwarf, the outward force due to the pressure p of the relativistic degenerate electron gas balances the gravitational pull

$$\frac{dp}{dr} = -\frac{GM(r)}{r^2},$$

where $G = 6.673 \cdot 10^{-11} \text{ Nm}^2/\text{kg}^2$ is the gravitational constant, and the mass at distance r from the star center is

$$M(r) = \frac{4}{3}\rho_m r^3,$$

where we assume that the density is constant within the star. By simple integration, we obtain the pressure at the center of the star

$$p = \frac{8\pi}{3}Gm_p^2\rho^2 R^2,$$

where we have assumed that the pressure vanishes at the surface.

On the other hand, the pressure can be calculated from the ground state properties of the relativistic electron gas. First of all, recall that the relativistic energy of the electron can be written as

$$\varepsilon_p = \sqrt{(mc^2)^2 + (cp)^2}.$$

This can be rewritten in terms of the wave vector as

$$\varepsilon_k = c\hbar\sqrt{k^2 + k_c^2},$$

where $p = \hbar k$ and $k_c = mc/\hbar$ is the *Compton wave vector*.

Now, in the relativistic limit $k_F \ll k_c$ and one can write the grand potential as

$$\begin{aligned}\Omega_G &= -k_B T \sum_{\varepsilon} \ln(1 + ze^{-\beta\varepsilon}) \\ &\approx -k_B T \int_0^\infty D(\varepsilon) \ln(1 + ze^{-\beta\varepsilon}) d\varepsilon \\ &= -k_B T \int_0^\infty 4\pi D(k) \ln(1 + ze^{-\beta\varepsilon(k)}) dk \\ &= -\frac{V}{3\pi^2} \int_0^\infty \frac{1}{z^{-1}e^{\beta\varepsilon(k)} + 1} k^3 \frac{d\varepsilon}{dk} dk \\ &\approx -\frac{V\hbar c}{12\pi^2} [k_F^4 - k_c^2 k_F^2 + \dots],\end{aligned}$$

where we have applied partial integration. The grand potential gives the pressure

$$p = -\left(\frac{\partial\Omega_G}{\partial V}\right)_{T,N} = \frac{\hbar c}{12\pi^2} [k_F^4 - k_c^2 k_F^2 + \dots].$$

This should obey the previous balance condition, resulting in

$$\frac{8\pi}{3}Gm_p^2\rho^2 R^2 = \frac{\hbar c}{12\pi^2} [k_F^4 - k_c^2 k_F^2 + \dots].$$

By substituting $k_F^3 = 3\pi^2\rho$, we obtain

$$\left(\frac{M}{M_c}\right)^{2/3} = 1 - \left(\frac{R}{R_c}\right)^2 \left(\frac{M_c}{M}\right)^{2/3},$$

where

$$\begin{aligned}M_c &= m_p \left(\frac{9\pi}{512}\right)^{1/2} \left(\frac{\hbar c}{Gm_p^2}\right)^{3/2} \approx 0.52 \cdot 10^{57} m_p \\ R_c &= \frac{\hbar}{mc} \left(\frac{9\pi}{8}\right)^{1/3} \left(\frac{M_c}{m_p}\right)^{1/3} \approx 4700 \text{ km}.\end{aligned}$$

For the radius of the star we obtain

$$R = R_c \left(\frac{M}{M_c}\right)^{1/3} \left[1 - \left(\frac{M}{M_c}\right)^{1/3}\right].$$

We see that the white dwarf has the maximum mass $M = M_c$, which it cannot exceed without collapsing to a neutron star or a black hole. This is called the *Chandrasekhar limit*, and even more detailed calculations give $M_c \approx 1.44M_\odot$, where M_\odot is the mass of the Sun.

9. Entropy, Information and Arrow of Time

Before the discussion on the main topic of this chapter, a word of caution is in order. The discussion in this chapter will be subjective and, at some places, somewhat speculative. Therefore, the contents should be read with an open mind and with a scientific doubt.

Entropy is perhaps the most influential, but at the same time controversial, concept that arises from statistical physics. So far, we have encountered three different ways to determine the entropy. It is hard to tell exactly *what entropy is*, but we should nevertheless try to develop an *intuition* for it⁴⁶. In this chapter, we will review the definitions, and also look closer at some of the myriad of implications that result.

9.1 Thermodynamic Entropy: Irreversibility and Arrow of Time

One really cannot overvalue the second law of thermodynamics. Heat cannot spontaneously flow from cold regions to hot regions. It is impossible to build a perpetual motion machine. After P. W. Bridgman, there have been nearly as many formulations of the second law as there have been discussions of it. But, possibly the deepest philosophical message is two-fold. The second law introduces the concept of entropy, that gives us the way to characterize all processes in nature; In an isolated system, only such processes are allowed that increase the entropy⁴⁷. As such, this is among the most fundamental statements we have about nature. But, it tells us more. It points us the *arrow of time*, separates the *future* from the *past*.

Consider the previous example on the mixing of two ideal gases (see p. 31). As we open the valve, the constraints for the two gases are changed. In the statistical physics language we have learned to use, the number of accessible microstates becomes increased and the systems are not in equilibrium. As the new equilibrium is achieved, one notices that the thing that characterizes this process is the second law, and the quantity that changes is the entropy⁴⁸.

The increase in entropy calls for *irreversibility*. For example, it is impossible to restore the initial equilibrium of the two gases, both in separate compartments, simply by closing the valve. In general, the ever-growing entropy tells us that one direction of time is preferred over the other. It defines, at least in the thermodynamical sense, the future

⁴⁶The same statement holds also for *energy* and *time*. Both of them are hard to define exactly, but we have learned to be fluent in using them to describe phenomena around us. This chapter is intended to be give guidelines to do the same in the case of entropy.

⁴⁷Strictly, processes that do not decrease the entropy.

⁴⁸As a side note, in this example it is important to notice the inadequacy of the thermodynamical definition of entropy

$$\Delta S \geq \frac{Q}{T},$$

where Q is the heat flow into the system. The circumstances we set up for the system were such that heat is not exchanged between the gas mixture and the surroundings, thus the change in entropy cannot result from such a process.

as the *direction* of time in which the entropy increases.

Loschmidt paradox

Let us dwell on this last statement. Consider a film first played forwards and then backwards. Normally, we can separate between the directions. A classic example is the falling of a drinking glass to the floor and it's shattering to small fragments. The reverse (the re-organization of the fragments) seems irrational to us. Our minds are formed so that we can separate between the directions of time in irreversible processes. But, the laws of nature are *time-reversible*. The motions of atoms look reasonable whether the time is running backwards or forwards. It is then quite a mystery how the irreversibility emerges from the reversible behaviour of things. Johann Loschmidt, a contemporary of Boltzmann, was the first to see this contradiction between the second law and the reversible laws of nature, hence it was named the *Loschmidt paradox*.

One attempt to resolve this issue is to state that the whole Universe was in a low entropy state right after the Big Bang. This way the familiar definitions of the past and the future come about in terms of the increasing entropy in the Universe. As a result, we would be witnessing a process where the Universe is moving towards equilibrium as time increases. In this equilibrium, in the words of Helmholtz (1854), all energy will become heat, all temperatures will become equal, and everything will be condemned to a state of eternal rest. This ultimate fate of the Universe has been named as the *heat death of the Universe*.

9.2 Disorder

Other way to interpret the entropy is as a measure of the disorder in the system. Consider again the entropy of mixing. We obtained (see Eq. (71) and the following discussion) that for the mixture of two gases of indistinguishable particles, the entropy of mixing can be written as (we assume that in the beginning $V_A = V_B = V/2$ and $N_A = N_B = N/2$)

$$\begin{aligned} \Delta S_{\text{mix}} &= S_{\text{mixed}} - S_{\text{unmixed}} \\ &= 2k_B \ln \left(\frac{V^{N/2}}{(N/2)!} \right) - 2k_B \ln \left(\frac{(V/2)^{N/2}}{(N/2)!} \right) \\ &= Nk_B \ln 2. \end{aligned}$$

We thus see that entropy increases by $k_B \ln 2$ every time we insert a particle into either of the two compartments, provided that we do not look which compartment we choose. This is a simplest possible instance of the *counting* entropy

$$S_{\text{counting}} = k_B \ln(\# \text{ of configurations}).$$

In the case of mixing entropy, each particle has two possible compartments, so N particles can be arranged into 2^N different configurations. Generally, if there are $m \geq 2$ possibilities for a particle, the added entropy is accordingly $k_B \ln m$. This idea will be exploited in the last section of this chapter, when we discuss about information entropy.

The above discussion holds in cases where there exists discrete amount of choices for the new particle. Discrete choice arises naturally in quantum statistics where the states are quantized. Generally, as we have already discussed, the equilibrium entropy measures the logarithm of the number of different states the system can be in. For example, the previous equilibrium statistical mechanics (Boltzmann) entropy

$$S = k_B \ln \Omega(E) \quad (93)$$

measures the number of microstates with macroscopic energy E , with phase-space volume h^{3N} allocated to each state.

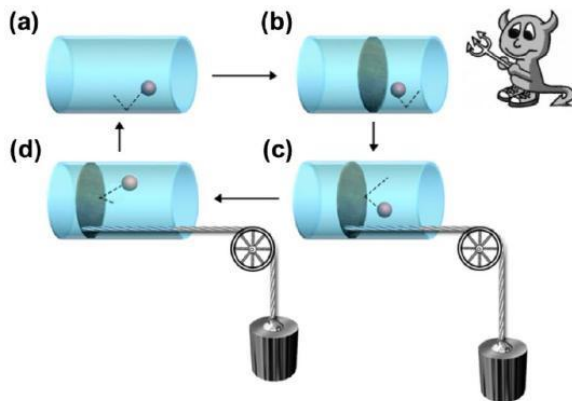
Maxwell's Demon

Is it possible to extract work from the mixing entropy? Consider a mixture of two ideal gases, A and B , and a clever creature, the *Maxwell's demon*, that works at the valve between two compartments of the container. Suppose that the demon is able to control the valve in such a way that it lets the A -molecules to pass through the valve to the right side of the valve, and the B -molecules to the left. Additionally, she would prevent the opposite processes. As a result, each transfer of a molecule through the valve decreases the entropy by $k_B \ln 2$, apparently violating the second law. The thing that is overlooked is that the demon itself consumes entropy, which then balances the entropy decrease as a whole. We will return to this in the following section.

Szilard's engine

But first, let us return to the initial goal of the demon, the extraction of work from entropy. Consider a simplified version of the Maxwell's demon, first presented in 1929 by the Hungarian L. Szilard. This was the first time that the significance of information in physics⁴⁹ was pointed out.

The Szilard's idea was to study one-molecule ideal gas in volume V . Note, that we can consider the single molecule as an ideal gas by forming an ensemble of one-molecule gases and calculating the averages of various quantities as if it was an ideal gas with a large number of molecules.



⁴⁹We will go deeper into this connection in the following section.

(Taken from K. Maruyama et al., Rev. Mod. Phys **81**, 1 (2009).)

The process goes as follows. (a) we insert a partition that divides the volume in two equal compartments. (b) the demon measures the position of the molecule, and *records* the measurement (whether the molecule is on the right or on the left of the partition). (c) the demon connects a load to the partition, working now as a piston, on the side where she has measured the molecule. (d) by keeping the system at constant temperature, the demon can let the gas expand isothermally and quasistatically (no change in internal energy), resulting in work

$$W = Q = \int_{V/2}^V p dV = k_B T \int_{V/2}^V V^{-1} dV = k_B T \ln 2.$$

Thus, we see that the Szilard's demonic engine extracts heat Q from the surroundings and converts it completely into mechanical work, apparently violating the second law. The explanation of this requires further insights on our concept of entropy.

9.3 Information Entropy

The Szilard's realisation of Maxwell's demon leads us to our last and the most profound interpretation of the entropy, as a measure of our *ignorance* about a system. As we have seen, in the entropy of mixing we add $k_B \ln 2$ to our ignorance of the system every time we add a particle into the container without looking where it went. Similarly, a cycle in the Szilard's engine transfers an amount $k_B T \ln 2$ of heat from surroundings into work, provided that one knows on which side of the partition the molecule is. Generally, this implies that the entropy is not a property of the system, but is related our knowledge about the system. Additionally, the second law then tells us that in the equilibrium, our knowledge becomes minimal, reflected in the equal probabilities of finding the system in one of its allowed states.

Shannon entropy

In 1948, C. Shannon laid the foundation to *information theory* by studying the information content of random variables. Interestingly, information and physics seem to have deep connections. Consider a random variable X with the possible outcomes $\{x_1, x_2, \dots, x_n\}$ ⁵⁰. Intuitively, we assume that the information content $I(X = x_i)$ related to an outcome x_i depends only on the probability of the event. By definition, we require that I is positive and additive. Also, if we assume that we have another random (independent on X) variable Y , the information content of measuring x_i and y_j simultaneously should be the sum of the individual information contents, i.e. $I(X \cap Y) = I(X) + I(Y)$. By taking these into account, one unavoidably ends up with the *information content* function

$$I(x_i) = -\log P(x_i),$$

⁵⁰We assume here that X is a discrete random variable. Generalization to continuous variables is straightforward.

where $P(x_i)$ is the probability of observing the outcome x_i . The base of the log determines the units of information. In the base of 2, the units are *bits*, and in the base of e the units are *nats*.

The average information content

$$S = \langle I(x_i) \rangle = - \sum_i P(x_i) \ln P(x_i)$$

is the cornerstone of the theory of information. Interestingly, the average information content is exactly of the form of Gibbs entropy

$$S = -k_B \langle \ln p_i \rangle = -k_B \sum_i p_i \ln p_i, \quad (94)$$

which was our generalization of the Boltzmann entropy! Thus, it was named as the *Shannon entropy*, according to its founder. Note, that in information theory the unit of entropy does not have the historical burden of having the units J/K .

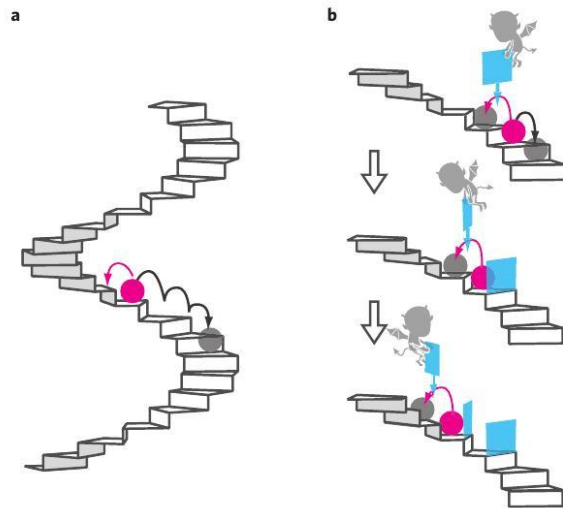
In 1951, L. Brillouin became the first to treat the entropy and information on the same footing. The Gibbs entropy can be seen as the maximum amount of information one can obtain from the system by making a measurement. To clarify this, consider an isolated system initially with entropy S_i . Assume then that we acquire information about the state of the system by making a measurement. The gain in information is linked to the change in the entropy by

$$I^{\text{meas}} = S_i - S_f > 0,$$

where S_f is the physical entropy after the measurement. This means that the gain in information decreases physical entropy. In the Szilard's engine, this occurs when we go from (a) to (b), and we see that the entropy in the system has turned into information inside the demon's head. This implies the *duality* of entropy, i.e. its having both thermodynamic and information theoretic aspects. If further measurements are not made, we see that the second law implies unavoidable decrease in the information about the system. Thus indeed, entropy can be seen as the lack of knowledge about the system.

Exorcized demon

What about our demon then? Well, as we have seen, the decrease in the entropy of Szilard's engine leads to an increase demon's knowledge about the location of the molecule. Further on, this information can be then converted into useful work, as shown recently by S. Toyabe et al. (2010).



(Taken from S. Toyabe et al., Nat. Phys. **6**, 988 (2010).)

In their experiment, Toyabe et al. studied system equivalent to particle stochastically jumping in a spiral-staircase due to the thermal fluctuations⁵¹. In order to prevent downward jumps, the demon places a block behind the particle whenever she observes an upward jump. Thus, the particle climbs up the stairs without direct energy injection.

Naturally, the energy for useful work is extracted from the heat bath setting up the temperature in the Szilard's engine. As we have already discussed, this apparently violates the second law because it does a perfect conversion of heat Q into work W . Szilard's original idea was that this energy is returned to the heat bath by the measurement device. However, Bennett showed in 1982 that the measurement can be done reversibly, i.e. without any change in entropy, at least in principle.

But, there is still one thing that we have overlooked. At the stage where the work is done, the gas expands isothermally and the demon loses her information about the location of the molecule. This is a *logically irreversible* process. In general, logical reversibility means injective, i.e. one-to-one, correspondence between logical and physical states. When we process information the reversibility guarantees that no information is lost. But, *erasure* of information is irreversible! The mapping is not one-to-one, instead it maps all physical states into the same logical state, which we refer to as the *standard state*. It corresponds to a blank sheet of paper. Landauer argued in his *erasure principle* that the logical irreversibility must involve dissipation. The minimum amount of dissipation equals the energy

$$k_B T \log 2$$

per bit. This being true, the demon's efforts to violate the second law have come to an end. In fact, just recently A. Bérut et al. (2012) have demonstrated experimentally that the Landauer limit of the information erasure can be reached but not exceeded.

⁵¹We will discuss more about fluctuations later when we consider non-equilibrium statistics.

Quantum entropy

What about the Universe then? Is it erasing information at the Landauer rate per bit towards the unavoidable heat death, or is it computing reversibly its following step? It is important to realise that non-zero entropies of isolated systems are related to our ignorance about the state of the system. The von Neumann definition,

$$S = -k_B \text{Tr}(\hat{\rho} \ln \hat{\rho}), \quad (95)$$

is the fundamental version of the entropy, since it is quantum mechanical, as is the nature. One should notice, however, that the von Neumann entropy is zero for pure states and non-zero for mixed states, but in both cases independent in time for isolated systems (exercise). This means that if we can follow the microscopic evolution of a closed system, our ignorance of the state does not change. The ignorance arises only from our lack of knowledge on how the system was build.

Consider then the whole Universe. By definition, it is isolated, and regardless of whether the Universe is in a pure or in a mixed state, its entropy has to be a constant! This makes the heat death of the Universe seem considerably exaggerated. However, it is impossible to determine the state of the whole Universe at microscopic level, let alone calculate its future evolution from the Schrödinger equation. Therefore, the constancy of its entropy should be out of reach of observations.

Consider then us making observations, which can only cover a tiny fraction of the whole. Moreover, since the small region of which we have knowledge, or want to study, is unavoidably interacting with a part that is unknown to us, we cannot accurately call it isolated. Thus, if we measure an increase in the thermodynamical entropy (in accordance with the second law) it unavoidably has to be consequence of a decrease in entropy of the environment. This is the only way to keep the total entropy of the Universe constant.

Finally, when we derived the H-theorem in the exercises we introduced the transition rates between different states. These would not be present in an isolated system, and can be seen as resulting from the (weak) coupling to the environment. In addition, the equations are *coarse-grained*, i.e. we have wiped out the quantum coherence between the states. As a result, we see that the apparent irreversibility that we observe in nature arises from our lack of exact knowledge on the coupling between the system and the rest of the Universe, and on the coherences between the quantum states, thus resolving Loschmidt's paradox.

References

This chapter is partly based on the following articles:

- K. Maruyama et al., Rev. Mod. Phys **81**, 1 (2009).
- S. Toyabe et al., Nat. Phys. **6**, 988 (2010).
- A. Bérut et al., Nature **483**, 187 (2012).
- S. Popescu et al., Nat. Phys. **2**, 754 (2006).

10. Interacting Systems

The discussion so far has been about systems consisting of non-interacting entities. For a real connection between the theory and experiments, one has to include the interparticle interactions operating in the system. The interactions are the reason why the equations in many-body physics become intractable. One possibility to include the interactions is in terms of the so-called *cluster expansion*. In low-density gases, the non-interacting system can serve as an approximation. The physical quantities can then be written in terms of series expansions in terms of two-particle potentials, which provide the corrections to the main terms arising from the non-interacting solution.

Another approach we are going to discuss is that of *second quantization*.

10.1 Cluster Expansions

Mayer et al. (1937 onwards) developed a systematic method to carry out a cluster expansion in real gases obeying classical statistics. The generalization to interacting quantum gases was initiated by Kahn and Uhlenbeck, and completed by Lee and Yang. The interplay between quantum statistical effects and the effects due to interactions leads to involved mathematics and will be omitted here.

Consider a gas obeying the Hamiltonian

$$H = \sum_{i=1}^N \frac{p_i^2}{2m} + \sum_{i<j} u_{ij},$$

where $u_{ij} = u(r_{ij})$ is the interaction potential between particles i and j , and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between the particles. One example of such interaction is the Lennard-Jones potential on p. 13.

In the classical limit, the canonical partition function is

$$\begin{aligned} Z(T, V, N) &= \frac{1}{N!} \int \frac{d\omega}{h^{3N}} e^{-\beta H} \\ &= \frac{1}{N! h^{3N}} \left[\int d^3N p e^{-\beta \sum_{i=1}^N \frac{p_i^2}{2m}} \right] \left[\int d^3N r e^{-\beta \sum_{i<j} u_{ij}} \right] \\ &= \frac{1}{N! \lambda_T^{3N}} Q_N(T, V), \end{aligned}$$

where we have carried out the momentum integration and defined the *configuration integral*

$$\begin{aligned} Q_N(T, V) &= Q(T, V, N) = \int d^3N r e^{-\beta \sum_{i<j} u_{ij}} \\ &= \int d^3N r \prod_{i<j} e^{-\beta u_{ij}}, \end{aligned}$$

where the product is taken over all pairs of particles (ij) (there are $N(N-1)/2$ such pairs). For ideal gas, $u_{ij} = 0$ and

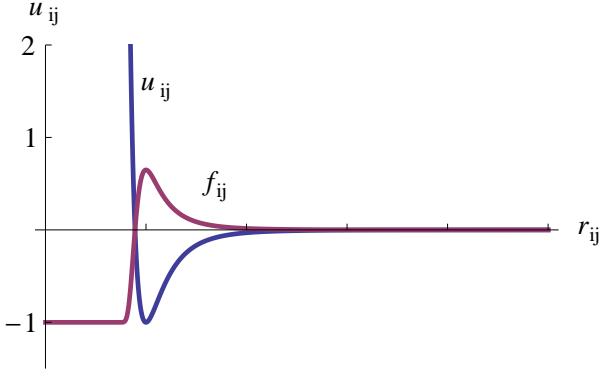
$$Z(T, V, N) = \frac{V^N}{N! \lambda_T^{3N}}$$

which is the same as Eq. (57) obtained previously.

Cluster expansion

We would like to calculate the configuration integral Q_N . For this, we expand it in terms of *Mayer functions*

$$f_{ij} = f(r_{ij}) = e^{-\beta u_{ij}} - 1.$$



We see that the Mayer functions f_{ij} are bounded everywhere and have the same range as the potential u_{ij} (we have used the Lennard-Jones potential). For the evaluation, we make the expansion

$$Q_N = \int d^3r_1 \cdots d^3r_N \prod_{i < j} (1 + f_{ij}).$$

Ursell showed that the grand partition function can be written in terms of *cumulant expansion* as

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} \frac{1}{N! \lambda_T^{3N}} e^{\beta N \mu} Q_N \\ &= \exp \left[\sum_{l=1}^{\infty} \frac{1}{l! \lambda_T^{3l}} e^{\beta l \mu} \int d^3r_1 \cdots d^3r_l U_l(\mathbf{r}_1, \dots, \mathbf{r}_l) \right], \end{aligned}$$

where $U_l(\mathbf{r}_1, \dots, \mathbf{r}_l)$ are called *cluster functions* or *Ursell functions*. The cluster functions can be obtained by expanding the cumulant expansion in powers of $\lambda_T^{-3} \exp(\beta \mu)$, leading into

$$\begin{aligned} U_1(\mathbf{r}_1) &= 1 \\ U_2(\mathbf{r}_1, \mathbf{r}_2) &= f_{12} \\ U_3(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= f_{12}f_{13} + f_{12}f_{23} + f_{13}f_{23} + f_{12}f_{13}f_{23} \\ &\vdots \end{aligned}$$

We thus see that U_l depends on all connected *clusters* of l particles.

Virial expansion

We define the *cluster integral* as

$$b_l = \frac{1}{l! V} \int d^3r_1 \cdots d^3r_l U_l(\mathbf{r}_1, \dots, \mathbf{r}_l).$$

Thus, the grand potential becomes explicitly proportional to V :

$$\Omega_G(T, V, \mu) = -k_B T \ln Z_G(T, V, \mu) = -V k_B T \sum_{l=1}^{\infty} \frac{b_l(T) e^{\beta l \mu}}{\lambda_T^{3l}}.$$

The grand potential allows us to calculate thermodynamic quantities, such as the pressure

$$p = - \left(\frac{\partial \Omega_G}{\partial V} \right)_{T, \mu} = k_B T \sum_{l=1}^{\infty} \frac{b_l(T) e^{\beta l \mu}}{\lambda_T^{3l}}$$

and the average density of particles

$$\rho = - \frac{1}{V} \left(\frac{\partial \Omega_G}{\partial \mu} \right)_{T, V} = \sum_{l=1}^{\infty} \frac{l b_l(T) e^{\beta l \mu}}{\lambda_T^{3l}}.$$

By recalling the *virial expansion* (39)

$$pV = k_B T [\rho + B_2(T) \rho^2 + B_3(T) \rho^3 + \cdots], \quad (96)$$

we can show that the *virial coefficients* B_k are now functions of the cluster integrals, e.g. (exercise)

$$\begin{aligned} B_2(T) &= -b_2(T) = \frac{1}{2} \int d^3r \left[1 - e^{-\beta u(r)} \right] \\ B_3(T) &= 4b_2^2 - 2b_3 \\ B_4(T) &= -20b_2^3 + 18b_2b_3 - 3b_4 \\ &\vdots \end{aligned}$$

van der Waals equation

For very dilute gases, two-body clusters give the dominant contribution for the interactions. The interactions between atoms in real gases can be well approximated by the Lennard-Jones potential, as we have discussed. Accordingly (see p. 13), we assume that

- the interaction is *hard core*, i.e. strongly repulsive when the distance between atoms $r \lesssim r_m$.
- on the average, the interaction is attractive when $r \gtrsim r_m$, but weak compared with the temperature, i.e. $\beta u(r) \ll 1$.

Thus,

$$e^{-\beta u(r)} = \begin{cases} 0 & \text{when } r \lesssim r_m \\ 1 - \beta u(r) & \text{when } r \gtrsim r_m \end{cases}$$

The second virial coefficient becomes

$$\begin{aligned} B_2 &= 2\pi \int_0^{\infty} dr r^2 \left[1 - e^{-\beta u(r)} \right] \\ &\approx b - \frac{a}{T}, \end{aligned}$$

where

$$\begin{aligned} b &= \frac{2\pi}{3} r_m^3 \\ a &= -2\pi \int_{r_m}^{\infty} dr r^2 u(r) > 0. \end{aligned}$$

In the above approximation, we have ended up with the van der Waals equation.

10.2 Second Quantization

Another way to introduce the interactions to the many-body systems is by a formalism known as the *second quantization*. In short, the distinction between the first and the second quantization is equivalent to that between the Hermite polynomials, and the creation and annihilation operators together with number (Fock) states in the single harmonic oscillator problem. The first quantization can be seen as the quantization of the motion of the elementary particles. The second quantization quantizes also the electromagnetic field, and even the particles. It is the starting point of the microscopic (BCS-)theory of superconductivity⁵², the quantum field theories and the quantum electrodynamics, the theory of interaction between matter and radiation.

We have already introduced most of the essential concepts. In a many-body system, the wave functions are (anti-)symmetric superpositions of products of single-particle wave functions (see Eqs. (80-81)). The first step in second quantization is to change from the wave function representation into the abstract number basis with the definition (82). This makes the calculations more convenient since the number basis does not carry along the unphysical information on the particle locations, which are not sensible quantities since the particles are indistinguishable.

Next, the *vacuum state* is defined as

$$|0, 0, \dots\rangle \equiv |0\rangle.$$

This is a unique state that contains no particles. With the application of the creation operators, one can write an arbitrary number state in terms of the vacuum state as

$$|n_1, n_2, \dots\rangle = \frac{1}{\sqrt{n_1!n_2!\dots}} (\hat{a}_1^\dagger)^{n_1} (\hat{a}_2^\dagger)^{n_2} \dots |0\rangle.$$

We have also defined before the number operator of the i^{th} single-particle state as

$$\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i.$$

The number states are eigenstates of the number operators

$$\hat{n}_i |n_1, \dots, n_i, \dots\rangle = n_i |n_1, \dots, n_i, \dots\rangle,$$

where the eigenvalue n_i is the number of particles in the corresponding single-particle state. Recall, that for bosons $n_i = 0, 1, 2, \dots$, and for the fermions $n_i = 0, 1$. The operator for the total number of particles in the system is given by

$$\hat{N} = \sum_i \hat{n}_i.$$

The second quantization equals to writing every operator in terms of the creation and annihilation operators⁵³. Especially, we assume that the interacting many-body Hamiltonian operator is

$$\hat{H} = \sum_{i=1}^N \hat{H}_0(\mathbf{r}_i) + \sum_{i<j} \hat{U}(\mathbf{r}_i - \mathbf{r}_j)$$

where \hat{H}_1 is the non-interacting single-particle Hamiltonian, and \hat{U} is the potential operator for the two-particle interactions. By applying the second quantization, we can write the Hamiltonian as⁵⁴

$$\hat{H} = \sum_{ij} \langle i | \hat{H}_0 | j \rangle \hat{a}_i^\dagger \hat{a}_j + \frac{1}{2} \sum_{ijkl} \langle ij | \hat{U} | kl \rangle \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k.$$

In the above, we have defined the matrix elements

$$\langle i | \hat{H}_0 | j \rangle = \int d^3r \psi_i^*(\mathbf{r}) \hat{H}_0(\mathbf{r}) \psi_j(\mathbf{r})$$

$$\langle ij | \hat{U} | kl \rangle = \int \int d^3r \int d^3r' \psi_i^*(\mathbf{r}) \psi_j^*(\mathbf{r}') \hat{U}(\mathbf{r}' - \mathbf{r}) \psi_l(\mathbf{r}') \psi_k(\mathbf{r}),$$

where ψ_i are the single-particle wave-functions defined in Eq. (78).

For convenience, one often defines the so-called *field operators*

$$\hat{\psi}(\mathbf{r}) = \sum_i \psi_i(\mathbf{r}) \hat{a}_i,$$

$$\hat{\psi}^\dagger(\mathbf{r}) = \sum_i \psi_i^*(\mathbf{r}) \hat{a}_i^\dagger.$$

This way the Hamiltonian operator in the second quantization formalism can be written as

$$\hat{H} = \int d^3r \hat{\psi}^\dagger(\mathbf{r}) \hat{H}_0(\mathbf{r}) \hat{\psi}(\mathbf{r})$$

$$+ \frac{1}{2} \int \int d^3r d^3r' \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}^\dagger(\mathbf{r}') \hat{U}(\mathbf{r}' - \mathbf{r}) \hat{\psi}(\mathbf{r}') \hat{\psi}(\mathbf{r}).$$

We notice that this resembles an expectation value of the Hamiltonian operator, but instead of wave functions we take calculate it with respect to the field operators. Hence the name second quantization.

The field operators obey the (anti-)commutation relations for all \mathbf{r} and \mathbf{r}'

$$[\hat{\psi}(\mathbf{r}), \hat{\psi}(\mathbf{r}')]_{\pm} = [\hat{\psi}^\dagger(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')]_{\pm} = 0$$

$$[\hat{\psi}(\mathbf{r}), \hat{\psi}^\dagger(\mathbf{r}')]_{\pm} = \delta(\mathbf{r} - \mathbf{r}'),$$

where minus (plus) sign stands for (anti-)commutator for bosons (fermions). A straightforward calculation gives the number operator in terms of the field operators as

$$\hat{N} = \int d^3r \hat{\psi}^\dagger(\mathbf{r}) \hat{\psi}(\mathbf{r}).$$

One can also show that

$$[\hat{H}, \hat{N}] = 0.$$

In general, an arbitrary many-body operator

$$\hat{A} = \sum_{i=1}^N \hat{A}(\mathbf{r}_i),$$

⁵²Further information, see Superconductivity course.

⁵³Recall, that creation and annihilation operators obey the (anti-)commutation relations for bosons (fermions).

⁵⁴We omit the lengthy derivation here. An interested reader can find it from Fetter and Walecka: *Quantum Theory of Many-Particle Systems*.

can be written in second quantization as

$$\begin{aligned}\hat{A} &= \sum_{ij} \langle i | \hat{A} | j \rangle \hat{a}_i^\dagger \hat{a}_j \\ &= \int d^3r \hat{\psi}^\dagger(\mathbf{r}) \hat{A}(\mathbf{r}) \hat{\psi}(\mathbf{r}).\end{aligned}$$

Plane wave expansion

If we assume that there is no external potential, the non-interacting part of the Hamiltonian is of the form

$$\hat{H}_0 = -\frac{\hbar^2}{2m} \nabla^2.$$

In the case of weak interactions, a good choice for the single-particle states are plane waves

$$\psi_{\mathbf{k}}(\mathbf{r}) = \frac{1}{\sqrt{V}} e^{i\mathbf{k} \cdot \mathbf{r}}.$$

This gives the kinetic energy matrix elements (now $i \rightarrow \mathbf{k}'$ and so on)

$$-\frac{\hbar^2}{2m} \langle \mathbf{k}' | \nabla^2 | \mathbf{k} \rangle = \frac{1}{V} \int d^3r e^{-i\mathbf{k}' \cdot \mathbf{r}} \frac{\hbar^2 k^2}{2m} e^{i\mathbf{k} \cdot \mathbf{r}} = \varepsilon_k \delta_{\mathbf{k}\mathbf{k}'},$$

where $\varepsilon_k = \frac{\hbar^2 k^2}{2m}$. Recall, that we have applied the periodic boundary conditions in the normalization of the plane wave, which implies that the wave vectors \mathbf{k} are quantized, and thus we can apply the Kronecker delta function.

Correspondingly, the matrix elements of the potential energy operator are

$$\begin{aligned}\langle \mathbf{k}_1 \mathbf{k}_2 | \hat{U} | \mathbf{k}_3 \mathbf{k}_4 \rangle &= \frac{1}{V^2} \int \int d^3r d^3r' e^{-i(\mathbf{k}_1 - \mathbf{k}_3) \cdot \mathbf{r}} \\ &\quad \times U(\mathbf{r}' - \mathbf{r}) e^{i(\mathbf{k}_4 - \mathbf{k}_2) \cdot \mathbf{r}'} \\ &= \frac{1}{V} \int d^3r e^{-i\mathbf{k} \cdot \mathbf{r}} U(\mathbf{r}) \\ &\equiv \frac{1}{V} U(\mathbf{k}),\end{aligned}$$

where

$$U(\mathbf{k}) = \int d^3r U(\mathbf{r}) e^{-i\mathbf{k} \cdot \mathbf{r}}$$

is the Fourier transformation whose inverse is defined as

$$U(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{k}} U(\mathbf{k}) e^{i\mathbf{k} \cdot \mathbf{r}}.$$

Note, that the matrix elements give the probability amplitudes for the collision between single particles in momentum states \mathbf{k}_3 and \mathbf{k}_4 . In the collision, total momentum has to be conserved, i.e. the momentum in the scattered states has to be equal to that of the colliding states

$$\mathbf{k}_1 + \mathbf{k}_2 = \mathbf{k}_3 + \mathbf{k}_4.$$

This has been applied in the last equality. In the above,

$$\mathbf{k} = \mathbf{k}_2 - \mathbf{k}_4 = -(\mathbf{k}_1 - \mathbf{k}_3)$$

is the momentum transfer during the collision.

In the case of half-integer spin particles, one has to label the annihilation and creation operators also with the z -component of the spin. Finally, the Hamiltonian can be written in the second quantization as

$$\begin{aligned}\hat{H} &= \sum_{\mathbf{k}\sigma} \varepsilon_k \hat{a}_{\mathbf{k}\sigma}^\dagger \hat{a}_{\mathbf{k}\sigma} + \frac{1}{2V} \sum_{\mathbf{k}_1\sigma} \sum_{\mathbf{k}_2\lambda} \sum_{\mathbf{k}_3} \sum_{\mathbf{k}_4} U(\mathbf{k}_2 - \mathbf{k}_4) \\ &\quad \times \delta_{\mathbf{k}_1 + \mathbf{k}_2, \mathbf{k}_3 + \mathbf{k}_4} \hat{a}_{\mathbf{k}_1\sigma}^\dagger \hat{a}_{\mathbf{k}_2\lambda}^\dagger \hat{a}_{\mathbf{k}_4\lambda} \hat{a}_{\mathbf{k}_3\sigma}.\end{aligned}\quad (97)$$

To simplify the notations, we have assumed that the potential does not change the spins of the particles. In such cases, one should include also the conservation of spin into the calculation of the matrix elements of the interaction.

Low-temperature behaviour of interacting Bose-gas

The Fourier transform of the potential defines the so-called *scattering amplitude*⁵⁵

$$a(\mathbf{k}) = \frac{m}{4\pi\hbar^2} U(\mathbf{k}),$$

which defines the length scale of the scattering process. Now, we are interested in the low-temperature behaviour, meaning that we assume

$$a \ll \lambda_T \quad \text{and} \quad \rho a^3 \ll 1.$$

In this limit, the particles are slow, $k \approx 0$ and the *scattering length* is

$$a = \frac{mU_0}{4\pi\hbar^2} \quad \text{where} \quad U_0 = \int d^3r U(\mathbf{r}).$$

For bosons, we can omit the spin-summations in Eq. (97) and we obtain in the low-temperature limit (recall $\mathbf{k} = \mathbf{k}_2 - \mathbf{k}_4 = -(\mathbf{k}_3 - \mathbf{k}_1) = 0$)

$$\begin{aligned}\hat{H} &= \sum_{\mathbf{k}} \varepsilon_k \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} + \frac{2\pi a \hbar^2}{mV} \left[\sum_{\mathbf{k}} a_{\mathbf{k}}^\dagger a_{\mathbf{k}}^\dagger a_{\mathbf{k}} a_{\mathbf{k}} \right. \\ &\quad \left. + \sum_{\mathbf{k} \neq \mathbf{k}'} (a_{\mathbf{k}}^\dagger a_{\mathbf{k}'}^\dagger a_{\mathbf{k}} a_{\mathbf{k}'} + a_{\mathbf{k}'}^\dagger a_{\mathbf{k}}^\dagger a_{\mathbf{k}'} a_{\mathbf{k}}) \right].\end{aligned}$$

By employing the properties of the annihilation and creation operators, one can write the Hamiltonian into form (exercise)

$$\hat{H} = \sum_{\mathbf{k}} \varepsilon_k \hat{n}_{\mathbf{k}} + \frac{2\pi a \hbar^2}{mV} \left[2\hat{N}^2 - \hat{N} - \sum_{\mathbf{k}} \hat{n}_{\mathbf{k}}^2 \right].$$

The energy eigenvalues are thus of form

$$E\{n_{\mathbf{k}}\} \simeq \sum_{\mathbf{k}} n_{\mathbf{k}} \varepsilon_k + \frac{2\pi a \hbar^2}{mV} \left[2N^2 - \sum_{\mathbf{k}} n_{\mathbf{k}}^2 \right],$$

where the dependence of the eigenvalue on the distribution between different wave vectors has been explicitly implicated.

⁵⁵Compare with Quantum Mechanics II.

For the *ground state*, the distribution set becomes

$$n_{\mathbf{k}} = \begin{cases} N & \text{for } \mathbf{k} = 0 \\ 0 & \text{for } \mathbf{k} \neq 0. \end{cases}$$

This gives the ground state energy

$$E_0 = \frac{2\pi a \hbar^2 N^2}{mV},$$

pressure

$$p_0 = - \left(\frac{\partial E_0}{\partial V} \right)_N = \frac{2\pi a \hbar^2 \rho^2}{m},$$

and chemical potential

$$\mu_0 = \left(\frac{\partial E_0}{\partial N} \right)_V = \frac{4\pi a \hbar^2 \rho}{m}.$$

We see that the ground state properties of the interacting Bose-gas at low temperatures are given by the scattering length and density.

Generally, the thermodynamics can be calculated from the partition function

$$\begin{aligned} Z(T, V, N) &= \sum_{\{n_{\mathbf{k}}\}} \exp(-\beta E\{n_{\mathbf{k}}\}) \\ &= \sum_{\{n_{\mathbf{k}}\}} \exp \left[-\beta \left(\sum_{\mathbf{k}} \varepsilon_{\mathbf{k}} n_{\mathbf{k}} + \frac{2\pi a \hbar^2 N^2}{mV} \left(2 - \frac{n_0^2}{N^2} \right) \right) \right]. \end{aligned}$$

The lowest order corrections to the ground state results, that arise from the low but finite temperatures, can be obtained by recalling the Eq. (90) for the ideal gas:

$$\begin{aligned} \frac{n_0}{N} &= 1 - \frac{\lambda_c^3}{\lambda_T^3} \quad \text{where } \lambda_c = \left(\frac{g_{3/2}(1)}{\rho} \right)^{1/3} \\ &= 1 - \frac{v}{v_c} \quad \text{where } v = \frac{1}{\rho} \quad v_c = \frac{\lambda_T^3}{g_{3/2}(1)}. \end{aligned}$$

By using this as an estimate for the relative population in the ground state, one obtains the Helmholtz free energy

$$\begin{aligned} A(T, V, N) &= -k_B T \ln Z \\ &= A_{\text{ig}}(T, V, N) + \frac{2\pi a \hbar^2 N}{m} \left(\frac{1}{v} + \frac{2}{v_c} - \frac{v}{v_c^2} \right), \end{aligned}$$

where A_{ig} is the free energy of the ideal Bose gas. The thermodynamics is obtained from the free energy

$$\begin{aligned} p &= - \left(\frac{\partial A}{\partial V} \right)_{T, N} = - \left(\frac{\partial(A/N)}{\partial v} \right)_T = p_{\text{id}} + \frac{2\pi a \hbar^2}{m} \left(\frac{1}{v^2} + \frac{1}{v_c^2} \right) \\ \mu &= \left(\frac{\partial A}{\partial N} \right)_{T, V} = \mu_{\text{ig}} + \frac{4\pi a \hbar^2}{m} \left(\frac{1}{v} + \frac{1}{v_c} \right). \end{aligned}$$

These should be compared with the ground state results when $v_c \rightarrow \infty$.

The condensation starts when $v = v_c$ and $\lambda_T = \lambda_c$, leading to

$$\begin{aligned} p_c &= p_{\text{ig}} + \frac{4\pi a \hbar^2}{m \lambda_c^6} g_{3/2}(1)^2 \\ \mu_c &= 4g_{3/2}(1) k_B T_c \frac{a}{\lambda_c}. \end{aligned}$$

This means that the interactions lead to non-zero chemical potential giving a correction to the fugacity

$$z_c \simeq 1 + 4g_{3/2}(1) \frac{a}{\lambda_c}.$$

11. Phase Transitions

The phenomena to which statistical physics has been applied can be divided roughly to two categories. So far, we have studied the first category, in which the constituting systems can be considered non-interacting. It includes specific heats of gases and solids, chemical reactions, Bose-Einstein condensation, blackbody radiation, paramagnetism, elementary electron theory of metals, to name a few. The most significant feature of this category is that the thermodynamic functions of the systems involved are smooth and continuous⁵⁶.

The second category can be characterised with discontinuous or singular thermodynamic functions, corresponding to *phase transitions*. It includes the condensation of gases, the melting of solids, coexistence of phases, the behaviour of mixtures and solutions, ferromagnetism, antiferromagnetism, the order-disorder transitions, superfluid transition from He-I to He-II, the normal to superconducting transition, and so on. These phenomena are characterized by the fact that the interactions between the constituents play an important role. The co-operation leads to *macroscopic significance* at a particular *critical temperature* T_c . We try here to capture the aspects of these co-operative processes in terms of simplified models.

11.1 General Remarks on Condensation

We study a system with N particles, and assume that the potential can be written as a sum of two-particle terms

$$\left. \begin{aligned} u(r) &= \infty & \text{for } r &\leq \sigma \\ 0 > u(r) &> -\varepsilon & \text{for } \sigma < r < r^* \\ u(r) &= 0 & \text{for } r &\geq r^*. \end{aligned} \right\}$$

Thus, each particle can be seen as a hard sphere with the diameter σ , surrounded with attractive potential of range r^* and depth $-\varepsilon$. Note, that the Lennard-Jones potential meets these requirements up to a very good approximation. We expect that the *exact* interparticle potentials lead to phenomena that are, at least qualitatively, similar to those obtained from our simplified form.

There are some generally accepted properties that the *exact* partition function $Z(T, V, N)$ has to display the following:

- In the thermodynamic limit, $f(v, T) = N^{-1} \ln Z = -A/Nk_B T$ tends to depend only on $v = V/N$ and T . Thus, it is an *intensive* variable, and the thermodynamic pressure is then given by

$$p(v, T) = - \left(\frac{\partial A}{\partial V} \right)_{T, N} = k_B T \left(\frac{\partial f}{\partial v} \right)_T,$$

turning out to be a strictly *non-negative* quantity.

- Function f is *concave*, implying⁵⁷

$$\left(\frac{\partial p}{\partial v} \right)_T \leq 0.$$

⁵⁶With a *sole* exception of BEC.

⁵⁷This is relevant for the mechanical stability.

The regions with zero slope correspond to coexistence of two or more phases with different densities, implying the onset of phase transition. The positive slopes in the (p, v) -isotherms are non-physical, and result always from the approximations made in the derivation of the equation of state⁵⁸. Such regions should be corrected with the *Maxwell construction of equal areas*.

- The exact partition function with *finite* N corrects the singularities resulting from such corrections made for the approximate isotherms at the limit $N \rightarrow \infty$.

11.2 Condensation of a van der Waals Gas

Let us start with the simplest model for an interacting gas, the van der Waals gas, obeying the equation of state

$$p = \frac{RT}{v - b} - \frac{a}{v^2},$$

where v is the *molar volume* of the gas, and thus a and b pertain to one mole of the gas. At the critical temperature, both $(\partial p / \partial v)_T$ and $(\partial^2 p / \partial v^2)_T$ vanish. This leads to the coordinates of the critical point

$$p_c = \frac{a}{27b^2}, \quad v_c = 3b, \quad T_c = \frac{8a}{27bR},$$

implying

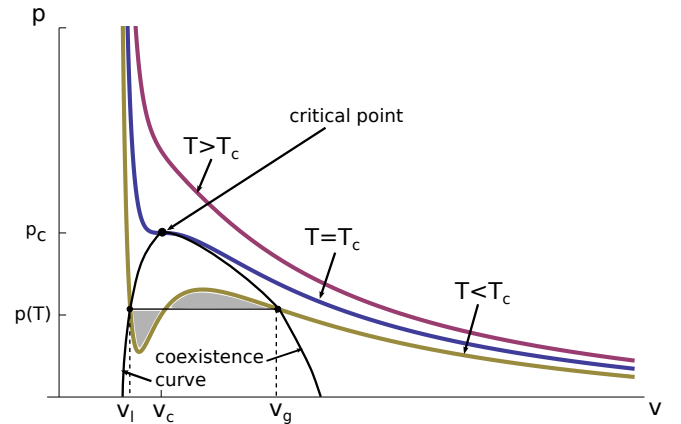
$$\mathcal{K} = \frac{RT_c}{p_c v_c} = \frac{8}{3}.$$

This allows us to write the equation of state in a *universal* form by using reduced variables

$$p_r = \frac{p}{p_c}, \quad v_r = \frac{v}{v_c}, \quad T_r = \frac{T}{T_c}.$$

This implies that the reduced equation of state becomes

$$\left(p_r + \frac{3}{v_r^2} \right) (3v_r - 1) = 8T_r.$$



Below the critical temperature, the van der Waals isotherms display regions with positive slope, i.e.

⁵⁸This and the following sentence will be clarified in the next section where we study these properties in terms of van der Waals gas.

$(\partial p/\partial v)_T > 0$. As stated, these regions should be corrected due to their non-physical nature. We replace them with horizontal lines. In the phase diagram ($((T, p)$ -plane), these indicate the locations of phase transitions. The replacements are made in terms of Maxwell's construction of equal areas. The requirement for equal areas follows simply from the requirements $dp = 0$ and $dT = 0$ for the corrected isotherm, that imply the change in the (molar) Gibbs energy

$$dg = -sdT + vdp = 0.$$

The same result should be obtained by integrating the theoretical isotherm, i.e. the quantity $dg = vdp$, from the state v_l to state v_g .

The locus of points $v_l(T)$ and $v_g(T)$ is called the *coexistence curve*. Within the region enclosed by this curve the gaseous and liquid phases mutually coexist.

The critical behaviour of the van der Waals system differs in several respects from that of real systems. Nevertheless, it is still used as the first approximation against which the more sophisticated theories are compared.

11.3 Ising Model of Phase Transitions

Next, we present a model that provides a unified theoretical basis for understanding (anti-)ferromagnetism, gas-liquid and liquid-solid transitions, order-disorder transitions in alloys, and so on. We will discuss the model in terms of ferromagnetism, but the formulation can be generalized to the languages of the other physical phenomena.

We assume that the atoms are bound to lattice sites, implying that the spin degrees of freedom are independent on other degrees of freedom

$$H \approx H_{\text{spin}} + H_{\text{other}}.$$

This implies that we can factorize the partition function

$$Z = Z_{\text{spin}} Z_{\text{other}}.$$

Because of the *exchange interaction*⁵⁹, there is an energy difference between the state of parallel spins and the one of antiparallel spins, given by

$$\varepsilon_{\uparrow\uparrow} - \varepsilon_{\uparrow\downarrow} = -2J_{ij}.$$

Because

$$\begin{aligned} \mathbf{s}_i \cdot \mathbf{s}_j &= \frac{1}{2} [(\mathbf{s}_i + \mathbf{s}_j)^2 - \mathbf{s}_i^2 - \mathbf{s}_j^2] \\ &= \frac{1}{2} S(S+1) - s(s+1) \\ &= \begin{cases} \frac{1}{4} & \text{if } S = 1 \\ -\frac{3}{4} & \text{if } S = 0 \end{cases} \end{aligned}$$

we can write the interaction energy as

$$\varepsilon_{ij} = \text{const.} - 2J_{ij}(\mathbf{s}_i \cdot \mathbf{s}_j).$$

⁵⁹The exchange interaction arises as manifestation of the interplay between the Coulomb repulsion and the Pauli principle between two electrons. Interested reader can refer to the course of Condensed matter physics.

Where the actual value of constant is unimportant, and can be used as the zero of energy. Interactions with $J_{ij} > 0$ favour the parallel alignment, and can display *ferromagnetism*. If $J_{ij} < 0$, the situation is reversed and we see the possibility of *antiferromagnetism*.

The most important interactions are assumed to occur between the nearest neighbours. As a result, the Hamiltonian of the N spin system can be written as

$$H = - \sum_{\langle ij \rangle} J_{ij} \mathbf{s}_i \cdot \mathbf{s}_j - \mu \mathbf{B} \cdot \sum_i \mathbf{s}_i$$

where $\langle ij \rangle$ indicates that the summation goes over all nearest neighbour pairs. Also, we have denoted the external magnetic field with \mathbf{B} . In the following, we simplify this even further by assuming that the strength of the interaction does not vary between the lattice sites, i.e. $J_{ij} = J$. Also, we will restrict the spin quantum number to the case $s_i = \frac{1}{2}$, and take into account only its z -component. This allows us to treat the system classically, since all variables in the Hamiltonian commute. We end up with the *Ising model*

$$H = -J \sum_{\langle ij \rangle} \sigma_i \sigma_j - \mu B \sum_i \sigma_i, \quad (98)$$

where $\sigma_i = +1$ for "up" spin and -1 for "down" spin.

The Ising model can be solved exactly in one and two dimensions. Systems analogous to Ising model are, e.g.

- *binary mixture* composed of two species of atoms, A and B , where each lattice point is occupied either type of atom.
- *lattice gas*, where each lattice point is either occupied by an atom or empty.

Mean field approach to the Ising model

Before we solve the Ising model exactly, let us introduce an approximative way to handle the pairwise interactions. We make the *mean field approximation*, which approximates the interaction energy between a lattice site and its nearest neighbours by assuming that the neighbours all have spin $\langle \sigma \rangle$, the average spin per site. This leads to the mean field Hamiltonian

$$H = -\frac{1}{2} \sum_{i=1}^N \nu J \langle \sigma \rangle \sigma_i - \mu B \sum_{i=1}^N \sigma_i = - \sum_{i=1}^N E(J, B) \sigma_i,$$

where $E(J, B) = \frac{1}{2} \nu J \langle \sigma \rangle + \mu B$, and ν is the number of nearest neighbours. Also, the prefactor $\frac{1}{2}$ guarantees that every pair is counted only once. The expectation value $\langle \sigma \rangle$ has to be determined in a self-consistent manner.

The partition function of a single spin is

$$Z = \sum_{\sigma_i = \pm 1} e^{\beta E \sigma_i} = 2 \cosh(\beta E).$$

Thus, the probability that the site i has the spin σ_i is

$$P(\sigma_i) = \frac{e^{\beta E \sigma_i}}{2 \cosh(\beta E)}.$$

Again, the average magnetization of the lattice is

$$M = N\mu\langle\sigma\rangle,$$

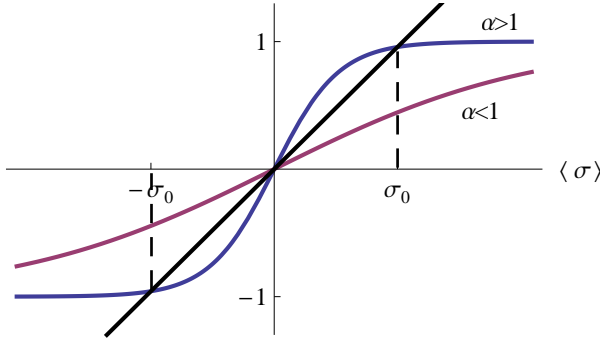
where

$$\langle\sigma\rangle = \frac{\sum_{\sigma_i=\pm 1} \sigma_i e^{\beta E \sigma_i}}{\sum_{\sigma_i=\pm 1} e^{\beta E \sigma_i}} = \tanh(\beta E).$$

If $B = 0$ we obtain the magnetization

$$\langle\sigma\rangle = \tanh\left(\beta \frac{1}{2} \nu J \langle\sigma\rangle\right) = \tanh\left(\frac{\nu J \langle\sigma\rangle}{2k_B T}\right).$$

We see that $\langle\sigma\rangle$ is the *order parameter* for the spin lattice. When temperature is high, the magnetization is zero since the spins are randomly oriented. At lower temperatures, the spins align spontaneously resulting in non-zero magnetization. We can determine the critical temperature T_c below which the lattice becomes more and more ordered. This is calculated graphically by plotting $\langle\sigma\rangle$ and $\tanh(\alpha\langle\sigma\rangle)$ into the same graph as a function of $\langle\sigma\rangle$. Here $\alpha = \nu J/2k_B T$.



The solutions are found at the intersections of the curves. We find that when $\alpha < 1$, there is only one intersection at $\langle\sigma\rangle = 0$. Similarly when $\alpha > 1$ there exists three intersections at $\langle\sigma\rangle = 0$ and $\pm\sigma_0$. One can show that the free energy per site is

$$\frac{G(J, 0)}{N} = \begin{cases} -k_B T \ln 2 & \text{if } \langle\sigma\rangle = 0 \\ -k_B T \ln[2 \cosh(\frac{1}{2}\beta\nu J\sigma_0)] & \text{if } \langle\sigma\rangle = \pm\sigma_0 \end{cases}$$

Thus, $\langle\sigma\rangle = \pm\sigma_0$ are the possible equilibrium states since they minimize the free energy. There is the phase transition (critical point) at $\alpha = 1$, implying the critical temperature

$$T_c = \frac{\nu J}{2k_B}.$$

We have thus shown that in the mean field approximation, there is a phase transition at finite temperature for d -dimensional lattice. In the following, we will show that the mean field theory leads to incorrect result in the case of one-dimensional lattice. In general, one can show that for $d \leq 3$, the mean field calculation gives too high estimates for the critical temperature. For dimensions $d \geq 4$, the estimates are good but have only mathematical, rather than physical, interest.

Exact calculation of one-dimensional Ising model

In practice, the only exactly solvable models are the one and two dimensional Ising models. The one-dimensional Ising model was first solved by Ising himself (1925). We assume periodic boundaries of the spin chain, i.e.

$$\sigma_{N+1} = \sigma_1.$$

This eliminates the undesired effects arising from the termination of the chain. Nevertheless, it does not affect the thermodynamics of the (infinitely long) chain. The Hamiltonian of the system can be written as

$$H = -J \sum_{i=1}^N \sigma_i \sigma_{i+1} - \mu B \sum_{i=1}^N \sigma_i$$

where $\sigma_i = \pm 1$. Now, the partition function can be written as

$$\begin{aligned} Z &= \sum_{\sigma_1} \sum_{\sigma_2} \cdots \sum_{\sigma_N} e^{\beta J \sum_{i=1}^N \sigma_i \sigma_{i+1} + \beta \mu B \sum_{i=1}^N \sigma_i} \\ &= \sum_{\sigma_1} \sum_{\sigma_2} \cdots \sum_{\sigma_N} \prod_{i=1}^N e^{\beta J \sigma_i \sigma_{i+1} + \frac{1}{2} \beta \mu B (\sigma_i + \sigma_{i+1})} \end{aligned}$$

We define the 2×2 transition matrix

$$T_{\sigma\sigma'} = e^{\beta J \sigma \sigma' + \frac{1}{2} \beta \mu B (\sigma + \sigma')},$$

where $\sigma, \sigma' = \pm 1$. The partition function is then

$$\begin{aligned} Z &= \sum_{\sigma_1} \sum_{\sigma_2} \cdots \sum_{\sigma_N} T_{\sigma_1 \sigma_2} T_{\sigma_2 \sigma_3} \cdots T_{\sigma_N \sigma_1} \\ &= \sum_{\sigma_1} (T^N)_{\sigma_1 \sigma_1} = \text{Tr } T^N. \end{aligned}$$

The transition matrix

$$T = \begin{pmatrix} e^{\beta(J+\mu B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-\mu B)} \end{pmatrix}$$

which is clearly symmetric. Its eigenvalues are given by

$$\lambda_{\pm} = e^{\beta J} \left[\cosh(\beta \mu B) \pm \sqrt{\sinh^2(\beta \mu B) + e^{-4\beta J}} \right].$$

The trace is independent on the basis, and thus

$$Z = \lambda_+^N + \lambda_-^N.$$

Also,

$$\ln Z = N \ln \lambda_+ + \ln \left[1 + \left(\frac{\lambda_-}{\lambda_+} \right)^N \right].$$

In the thermodynamic limit ($N \rightarrow \infty$),

$$\left(\frac{\lambda_-}{\lambda_+} \right)^N \rightarrow 0.$$

Thus,

$$\ln Z \approx N \ln \lambda_+,$$

meaning that the larger eigenvalue determines the thermodynamic properties.

The free energy per spin is given by

$$\begin{aligned} \frac{G}{N} &= -\frac{k_B T}{N} \ln Z \\ &= -J - k_B T \ln \left[\cosh(\beta \mu B) + \sqrt{\sinh^2(\beta \mu B) + e^{-4\beta J}} \right]. \end{aligned}$$

Again, thermodynamic quantities can be calculated from the free energy. In particular, the average spin per particle is

$$\sigma \equiv \langle \sigma \rangle = - \left(\frac{\partial G/N}{\partial \mu B} \right)_T = \frac{\sinh(\beta \mu B)}{\sqrt{\sinh^2(\beta \mu B) + e^{-4\beta J}}}.$$

We saw in the mean field approximation that the expectation value σ behaves as an order parameter: $\sigma = 0$ corresponds to completely stochastically oriented spins, whereas $|\sigma| = 1$ corresponds to the case of parallelly aligned spins. As previously, the magnetization is

$$M = N \mu \sigma.$$

Clearly, when $B \rightarrow 0$ also $M \rightarrow 0$ for finite temperatures T . This rules out the possibility of spontaneous break in the symmetry between the opposite spins. Thus, we cannot observe spontaneous magnetization, and hence the phase transition, contrary to the mean field approach!

From the low-field susceptibility

$$\chi_0(T) = \lim_{B \rightarrow 0} \left(\frac{\partial M}{\partial B} \right)_T = \frac{N \mu^2}{k_B T} e^{2J/k_B T},$$

we see immediately that a ferromagnetic ($J > 0$) system magnetizes strongly at low (but finite) temperatures. Nevertheless, when the external field is removed, the system returns to the disordered state with $\sigma = 0$. Thus, the one-dimensional Ising chain is a paramagnetic system without any phase transitions. However, it does not obey the Curie law ($M \propto 1/T$) so it is not a Curie paramagnet.

The low-field susceptibility χ_0 diverges when $T \rightarrow 0$. This means that there exists a phase transition at absolute zero, corresponding to complete alignment of the spins.

Two-dimensional Ising model

The two-dimensional Ising model can also be solved exactly by generalizing the transition matrix method (Onsager, 1944). The mathematics involved in the calculation are rather challenging, unfortunately. Therefore, we only give the results here. It turns out that in the two-dimensional case there is a phase transition at temperature

$$T_c = \frac{2J}{\ln(1 + \sqrt{2})} \approx 2.269 J.$$

Also, the specific heat diverges (logarithmically) at the critical point $T = T_c$, and the phase transition is continuous. The two-dimensional Ising model is perhaps the simplest exactly solvable model that displays a phase transition.

11.4 Critical Exponents

One way to classify phase transitions is the quantity m called the *order parameter*. As in the case of magnetization, the order parameter usually is assigned to the expectation value of some observable of the system (such as the average spin in the case of magnetization). Analogously to $\mu B/k_B T$ in the magnetization case, we identify h as the *ordering field* coupled to the observable.⁶⁰ In the low-field limit ($h \rightarrow 0$), m tends to limiting value m_0 which is 0 when $T \geq T_c$ and $\neq 0$ for $T < T_c$. We say that the phase transition is *continuous*, if the order parameter approaches the zero value continuously. On the other hand, if the change at the critical temperature is finite, the phase transition is *discontinuous*.⁶¹ We consider here continuous phase transitions.

It is typical for a continuous (second order) phase transition that the system goes from high temperature phase to a low temperature phase with lower symmetry. This is referred to as the *spontaneous symmetry breaking*. Examples include:

- Structural formation of crystal lattice.
- Ferromagnetic phase transition.
- Superconducting and superfluid transitions.
- Symmetry breaking of the electroweak interaction.

The order parameter, various response functions (such as the susceptibility) and correlation functions display non-analytic behaviour, such as singularities, near the critical point. The singular parts are with great accuracy proportional to some powers of the thermodynamic quantities h and $T - T_c$. The *critical exponents* are defined as follows

- The behaviour of the order parameter is determined by the exponent β

$$m(T) \sim \begin{cases} 0 & \text{when } T > T_c \\ (T_c - T)^\beta & \text{when } T < T_c. \end{cases}$$

- The divergence of the low-field susceptibility when $T \rightarrow T_c$ is given by the exponents γ (above) and γ'

⁶⁰For a gas-liquid transition, one can choose the density difference ($\rho_l - \rho_c$) as the order parameter and the pressure difference ($p - p_c$) as the ordering field.

⁶¹Typically, the order parameter is related to thermodynamic observables that are given in terms of first derivatives of the free energy. In our previous terminology then, the continuous transitions are second order transitions and discontinuous transitions first order transitions.

(below)

$$\chi_0(T) \sim \left(\frac{\partial m}{\partial h} \right)_{T, h \rightarrow 0} \sim \begin{cases} (T - T_c)^{-\gamma} & \text{when } T > T_c \\ (T_c - T)^{-\gamma'} & \text{when } T < T_c \end{cases}$$

Within experimental accuracy we have $\gamma = \gamma'$.

- The dependence of the order parameter on the field is given by the exponent δ

$$m|_{T=T_c} \sim h^{1/\delta} \quad \text{when } T = T_c, \quad h \rightarrow 0.$$

- The singular part of the specific heat is determined by the exponents α and α'

$$C_h(T) \sim -T \left(\frac{\partial^2 G}{\partial T^2} \right)_h \sim \begin{cases} (T - T_c)^{-\alpha} & \text{when } T > T_c \\ (T_c - T)^{-\alpha'} & \text{when } T < T_c \end{cases}$$

In practice, we have $\alpha = \alpha'$.

One can show that the critical exponents for van der Waals gas (exercise) and for the Ising model in the mean field approximation are

$$\beta = \frac{1}{2}, \quad \gamma = \gamma' = 1, \quad \delta = 3, \quad \alpha = \alpha' = 0.$$

In fact, *all* mean field theories give this set of exponents despite the possible structural differences. This suggests that they belong to the same *universality class*.

11.5 Landau's Theory of Phase Transitions

In a continuous phase transition, the slope of the free energy changes continuously and a symmetry is always broken. As we have discussed, this can be seen as an appearance of an order parameter in the phase with lower symmetry. All such transitions can be described in terms of *mean field theory* presented first by Landau (1937). We have already noticed that the mean field theory does not describe all features of continuous phase transitions correctly, but nevertheless gives a good starting point for understanding such transitions.

Landau argued that the critical behaviour of a given system can be determined by expanding its free energy in powers of the zero-field order parameter $m_0 = m(h = 0)$ as

$$\phi_0(t, m_0) = \phi_0(t, 0) + r(t)m_0^2 + s(t)m_0^4 + \dots$$

where $t = \frac{T - T_c}{T_c}$, $|t| \ll 1$.

The free energy should have the same symmetries as the Hamiltonian. Generally, the Hamiltonian remains unchanged (at zero field limit) when $m_0 \rightarrow -m_0$, which is

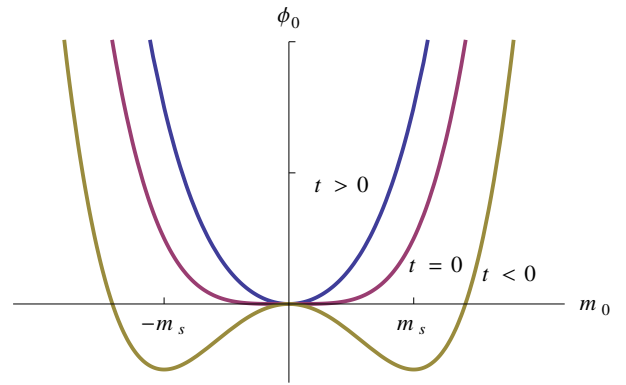
reflected as disappearance of the odd powers of m_0 . The coefficients may be written as

$$\phi_0(t, 0) = \sum_{k \geq 0} q_k t^k, \quad r(t) = \sum_{k \geq 0} r_k t^k, \quad s(t) = \sum_{k \geq 0} s_k t^k, \dots$$

The equilibrium occurs when the free energy is minimized. The thermodynamic stability requires $s(t) > 0$. We expect that the equilibrium is found at $m_0 = 0$ when $T > T_c$, and at $m_0 \neq 0$ when $T < T_c$. This is obtained by choosing $r(t) > 0$ and $r(t) < 0$, respectively. Thus, $r(0) = 0$. The lowest order approximation gives

$$\phi_0(t, m_0) \approx \phi_0(t, 0) + r_1 t m_0^2 + s_0 m_0^4 + \dots$$

Since we are considering phenomena near the critical point, we include only the terms written explicitly above.



We see that for $t > 0$ we have one minimum at $m_0 = 0$. For $t < 0$, we obtain two minima whose locations are given by

$$\frac{\partial \phi_0}{\partial m_0} = 0 \Rightarrow m_s = \pm \sqrt{r_1 |t| / 2s_0}.$$

Region between these values corresponds to unstable states. It should be corrected with a method similar to Maxwell's equal areas. Based on the above, the order parameter depends on the temperature as

$$m_0 \propto (T_c - T)^{1/2},$$

implying that the critical exponent $\beta = 1/2$ for the mean field theories.

We can thus write near the critical point

$$\phi_0(t) \approx \begin{cases} \phi_0(t, 0) & \text{for } T > T_c \\ \phi_0(t, 0) - \frac{r_1^2}{4T_c^2 s_0} (T_c - T)^2 & \text{for } T < T_c \end{cases}$$

This implies that the critical exponents $\alpha = \alpha' = 0$. Also, it implies a jump $r_1^2 / 2s_0 T_c$ in the specific heat at the critical point.

Effects of the field and universality

When the field is non-zero, one benefits of making Legendre transformation

$$\phi_h(t, m) = -hm + \phi_0(t, 0) + r_1 t m_0^2 + s_0 m_0^4 + \dots$$

With a straightforward calculation, one can show that the critical exponents $\gamma = \gamma' = 1$. Quick inspection reveals that the introduction of the field breaks the symmetry in the system when $T > T_c$. This means that there is no more continuous phase transition.

Thus, we have shown that the critical exponents of the Landau theory, build only upon the analyticity of the free energy and the symmetries of the Hamiltonian, result in exactly same critical exponents as the previous calculations made with mean field theory! This means that while the free energy itself depends on the structure of the system, the critical exponents do not. This *universality* of the critical exponents suggests that we are dealing with a class of systems that display qualitatively same critical behaviour, despite the structural differences.

Breakdown of Landau theory

The mean field (Landau) theory neglects the fluctuations of the order parameter in large length scales. However, when one approaches the critical point these fluctuations diverge. One often refers them as *critical fluctuations*. When the fluctuations become larger than the mean value for the order parameter, the mean field theory breaks down. In other words, a mean field theory cannot be valid arbitrarily close to the critical point.

To see this, consider *correlations* between particles in the high temperature phase of the Ising model. They can be characterized with the spin-spin *correlation function*

$$g(i, j) = \langle \sigma_i \sigma_j \rangle - \langle \sigma_i \rangle \langle \sigma_j \rangle, \quad (99)$$

defined for the pair of spins at sites i and j . When $i = j$, one obtains the fluctuations (standard deviation) of the spin σ at site i . When the separation between i and j increases, one expects that the corresponding spins get uncorrelated and $\langle \sigma_i \sigma_j \rangle \rightarrow \langle \sigma_i \rangle \langle \sigma_j \rangle$ leading to $g(i, j) \rightarrow 0$.

By applying the partition function (98), we obtain

$$\frac{\partial}{\partial B} \ln Z = \beta \mu \langle \sum_i \sigma_i \rangle = \beta \langle M \rangle,$$

where $\mu \sum_i \sigma_i$ is the net magnetization of the system. Similarly,

$$\frac{\partial^2}{\partial B^2} \ln Z = \beta^2 (\langle M^2 \rangle - \langle M \rangle^2).$$

Thus, we obtain for the susceptibility

$$\chi = \frac{\partial \langle M \rangle}{\partial B} = \beta (\langle M^2 \rangle - \langle M \rangle^2) = \beta \mu^2 \sum_i \sum_j g(i, j).$$

This is an important result! It relates the susceptibility with the fluctuations in the order parameter (magnetization). Also, it tells that the the fluctuations are intrinsically related to the correlations between the individual spins.

The mean field theory is assumed to be valid if it obeys the so-called *Ginzburg criterion*

$$\frac{\sigma_M^2}{\langle M \rangle^2} \ll 1 \Rightarrow k_B T \chi \ll \langle M \rangle^2,$$

where $\sigma_M^2 = \langle M^2 \rangle - \langle M \rangle^2$. In the mean field theory, the susceptibility diverges at the critical temperature. Implying the breakdown of the theory. In some systems, however, the breakdown occurs only for very small $|t|$. Examples of such systems include

- *Superfluids*: The normal-superfluid transition in He^4 is a continuous phase transition which involves broken gauge symmetry. Order parameter is the macroscopic wave function Ψ of the condensed phase, and the free energy is given as

$$\phi(T, P) = \phi_0(T, P) + r(T, P)|\Psi|^2 + s(T, P)|\Psi|^4 + \dots,$$

where $r(T, P) = r_0(T, P)(T - T_c)$ and r_0 and s are slowly varying functions of T and P .

- *Superconductors*: The condensed phase is a macroscopically occupied quantum state Ψ , serving as the order parameter. In the normal-superconductor transition the phase of the order parameter changes. The free energy is given by

$$\psi(T) = \int d^3r \left[\psi_n(T) + r(T)|\Psi(\mathbf{r})|^2 + s(T)|\Psi(r)|^4 + \frac{1}{2m} |i\hbar \nabla_{\mathbf{r}} \Psi(\mathbf{r})|^2 \right].$$

The *Ginzburg-Landau theory* of superconductivity is obtained by extremizing the free energy with respect to variations is Ψ^* .

11.6 Scaling, Universality and Renormalization Group Approach

Even though the neglect of fluctuations in the mean field theory leads into quantitatively incorrect results, it nevertheless produces the correct result that every physical system belongs in a specific universality class, based on its critical behaviour. In spite of the prediction, the mean field theory does not provide any reason for universality. For that, we need the scaling theory.

Widom scaling

We have described the singular behaviour in the vicinity of the critical point in terms of the critical exponents. The free energy can be written in terms of the regular and singular terms as

$$\phi(t, h) = \phi_r(t, h) + \phi_s(t, h),$$

implying that the critical exponents must come from the singular part ϕ_s .

The idea behind the *scaling hypothesis* is that when the system approaches the critical point, the *correlation length*⁶² ξ becomes infinitely large resulting in diminished

⁶²Correlation length is the measure of the range of the correlations in the system.

sensitivity to length transformations (changes of scale). In other words, we say that there is no natural measure of length in the system, except the atomic length. The system looks similar no matter what scale is used, and it is called *self-similar*. This suggests that the singular part of the free energy is a *generalized homogeneous function*, i.e. it *scales* like

$$\phi_s(\lambda^p t, \lambda^q h) = \lambda \phi_s(t, h),$$

when $\lambda > 0$ and p and q are independent on the system. Direct calculations give the critical exponents in terms of p and q (exercise):

$$\begin{aligned}\beta &= \frac{1-q}{p} \\ \delta &= \frac{q}{1-q} \\ \gamma &= \gamma' = \frac{2q-1}{p} \\ \alpha &= \alpha' = 2 - \frac{1}{p}.\end{aligned}$$

These imply the scaling laws

$$\begin{aligned}\alpha + 2\beta + \gamma &= 2 \\ \beta(\delta - 1) &= \gamma.\end{aligned}$$

We thus see that only two of the critical exponents are independent.

Kadanoff scaling and renormalization group

Unlike the Widom scaling hypothesis, the method presented by Kadanoff and developed by Wilson is based on the microscopic properties of matter. We will present the only the outlines of method:

- Combine the original microscopic state variables blockwise into *block variables*.
- Determine the effective interactions between the blocks. This is *coarse graining* of the system called the *block transform*.
- The block transforms form a semigroup, so called *renormalization group*. One can perform transforms sequentially.
- Because in a system at the critical point has no natural length scale due to diverging correlation length, the transformed systems look like copies of each other. Thus, the critical point corresponds to a *fixed point* of the transformations.

In general, the renormalization is a collection of techniques to remove the infinities arising from the theory. It was first developed to deal with the infinite terms in the perturbative expansions of the quantum electrodynamics. Such infinities typically arise because when making the theory we have chosen to describe phenomena occurring at a chosen scale of length, ignoring the degrees of freedom at

shorter scales. This kind of theories can be described in terms of renormalization group, and are called *effective field theories*. For example, practically all of the condensed matter physics consists of effective field theories. Also, the general relativity can be seen as an effective field theory of (yet unknown) complete theory of quantum gravity.

12. Non-Equilibrium Statistics

This course has so far dealt with only *statistical averages* of physical quantities. We have shown that up to very high accuracy these averages correspond to the relevant measurements made on systems in *equilibrium*. Nevertheless, there do exist deviations from the average values called *fluctuations*. These fluctuations are generally small in magnitude, but are important for several reasons:

- In the onset of the high degree spatial correlations in the vicinity of the critical point.
- In understanding how a non-equilibrium state approaches equilibrium, and how the equilibrium persists.
- In relating the dissipative properties of a system with the microscopic properties of the equilibrium (*fluctuation-dissipation* theorem).

In this final chapter, we wish to gain understanding on the fluctuations in thermal equilibrium, the approach of the equilibrium, and dissipation. It turns out that all these are intimately connected.

12.1 Brownian Motion

Brownian motion of a relatively massive particle immersed in fluid will provide us the paradigm for describing the equilibrium and non-equilibrium processes. Even in thermal equilibrium, the Brownian particle exhibits random-like motion. This randomness is a consequence of the unavoidable interactions between the particle and its surroundings.

The situation can be modelled with the Hamiltonian⁶³

$$H = H_{\text{sys}} + H_{\text{bath}} + H_{\text{int}},$$

where H_{sys} , H_{bath} and H_{int} are the Hamiltonians for the system (the particle), its surroundings, and the interaction between them. We will consider, for simplicity, a one dimensional particle, and we focus our attention on the location x of the particle. Hamilton's equations of motion give

$$\begin{aligned} \frac{dx}{dt} &= v \\ m \frac{dv}{dt} &= -\frac{dH_{\text{sys}}}{dx} - \frac{dH_{\text{int}}}{dx} \equiv \mathcal{F}(t) + F(t). \end{aligned}$$

We have assumed above that the system and the bath interact (linearly) via the variable x . We denote with \mathcal{F} the force exerted on the massive particle by some external potential, and with F the force that describes the interactions with the many other degrees of freedom in the surrounding bath⁶⁴ (we will use the words bath and fluid

interchangeably in the following). Since the studied particle is large compared to the vast number of smaller ones constituting the fluid, its movement is very slow, and $F(t)$ is very rapidly fluctuating function in particle's time scale. Thus, we are not interested in the explicit time evolution of the bath degrees of freedom. What we assume instead is that in the above equation of motion those are averaged (traced out) quantities. As a consequence, the force $F(t)$ can be treated as a random function of time t . Additionally, we assume that in equilibrium ($v = 0$) there is no preferred direction in space, and the ensemble average vanishes, $\langle F \rangle = 0$.

It is beneficial to write the random force as

$$F = \langle F \rangle_{\xi} + \xi,$$

where ξ denote the rapidly fluctuating part around the ensemble average⁶⁵ $\langle F \rangle_{\xi}$. Naturally, $\langle \xi \rangle = 0$. In the limit of small $\langle v \rangle$, one can expand $\langle F \rangle$ as

$$\langle F \rangle_{\xi} = -m\gamma \langle v \rangle_{\xi}.$$

We will show later that the coefficient γ is determined by the properties of the function $F(t)$ itself. For the moment, it will be taken as a phenomenological constant of *friction*. One sees that the effect of the slowly varying component $\langle F \rangle_{\xi}$ of the fluctuating force is causing relaxation of a particle with an initial velocity $v(0)$ towards the equilibrium $v = 0$.

Due to the rapidly fluctuating force, the velocity of the particle also fluctuates in time. Similarly to the force, the velocity can be represented in terms rapid, zero-mean fluctuations v' around the ensemble average velocity $\langle v \rangle_{\xi}$ as

$$v = \langle v \rangle_{\xi} + v'.$$

The variations in $\langle v \rangle_{\xi}$ are much slower than those in v . Altogether, the equation of motion is written as

$$m \frac{dv}{dt} = \mathcal{F}(t) - m\gamma v + \xi(t),$$

where we have assumed that $\gamma v = \gamma \langle v \rangle_{\xi}$ ⁶⁶. The above equation of motion is called the *Langevin equation*. The Langevin equation is a *stochastic* differential equation, meaning that it contains a random term. The random solutions of such equations are generally called as *stochastic variables*.

The Langevin equation can be shown to hold also for the quantum mechanical operator \hat{x} , in addition to the classical expectation value $x = \langle \hat{x} \rangle$ discussed above. For further details on quantum Langevin equation, refer e.g. Cardiner and Zoller, *Quantum Noise*.

⁶³Note that this Hamiltonian describes a very general *open* system, i.e. system interacting with a much larger bath.

⁶⁴The term *noise* is often used to describe the fluctuating force F , since it describes the typically *undesired* deviations from the behaviour of an isolated system.

⁶⁵In the present case, the non-equilibrium ensemble is formed as mental copies of the Brownian particle with different realisations of the random variable ξ . We will continue of using the notation $\langle \cdot \rangle$ for the equilibrium (microcanonical, canonical, grand canonical) averages.

⁶⁶Fluctuations in v are slower than those of the force F that produces them. This is because of the large mass of the particle.

The informal derivation of the Langevin equation presented above shows explicitly how irreversibility arises in Hamiltonian (quantum) mechanics. Consider the combined system of the particle and the bath (fluid). This constitutes an isolated system, whose governing equations are reversible and the energy should be conserved. Now, the Langevin equation covers only the degrees of freedom of the particle, neglecting the time evolution of the bath. The effect of the bath on the particle is described with frictional force $-m\gamma v$, leading to dissipation of energy from the particle into the bath, and of the random force $\xi(t)$ that causes fluctuations in the particle degrees of freedom even at the equilibrium.

Equilibrium fluctuations

We assume for now that the external potential is zero, and that the particle is in thermal equilibrium with the bath. Thus, we the Langevin equation (multiplied with x) gives

$$mx \frac{d\dot{x}}{dt} = m \left[\frac{d}{dt}(x\dot{x}) - \dot{x}^2 \right] = -m\gamma x\dot{x} + x\xi(t).$$

Next, we take an ensemble average of both sides, and employ the equipartition theorem to $\frac{1}{2}m\langle\dot{x}^2\rangle = \frac{1}{2}k_B T$. The resulting differential equation is easily solvable, leading into

$$\langle x\dot{x} \rangle_\xi = C e^{-\gamma t} + \frac{k_B T}{m\gamma},$$

where C is a constant of integration. We thus see that γ^{-1} is a characteristic time constant of the system. If we assume that initially at $t = 0$ each particle in the ensemble is at $x(0) \equiv x_0 = 0$, we obtain $C = -k_B T/m\gamma$. Thus,

$$\langle x\dot{x} \rangle_\xi = \frac{k_B T}{m\gamma} (1 - e^{-\gamma t}),$$

and by integrating once more we have obtained the equilibrium fluctuations (variance, $x_0 = 0$)

$$\langle (x(t) - x_0)^2 \rangle_\xi = \langle x^2 \rangle_\xi = \frac{2k_B T}{m\gamma} (t - \gamma^{-1}(1 - e^{-\gamma t})). \quad (100)$$

When $t \ll \gamma^{-1}$, we have

$$\langle x^2 \rangle_\xi \approx \frac{k_B T}{m} t^2.$$

This means that at short time intervals the particle behaves as a free particle with the constant thermal velocity $v = \sqrt{k_B T/m}$. In the opposite limit, $t \gg \gamma^{-1}$, we have

$$\langle x^2 \rangle_\xi \approx \frac{2k_B T}{m\gamma} t.$$

Here one should recall the result from the very beginning of the course, which stated that the variance of a random walk grows as the number of steps N . Thus, in the long time limit the particle executes a continuous random walk and *diffuses* to the fluid with the *diffusion constant*

$$D = \frac{k_B T}{m\gamma},$$

which is referred to as the *Einstein relation*.

We stress that the variance $\langle x^2 \rangle_\xi$ is caused by the coupling to the environment. In the absence of the coupling, the particle located initially at x_0 stays there forever. However, the constant bombardment of the fluid molecules cause random "kicks" to the particle, thus causing the diffusive behaviour.

Correlation functions

Instead of assuming the thermal equilibrium, we can calculate directly the mean square velocity of the Brownian particle (assume again that $\mathcal{F} = 0$). First of all, by integrating the Langevin equation we obtain

$$v(t) = v(0)e^{-\gamma t} + \frac{1}{m} \int_0^t e^{-\gamma(t-t')} \xi(t') dt' \quad (101)$$

It should be noted that the solution of the above equation holds for a single realisation of ξ . Because $\xi(t)$ is a random variable, also $v(t)$ and $x(t)$ are random variables whose properties are determined by $\xi(t)$. This leads into

$$\begin{aligned} \langle v^2(t) \rangle_\xi &= v^2(0)e^{-2\gamma t} + \frac{2}{m} e^{-2\gamma t} \int_0^t e^{\gamma t'} \langle \xi(t') \rangle_\xi dt' \\ &\quad + \frac{1}{m^2} e^{-2\gamma t} \int_0^t \int_0^t dt' dt'' e^{\gamma(t'+t'')} \langle \xi(t') \xi(t'') \rangle_\xi. \end{aligned}$$

The second term in the above equation is zero because $\langle \xi(t) \rangle_\xi = 0$ for all t .

The function $\langle \xi(t') \xi(t'') \rangle_\xi$ appearing in the above equation is of greatest importance to us. It can be seen as the *temporal* equivalent to the spin-spin correlation function defined in Eq. (99). It is a measure of the statistical correlation between the value of ξ at time t' and its value at time t'' . We will refer to it as the (time) *autocorrelation function* of the variable ξ . Generally, we will denote the temporal correlation function of the variables A and B with

$$C_{AB}(t, t') = \langle A(t) B(t') \rangle_\xi.$$

In order to proceed with the calculation, one needs thus information over the time correlations of ξ . Typically in the case of Brownian particle, the fluctuating force is assumed to be independent on its previous values⁶⁷. In addition, one often assumes that $\xi(t)$ is a stationary process, meaning that its mean and variance does not change in shifts in time and space. Mathematically this is called *Gaussian white noise* process⁶⁸ and expressed as delta-correlation

$$C_\xi(t, t') = g \delta(t - t'),$$

where $\delta(t)$ is Dirac's delta function and g is a constant given by the equipartition theorem.

This assumption results in

$$\langle v^2(t) \rangle_\xi = v^2(0)e^{-2\gamma t} + \frac{g}{2m^2\gamma} (1 - e^{-2\gamma t}).$$

⁶⁷This assumption of *short memory* is referred as the *Markovian approximation*.

⁶⁸This terminology will become clear later when we discuss about spectral density

When $t \rightarrow \infty$ we have that $\langle v^2(t) \rangle \rightarrow g/2m^2\gamma$. In this limit, no matter what the initial velocity $v(0)$ is, we expect that the particle is in equilibrium with the bath. Thus, the equipartition theorem gives

$$g = 2m\gamma k_B T.$$

We obtain for delta-correlated noise

$$\langle v^2(t) \rangle_\xi = v^2(0)e^{-2\gamma t} + \frac{k_B T}{m}(1 - e^{-2\gamma t}).$$

By integrating Eq. (101), we obtain

$$x(t) = x(0) + \frac{v(0)}{\gamma}(1 - e^{-\gamma t}) + \frac{1}{m\gamma} \int_0^t (1 - e^{-\gamma(t-t')}) \xi(t') dt'.$$

Similar to v , we obtain (we set the origin at $x(0) = 0$)

$$\begin{aligned} \langle x^2(t) \rangle_\xi &= \left(\frac{v(0)}{\gamma}(1 - e^{-\gamma t}) \right)^2 \\ &\quad + \frac{2k_B T}{m\gamma} \int_0^t ds (1 - e^{-\gamma(t-s)})(1 - e^{-\gamma(t-s)}) \\ &= \left(\frac{v(0)}{\gamma}(1 - e^{-\gamma t}) \right)^2 \\ &\quad + \frac{2k_B T}{m\gamma} \left(t - \frac{2}{\gamma}(1 - e^{-\gamma t}) + \frac{1}{2\gamma}(1 - e^{-2\gamma t}) \right) \\ &= \left(\frac{v(0)}{\gamma}(1 - e^{-\gamma t}) \right)^2 \\ &\quad + \frac{2k_B T}{m\gamma} t - \frac{k_B T}{m\gamma^2} ((1 - e^{-\gamma t})(3 - e^{-\gamma t})) \end{aligned}$$

When the initial velocity has the thermal equilibrium value $v^2(0) = k_B T/m$, we recover the result (100).

Fokker-Planck equation

Instead of solving the average behaviour of ensemble of Brownian particles, a more generalized way is to find the rate at which the *distribution* of Brownian particles approaches the thermal equilibrium. The equation which governs the decay of the non-equilibrium distribution is called the *master equation*. Here, we present a simplified version of it, named the *Fokker-Planck equation*.

Consider the displacement $x(t)$ of a Brownian particle⁶⁹. Let $f(x, t)dx$ be the probability that an arbitrary particle from the ensemble is found with the displacement between x and $x + dx$ at any given time t . Naturally,

$$\int_{-\infty}^{\infty} f(x, t) dx = 1.$$

The time-evolution of a probability density, such as $f(x, t)$, is typically governed by the master equation

$$\frac{\partial f}{\partial t} = \int_{-\infty}^{\infty} \left(f(x', t)W(x', x) - f(x, t)W(x, x') \right) dx',$$

⁶⁹The following discussion can be done also for any other dynamical random variable, such as v .

where $W(x, x')\delta t$ is the probability that within a small time interval δt the particle makes a transition from the displacement x into x' . Note, that here we also make a Markovian approximation by assuming that the rate $W(x, x')$ does not depend on the *history* of the particle.

In the case of Brownian motion, we assume that only those transitions are possible that are between nearby states. This means that $W(x, x')$ is sharply peaked at $x' = x$, and rapidly goes to zero away from x . By writing the master equation in terms of $\eta = x - x'$, and making a second order Taylor expansion around $\eta = 0$, one ends with (exercise)

$$\frac{\partial f}{\partial t} = -\frac{\partial}{\partial x}(\mu(x)f(x, t)) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(D(x)f(x, t)),$$

which is called the *Fokker-Planck equation*. In the above, the first and second moments of the transition rate $W(x, x')$ are given by

$$\begin{aligned} \mu(x) &= \int_{-\infty}^{\infty} \eta W(x, x - \eta) d\eta = \frac{\langle \delta x \rangle}{\delta t} = \langle v \rangle \\ D(x) &= \int_{-\infty}^{\infty} \eta^2 W(x, x - \eta) d\eta = \frac{\langle \delta x^2 \rangle}{\delta t}. \end{aligned}$$

In the case of Brownian particle, we obtained

$$\begin{aligned} \mu(x) &= \langle v \rangle = 0 \\ D(x) &= 2 \frac{k_B T}{m\gamma} \equiv 2D. \end{aligned}$$

Thus, the Fokker-Planck equation reduces in

$$\frac{\partial f}{\partial t} = D \frac{\partial^2 f}{\partial x^2},$$

which is the one-dimensional *diffusion equation*. Its solution is a Gaussian

$$f(x, t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{(x-x_0)^2}{4Dt}},$$

with the variance $\sigma^2 = 2Dt$. Again, the Gaussian solution suggest the interpretation of the diffusion process in terms of continuous random walk.

12.2 Fluctuation-Dissipation Theorem

In order to complete the discussion, we would still like a method how to relate the constant of friction γ with the fluctuations. This is done with the fluctuation-dissipation theorem, which basically tells that the decay of the equilibrium fluctuations obey the same law than that for a non-equilibrium state⁷⁰. We would like study how a thermodynamic system responds when it is (weakly) forced out of equilibrium. Somewhat surprisingly, the calculation turns out to be the simplest using the quantum mechanical density operator approach.

⁷⁰This way of stating the fluctuation-dissipation theorem is the *Onsager regression hypothesis*.

Let us consider the changes to an observable \hat{A} made by a weak external field $F(t)$. The nature of the field is arbitrary, it can be stochastic or deterministic. We assume that the field is coupled to the system via the observable \hat{B} . The Hamiltonian for the system is

$$\hat{H} = \hat{H}_0 - F(t)\hat{B},$$

where \hat{H}_0 is the equilibrium state Hamiltonian. Correspondingly, the density operator of the equilibrium state is given by

$$\hat{\rho}_{\text{eq}} = \frac{\exp(-\beta\hat{H}_0)}{\text{Tr}(\exp(-\beta\hat{H}_0))}.$$

Also, the equilibrium expectation value of an arbitrary (hermitian) operator \hat{A} is given by

$$\langle \hat{A} \rangle_{\text{eq}} = \text{Tr}(\hat{A}\hat{\rho}_{\text{eq}}).$$

The application of the small time-dependent field induces a small change to the density operator

$$\hat{\rho}(t) = \hat{\rho}_{\text{eq}} + \delta\hat{\rho}(t).$$

The time-evolution is given by the von Neumann equation, which can be approximated as

$$\frac{\partial \hat{\rho}}{\partial t} = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}(t)] \approx -\frac{i}{\hbar}([\hat{H}_0, \delta\hat{\rho}(t)] - F(t)[\hat{B}, \hat{\rho}_{\text{eq}}]).$$

We see that in this approximation, the response is linear with respect to the field. In the interaction picture⁷¹ one can write the solution as

$$\delta\hat{\rho}(t) = \frac{i}{\hbar} \int_{-\infty}^t dt' F(t') e^{-i\hat{H}_0(t-t')/\hbar} [\hat{B}, \hat{\rho}_{\text{eq}}] e^{i\hat{H}_0(t-t')/\hbar}.$$

The change in the expectation value (the response) of the observable is then

$$\langle \delta\hat{A}(t) \rangle = \langle \hat{A} \rangle - \langle \hat{A} \rangle_{\text{eq}} = \text{Tr}(\hat{A}\delta\hat{\rho}(t)).$$

One can show (exercise) that the response can be written as the convolution

$$\langle \delta\hat{A}(t) \rangle = \frac{i}{\hbar} \int_{-\infty}^t dt' F(t') \langle [\hat{A}(t), \hat{B}(t')] \rangle_{\text{eq}},$$

where $\hat{A}(t) = e^{i\hat{H}_0 t/\hbar} \hat{A} e^{-i\hat{H}_0 t/\hbar}$. Thus, we can identify the linear response function (see the end of this chapter for more information on linear response theory)

$$\chi_{AB}(\tau) = \frac{i}{\hbar} \langle [\hat{A}(\tau), \hat{B}(0)] \rangle_{\text{eq}},$$

where we have employed the knowledge that the equilibrium correlation depends only on the time interval τ .

We see that the Fourier transform of the response becomes

$$\langle \delta\hat{A}(\omega) \rangle = \int_{-\infty}^{\infty} dt e^{i\omega t} \langle \delta\hat{A}(t) \rangle = (\chi'_{AB}(\omega) + i\chi''_{AB}(\omega)) F(\omega),$$

⁷¹Unitary transformation $\hat{\rho}_I = \hat{U}^\dagger \hat{\rho} \hat{U}$ into the interaction picture is obtained with $\hat{U} = \exp(-i\hat{H}_0 t/\hbar)$.

where

$$\chi''_{AB}(\omega) = \frac{1}{2\hbar} \int_{-\infty}^{\infty} dt \langle [\hat{A}(t), \hat{B}(0)] \rangle_{\text{eq}} e^{i\omega t},$$

and $\chi'_{AB}(\omega)$ is given by the Kramers-Krönig relation (defined later).

On the other hand, the power spectrum related to temporal correlations between the observables \hat{A} and \hat{B} is given by Fourier transform of the correlation function (this is a generalization of the Wiener-Khintchine theorem presented below)

$$S_{AB}(\omega) = \int_{-\infty}^{\infty} dt \langle \delta\hat{A}(t) \delta\hat{B}(0) \rangle_{\text{eq}} e^{i\omega t},$$

where the fluctuations are defined as $\delta\hat{A}(t) = \hat{A}(t) - \langle \hat{A} \rangle_{\text{eq}}$. By employing the result (exercise)

$$\int_{-\infty}^{\infty} dt \langle \hat{B}(0) \hat{A}(t) \rangle_{\text{eq}} e^{i\omega t} = e^{-\beta\hbar\omega} \int_{-\infty}^{\infty} dt \langle \hat{A}(t) \hat{B}(0) \rangle_{\text{eq}} e^{i\omega t},$$

we obtain

$$\chi''_{AB}(\omega) = \frac{1}{2\hbar} (1 - e^{-\beta\hbar\omega}) S_{AB}(\omega).$$

This is the celebrated *fluctuation-dissipation theorem* that relates the dissipation due to the perturbative field with the spontaneous thermal fluctuations that occur in the equilibrium.

The classical limit of the fluctuation-dissipation theorem is given when $\hbar\omega \ll k_B T$:

$$\chi''_{AB}(\omega) = \frac{\omega}{2k_B T} S_{AB}(\omega).$$

Linear response theory

Generally, the linear response theory studies how the system behaves when it is displaced from the equilibrium. It assumes that the force $F(t)$ driving the system out of equilibrium is *weak*, leading to

$$\begin{aligned} \langle \delta\hat{x}(t) \rangle &\equiv \langle \hat{x}(t) - \langle \hat{x} \rangle_{\text{eq}} \rangle = \int_{-\infty}^t dt' \chi(t-t') F(t') \\ &= \int_0^\infty d\tau \chi(\tau) F(t-\tau), \end{aligned}$$

where $\langle \hat{x} \rangle_{\text{eq}}$ is the equilibrium value of the observable \hat{x} , and χ is the *response function* that characterises the variable's response to the force field. The observable \hat{x} can here be taken as the position of the Brownian particle, but can be in general any relevant observable that is randomized by the force. The integral can be extended to $-\infty$ by assuming *causality*, i.e. $\chi(t) = 0$ for $t < 0$. Thus, we see that the response of x is a *convolution* of the response function and the force. By taking a Fourier transform, we obtain the linear response in frequency space

$$\langle \delta\hat{x}(\omega) \rangle = \chi(\omega) F(\omega).$$

The Fourier transform of the response function is often referred to as the *susceptibility*. Note, that the response function is real since the response $\delta\hat{x}$ is observable quantity. Nevertheless, the susceptibility is complex

$$\chi(\omega) = \chi'(\omega) + i\chi''(\omega),$$

where

$$\begin{aligned}\chi'(\omega) &= \int_{-\infty}^{\infty} \chi(t) \cos(\omega t) \\ \chi''(\omega) &= \int_{-\infty}^{\infty} \chi(t) \sin(\omega t).\end{aligned}$$

The real part of the susceptibility is sometimes (especially when one studies linear response of electric circuits) called the *reactive* response, and the imaginary part the *dissipative* response. Especially, in the case of Brownian particle subjected to white noise, one can show that the susceptibility is

$$\chi(\omega) = (m(\gamma - i\omega))^{-1},$$

One can show (omitted here) that the causality implies that the reactive part can be written in terms of the dissipative response, or vice versa, as

$$\begin{aligned}\chi'(\omega) &= \frac{2}{\pi} \int_0^{\infty} \chi''(\omega') \frac{\omega'}{\omega'^2 - \omega^2} d\omega', \\ \chi''(\omega) &= -\frac{2\omega}{\pi} \int_0^{\infty} \chi'(\omega') \frac{\omega'}{\omega'^2 - \omega^2} d\omega',\end{aligned}$$

which are the *Kramers-Krönig relations*.

Power spectral density

Consider fluctuation $\delta\hat{x}(t) = \hat{x}(t) - \langle\hat{x}\rangle$ of an observable \hat{x} . It can be split into spectral components by taking the (truncated) Fourier transform

$$\delta\hat{x}(\omega) = \int_{-T/2}^{T/2} dt \delta\hat{x}(t) e^{i\omega t}.$$

This describes the fraction of the fluctuation $\delta\hat{x}(t)$ that is oscillating with the frequency ω . Typically, the power of the oscillations is related to the square of the variable, which persuades us to define the equilibrium *power spectral density* as

$$S(\omega) = \lim_{T \rightarrow \infty} \frac{1}{T} \langle |\delta\hat{x}(\omega)|^2 \rangle.$$

Now, we can write⁷²

$$\begin{aligned}S(\omega) &= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} \int_{-T/2}^{T/2} dt dt' e^{i\omega(t'-t)} \langle \delta\hat{x}(t') \delta\hat{x}(t) \rangle \\ &= \lim_{T \rightarrow \infty} \int_{-T}^T d\tau \frac{T - |\tau|}{T} e^{i\omega\tau} \langle \delta\hat{x}(\tau) \delta\hat{x}(0) \rangle \\ &= \int_{-\infty}^{\infty} d\tau e^{i\omega\tau} C_{\delta x}(\tau).\end{aligned}$$

This is the *Wiener-Khintchine theorem*, that relates the power spectral decomposition of the fluctuations in an observable with the Fourier transform of the autocorrelation function of the fluctuations. In the above, we have assumed that the state is *stationary*, meaning that

$$C_{\delta x}(t + \tau, t) = C_{\delta x}(\tau, 0) \equiv C_{\delta x}(\tau).$$

Conclusions

This concludes the course of *Statistical Physics*. The first part of the course lays the microscopic foundation to the classical theory of thermodynamics. The quantum mechanical indistinguishability of identical particles were discussed in the beginning of the latter half of the course. New phenomena that result from the quantum nature of particles, such as Bose-Einstein condensation and relativistic electron gas, were introduced. The inclusion of unavoidable interactions between particles was studied in terms of cluster expansions and second quantization. Finally, an introduction to the more elaborate theories of phase transitions and non-equilibrium statistics was given. Throughout the course, the aim has been in learning the fundamental statistical concepts that are needed in the more specialised advanced courses, such as Superconductivity, Condensed Matter Physics, Quantum Optics in Electric Circuits, and so on.

⁷²Recall how the change of variables is made with multi-dimensional integrals!