HEIDELBERG UNIVERSITY

DEPARTMENT OF PHYSICS AND ASTRONOMY

Theoretical Statistical Physics

Prof. Ulrich Schwarz
Winter term 2022/23
Last update: April 11, 2025

Foreword

This script is written for the course *Theoretical Statistical Physics* which is one of the core courses for the 1st semester master students of physics at Heidelberg University, although in practise it is also attended by many 5th semester bachelor students. I have been giving this course several times now namely in the winter terms of 2012, 2015, 2017, 2020 and 2022. It is my experience that a script helps to correct the unavoidable errors made at the blackboard, to introduce additional material that complements the material discussed during the lecture, to solidify the new knowledge through a coherent presentation and to prepare for the final exam and also for the oral master exam. There exist many very good textbooks on statistical physics and the purpose of this script is soley to document my personal choice from the vast range of subjects covered by statistical physics.

Statistical physics provides the basis for many important parts of physics, including atomic and molecular physics, solid state physics, soft matter physics, biophysics, astrophysics, environmental and socioeconomic physics. For example, you cannot understand the greenhouse effect or the cosmic microwave background without the Planck formula for the statistics of photons at a given temperature (black body radiation) or the electrical conduction of solids without the concept of a Fermi sphere (the ground state of a fluid of electrons). Equally important, however, statistical physics provide the basis for our understanding of phase transitions, which are truely collective effects and often do not depend much on microscopic details. As you will learn in this course, at the heart of statistical physics is the *art of counting*, which is formalized in the concept of a partition sum. The details of how this has to be done in different systems can be quite challenging and only few exact solutions exist (most prominently for the 2D Ising model). Therefore statistical physics if full of approximations (like the virial expansion for a real gas) and schemes that complement a full analytical solution (like computer simulations or the renormalization group theory). Due to these technical challenges, but also due to the addition of new and interesting areas of application, statistical physics is still a very active research area, continuously expanding into new applications and developing new methods.

Several guiding principles and helpful books determined the design of this course. First I completely agree with Josef Honerkamp, who in his book *Statistical Physics* notes that statistical physics is much more than statistical mechanics. A similar notion is expressed by James Sethna in his book *Entropy, Order Parameters, and Complexity*. Indeed statistical physics teaches us how to think about the world in terms of probabilities. This is particularly relevant when one deals with complex systems and real world data. Therefore applications of statistical physics can also be found in data-intensive research areas, such as astrophysics, environmental physics, biophysics, socioeconophysics and physics of information (including machine learning). As instructive examples, consider the models for the spread of rumours or viruses on networks, or the algorithms used for segmentation and object recognition in image processing. If you investigate how these models work, you will realize that they often relate to the Ising model for ferromagnets, arguably the most important model of statistical physics and an important subject for this course.

Second a course on statistical physics certainly has to make the connection to thermodynamics. Thermodynamics can be quite cumbersome and hard to digest at times, so a pedagogical approach is highly appreciated by most students. Here I am strongly motivated by the axiomatic and geometrical approach to thermodynamics as layed out in the beautiful book *Thermodynamics and an introduction to thermostatistics* by Herbert Callen. The same approach has been taken by Luca Peliti in his book *Statistical mechanics in a nutshell*. Historically thermodynamics developed as a phenomenological theory of heat transfer, but when being approached from the axiomatic and geometrical side, it becomes the convincing and universal theory that it actually is. The book by Callen also draws heavily on the work by Edwin Jaynes on the relationship between statistical physics and information theory as pioneered by Claude Shannon. Although somehow debated, this link shows once again that statistical physics is more than statistical mechanics. Information theory provides very helpful insight into the concept of entropy, which is the cornerstone of statistical mechanics. Recently this area has been revived by the advent of stochastic thermodynamics, which shows that entropy is not only an ensemble property, but can also be defined for single trajectories. By now the first textbook *Stochastic thermodynamics: an introduction* has been written by Luca Peliti and Simone Pigolotti.

Third a comprehensive course on statistical physics should also include some numerical component, because modern statistical physics cannot be practised without computational approaches, as nicely argued also by Josef Honerkamp, James Sethna and Luca Peliti. Moreover statistical physics is much more than thermodynamic equilibrium and if time permits, a course on statistical physics should also cover some aspects of non-equilibrium physics, for example the exciting recent developments in stochastic thermodynamics. Although it is hard to fit all of these aspects into a one-semester course, some of them are included here.

Together, these considerations might explain the structure of this script. We start with an introduction to the concepts of probability theory, which should be useful also in other contexts than only statistical mechanics. Here it is important to think clearly about probability as a measure for volume in event space and to practise with a few important examples like the binomial distribution. It also allows us to introduce the concept of entropy (from the viewpoint of information theory) and random walks (as extension of random variables into the time domain, eventually leading to the Langevin and Fokker-Planck equations). We then introduce the fundamental postulate of equilibrium physics, namely that each microstate is equally probable, leading to the microcanonical ensemble and the principle of maximal entropy. We next discuss the canoncial and grandcanonical ensembles, when reservoirs exist for exchange of heat and particle number, respectively. We then apply these concepts to quantum fluids, in particular the

Fermi fluid (e.g. electrons in a solid) and the Bose gas (e.g. black body radiation with photons or the Debye model for crystal vibrations). These systems do not have direct interactions, but the particles interact indirectly through the parity rules, which has to be accounted for by the correct way of counting. This then also leads to our first phase transition, namely the Bose-Einstein condensation. We then introduce the concept of phase transitions emerging from direct interactions through the example of the Ising model. In particular, it is here that we introduce one of the most important advances of theoretical physics of the 20th century, namely the renormalization group. We then continue to discuss phase transitions, now for complex fluids, starting with the van der Waals fluid and the virial expansion. We close with a discussion of thermodynamics, from which we see that statistical physics and thermodynamics essentially lead to the same formal structure of ensembles and thermodynamic potentials, but that they also complement each other in a unique manner: statistical physics focuses on the emergence of macroscopic properties from microscopic mechanisms, and thermodynamics on the macroscopic principles that necessarily have to be valid in the thermodynamic limit of very large system size, independent of the microscopic details.
Finally one should note some subjects which are not covered in the script due to space reasons. We do not cover kinetic and transport theories, which would also include the Boltzmann equation. The very important subject of fluctuations and correlations (including the fluctuation-dissipation theorem) is mentioned only in passing. We also cannot treat much out-of-equilibrium physics here, in particular we do not cover Green-Kubo relations, Onsager's reciprocity theorem, Kramers-Krönig relations or linear response theory. From the subject side, we will not have time to cover such interesting subjects as liquid crystals, percolation, disordered and glassy systems (including the replica method), nucleation, coarsening and Ostwald ripening, or the dynamics of chemical reactions and populations.

Heidelberg, winter term 2022/23 Ulrich Schwarz

Contents

1 Introduction to probability theory

1.1 Probability in physics

Classical physics (classical mechanics and electrodynamics) is deterministic, that means the governing equations (Newton's and Maxwell's equations, respectively) are differential equations that have a unique solution once we know the initial conditions (and boundary conditions for the case of Maxwell's equations, which are partial differential equations). Probability might enter only in the sense that the intial and boundary conditions are not known with certainty. Quantum mechanics of course introduces probability into physics in the form of the statistical (Kopenhagen) interpretation, that is experiments lead to the collapse of the wavefunction with probabilistic outcomes. Yet here we still solve a deterministic differential equation (Schrödinger's equation for the wavefunction) and then probability for the outcome follows only at the end as the squared modulus of the complex wavefunction.

In marked contrast, statistical physics directly brings the concept of probability into physics. Now the central concept is to calculate the probability of a certain macroscopic state, thus probability is not a derived quantity, but the most elementary concept. For example, in the canonical ensemble the relevant statistics will be the Boltzmann distribution. Therefore we start our course on statistical physics with an introduction into probability theory. Later of course we have to ask how the probabilistic nature of statistical physics emerges from more microscopic descriptions, and we will see that both classical and quantum mechanics provide some justification for this.

1.2 Frequentist approach

The history of probability theory is long and twisted. Yet everybody has an intuitive notion of probability that is related to frequencies of certain outcomes. We start with a simple example (throwing dice) to illustrate what this means and what one would expect from a theory of probability. Possible outcomes for a die are $\{1,2,3,4,5,6\}$. For N throws the event $\{i\}$ occurs N_i times. We then identify the probability p_i for event $\{i\}$ with its frequency:

$$p_i = \frac{\text{\# favorable outcomes}}{\text{\# possible outcomes}} = \frac{N_i}{N} \quad \text{in the limit } N \to \infty$$

For an ideal die we expect $p_i = \frac{1}{6} \approx 0.167$. Hence for 1000 throws $\{6\}$ should occur around 167 times.

We first note that our definition is normalized:

$$\sum_{i=1}^{6} N_i = N \quad \overset{1/N}{\Rightarrow} \quad \sum_{i=1}^{6} p_i = 1$$

We next consider events that are not directly an experimental outcome, but a more complicated question to ask about the system. E.g. what is the probability to get an odd outcome?

$$p_{odd} = \frac{\#\text{ favorable outcomes}}{\#\text{ possible outcomes}} = \frac{N_1 + N_3 + N_5}{N} = p_1 + p_3 + p_5$$

⇒ sum rule: summation of probabilities for simultaneous disjunct events

What is the probability to get twice $\{6\}$ when throwing two times? We first throw N times and find N_6 times a 6. We then throw M times and find M_6 times a 6. Thus we count

$$p_{66} = \frac{\#\text{ favorable outcomes}}{\#\text{ possible outcomes}} = \frac{N_6 \cdot M_6}{N \cdot M} = \frac{N_6}{N} \cdot \frac{M_6}{M} = p_6 \cdot p_6 = \frac{1}{36}$$

⇒ product rule: multiplication of probabilities for subsequent independent events

Finally we note that we could either throw N dice at once or the same die N times - the result should be the same

⇒ ergodic hypothesis of statistical physics: ensemble average = time average

Identifying probability with frequency is called the *classical* or *frequentist* interpretation of probability. There are two problems with this. First there are some examples for which naive expectations of this kind fail and a more rigorous theory is required. Second there are many instances in which an experiment cannot be repeated. Consider e.g. the statistical distribution of galaxy sizes in the universe, for which we have only one realization in our hands. In order to address these problems, it is better to assign some *a priori* probabilities to events and to include new knowledge about outcomes as it becomes available. This is called the *Bayesian* interpretation of probability. In order to avoid certain paradoxa, this approach has to be formalized with the correct choice of axioms.

1.3 Axiomatic approach

Above we described an empirical approach to measure probability for the dice throwing experiment and this sharpened our intuition what we expect from a theory of probability. We now construct a mathematical theory of probability by introducing an axiomatic system (*Kolmogorov 1933*). It has been shown that this approach allows to describe also complex systems without generating contradictions[1].

---

[1] For an introduction into probability theory, we recommend Josef Honerkamp, *Stochastische Dynamische Systeme*, VCH 1990; and Geoffrey Grimmett and Dominic Welsh, *Probability: an introduction*, 2nd edition 2014, Oxford University Press.

Let $\Omega = \{\omega_i\}$ be the set of elementary events. The complete set of possible events is the event space $\mathcal{B}$ defined by:

① $\Omega \in \mathcal{B}$

② if $A \in \mathcal{B}$, then $\overline{A} \in \mathcal{B}$

③ if $A_1, A_2, \cdots \in \mathcal{B}$, then $\cup_{i=1}^{\infty} A_i \in \mathcal{B}$

By setting all A_i with i larger than a certain value to empty sets, the last point includes unions of a finite number of sets. We see that the event space is closed under the operations of taking complements and countable unions. This concept is also known as σ-algebra. In our case we actually have a Borel-algebra, because the σ-algebra is generated by a topology. The most important point is that we have to avoid non-countable unions, because this might lead to pathological situations of the nature of the Banach-Tarski paradoxon (which states that a sphere can be disassembled into points and that they then can be reassembled into two spheres because the set of real numbers is non-countable). Later we will see that probability is some kind of volume in event space and that it has to be conserved, so we have to make sure that such changes in volume are not possible.

Corollaries

① $\emptyset \in \mathcal{B}$

② $A \cap B = \overline{\overline{A} \cup \overline{B}} \quad \in \mathcal{B}$

Examples

① $\Omega = \{1, ..., 6\}$ for the ideal die. This set of elementary events is complete and disjunct ($\omega_i \cap \omega_j = \emptyset$ if $i \neq j$, $\bigcup_{i=1}^{6} \omega_i = \Omega$). This event space is <u>discrete</u>.

② All intervals on the real axis, including points and semi-infinite intervals like $x \leq \lambda$. Here x could be the position of a particle. This event space is <u>continuous</u>.

We now introduce the concept of probability. For each event A in the event space $\mathcal{B}$ we assign a real number $p(A)$, such that

① $p(A) \geq 0 \quad \forall A \in \mathcal{B}$

② $p(\Omega) = 1$

③ $p(\bigcup_i A_i) = \sum_i p(A_i)$ if $A_i \cap A_j = \emptyset$ for $i \neq j$

Note that the last assumption is the sum rule. Kolmogorov showed that these rules are sufficient for a consistent theory of probability.

Corollaries

① $p(\emptyset) = 0$

② $p(A) + p(\overline{A}) = p(\Omega) = 1 \quad \Rightarrow p(A) = 1 - p(\overline{A}) \quad \Rightarrow 0 \leq p(A) \leq 1$

③ Consider $A_1, A_2 \in \mathcal{B}$:

$$p(A_1) = p(\underbrace{A_1 \cap A_2}_{:=I}) + p(\underbrace{A_1 \cap \overline{A_2}}_{:=C_1})$$

$$p(A_2) = p(\underbrace{A_2 \cap A_1}_{:=I}) + p(\underbrace{A_2 \cap \overline{A_1}}_{:=C_2})$$

$$\begin{aligned} \Rightarrow \quad p(A_1) + p(A_2) &= p(C_1) + p(C_2) + 2p(I) \\ &= p(A_1 \cup A_2) + p(I) \\ \Rightarrow \quad p(A_1 \cup A_2) &= p(A_1) + p(A_2) - p(A_1 \cap A_2) \end{aligned}$$

1.4 Continuous distributions and distribution function

Consider the event space containing the intervals and points on the real axis. $p(x \leq \lambda)$ is the probability that x is smaller or equal to a given λ (eg the position of a particle in 1D):

$$P(\lambda) := p(x \leq \lambda) \quad \text{cumulative distribution function}$$

If $P(\lambda)$ is differentiable, then

$$P(\lambda) = \int_{-\infty}^{\lambda} p(x) dx$$

where

$$p(\lambda) = \frac{dP(\lambda)}{d\lambda} \quad \text{probability density or distribution function}$$

We now can write the probability for $x \in [x_1, x_2]$ as $\int_{x_1}^{x_2} p(x) dx$. With $x_2 = x_1 + dx_1$, we can approximate the integral by a product and thus find that $p(x_1)dx_1$ is the probability to have $x \in [x_1, x_1 + dx_1]$. Thus $p(x)$ is the probability density and $p(x)dx$ is the probability to find a value around x. Note that the physical dimension of $p(x)$ is 1/m, because you still have to integrate to get the probability.

1.5 Joint, marginal and conditional probabilities

A multidimensional distribution $\vec{x} = (x_1, ...x_n)$ is called a *multivariate* distribution, if

$$p(\vec{x})\ \mathrm{d}x_1\ ...\ \mathrm{d}x_n \text{ is the probability for } x_i \in [x_i, x_i + \mathrm{d}x_i]$$

We also speak of a *joint* distribution. Note that in principle we have to distinguish between the random variable and its realization, but here we are a bit sloppy and do not show this difference in the notation.

Examples

① A classical system with one particle in 3D with position and momentum vectors has six degrees of freedom, thus we deal with the probability distribution $p(\vec{q}, \vec{p})$. For N particles, we have $6N$ variables.

② We measure the probability $p(a, i)$ for a person to have a certain age a and a certain income i. Then we can ask questions about possible correlations between age and income.

③ Consider a collection of apples (a) and oranges (o) distributed over two boxes (left l and right r). We then have a discrete joint probability distribution $p(F, B)$ where $F = a, o$ is fruits and $B = l, r$ is boxes.

Marginal probability: now we are interested only in the probability for a subset of all variables, e.g. of x_1:

$$p(x_1) = \int \mathrm{d}x_2 \, ... \, \mathrm{d}x_n \; p(\vec{x})$$

is the probability for $x_1 \in [x_1, x_1 + \mathrm{d}x_1]$ independent of the outcome for $x_2, ..., x_n$.

Examples

① We integrate out the momentum degrees of freedom to focus on the positions.

② We integrate $p(a, i)$ over i to get the age structure of our social network.

③ We sum over the two boxes to get the probability to have an orange

$$p(o) = p(o, l) + p(o, r)$$

This example shows nicely that the definition of the marginal probability essentially implements the sum rule.

Conditional probability: we start with the joint probability and then calculate the marginal ones. From there we define the conditional ones. Consider two events $A, B \in \mathcal{B}$. The conditional probability for A given B, $p(A|B)$, is defined by

$$\underbrace{p(A, B)}_{\text{joint probability}} = \underbrace{p(A|B)}_{\text{conditional probability for A given B}} \cdot \underbrace{p(B)}_{\text{marginal probability for B}}$$

Thus the definition of the conditional probability essentially introduces the product rule.

Example

Consider a fair die and the events $A = \{2\}$ and $B = \{2,4,6\}$.

$$p(A|B) = \frac{p(A,B)}{p(B)} = \frac{p(A)}{p(B)} = \frac{1}{3}$$

$$p(B|A) = \frac{p(A,B)}{p(A)} = \frac{p(A)}{p(A)} = 1$$

Statistical independence: $p(A_1|A_2) = p(A_1)$ $\quad A_1$ is independent of A_2

$$\Rightarrow p(A_1, A_2) = p(A_1|A_2)\, p(A_2) = p(A_1)\, p(A_2)$$

Thus we get the product rule (multiplication of probabilities) that we expect for independent measurements, compare the example of throwing dice discussed above. We also see that

$$\Rightarrow p(A_2|A_1) = \frac{p(A_1, A_2)}{p(A_1)} = p(A_2)$$

Statistical independence is mutual.
Bayes' theorem: $p(A,B) = p(A|B) \cdot p(B) = p(B,A) = p(B|A) \cdot p(A)$

$$\Rightarrow \boxed{p(B|A) = \frac{p(A|B) \cdot p(B)}{p(A)} = \frac{p(A|B) \cdot p(B)}{\sum_{B'} p(A|B') \cdot p(B')}} \quad \text{Bayes' theorem}$$

where for the second form we have used the sum rule. Despite of its simplicity, this formula named after Thomas Bayes (1701-1761) is of extremely large practical relevance. It allows to ask questions about the data that are not directly accessible by measurements.

Examples

① Consider again the fruits ($F = a, o$) in the boxes ($B = l, r$). We assume that left and right are selected with probabilites $p(l) = 4/10$ and $p(r) = 6/10$ (they sum to 1 as they should). We next write down the known conditional probabilities by noting that there are two apples and six oranges in the left box and three apples and one orange in the right box:

$$p(a|l) = 1/4, p(o|l) = 3/4, p(a|r) = 3/4, p(o|r) = 1/4$$

We now ask: what is the probability of choosing an apple ?

$$p(a) = p(a|l)p(l) + p(a|r)p(r) = 11/20$$

Note that the result is not 5/12 that we would get if there was no bias in choosing boxes. The probability of choosing an orange is

$$p(o) = 1 - p(a) = 9/20$$

We next ask a more complicated question: if we have selected an orange, what is the probability that it did come from the left box ? The answer follows by writing down the corresponding conditional probability:

$$p(l|\,o) = \frac{p(o|\,l)p(l)}{p(o)} = 2/3$$

Therefore

$$p(r|\,o) = 1 - 2/3 = 1/3$$

Above we have formulated the probability $p(F|\,B)$ for the fruit conditioned on the box. We now have reverted this relation to get the probability $p(B|\,F)$ for the box conditioned on the fruit. Our prior probability for the left box was $p(l) = 4/10 < 0.5$. Our posterior probability for the left box, now that we know that we have an orange, is $p(l|\,o) = 2/3 > 0.5$. Thus the additional information has reverted the bias for the two boxes.

② We discuss the statistics of medical testing. Imagine a test for an infection with the new Corona virus Sars-CoV-2. The standard test is based on the polymerase chain reaction (PCR), but now there new tests that are cheaper and faster, but not as reliable (e.g. the LAMP-test from ZMBH Heidelberg or the rapid antigen test by Roche). At any rate, such a test always has two potential errors: false positives (test is positive, but patient is not infected) and false negatives (test is negative, but patient is infected). We have to quantify these uncertainties. Let's assume that the probability that the test is positive if someone is infected is 0.95 (so the probability for false negatives is 0.05) and that the probability that the test is positive if someone is not infected is 0.01 (false positives). Actually these numbers are quite realistic for antigen tests against Sars-CoV-2 (PCR-tests are much more reliable).

Let A be the event that someone is infected and B the event that someone is tested positive. Our two statements on the uncertainties are then conditional probabilities:

$$p(B|A) = 0.95, \quad p(B|\bar{A}) = 0.01 \; .$$

We now ask what is the probability $p(A|B)$ that someone is infected if the test was positive. As explained above, this question corresponds to the kind of change of viewpoint that is described by Bayes' theorem. We will answer this question as a function of $p(A) = x$, because the answer will depend on which fraction of the population is infected.

According to Bayes' theorem, the conditional probability $p(A|B)$ is determined by

$$p(A|B) = \frac{p(B|A)x}{p(B)} = \frac{p(B|A)x}{p(B|A)x + p(B|\bar{A})p(\bar{A})} \; . \tag{1.1}$$

Using $x + p(\bar{A}) = 1$, we get

$$p(A|B) = \frac{p(B|A)x}{[p(B|A) - p(B|\bar{A})]\, x + p(B|\bar{A})} = \frac{x}{\left[1 - \frac{p(B|\bar{A})}{p(B|A)}\right] x + \frac{p(B|\bar{A})}{p(B|A)}} \,. \tag{1.2}$$

Introducing the ratio of false positive test results to correctly positive ones, $c := p(B|\bar{A})/p(B|A)$, we have our final result

$$p(A|B) = \frac{x}{[1-c]\, x + c} \,. \tag{1.3}$$

Thus the probability $p(A|B)$ that someone is in fact infected when tested positive vanishes for $x = 0$, increases linearly with x for $x \ll c$ and eventually saturates at $p(A|B) = 1$ as $x \to 1$. This type of saturation behaviour is very common in many applications, e.g. for adsorption to a surface (Langmuir isotherm) or in the Michaelis-Menten law for enzyme kinetics.

Putting in the numbers from above gives $c = 0.01 \ll 1$. Therefore we can replace the expression for $p(A|B)$ from above by

$$p(A|B) \approx \frac{x}{c + x} \,. \tag{1.4}$$

For a representative x-value below c, we take $x = 1/1000$ (one out of 1000 people is infected). Then $p(A|B) = 0.1$ and the probability to be infected if the test is positive is surprisingly small. It only becomes $1/2$ if $x = c$ (one out of 100 people is infected). Thus the test only becomes useful when the fraction of infected people x is larger than the fraction of false positives c.

③ A company produces computer chips in two factories:

$$\text{factory: events } A \text{ and } \overline{A} \begin{cases} 60\% \text{ come from factory A} \\ 40\% \text{ come from factory B} \end{cases}$$

$$\text{defect or not: events } d \text{ and } \overline{d} \begin{cases} 35\% \text{ from factory A} \\ 25\% \text{ from factory B} \end{cases}$$

What is the probability that a defect chip comes from factory A?

$$p(A|\, d) = \frac{p(d|\, A)\; p(A)}{p(d)}$$

$$p(d) = p(d|\, A)\; p(A) + p(d|\, B)\; p(B)$$

$$p(A) = 0.6,\;\; p(B) = 0.4,\;\; p(d|\, A) = 0.35, p(d|\, B) = 0.25$$

$$\Rightarrow p(A|\, d) = 0.68$$

④ We can design a webpage that makes offers to customers based on their income. However, the only data we are allowed to ask them for is age. So we buy the correlation data $p(a,i)$ from the tax office and then estimate the income of our users from their age information. The more multivariate data sets we can use for this purpose, the better we will be with these estimates and the more accurate our offers will be.

1.6 Expectation and covariance

Both for discrete and continuous probability distributions, the most important operation is the calculation of the expectation of some function f of the random variable:

$$\langle f \rangle = \sum_i f(i) p_i \quad \text{or} \quad \langle f \rangle = \int f(x) p(x) dx$$

In particular, the average of the random variable itself is

$$\mu = \langle i \rangle = \sum_i i p_i \quad \text{or} \quad \mu = \langle x \rangle = \int x p(x) dx$$

Examples

① Throwing the dice: $\langle i \rangle = 21/6 = 3.5$

② Particle with uniform probability for position $x \in [-L, L]$: $\langle x \rangle = 0$

The next important operation is the calculation of the mean squared deviation (MSD) or variance, which tells us how much the realization typically deviates from the average (now only for the discrete case):

$$\begin{aligned}\sigma^2 = \left\langle (i - \langle i \rangle)^2 \right\rangle &= \left\langle (i^2 - 2i \langle i \rangle + \langle i \rangle^2) \right\rangle \\ &= \left\langle i^2 \right\rangle - 2\langle i \rangle^2 + \langle i \rangle^2 = \left\langle i^2 \right\rangle - \langle i \rangle^2\end{aligned}$$

Here we have used the fact that averaging is a linear operation. σ is called the standard deviation.
For two random variables, the covariance is defined as

$$\sigma_{ij}^2 = \langle (i - \langle i \rangle)(j - \langle j \rangle) \rangle = \langle ij \rangle - \langle i \rangle \langle j \rangle$$

where the average has to be taken with the joint probability distribution if both variables are involved. If i and j are independent, then their covariance vanishes.

Examples

① Throwing the dice: $\sigma^2 = 35/12 = 2.9$

② Particle with uniform probability for position $x \in [-L, L]$: $\sigma^2 = L^2/3$

1.7 Binomial distribution

The binomial distribution is the most important discrete distribution.
We consider two possible outcomes with probabilities p and q ($p + q = 1$, binary process), respectively, and repeat the process N times.

Examples

① flipping a coin N times, outcomes head or tail

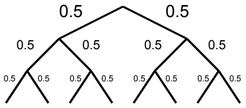

Figure 1.1: lineage tree for the ideal coin experiment

② following a ball falling through an 'obstacle array'

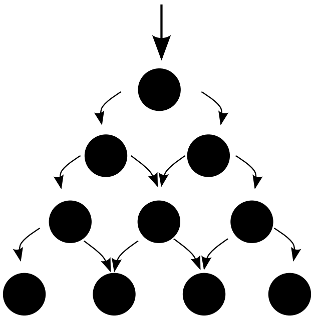

Figure 1.2: obstacle array

③ stepping N times forward or backward along a line ⇒ 1D Brownian random walk ('drunkard's walk')

④ throwing the dice N times and counting # {6} $\Rightarrow p = \frac{1}{6}, q = \frac{5}{6}$

⑤ N gas atoms are in a box of volume V which is divided into subvolumes pV and qV. On average $\langle n \rangle = p \cdot N$ atoms are in the left compartment. What is the probability for a deviation Δn? Or the other way round: Can one measure N by measuring the frequencies of deviations Δn ?

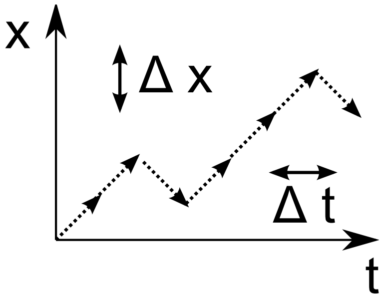

Figure 1.3: random walk: one possible path out of many

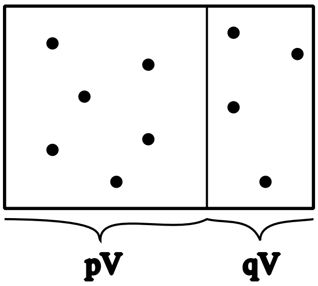

Figure 1.4: gas box with two compartments

In the following we use the terms of example 1. What is the probability to get $i \in \{0, 1, ..., N\}$ heads? Because every throw is independent we have to multiply the probabilities:

$$p^i \cdot q^{N-i}$$

However, there is more than one sequence of head and tails to arrive at i heads. The number of realizations is given by the *binomial coefficient*:

$$\frac{N(N-1)\dots(N-(i-1))}{i(i-1)\dots 1} = \frac{N!}{(N-i)!i!} = \binom{N}{i} \quad \text{binomial coefficient}$$

The binomial coefficient is the number of ways one can choose i objects out of N. The numerator is the number of ways we can select i objects out of N without putting them back. However, here we assume that we have numbered the objects and care for their identify. We therefore have to correct for the number of ways to rearrange the i objects, which is the denominator.

Another way to understand that the binomial coefficient has to appear here is to write our procedure of repeating the experiment with the two possible outcomes N times:

$$(p+q)^N = \underbrace{(p+q)(p+q)...(p+q)}_{\text{N times}} = \sum_{i=0}^{N} \binom{N}{i} \cdot p^i q^{N-i}$$

where in the last step we have used the binomial formula $(x+y)^N = \sum_{i=0}^{N} \binom{N}{i} x^i y^{N-i}$. Thus the binomial coefficient simply gives the multiplicity that arises from multiplying out this expression. Remember Pascal's triangle (Figure 1.5):

$$\binom{N}{i} = \binom{N-1}{i-1} + \binom{N-1}{i}$$

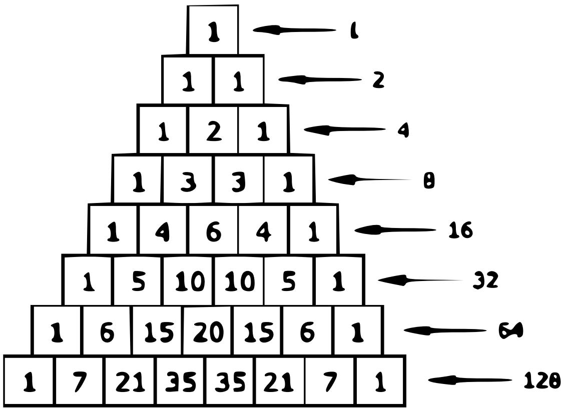

Figure 1.5: Pascal's triangle with sums on the right

In summary, we now have for the binomial distribution:

$$\boxed{p_i = \binom{N}{i} \cdot p^i \cdot q^{N-i}}$$

Note that normalization is ensured:

$$\sum_{i=0}^{N} p_i = \sum_{i=0}^{N} \binom{N}{i} \cdot p^i q^{N-i} = (p+q)^N = 1^N = 1$$

Obviously we have 2^N possible outcomes of the experiment, which we also see by considering

$$\sum_{i=0}^{N} \binom{N}{i} = \sum_{i=0}^{N} \binom{N}{i} 1^i \cdot 1^{N-i} = (1+1)^N = 2^N$$

Example

We plot the binomial distribution in Figure 1.6 for $p = \frac{1}{6}$ and $N = 10$. This is the probability to get i times a 6 when we throw the dice 10 times. The average of this distribution is $N \cdot p = 10/6 = 1.67$ and close to the peak.

Next we want to characterise the binomial distribution. It typically looks like in Figure 1.6 with one clear peak. We first calculate the average μ:

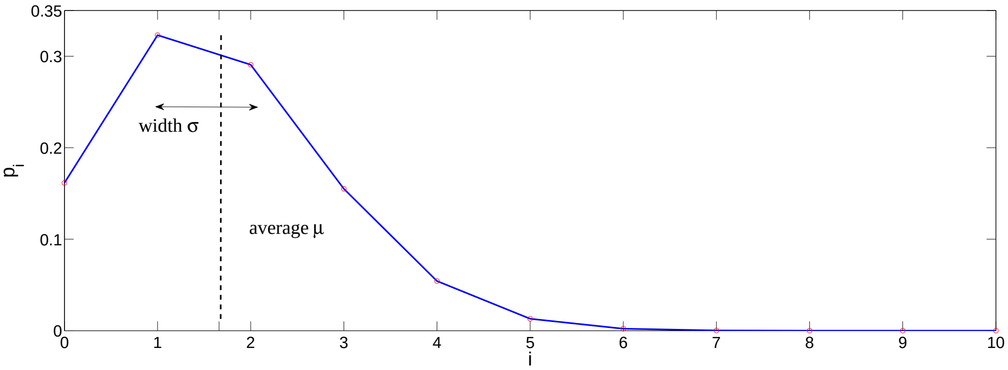

Figure 1.6: Binomial distribution for $p = \frac{1}{6}, \quad N = 10$. Points are joined by lines to better show the shape of the distribution. The average is $\mu = 1.67$ and the width is $\sigma = 1.18$.

$$
\begin{aligned}
\mu = \langle i \rangle = \sum_{i=0}^{N} i \cdot p_i &= \sum_{i=0}^{N} i \cdot \frac{N!}{(N-i)!i!} p^i \cdot q^{N-i} \\
&= N \cdot p \sum_{i=1}^{N} \frac{(N-1)!}{(N-i)!(i-1)!} p^{i-1} \cdot q^{N-i} \\
&= N \cdot p \underbrace{\sum_{j=0}^{M} \frac{M!}{j!(M-j)!} p^j \cdot q^{M-j}}_{=1} = N \cdot p
\end{aligned}
$$

where we relabeled according to $M = N - 1, \quad j = i - 1$ and used $(N - i) = (N - 1) - (i - 1) = M - j$.

A more elegant way to get the same result is:

$$
\begin{aligned}
\langle i \rangle &= \sum_{i=0}^{N} \binom{N}{i} (p \cdot \frac{d}{dp}) p^i \cdot q^{N-i} \\
&= (p \frac{d}{dp})(p+q)^N = N \cdot p(p+q)^{N-1} = N \cdot p
\end{aligned}
$$

The trick is to consider p as a variable before using $p + q = 1$ in the final step.

We now use the same trick to calculate the second moment:

$$
\begin{aligned}
\langle i^2 \rangle &= \sum_{i=0}^{N} \binom{N}{i} (p\frac{d}{dp})^2 p^i \cdot q^{N-i} \\
&= (p\frac{d}{dp})^2 (p+q)^N = (p\frac{d}{dp}) p \cdot N(p+q)^{N-1} \\
&= p \cdot N(p+q)^{N-1} + p^2 \cdot N(N-1)(p+q)^{N-2} \\
&= p \cdot N + p^2 \cdot N \cdot (N-1)
\end{aligned}
$$

The mean squared deviation (MSD or variance) follows as:

$$
\begin{aligned}
\sigma^2 &= \langle (i - \langle i \rangle)^2 \rangle = \langle i^2 \rangle - \langle i \rangle^2 \\
&= p \cdot N + p^2 \cdot N(N-1) - p^2 \cdot N^2 = p \cdot N - p^2 \cdot N \\
&= p \cdot N(1-p) = N \cdot p \cdot q \quad \text{vanishes for } p = 0 \text{ or } q = 0 \\
\Rightarrow &\quad \boxed{\langle i \rangle = p \cdot N, \quad \sigma^2 = N \cdot p \cdot q}
\end{aligned}
$$

σ is called *'width'* or *'standard deviation'* (SD). The *'relative width'* or *'coefficient of variation'* is then given as

$$
\frac{\sigma}{\mu} = \frac{\sqrt{N \cdot p \cdot q}}{N \cdot p} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}} \overset{N \to \infty}{\to} 0
$$

This is an example of the *'law of large numbers'*: For large N the distribution becomes very sharp.

Examples

① 10^{24} gas atoms in a box, divided into two compartments of equal size

$$
\Rightarrow \quad p = q = 0.5
$$

$$
\Rightarrow \quad \mu = N \cdot p = N/2 \quad \text{atoms on average on left side}
$$

The actual number deviates by

$$
\sigma = \sqrt{N \cdot p \cdot q} = 0.5 \cdot 10^{12}
$$

$$
\Rightarrow \quad \frac{\sigma}{\mu} = 10^{-12} \quad \text{The relative deviation is tiny.}
$$

This is the reason why thermodynamics works; N represents the Avogadro number and for so many particles, only the mean is relevant.

② We can use the result for the standard deviation to check experimentally if a die is ideal. We throw $N = 10^5$ times and find $n = 1.75 \cdot 10^4$ times a $\{6\}$

$$\Rightarrow \quad p_6 = \frac{n}{N} = 0.175$$

This is not in agreement with the ideal value $p_6 = 0.167$. However, how relevant is the deviation? We calculate the expected deviation for p_6:

$$\frac{\Delta n}{N} = \frac{\sqrt{N \cdot q \cdot p}}{N} = \frac{\sqrt{5}}{6\sqrt{N}} = 0.001$$

where we have used $p = 1/6$ and $q = 5/6$ for the ideal die. Because the measured value is farther away from the ideal p_6 than this, we conclude that the die is not ideal. Note that for $N = 10^3$ the width $\sigma = 0.012$ would have been too large to draw this conclusion.

1.8 Gauss distribution

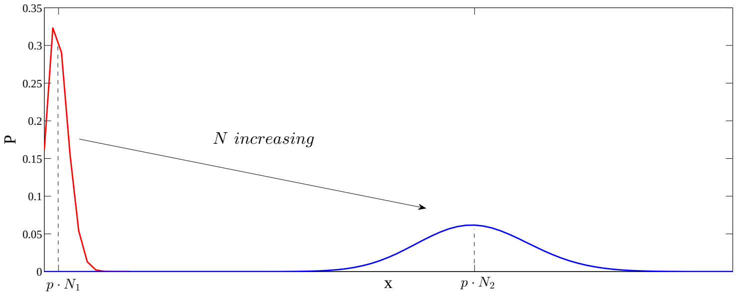

Figure 1.7: change of the binomial distribution with increasing N

We now consider the case that we perform infinitely many realizations of the binary process. For $p = \text{const}$ the limit $N \to \infty$ implies:

$$\mu = p \cdot N \to \infty \quad \text{for } N \to \infty.$$

However, the relative width $\frac{\sigma}{\mu}$ becomes smaller and smaller:

$$\frac{\sigma}{\mu} = \sqrt{\frac{1-p}{p}} \frac{1}{\sqrt{N}} \to 0 \quad \text{for } N \to \infty.$$

Where is the peak of the distribution?

$$p_i = \binom{N}{i} p^i \cdot q^{N-i}$$

A Taylor expansion around the peak is problematic, because $p_i = \binom{N}{i} p^i q^{N-i}$ is a sensitive function of i. Therefore it is better to expand its logarithm $\ln p_i$. In regard to the normalization, our method is an example of the *saddle-point approximation* or *method of steepest descent*.
We start by taking the logarithm:

$$\ln p_i = \ln N! - \ln i! - \ln(N-i)! + i \ln p + (N-i) \ln q$$

We need derivatives:

$$\begin{aligned} \frac{d \ln i!}{di} &\approx \frac{\ln(i+1)! - \ln i!}{1} \\ &= \ln(i+1) \approx \ln i \quad \text{for } i \gg 1 \end{aligned}$$

Integration yields:

$$\ln i! \approx i(\ln i - 1)$$

This agrees with Stirling's formula:

$$\ln i! = i \ln i - i + O\left(\ln(i)\right)$$

Back to p_i:

$$\begin{aligned} \Rightarrow \quad \frac{d \ln p_i}{di} &= -\ln i + \ln(N-i) + \ln p - \ln q \\ &= \ln \underbrace{\frac{(N-i)p}{i \cdot q}}_{=1} \overset{!}{=} 0 \quad \text{at peak } i_m \end{aligned}$$

$$\begin{aligned} \Rightarrow \quad (N - i_m)p &= i_m(1-p) \\ \Rightarrow \quad &\boxed{i_m = p \cdot N = \mu} \end{aligned}$$

We note that peak and average of the binomial distribution are the same in this limit.
We next consider the second derivative

$$\frac{d^2 \ln(p_i)}{di^2} = -\frac{1}{i} - \frac{1}{N-i} = -\frac{N}{i(N-i)}$$

which we then evaluate at the peak position as

$$-\frac{1}{N \cdot p \cdot q} = -\frac{1}{\sigma^2} < 0 \quad \text{therefore it is a maximum}$$

$$\Rightarrow \ln p_i = \ln p_m - \frac{1}{2}\frac{1}{\sigma^2}(i-\mu)^2 + ...$$

$$\Rightarrow \boxed{p_i = p_m \cdot e^{-\frac{(i-\mu)^2}{2\sigma^2}}}$$

This is a Gaussian with average $\mu = p \cdot N$ and variance $\sigma^2 = N \cdot p \cdot q$.
We now change from index i to a spatial variable $x = N \cdot \Delta x$ with $\mu_x = \mu \cdot \Delta x$ and $\sigma_x = \sigma \cdot \Delta x_i$.

$$\Delta x \propto \frac{1}{N} \to 0 \quad \text{for } N \to \infty$$
$$\Rightarrow \mu_x = p \cdot N \cdot \Delta x = \text{const}$$
$$\Rightarrow p(x) = p_m \cdot e^{-(x-\mu_x)^2/(2\sigma_x^2)} \quad \text{continuous probability density}$$

The probability for the <u>continuous</u> random variable x $\in [x_1, x_2]$ is given by $\int_{x_1}^{x_2} \mathrm{d}x\, p(x)$ where we integrated over $p(x)\mathrm{d}x$, the probability to find x between x and $x + \mathrm{d}x$. The continuous probability distribution $p(x)$ is actually a probability density, because we have to integrate to turn it into a probability. For simplicity, in the following we drop the subscript x again.
Regarding the normalization, we first note that it is independent of the average μ, so we can take $\mu = 0$:

$$\int p(x)\mathrm{d}x = 1 = p_m \int_{-\infty}^{\infty} \mathrm{d}x\, e^{-x^2/(2\sigma^2)}$$

Thus we have to deal with the Gauss integral:

$$\begin{aligned}
\int \mathrm{d}x\, e^{-ax^2} &= \left(\int \mathrm{d}x\, e^{-ax^2} \int \mathrm{d}y\, e^{-ay^2} \right)^{\frac{1}{2}} \\
&= \left(\int_0^{2\pi} \mathrm{d}\phi \int_0^{\infty} \mathrm{d}r\, r \cdot e^{-ar^2} \right)^{\frac{1}{2}} \\
&= \left(2\pi \int_0^{\infty} \frac{1}{2}\, \mathrm{d}u\, e^{-au} \right)^{\frac{1}{2}} \\
&= \boxed{\left(\frac{\pi}{a} \right)^{\frac{1}{2}}}
\end{aligned}$$

In the following, we often need the second moment of this distribution:

$$\begin{aligned}
\sigma^2 = \left(\frac{a}{\pi} \right)^{\frac{1}{2}} \int \mathrm{d}x\, x^2 e^{-ax^2} &= -\left(\frac{a}{\pi} \right)^{\frac{1}{2}} \frac{d}{da} \int \mathrm{d}x\, e^{-ax^2} \\
&= -\left(\frac{a}{\pi} \right)^{\frac{1}{2}} \cdot \frac{d}{da} (\frac{\pi}{a})^{\frac{1}{2}} = \frac{1}{2a}
\end{aligned}$$

In our case $a = 1/(2\sigma^2)$ and thus $p_m = (2\pi\sigma^2)^{-\frac{1}{2}}$

$$\Rightarrow \quad \boxed{p(x) = \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \cdot e^{-(x-\mu)^2/(2\sigma^2)}} \quad \text{Gauss or normal distribution}$$

Its variance is $\langle x^2 \rangle = \frac{1}{2a} = \sigma^2$. Thus the factor of 2 in the Gaussian is needed to get the correct variance.

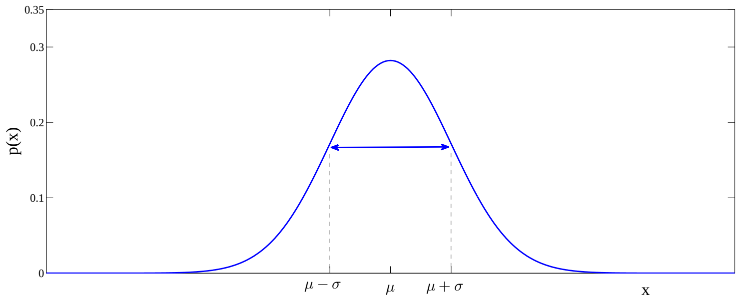

Figure 1.8: Gaussian with indicated 1σ-interval

The probability to find x within an interval of one, two or three standard deviations around μ is

$$\int_{\mu-m\sigma}^{\mu+m\sigma} \mathrm{d}x\, p(x) = \begin{cases} 0.683, & m=1 \\ 0.954, & m=2 \\ 0.997, & m=3 \end{cases}$$

1.9 Poisson distribution

Note that the continuum limit to the Gauss distribution does not work for rate events, e.g. when $p \to 0$. For the above arguments to work we need $p = \text{const}$, hence $\mu = p \cdot N \to \infty$. If we take the alternative limit

$$N \to \infty, \quad p \to 0, \quad p \cdot N = \mu = \text{const}$$

we get a different result called the *'Poisson distribution'*. We now have

$$\begin{aligned} p_i = \frac{N!}{(N-i)!i!} p^i q^{N-i} &= \frac{N(N-1)...(N-i+1)}{i!} \left(\frac{\mu}{N}\right)^i \left(1-\frac{\mu}{N}\right)^N \left(1-\frac{\mu}{N}\right)^{-i} \\ &= 1\left(1-\frac{1}{N}\right)...\left(1-\frac{i-1}{N}\right)\left(\frac{\mu^i}{i!}\right)\left(1-\frac{\mu}{N}\right)^N\left(1-\frac{\mu}{N}\right)^{-i} \to \frac{\mu^i}{i!}e^{-\mu} \end{aligned}$$

where we have used that in the limit $N \to \infty$ the first and the last parts go to 1, and that $(1-\frac{\mu}{N})^N \to e^{-\mu}$. Thus we get

$$\boxed{p_i = \frac{\mu^i}{i!} \cdot e^{-\mu}\,.}$$

Normalization can be checked as follows

$$\sum_{i=0}^{\infty} p_i = \sum_{i=0}^{\infty} \frac{\mu^i}{i!} e^{-\mu} = e^{-\mu} e^{\mu} = 1 \tag{1.5}$$

The first moment is:

$$\langle i \rangle = \sum_{i=0}^{\infty} i \frac{\mu^i}{i!} e^{-\mu} = \mu e^{-\mu} \sum_{i=1}^{\infty} \frac{\mu^{i-1}}{(i-1)!} = \mu \tag{1.6}$$

Higher moments can be derived recursively:

$$\mu \frac{d}{d\mu} \langle i^n \rangle = \sum_{i=0}^{\infty} \frac{i^n}{i!} e^{-\mu} \left(i\mu^i - \mu^{i+1} \right) = \langle i^{n+1} \rangle - \mu \langle i^n \rangle \tag{1.7}$$

For $n = 1$, this results in $\langle i^2 \rangle = \mu + \mu^2$, which means that $\sigma^2 = \mu$. Thus variance and average are identical. For $n = 2$ we get $\langle i^3 \rangle = \mu(\frac{d}{d\mu} + 1)(\mu + \mu^2) = \mu + 3\mu^2 + \mu^3$. In general, the Poisson distribution is completely determined by its first moment μ, this distribution has only one parameter.

1.10 Random walks

As indicated in Figure 1.3, a random walk is the trajectory in which for each step we draw a random number to decide whether to step to the left or to the right. This has been compared to a drunkard walking home along the pavement.
To make this more precise, we define the following quantities:

$$\text{\#right steps: } n_r = i$$

$$\text{\#left steps: } n_l = N - i$$

$$\text{time: } t = N \cdot \Delta t$$

$$\text{position: } x = m \cdot \Delta x$$

$$m = n_r - n_l = i - (N - i) = 2i - N$$

Note that at a given time step the particle can be only at an even <u>or</u> odd position. Our terms yield the following average:

$$\begin{aligned} \Rightarrow \quad \langle x \rangle &= (2 \langle i \rangle - N)\Delta x = (2p - 1)N \cdot \Delta x \\ &= \underbrace{(2p-1)\frac{\Delta x}{\Delta t}}_{\text{drift velocity v}} \cdot t \end{aligned}$$

$p = q = 1/2$ (symmetric random walk) $\Rightarrow v = 0$.

$$\begin{aligned} \langle x^2 \rangle &= \langle (2i - N)^2 \rangle \Delta x^2 \\ &= (4 \langle i^2 \rangle - 4 \langle i \rangle N + N^2)\Delta x^2 \\ &= (4Np(1-p) + N^2(4p^2 - 4p + 1))\Delta x^2 \\ &= 2(4pq\frac{\Delta x^2}{2\Delta t}t + ((2p-1)\frac{\Delta x}{\Delta t})^2 t^2) \end{aligned}$$

The second term quadratic in t is the drift term with the same drift velocity v above, which vanishes for the symmetric random walk. Then only the first term remains, which is linear in t. This is the famous diffusion part of the random walk. The term in brackets is called the diffusion constant D. For the symmetric random walk, $4pq = 1$ and we have the standard definition of D. We now have the most important result for random walks:

$$\boxed{\sigma^2 = \langle x^2 \rangle = 2 \cdot D \cdot t}$$

$$\Rightarrow \quad \sigma \propto \sqrt{t}$$

The *'mean squared displacement'* (MSD) grows $\propto t$, the *'root mean squared displacement'* (RMSD) grows as $\propto t^{0.5}$.

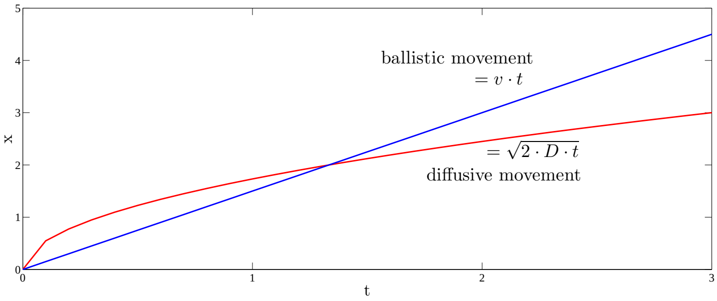

Figure 1.9: ballistic versus diffusive movement

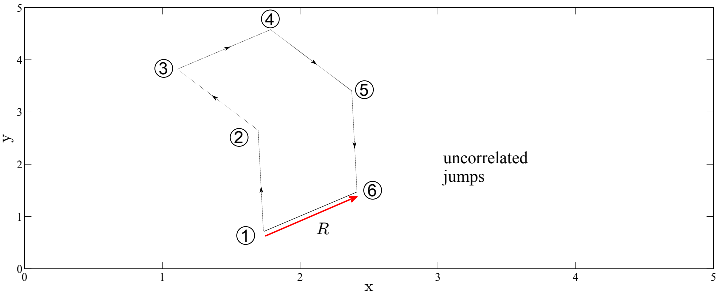

Figure 1.10: The uncorrelated jumps of a particle starting at position 1 and ending at 6. $\vec{R}$ is the resulting end-to-end vector.

A more general derivation for arbitrary dimensions d is as follows: We consider the end-to-end vector defined by

$$\vec{R} = \sum_{i=1}^{N} \vec{r}_i$$

$$\Rightarrow \left\langle \vec{R} \right\rangle = 0$$

For the mean squared displacement we get

$$\begin{aligned}
\left\langle \vec{R}^2 \right\rangle &= \left\langle (\sum_i^N \vec{r}_i) \cdot (\sum_j^M \vec{r}_j) \right\rangle \\
&= \sum_i \sum_j \langle \vec{r}_i \vec{r}_j \rangle \\
&= \sum_{i=1}^N \langle r_i^2 \rangle = \sum_{i=1}^N d \cdot \Delta x^2 = N \cdot d \cdot \Delta x^2 = 2 \cdot d \cdot \underbrace{\frac{\Delta x^2}{2\Delta t}}_{=D} \cdot t
\end{aligned}$$

$$\Rightarrow \boxed{\left\langle \vec{R}^2 \right\rangle = 2 \cdot d \cdot D \cdot t}$$

The most important aspect here is that two random vectors are uncorrelated if $i \neq j$, so all off-diagonal terms vanish and only the diagonal terms survive. This creates the linear scaling with N and therefore also with t. d is the spatial dimension and we use Pythagoras and the fact that all dimensions are equivalent (in each dimension, the walker must make a step a size Δx). In one dimension, $d = 1$, this result becomes the same as above. Note that if you identify time with contour length, this would be a model for a polymer (*freely jointed chain*).

Rate equation approach

Above we have derived the global properties of the random walk. Now we address its local properties and see that in the end we essentially get the same results again. We ask how the equation of motion looks like for a random walk particle. We introduce the probabilities p and q to jump to the right or the left, respectively, with $p + q = 1$. We then write a self-consistency equation for the probability:

$$p(x, t + \Delta t) = pp(x - \Delta x, t) + qp(x + \Delta x, t)$$

We next Taylor-expand both in time and space:

$$p(x, t + \Delta t) = p(x, t) + \dot{p}(x, t)\Delta t + \dots$$

$$p(x + \Delta x, t) = p(x, t) + p'(x, t)\Delta x + \frac{1}{2} p''(x, t)\Delta x^2 + \dots$$

$$p(x - \Delta x, t) = p(x, t) - p'(x, t)\Delta x + \frac{1}{2} p''(x, t)\Delta x^2 + \dots$$

Inserting into the jump equation give

$$p(x, t) + \dot{p}(x, t)\Delta t = (p + q)p(x, t) + (-p + q)p'(x, t)\Delta x + \frac{1}{2} p''(x, t)(p + q)\Delta x^2$$

The first terms on both sides cancel and we can divide by Δt and take the continuum limit:

$$\Delta x, \Delta t \to 0 \quad \Rightarrow \quad \boxed{\dot{p}(x,t) = -vp'(x,t) + Dp''(x,t)} \quad \text{Fokker-Planck equation}$$

Here the drift velocity $v = (p-q)\Delta x/\Delta t$ and the diffusion constant $D = (p+q)\Delta x^2/2\Delta t = \Delta x^2/2\Delta t$ are defined as above. The general form of the Fokker-Planck equation is

$$\boxed{\dot{p}(x,t) = -\partial_x[(v(x)p(x,t) - \partial_x D(x))p(x,t)]}$$

where we now also allow for position-dependent drift velocity and diffusion constant. This form shows that the Fokker-Planck equation really is a continuity equation $\dot{p} + \partial_x J = 0$, with the probability flux J being the expression in square brackets. Note that here we give the so-called overdamped form of the Fokker-Planck equation (because we stay in one dimension for the phase space, to include inertia we needed an additional phase space dimension for momentum), thus drift velocity can be equated with force, $\xi v(x) = F(x) = -\partial_x U(x)$, with some potential U and the friction coefficient ξ.
In general, the Fokker-Planck equation is a PDE of second order that has a similar character as does the Schrödinger equation. For $v = 0$ it is the time-dependent diffusion equation. For the initial condition

$$p(x, t=0) = \delta(x)$$

the solution is given by a Gaussian

$$p(x,t) = \frac{1}{\sqrt{4\pi D t}} \cdot e^{-(x-vt)^2/(4Dt)}$$

as you can check by reinserting it into the FPE. This is the same solution as above, with $\sigma = \sqrt{2 \cdot D \cdot t}$. So as the global analysis, also the local considerations give us again the square root scaling of a random walk.

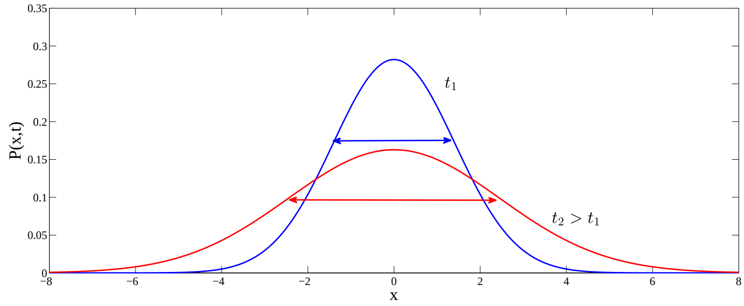

Figure 1.11: Two distributions with different t values (without drift). The distribution's width $\sigma = \sqrt{2 \cdot D \cdot t}$ increases with the root of time.

Mean first passage time

We finally consider the following question. A particle is performing a symmetric random walk in $d = 1$ with step size δ and jump time τ in the interval $x \in [0, b]$. We ask how long it will take to leave this boundary. Obviously this time will depend on the starting position x and we call it $T(x)$, the mean first passage time. We now use a similar consideration as for the rate equation to write

$$T(x) = \tau + \frac{1}{2}[T(x+\delta) + T(x-\delta)]$$

We rearrange to get

$$\frac{1}{\delta^2}[T(x+\delta) + T(x-\delta) - 2T(x)] + \frac{2\tau}{\delta^2} = 0$$

We identify again a second spatial derivate and perform the continuum limit to get

$$T''(x) + \frac{1}{D} = 0$$

Thus we have obtained an ordinary differential equation for $T(x)$. The solution has to be a polynomial of order 2 that depends on boundary conditions. For two absorbing boundaries we have $T(0) = T(b) = 0$ and the solution is

$$T(x) = \frac{1}{2D}(bx - x^2)$$

We assume that the position is released at a random position and therefore we average over x:

$$T = \frac{1}{b}\int_0^b T(x)dx = \frac{b^2}{12D}$$

Again we see that the time scales with the distance b squared. The inverse scaling with D is expected for dimensional reasons. The prefactor 12 can only be obtained by doing the full calculation. For a reflecting boundary at $x = 0$, one has to use the boundary condition $T'(0) = 0$.

1.11 Computation with random variables

Let x be some random variable with a continuous distribution $p(x)$. We consider a coordinate transformation $x \to y(x)$. Assuming that also the inverse transformation $y \to x(y)$ exists, we want to know what the probability distribution $p_y(y)$ is. This result is easily obtained using the transformation theorem.

$$1 = \int \mathrm{d}x \; p_x(x) = \int \mathrm{d}y \; \underbrace{|\frac{dx}{dy}| \cdot p_x(x(y))}_{=p_y(y)}$$

Examples

①

$$\begin{aligned} y &= c \cdot x \\ \Rightarrow \quad p_y(y) &= \frac{1}{c} \cdot p_x(\frac{y}{c}) \end{aligned}$$

Another way to see this:

$$\begin{aligned} p_y(y) &= \int \mathrm{d}x\; p_x(x)\delta(y - cx) \\ &= \int \frac{\mathrm{d}z}{c} p_x\left(\frac{z}{c}\right)\delta(z - y), \quad \text{substituting } z = c \cdot x \\ &= \frac{1}{c} p_x\left(\frac{y}{c}\right) \end{aligned}$$

Moments:

$$\begin{aligned} \langle y \rangle &= \int \mathrm{d}y\; y \cdot p_y(y) = \int (c\,\mathrm{d}x)(c \cdot x)\frac{1}{c} \cdot p_x(x) = c\,\langle x \rangle \\ \langle y^2 \rangle &= \int \mathrm{d}y\; y^2 \cdot p_y(y) = c^2\,\langle x^2 \rangle \\ \sigma_y^2 &= \langle (y - \langle y \rangle)^2 \rangle = \langle y^2 \rangle - \langle y \rangle^2 = c^2 \cdot \sigma_x^2 \end{aligned}$$

②

$$\begin{aligned} y &= -\frac{1}{\omega}\ln(1 - x) \\ \Rightarrow x &= 1 - e^{-\omega y} \\ \Rightarrow p_y(y) &= \omega \cdot e^{-\omega y} \cdot p_x(x(y)) \end{aligned}$$

Let's take x to be a uniformly distributed variable in $[0, 1]$ with $p_x(x) := 1$. We will need this later, eg for Monte Carlo simulations of the Boltzmann distribution.

$$\Rightarrow p_x = 1 \Rightarrow \text{y is exponentially distributed}$$

③ A similar procedure exists in higher dimensions. We have a look at an n-tupel $\vec{x} = (x_1, ..., x_n)$. This tupel is transformed to another tupel $y_i(\vec{x})$ with the index i in the set $[1, ..., n]$:

$$p_{\vec{y}}(\vec{y}) = \underbrace{|\frac{\partial(x_1, ..., x_n)}{\partial(y_1, ..., y_n)}|}_{\text{Jacobian}} \cdot p_{\vec{x}}(\vec{x}(\vec{y}))$$

A standard method to obtain Gaussian random variables is the Box-Müller procedure:

$$\begin{aligned}
y_1 &= \sqrt{-2\ln x_1}\cos 2\pi x_2 \\
y_2 &= \sqrt{-2\ln x_1}\sin 2\pi x_2 \\
\Rightarrow \quad x_1 &= e^{-\frac{1}{2}(y_1^2+y_2^2)} \\
\Rightarrow \quad x_2 &= \frac{1}{2\pi}\arctan\frac{y_2}{y_1} \\
\Rightarrow \quad p_{\vec{y}}(\vec{y}) &= \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}y_1^2}\cdot\frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}y_2^2}\cdot\underbrace{p_{\vec{x}}(\vec{x})}_{=1}
\end{aligned}$$

Both y_1 and y_2 are Gaussian distributed with $\sigma = 1$, if x_1 and x_2 are uniformly distributed in $[0,1]$.

1.12 Addition of random variables

We next consider the addition of two random variables x and y. Given $p_x(x), p_y(y)$, we define $z = x + y$ and write

$$p_z(z) = \int\int dxdy\delta(z-(x+y))p(x,y) = \int dx\; p(x,z-x) = \int dx\; p_x(x)p_y(z-x)$$

where in the last step we have assumed that the two variables are independent. We see that the resulting probability distribution is a convolution. This suggests that it might be helpful to use Fourier transforms.

For a given probability distribution $p(x)$, we define the following *characteristic* or *generating* function:

$$G(k) = \left\langle e^{ikx}\right\rangle = \int \mathrm{d}x\; p(x)e^{ikx}$$

which is the Fourier transform of $p(x)$.

If all moments exist and grow sufficiently slowly, we can make a Taylor expansion:

$$G(k) = \sum_{n=0}^{\infty}\frac{(ik)^n}{n!}\left\langle x^n\right\rangle$$

$$\Rightarrow \quad \left\langle x^n\right\rangle = \mu_n = \frac{1}{i^n}\cdot\left.\frac{d^nG(k)}{dk^n}\right|_{k=0}$$

Hence if we know the characteristic function we can generate all moments μ_n.

Now let us do the same thing for the logarithm:

$$\ln G(k) = \sum_{n=0}^{\infty}\frac{(ik)^n}{n!}\cdot\underbrace{\kappa_n}_{\text{cumulants of order n}}$$

The cumulants can be generated using the same trick as above:

$$\kappa_n = \frac{1}{i^n} \cdot \left. \frac{d^n \ln G(k)}{dk^n} \right|_{k=0}$$

Looking at the first few cumulants, we see that they are interesting quantities:

$$\begin{aligned}
\kappa_0 &= \ln G(k)|_{k=0} = \ln 1 = 0 \\
\kappa_1 &= \frac{1}{i} \cdot \left. \frac{G'}{G} \right|_{k=0} = \langle x \rangle \\
\kappa_2 &= \frac{1}{i^2} \left. \frac{G''G - G'^2}{G^2} \right|_{k=0} = \left\langle x^2 \right\rangle - \langle x \rangle^2 = \sigma^2
\end{aligned}$$

The third cumulant looks more complicated:

$$\kappa_3 = \frac{1}{i^3} \left. \left(G''' - 3G'G'' + 2G'^3 \right) \right|_{k=0} = \mu_3 - 3\mu_1\mu_2 + 2\mu_1^3$$

The cumulants are specific combinations of the moments. The relation can also be reversed:

$$\begin{aligned}
\mu_1 &= \kappa_1 \\
\mu_2 &= \kappa_2 + \kappa_1^2 \\
\mu_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3
\end{aligned}$$

You can characterize a probability distribution either by its moments or by its cumulants (except if its moments are not finite, as for the Lorentz distribution, or if they grow too fast, then you have to know the distribution directly). A theorem of Marcinkiewicz from 1939 states that either all but the first two cumulants vanish or there are an infinite number of non-vanishing cumulants.

Examples

The most important example is the Gauss distribution. Its characteristic function is given as:

$$\begin{aligned}
G(k) &= \int \mathrm{d}x \, e^{ikx} \cdot \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \cdot e^{-(x-\mu)^2/(2\sigma^2)} \\
&= \int \mathrm{d}x \, \frac{1}{(2\pi\sigma^2)^{\frac{1}{2}}} \cdot e^{-\frac{(x-\mu-ik\sigma^2)^2+k^2\sigma^4-2ik\sigma^2\mu}{(2\sigma^2)}} \\
&= e^{-\frac{1}{2}k^2\sigma^2+ik\mu} \\
\Rightarrow \ln G(k) &= ik\mu - \frac{1}{2}k^2\sigma^2 \\
\Rightarrow \kappa_1 &= \frac{1}{i} \left. \frac{d \ln G}{dk} \right|_{k=0} = \mu \\
\kappa_2 &= \frac{1}{i^2} \left. \frac{d^2 \ln G}{dk^2} \right|_{k=0} = \sigma^2
\end{aligned}$$

All other κ_i vanish. The Gauss distribution is the only one having just two non-zero cumulants.

Back to addition of two random variables x and y:

$$z = x + y$$

$$\Rightarrow \quad p_z(z) = \int \mathrm{d}x \; p_x(x) \cdot p_y(x - z)$$

Due to the convolution theorem this becomes

$$G_z(k) = G_x(k) \cdot G_y(k)$$

$$\Rightarrow \ln G_z(k) = \ln G_x(k) + \ln G_y(k)$$

Having in mind the definition of the cumulants this yields

$$\kappa_n^{(z)} = \kappa_n^{(x)} + \kappa_n^{(y)}$$

By iteration, hence $z = x_1 + ... + x_N$, it is possible to construct the n-th cumulant:

$$\kappa_n^z = \sum_{i=1}^{N} \kappa_n^{(i)}$$

We are now in a state to formulate the law of large numbers:

> *The average of many realizations approaches the expectation value.*

Proof

$$Z = \frac{1}{N}(x_1 + ... + x_N)$$

$$\Rightarrow \quad \langle z \rangle = \frac{1}{N} \sum_i^N \langle x_i \rangle \underbrace{=}_{x_i := x} \langle x \rangle$$

$$\Rightarrow \quad \sigma_z^2 = (\frac{1}{N})^2 \sum_{i=1}^N \sigma_{x_i}^2 \underbrace{=}_{x_i := x} \frac{1}{N}\, \sigma_x^2 = \frac{\sigma_x^2}{N}$$

$$\kappa_n^z \underbrace{=}_{n>2} (\frac{1}{N})^n \sum_{i=1}^N \kappa_n^{x_i} \underbrace{=}_{x_i := x} \frac{1}{N^{n-1}}\, \kappa_n^x$$

Thus the variance and all higher order moments / cumulants vanish in the large N limit.
In order to get a finite variance, we have to sum differently. This leads to the central limit theorem:

The sum of many independent variables has a Gauss distribution after appropriate rescaling.

From above we see how we have to rescale to get this result:

$$Z = \frac{1}{\sqrt{N}}((x_1 - \langle x_1 \rangle) + ... + (x_N - \langle x_N \rangle))$$

Then $\langle Z \rangle = 0$, $\sigma_z^2 = \sigma_x^2$ and all higher cumulants vanish in the limit of very large N. Therefore this sum indeed has a Gauss distribution (normal distribution if we also scale the variance to 1).

1.13 Information entropy

In 1948 Claude Shannon introduced a measure for uncertainty of a probability distribution which agrees with the thermodynamic definition of entropy. The relation between information theory and statistical physics was further discussed by Edwin Jaynes in 1957, who introduced the maximum entropy principle, which today is widely used in statistical data analysis (eg image processing)[2].
For a given discrete probability distribution p_i, we introduce a dimensionless measure for its uncertainty (information theory) or disorder (physics), the entropy:

$$\boxed{S = -\sum_i p_i \, \ln p_i}$$

Comments

① One can interpret $\ln(1/p_i) = -\ln(p_i)$ is the surprise or uncertainty about state i being the outcome of an experiment: the smaller p_i, the more surprised we are. The entropy is then the average over these individual surprises.

② $0 \leq p_i \leq 1 \;\Rightarrow\; S \geq 0$

③ $\exists j: \; p_j = 1 \;\Rightarrow\; p_i = 0 \, for \; i \neq j \;\Rightarrow\; S = 0$
The entropy is minimal, if outcome is certain.

④ For a given homogeneous distribution $p_i = \text{const} = \frac{1}{\Omega}$, with $\Omega = \#states$:

$$\sum_i p_i = \sum_{i=1}^{\Omega} \frac{1}{\Omega} = \Omega \cdot \frac{1}{\Omega} = 1$$

---

[2] For an introduction into information science, we recommend Christopher M. Bishop, Pattern recognition and machine learning, Springer 2006. A very pedagogical introduction is given by Henry Pinkard and Laura Waller, A visual introduction to information theory, `https://doi.org/10.48550/arXiv.2206.07867`

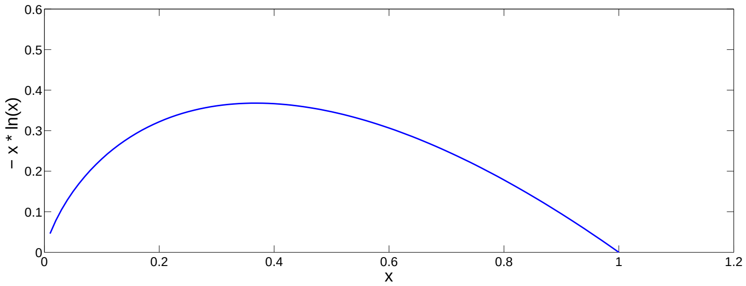

Figure 1.12: $-x \cdot \ln x$ as a function of x

$$S_h = -\sum_{i=1}^{\Omega} \frac{1}{\Omega} \ln \frac{1}{\Omega} = \ln \Omega$$

Hence we see that entropy increases with the number of states.

⑤ Entropy is maximal for the homogeneous distribution.

$$\begin{aligned} S_h - S &= \sum_{i=1}^{\Omega} p_i \ln p_i + \ln \Omega \\ &= \sum_{i=1}^{\Omega} p_i \ln p_i + \sum_{i=1}^{\Omega} p_i \ln \Omega \\ &= \frac{1}{\Omega} \sum_{i=1}^{\Omega} \underbrace{(p_i \, \Omega)}_{=x_i} \ln(\underbrace{p_i \, \Omega}_{x_i}) \end{aligned}$$

Our aim is to minimize $S_h - S$. Direct minimization gives us

$$\delta(S_h - S) = \frac{1}{\Omega} \sum_{i=1}^{\Omega} (\ln x_i + 1) \, \delta x_i \overset{!}{=} 0 \Rightarrow x_i = e^{-1}$$

because $x = 1/e$ is the position of the minimum of $x \ln x$ and because each term has to be minimized by itself independent of the others. However, this result does not satisfy the normalization constraint $\sum_{i=1}^{\Omega} x_i = \Omega$. We add this constraint to the target function with a Lagrange parameter λ and thus have

$$\frac{1}{\Omega} \sum_{i=1}^{\Omega} (\ln x_i + 1 + \lambda) \, \delta x_i \overset{!}{=} 0 \Rightarrow x_i = e^{-(1+\lambda)}$$

Implementing the constraint specifies $\lambda = -1$ and thus $x_i = 1$ will be the correct solution. Hence the entropy (or disorder) is maximal for the (normalized) homogeneous distribution and all other distributions have smaller entropies.

The same result can also be obtained by using a small trick rather than a Lagrange parameter. Rather than introducing λ, we add zero to the target function in a way that only works for the properly normalized distribution:

$$S_h - S = \frac{1}{\Omega}\sum_{i=1}^{\Omega} x_i \ \ln x_i + \underbrace{\frac{1}{\Omega}\sum_{i=1}^{\Omega}(1-x_i)}_{=0}$$

$$= \frac{1}{\Omega}\sum_{i=1}^{\Omega}(x_i \ \ln x_i + 1 - x_i)$$

$\Rightarrow$ The only minimum (Figure 1.13) is at $x_i = 1$.

Thus we get the same result as for the calculation with the Lagrange parameter.

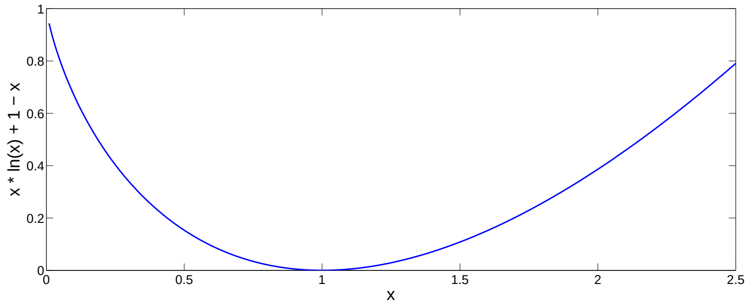

Figure 1.13: $x \cdot \ln x + 1 - x$ as a function of x

⑥ A simple example to illustrate this point: in the example of apples and oranges in boxes, we had $p_l = 4/10$ and $p_r = 6/10$. This gives the entropy $S = 4/10 \ln(10/4) + 6/10 \ln(10/6) = 0.673$. We compare with the homogeneous distribution, which has entropy $S = 1/2 \ln(2) + 1/2 \ln(2) = \ln(2) = 0.693$, thus the homogeneous distribution has indeed the larger entropy. If we make the distribution more heterogeneous, let say $S = 1/10 \ln(10) + 9/10 \ln(10/9) = 0.325$, then entropy becomes even smaller. Because we use the natural logarithm, 1 bit = $\ln(2)$ and we should measure everything in this unit. So the three example would be 1, 0.971 and 0.469 bits.

⑦ We consider two independent subsystems 1 and 2:

$$p_{ij} = p_i^{(1)} \cdot p_j^{(2)}$$

$$
\begin{aligned}
\Rightarrow S &= -\sum_{i,j} p_{ij} \, \ln p_{ij} \\
&= \sum_{ij} p_i^{(1)} \, p_j^{(2)} \, \left[\ln p_i^{(1)} + \ln p_j^{(2)} \right] \\
&= -\sum_i \left(\sum_j p_j^{(2)} \right) \, p_i^{(1)} \, \ln p_i^{(1)} - \sum_j \left(\sum_i p_i^{(1)} \right) \, p_j^{(2)} \, \ln p_j^{(2)} \\
&= S_1 + S_2
\end{aligned}
$$

Entropy is additive for independent subsystems.

⑧ A bit of information theory: How many ‘yes-or-no’ questions do we have to ask to find out which state j is realized? We divide Ω into two disjunct domains Ω_1 and Ω_2 such that

$$
\sum_{i \in \Omega_1} p_i = \sum_{i \in \Omega_2} p_i = \frac{1}{2}
$$

Firstly we ask: Is j in Ω_1? We then choose the correct domain and scale the probability by 2 and repeat the dividing procedure. This has to be done k-times while k is given by

$$
2^k \, p_j = 1
$$

$$
\Rightarrow \boxed{k = -\log_2 p_j}
$$

On average, the number of questions required is

$$
\sum_{i=1}^{\Omega} p_i \, (-\log_2 p_i) = S / \ln 2
$$

because change of base gives a constant factor, $\log_2 x = (\ln x)/(\ln 2)$. In general it does not matter to which base we take the logarithm, because this only changes the unit in which we have to measure entropy; while in physics we always take the natural logarithm, in information theory the binary logarithm is used. We conclude again that entropy is a measure for uncertainty in our expectation of the result.

⑨ Information theory is closely related to coding messages. Shannon proved the noiseless coding theorem: entropy is a lower bound on the number of bits needed to transmit the state of a random variable.

⑩ We now aim to maximize entropy with the additional constraint:

$$
U = \sum_i p_i \, E_i = \text{const}
$$

Here E_i is some scalar property of each state (later this will be energy, thus U will be the average energy).

$$\Rightarrow \quad \delta S = -\sum_{i=1}^{\Omega} (\ln p_i + 1)\, \delta p_i \overset{!}{=} 0$$

Auxiliary conditions:

① $\sum_i p_i = 1 \;\Rightarrow\; \sum_i \delta p_i = 0$

② $\sum_i p_i\, E_i = U \;\Rightarrow\; \sum_i E_i\, \delta p_i = 0$

Method of Lagrange multipliers:

$$-\sum_i \underbrace{(\ln p_i + 1 + \lambda_1 + \lambda_2\, E_i)}_{=0} \cdot \delta p_i = 0$$

$$\Rightarrow \quad \boxed{p_i = e^{-(1+\lambda_1+\lambda_2\, E_1)}}$$

The auxiliary conditions then yield

①

$$e^{-(1+\lambda_1)} \sum_i e^{-\lambda_2\, E_i} = 1$$

$$e^{-(1+\lambda_1)} = \frac{1}{\sum_i e^{-\lambda_2\, E_i}} := \frac{1}{Z}$$

②

$$\frac{\sum_i E_i\, e^{-\lambda_2\, E_i}}{\sum_i e^{-\lambda_2\, E_i}} = U$$

$$\Rightarrow \lambda_2 := \beta \text{ is a function of U}$$

$$\Rightarrow \quad \boxed{p_i = \frac{1}{Z} \cdot e^{-\beta E_i}}$$

This exponential distribution later becomes the Boltzmann distribution.

Continuous distribution

In the case of a continuous distribution the entropy becomes:

$$S = -\int \mathrm{d}x\; p(x)\; \ln p(x)$$

Mathematically this case is a bit tricky as you can see that we have a problem with dimensions, so a more rigorous derivation gives some correction factors. If we do a similar variation like above and enforce conservation of probability, fixed first moment

μ and fixed variance σ^2 using Lagrange multipliers, then the result would be the Gaussian distribution. The corresponding entropy is

$$\begin{aligned} S &= \int \mathrm{d}x \, \frac{1}{\sqrt{2\pi\sigma}} \, e^{-(x-\mu)^2/(2\sigma)} \, \left(\frac{(x-\mu)^2}{2\sigma^2} + \frac{1}{2} \, \ln 2\pi \, \sigma^2 \right) \\ &= \frac{1}{2} \, (1 + \ln(2\pi \, \sigma^2)) \end{aligned}$$

The entropy increases with the variance: the broader the distribution, the larger the disorder or uncertainty. This is a similar conclusion as above for the discrete distribution, where the homogeneous distribution is optimal.

1.14 Mutual information

We finally discuss other important definitions which follow from the concept of information entropy and which help to understand it better. The first one is the *Kullback–Leibler divergence* (also called *relative entropy*), which describes how different two probability distributions are from each other:

$$D(p\|q) := \sum_x p_x \ln \frac{p_x}{q_x}$$

One immediately notes that $D(p\|q) \neq D(q\|p)$. Further one can show that $D(p\|q) \geq 0$, with the equality being valid if the two distributions are identical.
For the proof, we need *Jensen's inequality*. For any convex function, we know that

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0)$$

because it always has to lie above the linear approximation. If we now choose $x_0 = \langle x \rangle$ and average on both sides, we see that the linear term vanishes and we get

$$\langle f(x) \rangle \geq f(\langle x \rangle)$$

Thus the averaged value of f is larger or equal than f evaluated at the average value of x. A typical convex functions for which we can apply this would be $f(x) = e^{-x}$.
The desired proof now amounts to applying Jensen's inequality to the function $f(x) = -\ln x$:

$$D(p\|q) = \sum_x p_x \ln \frac{p_x}{q_x} = \left\langle -\ln \frac{q_x}{p_x} \right\rangle \geq -\ln \left\langle \frac{q_x}{p_x} \right\rangle = -\ln \sum_i p_x \frac{q_x}{p_x} = -\ln 1 = 0$$

Next we consider two random variables with joint probability $p_{x,y}$. One example would be the apples and oranges (x) in the two boxes left and right (y). The *mutual information* between the two random variables is defined as

$$I(x;y) := D(p_{x,y}\|p_x p_y) = \sum_{x,y} p_{x,y} \ln \frac{p_{x,y}}{p_x p_y}$$

where p_x and p_y are the marginal probabilities. One immediately notes that now $I(x;y) = I(y;x)$. Moreover $I = 0$ if the two random variables are independent. Because this is a Kullback–Leibler divergence, we also have $I \geq 0$. Mutual information I compares a joint distribution of two random variables to one of two independent random variables. The larger I, the more information is shared between the two random variables.

Using the definitions of conditional probability, we can rewrite the mutual information

$$I(x;y) = \sum_{x,y} p_{x,y} \ln \frac{p_{x|y} p_y}{p_x p_y} = \sum_{x,y} p_{x,y} \ln \frac{p_{x|y}}{p_x} = S(x) - S(x|y)$$

where the definition of the conditional entropy is

$$S(x|y) = -\sum_{x,y} p_{x,y} \ln p_{x|y}$$

In the same manner we get

$$I(x;y) = S(y) - S(y|x)$$

We next consider the joint entropy and again use the definitions of conditional probabilities to get

$$S(x,y) = -\sum_{x,y} p_{x,y} \ln p_{x,y} = -\sum_{x,y} p_{x,y} \ln(p_{x|y} p_y) = S(x|y) + S(y)$$

In the same manner we have

$$S(x,y) = S(y|x) + S(x)$$

Thus one can use the conditional entropies to connect joint entropy and mutual information:

$$S(x|y) = S(x,y) - S(y) = S(x) - I(x;y) \Rightarrow \boxed{S(x,y) = S(x) + S(y) - I(x;y)}$$

Thus mutual information quantifies the reduction of uncertainty about the pair (x,y) due to their mutual dependence. Its upper bound is set by the random variable containing less information, that is by the smaller of $S(x)$ or $S(y)$. This means that a random variable can not convey more information about another random variable as it contains itself.

To summarize, we have introduced the following concepts. Entropy is the average uncertainty in a single random event. Mutual information is uncertainty that is in common with another random event. Conditional entropy is the remaining uncertainty in one random event after subtracting the mutual information shared with another random event. Joint entropy is the average uncertainty when describing both random events together.

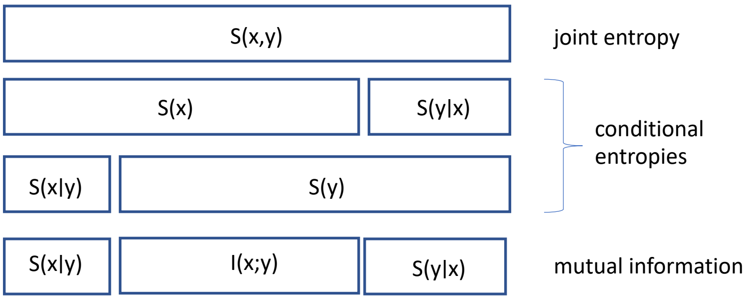

Figure 1.14: The joint entropy can be decomposed in several ways. Mutual information is the difference to the sum of the entropies of the two marginal distributions, $S(x, y) = S(x) + S(y) - I(x; y)$.

2 The microcanonical ensemble

2.1 Thermodynamic equilibrium

It is a matter of experience that after some time (*relaxation time*) an isolated physical system arrives at a state of macroscopic rest (*thermodynamic equilibrium*). Then a few macroscopic quantities (such as volume V, pressure p and temperature T) are sufficient to completely characterise the system (*state variables*). If external conditions are changed, the system adopts a new equilibrium state. If we change external conditions very slowly, the system is going through a series of equilibrium states (*quasi-stationary process*). Even if a system is in macroscopic equilibrium, its constituents move on a microscopic level as long as temperature is finite. This becomes evident when observing dust particles in gas or liquid, or to be more specific, aerosols in the atmosphere or colloids in a fluid solvent: they all undergo Brownian motion. Therefore many microscopic states correspond to the macrosopic state. The essence of equilibirum statistical physics is to calculate the macroscopic properties of equilibirum states from our knowledge of the underlying microscopic physics. As long as the system is not completely at rest (that means at finite temperature), it will always explore as many microstates as possible.

Examples

① Consider a gas container with two compartments, of which only one is filled. If the wall is removed, the gas expands until it fills the whole space (Figure 2.1).

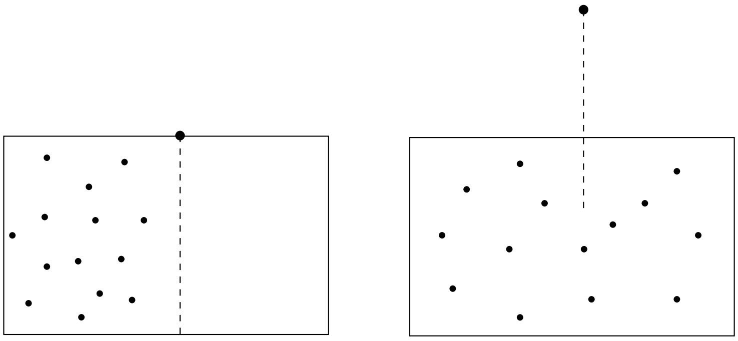

Figure 2.1: A gas expands into empty space if the wall is removed. In a larger volume, the gas has more configurations.

② Again we consider a box with two compartments, this time with two different fluids (black and white). After the separating wall is removed, the fluids mix due to diffusion. A grey final state is established (Figure 2.2).

Figure 2.2: Two fluids mix by diffusion. Like in the first example, the mixed state has more microstates.

③ Stretch a polymer to its maximal extension and let it shrink afterwards. The following movement results in a coiled equilibrium state (Figure 2.3).

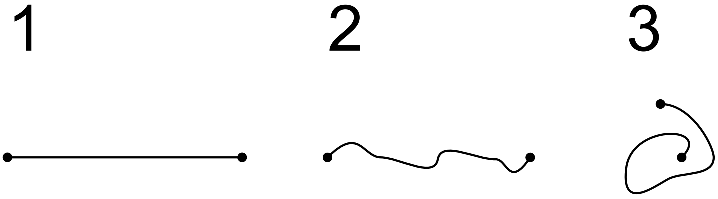

Figure 2.3: The equilibrium state of a polymer is a random coil. In a coiled state it can explore many more microstates than in a stretched state.

2.2 Micro- and macrostates

Talking about many ($\approx 10^{23}$) particles, it is neither possible nor desirable to follow their trajectories. Moreover we could be never certain about the initial conditions to assume for such a calculation. Therefore the statistical description aims at averaging or projecting the microstates onto the macrostates without going into the issue of time evolution. The solution to this challenge is a central postulate which will allow us to perform such a procedure. For the averaging procedure, we need a statement about the probability p_i that a certain microstate i occurs.

In thermodynamic equilibrium, the microstates are visited one after the other (*'time average'*) or occur simultaneously in equivalent realizations (*'ensemble average'*). The equivalence of time and ensemble averages is called *'ergodicity'*.

In statistical physics, we first have to identify the relevant microstates. In general, they are characterised by their quantum numbers resulting from the solution of Schrödinger's equation. Each state i then has a distinct probability p_i.

Examples

① Consider a spin-$\frac{1}{2}$ particle with spin $|\uparrow\rangle$ and $|\downarrow\rangle$, hence this is a two state system. For an electron we have the magnetic moment $\mu = e\hbar/2mc$, energy $E = -2\mu H m_z$ and spin $m_z = \pm 1/2$. H is the external magnetic field.

② N non-interacting spin-$\frac{1}{2}$ particles. Now the system has 2^N possible states. The energy of the system is then given as $E = \sum_{i=1}^{N} E_i$.

③ 1D harmonic oscillator $E_n = \hbar\,\omega\,(n + \frac{1}{2}), \quad n = 0, 1, ...$ This system has infinitely many but countable states.

④ N atoms in a crystal, each vibrating around its mechanical equilibrium state ("harmonic crystal"). In three dimensions, a simple model is to consider this is a collection of 3N non-interacting harmonic oscillators ("Einstein model"). Although these are infinitely many states, they can be counted easily.

⑤ A particle in a box. The Schrödinger equation then is:

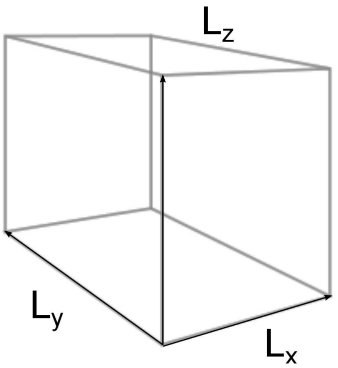

Figure 2.4: 3D potential well

$$\frac{-\hbar^2}{2m}\Delta\Psi = E\,\Psi$$

$$\Rightarrow \; \Psi(t) = \sin\frac{\pi n_x x}{L_x} \cdot \sin\frac{\pi n_y y}{L_y} \cdot \sin\frac{\pi n_z z}{L_z}$$

$$E = \frac{\pi^2\hbar^2}{2\,m}\,(\frac{n_x^2}{L_x^2} + \frac{n_y^2}{L_y^2} + \frac{n_z^2}{L_z^2})$$

This system has infinitely many but countable states (n_x, n_y, n_z).

⑥ Now consider N such particles in a box without interaction ("ideal gas"). Like for the harmonic crystal, this will be infinitely many states, but they can be counted.

2.3 Density of states

We see that in quantum mechanics we can easily count the number of states. The question now is how to deal with classical systems. We clarify this issue using the 1D harmonic oscillator.
In quantum mechanics the energy levels are given as:

$$E_n = \hbar\,\omega\,(n + \frac{1}{2})$$

The number of states with $E_n \leq E$ is then obtained as

$$N_E \approx \frac{E}{\hbar\omega}$$

In the classical case it is a bit more complicated:

$$m\ddot{x} + kx = 0 \;\Rightarrow\; \omega^2 = \frac{k}{m}$$

$$p = \dot{x}m,\; q = x$$

$$\Rightarrow \quad H(q,p) = \frac{1}{2}\,m\,\omega^2\,q^2 + \frac{p^2}{2m} \overset{!}{=} E$$

Defining the quantities:

$$a := \sqrt{2mE} \quad b := \sqrt{\frac{2E}{m\omega^2}}$$

$$\Rightarrow \quad \left(\frac{p}{a}\right)^2 + \left(\frac{q}{b}\right)^2 = 1$$

This is the equation of an ellipse in phase space with area $A = \pi\,a\,b$ which represents the number of states with energy $\leq E$.

$$\Rightarrow \quad V_E = \pi\,a\,b = \frac{2\pi E}{\omega}$$

Comparing this result with the quantum mechanics one gets

$$\boxed{N_E = \frac{V_E}{h}}$$

Hence we see that the number of classical states is obtained by dividing the phase space volume through the Planck constant h.

This example suggests to count states in a classical system by dividing phase space into boxes of size h. Another argument leading to the same conclusion comes from wave mechanics. According to de Broglie, for a single particle with mass m we have $p = \hbar k = mv = h/\lambda$, therefore $v = h/(m\lambda)$, where λ is the wavelength of the particle. For a wave packet with size l, the uncertainty in position is $\delta x = l$. The relative uncertainty in

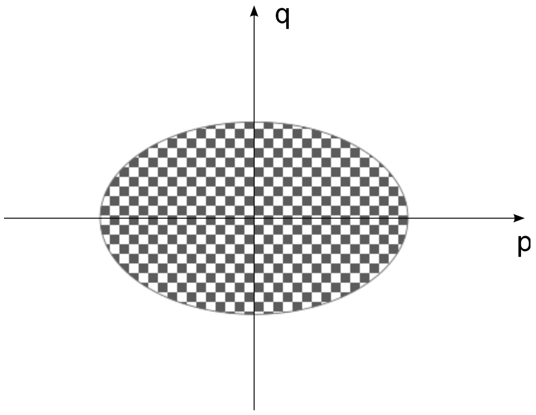

Figure 2.5: The classical 1D harmonic oscillator corresponds to a movement in 2D phase space on the elliptical path that corresponds to energy conservation. The enclosed area A represents all states with equal or lower energy. In quantum mechanics, state space is divided into small boxes of size h. Thus A/h is the number of states corresponding to the classical system.

wavelength is $\delta\lambda/\lambda = \lambda/l$. Therefore $\delta p = m\delta v = h\delta\lambda/\lambda^2 = h/l$ and we finally get the Heisenberg uncertainty relation $\delta x \delta p = h$, because l drops out. Therefore in quantum mechanics it is meaningless to choose a smaller box size in phase space and it becomes possible to count the number of states by dividing classical phase space volume by the Planck constant h (for one particle).

2.4 The fundamental postulate

For both classical and quantum mechanical systems, energy E is a conserved quantity for an isolated system. For a system at rest, E is the only relevant integral of motion, as suggested by Noether's theorem (energy conservation follows from invariance to translation in time, and this is always valid). We now consider the ensemble of microstates corresponding to a macrostate with state variables E, V, N given (*'microcanonical ensemble'*). For a system at rest, there is no physical reason to single out any other features of the system as being relevant on the macroscopic level. Let's assume that the value of E can be determined with an uncertainty δE. We look at the quantum mechanical number of microscopic states between $E - \delta E$ and E (*'microcanonical partition sum'*):

$$\Omega(E) = \sum_{n:\ E-\delta E \leq E_n \leq E} 1$$

Later we will see that the exact value of δE does not matter, so we do not introduce any dependance on it. For a classical multiparticle system, one has to correct not only for the microscopic box size as explained above for a single particle, but also for the fact that in quantum mechanics particles are indistinguishable (the wavefunction is invariant under particle exchange), which yields a factor of $1/N!$ due to $N!$ being the number of possible permutations:

$$\Omega(E) = \frac{1}{h^{3N}\,N!}\int_{E-\delta E\leq\mathcal{H}(\vec{q},\vec{p})\leq E} \mathrm{d}\vec{q}\,\mathrm{d}\vec{p}$$

if we count the number of microscopic states in classical phase space (counting from quantum mechanics is straight forward and does not need any further comments).

What is the probability density ρ for the microcanonical ensemble? First we know that ρ has to vanish outside the energy hypersurface in phase space. Second we expect that the system smears out its presence in phase space as it goes towards equilibrium. Several reasons have been evoked for this expectation, including deterministic chaos and spreading of the wavefunction in a system that is not perfectly isolated. Also one can use Liouville's and von Neumann's equations to motivate this (see below). Here however we take this expectation as our fundamental postulate and show that the whole machinery of statistical mechanics and thermodynamics follows from this one postulate:

For an isolated system at equilibrium, all microstates are equally probable.

$$\Rightarrow \quad p_i(E) = \begin{cases} \dfrac{1}{\Omega(E)} = \text{const} & E-\delta E \leq E_i \leq E \\ 0 & \textit{otherwise} \end{cases}$$

Due to its huge success, there is no doubt that the fundamental postulate is correct.

We immediately note with our knowledge from information theory that the homogeneous distribution maximizes entropy. Moreover the Shannon entropy now simplifies to the Boltzmann formula:

$$S = -k_B \sum_{i=1}^{\Omega} p_i\,\ln p_i = k_B\,\ln\Omega$$

This equation is in fact written on his gravestone at the Zentralfriedhof at Vienna. The physical constant k_B is the Boltzmann constant and later gives us a temperature scale. The fundamental postulate immediately suggests an extremum principle, which in many physics theories is the most convenient and elegant way to solve problems (e.g. Hamilton's principle of least action in mechanics or Fermat's principle in optics). Note that the maximum entropy principle does not require any more physical reasoning, it is simply a mathematical consequence of the fundamental postulate, as shown in the section on information entropy in the first chapter. Our physical understanding for this process is that due to unavoidable microscopic interactions, more and more regions of phase space are explored as the systems moves towards equilibrium.

Examples for the microcanonical ensemble

① **three spin-$\frac{1}{2}$ particles**

Possible states are:

n	m_1	m_2	m_3	*energy E*
1	+	+	+	$-3\,\mu H$
2	+	+	-	$-1\,\mu H$
3	+	-	+	$-1\,\mu H$
4	-	+	+	$-1\,\mu H$
5	+	-	-	$+1\,\mu H$
6	-	+	-	$+1\,\mu H$
7	-	-	+	$+1\,\mu H$
8	-	-	-	$+3\,\mu H$

If we know that $E = -\mu\ H$, then the corresponding microcanonical ensemble is

$$\{(++-),(+-+),(-++)\} \ \Rightarrow\ \Omega = 3.$$

Each state is equally likely with

$$p_n = \frac{1}{\Omega} = \frac{1}{3}$$

② **ideal gas**

$$\Omega(E) = \frac{1}{h^{3N}\ N!} \int_{E-\delta E \le \mathcal{H}(\vec{q},\vec{p}) \le E} \underbrace{\mathrm{d}\vec{q}\ \mathrm{d}\vec{p}}_{\text{6N-dimensional integral}}$$

$$\mathcal{H} = \sum_{i=1}^{N} \frac{\vec{p_i}^{\,2}}{2m_i} = \frac{\vec{p}^2}{2m} \overset{!}{=} E$$

$$\Rightarrow \quad \Omega = \frac{1}{h^{3N}\ N!}\ V^N\ f(R,\delta R)$$

where $f(R,\delta R)$ is the volume of a spherical shell with radius $R = \sqrt{2mE}$ and thickness $\delta R = \frac{1}{2}\ \sqrt{(2m)/E}\ \delta E$ in 3N-dimensional space. ($D = 3N$)

From analysis we know the expression for the volume of a sphere with radius R in D-dimensional space:

$$V(R) = \frac{\pi^{D/2}}{(D/2)!}\ R^D$$

Taking $D = 3$ we recover the well known formula:

$$V = \frac{4\pi}{3}\, R^3$$

because $(3/2)! = \Gamma(5/2) = 3\sqrt{\pi}/4$.

$$\begin{aligned}
\Rightarrow \quad f(R, \delta R) &= V(R) - V(R - \delta R) \\
&= \frac{\pi^{D/2}}{(D/2)!} \left[\underbrace{R^D - (R - \delta R)^D}_{R^D\,\left[1 - (1 - \frac{\delta R}{R})^D\right]} \right]
\end{aligned}$$

For $D = \text{const}$, we could expand this expression in $\frac{\delta R}{R}$ and got $\frac{D\delta R}{R}$ for the square bracket. Here however we are interested in the limit $D \to \infty$ for $\frac{\delta R}{R} \ll 1$:

$$(1 - \frac{\delta R}{R})^D = e^{D\ \ln(1 - \frac{\delta R}{R})} \approx e^{-D\ \frac{\delta R}{R}} \overset{D \to \infty}{\to} 0$$

$$\Rightarrow \quad f(R, \delta R) = \frac{\pi^{D/2}}{(D/2)!}\, R^D = V(R)$$

δR drops out; in high dimensions the volume of the sphere is mainly at its surface.

$$\Rightarrow \quad \boxed{\Omega(E, V, N) = \frac{V^N\ \pi^{3N/2}\ (2mE)^{3N/2}}{h^{3N}\ N!\ (\frac{3N}{2})!}} \text{microcanonical partition sum for an ideal gas}$$

The entropy of an ideal gas then becomes:

$$S(E, V, N) = k_B \ln \Omega = k_B \left\{ N \ln V + \frac{3N}{2}\ \ln(\frac{2\pi m E}{h^2}) \underbrace{- \ln N! - \ln(\frac{3N}{2})!}_{:=\ \text{expression A}} \right\}$$

Considering $N \gg 1$ and applying two times Stirling's formula, expression A becomes:

$$\begin{aligned}
- \ln N! - \ln(\frac{3N}{2})! &= - \left[N\ (\ln N - 1) + \frac{3N}{2}\ (\ln \frac{3N}{2} - 1) \right] \\
&= -N \left[-\frac{5}{2} + \frac{3}{2}\ \ln \frac{3}{2} + \frac{5}{2}\ \ln N \right] \\
&= N \left[\frac{5}{2} - \ln \left(\left(\frac{3}{2} \right)^{3/2}\ N^{5/2} \right) \right]
\end{aligned}$$

$$\Rightarrow \quad \boxed{S = k_B\ N\ \left\{ \ln \left[\left(\frac{V}{N} \right)\ \left(\frac{4\pi m E}{3h^2 N} \right)^{3/2} \right] + \frac{5}{2} \right\}}$$

Comments

① S, V, E, N are all extensive and occur only in the combinations $s = \frac{S}{N}$, $v = \frac{V}{N}$, $e = \frac{E}{N}$ (*'specific quantities'*).

② The exponent $\frac{3}{2}$ reflects the two facts that each particle has 3 degrees of freedom and that momentum and energy are related by a square root.

③ With $\lambda := \left(\frac{3h^2}{4\pi m e}\right)^{\frac{1}{2}}$ (*'thermal de Broglie' wavelength*) we get:
$s = k_B \ \left\{\ln\left(\frac{v}{\lambda^3}\right) + \frac{5}{2}\right\}$

④ This first non-trivial result from statistical physics has been known before in thermodynamics as *'Sackur-Tetrode-equation'*. It has been impressively verified in experiments.

2.5 Equilibrium conditions

We consider two isolated systems to form a composite isolated system. The entropy then is given as:

$$\begin{aligned} S &= k_B \ \ln(\Omega_1 \cdot \Omega_2) \\ &= k_B \ \ln \Omega_1 + k_B \ \ln \Omega_2 \\ &= S_1 + S_2 \end{aligned}$$

The quantities E, V and N are also additive (*'extensive'*).

Thermal contact

We next bring the systems in thermal contact, such that they can exchange energy, but not volume or particles:

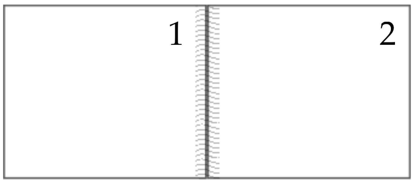

Figure 2.6: Thermal contact allows for heat exchange between the two systems.

In the new equilibrium state, energy will be E_1' and E_2', with:

$$E = E_1 + E_2 = E_1' + E_2'$$

In equilibrium, $S = k_B\ \ln \Omega$ must be maximal:

$$\begin{aligned} \Rightarrow\ \mathrm{d}S &= \frac{\partial S_1}{\partial E_1'}\,\mathrm{d}E_1' + \frac{\partial S_2}{\partial E_2'}\ \underbrace{\mathrm{d}E_2'}_{=-\mathrm{d}E_1'} \overset{!}{=} 0 \\ &= \underbrace{\left(\frac{\partial S_1}{\partial E_1'} - \frac{\partial S_2}{\partial E_2'}\right)}_{=0}\mathrm{d}E_1' = 0 \end{aligned}$$

We define a new state variable T (*'temperature'*) by:

$$\boxed{\left.\frac{\partial S(E,V,N)}{\partial E}\right|_{V,N} := \frac{1}{T(E,V,N)}}$$

$$\Rightarrow\ \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\ \mathrm{d}E_1' = 0$$

$$\Rightarrow\ \boxed{T_1 = T_2}$$

The two systems exchange energy until their temperatures are the same.

Usually the number of states Ω and therefore entropy $S = k_B \ln \Omega$ increases with energy E and therefore $1/T$ and with this also T will be positive (e.g. $S \sim \ln E^{3/2}$ for the ideal gas, see above). There are, however, completely reasonable models in statistical physics in which Ω can go down with E, namely if the number of states has an upper limit, e.g. in a finite-sized spin system or the two-state system from below. Then we formally get a negative temperature. Although this does not agree with our everyday life intuition about temperature, there is nothing wrong with such a result.

$S(E)$ usually flattens with increasing energy (e.g. $S \sim \ln E^{3/2}$ for the ideal gas, see above). This implies that high energy corresponds to high temperature, in agreement with our everyday life intuition about temperature.

In general, temperature T describes the coupling between energy and entropy. Inverse temperature is the cost in entropy when buying a unit of energy from the environment. Due to the equipartition theorem (compare the chapter on the canonical ensemble), temperature is often identified with kinetic energy; however, the temperature definition of statistical physics from above is much more general.

Before equilibrium is reached, entropy grows:

$$\mathrm{d}S = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\ \mathrm{d}E_1 > 0$$

$$T_1 > T_2\ \Rightarrow \mathrm{d}E_1 < 0$$

Hence we see that energy flows to the cooler system. Temperature defined this way agrees with our intuitive understanding of temperature.

If the two systems only exchange energy, then:

$$\begin{aligned}
&\mathrm{d}E_i = T_i\,\mathrm{d}S_i \\
\Rightarrow\ &\mathrm{d}S_2 = \frac{\mathrm{d}E_2}{T_2} = -\frac{\mathrm{d}E_1}{T_2} = -\frac{T_1}{T_2}\,\mathrm{d}S_1 \\
\Rightarrow\ &\mathrm{d}S = \mathrm{d}S_1 + \mathrm{d}S_2 \\
&\quad = \mathrm{d}S_1\left(1 - \frac{T_1}{T_2}\right) > 0 \quad \text{before equilibrium is reached}
\end{aligned}$$

$$T_1 > T_2 \ \Rightarrow\ \mathrm{d}S_1 < 0,\ \mathrm{d}S_2 > 0,\ |\mathrm{d}S_2| > |\mathrm{d}S_1|$$

The warmer system loses entropy, the cooler system gains entropy. Overall more entropy is generated.

$$\Rightarrow \quad \text{entropy is not conserved (unlike energy } E)$$

Contact with volume exchange

We now assume that the wall is also mobile, thus volume can be exchanged:

$$\mathrm{d}S = \frac{\partial S_1}{\partial E_1'}\,\mathrm{d}E_1' + \frac{\partial S_2}{\partial E_2'}\,\underbrace{\mathrm{d}E_2'}_{=-\mathrm{d}E_1'} + \frac{\partial S_1}{\partial V_1}\,\mathrm{d}V_1 + \frac{\partial S_2}{\partial V_2}\,\underbrace{\mathrm{d}V_2}_{=-\mathrm{d}V_1}$$

We define another new state variable (*'pressure'*) by:

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

$$\Rightarrow\ \mathrm{d}S = \underbrace{\left(\frac{1}{T_1} - \frac{1}{T_2}\right)}_{=0}\mathrm{d}E_1 + \underbrace{\left(\frac{p_1}{T_1} - \frac{p_2}{T_2}\right)}_{=0}\,\mathrm{d}V_1 \overset{!}{=} 0$$

$$T_1 = T_2,\ p_1 = p_2$$

Volume is exchanged until the pressures are the same.

If temperatures are equal:

$$\mathrm{d}S = \frac{p_1 - p_2}{T}\mathrm{d}V_1 > 0$$

The system with larger pressure increases its volume.
This definition of pressure might seem a bit odd to you because it relates to entropy and not to energy, as you might have expected. Therefore we aim to rewrite it in terms

of energy. We start with a mathematical identity for any smooth function $f(x, y)$ that is explained in the appendix:

$$\left.\frac{\partial f}{\partial x}\right|_y \left.\frac{\partial x}{\partial y}\right|_f \left.\frac{\partial y}{\partial f}\right|_x = -1$$

Note that this result might look odd if you like to think about differentials as real numbers that you can cancel like in fractions. This result teaches you not to do this, and reflects the fact that for two positive changes you need one negative change to close a loop in (f, x, y)-space. We now apply this formula to $S(E, V, N)$:

$$\left.\frac{\partial S}{\partial V}\right|_{E,N} \left.\frac{\partial V}{\partial E}\right|_{S,N} \left.\frac{\partial E}{\partial S}\right|_{V,N} = -1$$

$$\Rightarrow \; \frac{p}{T} \left(\left.\frac{\partial E}{\partial V}\right|_{S,N} \right)^{-1} \left(\left.\frac{\partial S}{\partial E}\right|_{V,N} \right)^{-1} = -1$$

Noting that the last term simply gives T, we finally have

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

Therefore pressure p can also be interpreted as the increase in energy when reducing volume. This is closer to our intuition on pressure, but note that this should be done at constant entropy, which basically means without heat flux.

Contact with exchange of particle number

Finally let's assume a permeable membrane and define a new state variable μ (*'chemical potential'*) by:

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

The equilibrium condition becomes:

$$\mu_1(E_1, V_1, N_1) = \mu_2(E_2, V_2, N_2)$$

Assume $T_1 = T_2$, but $\mu_2 > \mu_1$:

$$\Rightarrow \; \mathrm{d}S = (-\mu_1 + \mu_2)\, \frac{\mathrm{d}N_1}{T} > 0$$

$$\mathrm{d}N_1 > 0 \;\Rightarrow\; \mu_2 > \mu_1$$

$\Rightarrow$ particles flow from large to small chemical potential

Again it might be more intuitive to rewrite the definition of chemical potential in terms of energy rather than entropy. We now apply our mathematical relation to $S(E,N,V)$:

$$\left.\frac{\partial S}{\partial N}\right|_{E,V} \left.\frac{\partial N}{\partial E}\right|_{S,V} \left.\frac{\partial E}{\partial S}\right|_{N,V} = -1$$

$$\Rightarrow \; -\frac{\mu}{T}\left(\left.\frac{\partial E}{\partial N}\right|_{S,V}\right)^{-1}\left(\left.\frac{\partial S}{\partial E}\right|_{V,N}\right)^{-1} = -1$$

The last term again cancels the T in the first term and we thus get the final result

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

This means that the chemical potential is the energy cost when increasing particle number (at constant S and V, that is without heat flux and with a constant volume).

Equations of state

We note that the three newly introduced variables:

$$T = T(E,V,N), \quad p = p(E,V,N), \quad \mu = \mu(E,V,N)$$

defined by

$$\mathrm{d}S = \frac{1}{T}\,\mathrm{d}E + \frac{p}{T}\,\mathrm{d}V - \frac{\mu}{T}\mathrm{d}N$$

are *intensive*, that is their values do not change if the system is doubled, because otherwise S could not be extensive.

Rearranging the equation above for dE gives:

$$\mathrm{d}E = T\,\mathrm{d}S - p\,\mathrm{d}V + \mu\,\mathrm{d}N$$

The pairs (T,S), $((-)\,p,V)$ and (μ,N) are *'conjugate'* variables in regard to energy. We identify the three types of energies as heat, mechanical energy and chemical energy.

$S = S(E,V,N)$ is the *'fundamental equation'*, containing the complete information on the system. The three equations for T, p and μ are *'equations of state'*. Each by itself contains only incomplete information on the system. Typically the equations of state are experimentally accessible and thus ground our theory in experiments. If only some of them are known, the others have to be guessed based on some additional information (e.g. a model). Moreover thermodynamic relations give strong additional constraints on possible equations of state (see other chapter).

2.6 Equations of state for ideal gas

$$S = k_B\, N \left[\ln \left(\frac{V}{N\lambda^3} \right) + \frac{5}{2} \right] \quad \text{fundamental equation}$$

$$\lambda \propto \left(\frac{N}{E} \right)^{\frac{1}{2}}$$

$$\Rightarrow \quad \frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{V,N} = k_B\, N \, \frac{\partial}{\partial E} \ln E^{\frac{3}{2}} = \frac{3}{2}\, k_B\, N\, \frac{1}{E}$$

$$\Rightarrow \quad \boxed{E = \frac{3}{2}\, N\, k_B\, T} \quad \text{caloric equation of state}$$

$$\Rightarrow \quad e = \frac{E}{N} = \frac{3}{2}\, k_B\, T$$

Hence each degree of freedom carries energy $\frac{1}{2}\, k_B\, T$. We also note for the specific heat:

$$c_v = \frac{dE}{dT} = \frac{3}{2} N k_B$$

thus the specific heat is independent of temperature. In fact this result is the classical high-temperature limit for the quantum mechanical fluids we are discussing later.

$$\frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E,N} = k_B\, N\, \frac{1}{V} \Rightarrow \boxed{p\, V = N\, k_B\, T} \quad \text{thermal equation of state}$$

The thermal equation of state is also known as the *ideal gas law* (combining the two laws of Gay-Lussac and Boyle-Mariotte).

$$-\frac{\mu}{T} = \left. \frac{\partial S}{\partial N} \right|_{E,V} = k_B \left[\ln \left(\frac{V}{N\lambda^3} \right) + \frac{5}{2} \right] + k_B\, N\, \frac{\partial}{\partial N}\ \ln N^{-\frac{5}{2}}$$

$$\Rightarrow \quad \boxed{\mu = k_B\, T\ \ln \left(\frac{N\lambda^3}{V} \right)} \quad \text{equation of state for chemical potential}$$

Defining density $\rho = \frac{N}{V}$, p becomes:

$$p = \rho\, k_B\, T$$

$$[p] = \frac{J}{m^3} = \frac{N}{m^2} = Pa$$

Pressure can be interpreted as force per area or as energy density. For the chemical potential we can write

$$\mu = k_B\, T\ \ln \rho\lambda^3 = k_B\, T\ \ln \frac{p}{p_0} \quad \text{with } p_0 = \frac{k_B\, T}{\lambda^3}$$

Thus for an ideal (=dilute) gas chemical potential grows logarithmically with pressure. The following numbers are quite instructive:

At room temperature: $T \approx 300K$.

$$\Rightarrow \quad k_B\, T = 4.1 \cdot 10^{-21}\, J = 4.1\, pN \cdot nm = \frac{1}{40}\, eV = 25\, meV \quad \text{thermal energy}$$

$$p_0 = \frac{4.1 \cdot 10^{-21}\, J}{(10^{-10}\, m)^3} = 4.1\, GPa$$

Because ambient pressure $p = 1atm = 10^5\, Pa$, the chemical potential μ is usually negative. We calculate the volume of one mole ideal gas at ambient pressure:

$$V = \frac{N\, k_B\, T}{p} = \frac{6.022 \cdot 10^{23} \cdot 4.1 \cdot 10^{-21}\, J}{10^5\, Pa} \approx 24 \cdot 10^{-3}\, m^3 = 24\, l$$

2.7 Two-state system

We now consider other important examples for the microcanonical ensemble. We first consider an 'atom' which can either be in its ground (zero energy) or excited state (energy ϵ).

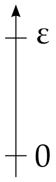

Figure 2.7: Two state system with corresponding energies $E_{ground} = 0$ and $E_{excited} = \epsilon$.

As the total energy E and the number of particles N are fixed, we know exactly how many atoms are in the excited state:

$$N_\epsilon = \frac{E}{\epsilon} \quad \text{number of atoms in excited state}$$

$$N_0 = N - N_\epsilon \quad \text{atoms in ground state}$$

The number of microstates is the number of ways one can choose N_ϵ out of N:

$$\Omega = \frac{N!}{N_\epsilon!\ (N - N_\epsilon)!}$$

The entropy then becomes, using Stirling's formula:

$$\begin{aligned} S &= k_B \ \ln \Omega \\ &\approx k_B \left[(N \ \ln N - N) - (N_\epsilon \ \ln N_\epsilon - N_\epsilon) - ((N - N_\epsilon) \ \ln (N - N_\epsilon) - (N - N_\epsilon)) \right] \\ &= k_B \ \left[-N_\epsilon \ \ln \frac{N_\epsilon}{N} - (N - N_\epsilon) \ \ln \left(1 - \frac{N_\epsilon}{N} \right) \right] \end{aligned}$$

$$\Rightarrow \quad \boxed{s = -k_B \left[\rho_\epsilon \ \ln \rho_\epsilon + (1 - \rho_\epsilon) \ \ln (1 - \rho_\epsilon) \right]}$$

where $\rho_\epsilon = \frac{N_\epsilon}{N} = \frac{E}{\epsilon \, N}$, $0 \leq \rho_\epsilon \leq 1$ is the density of the excited state and $1 - \rho_\epsilon = \frac{N - N_\epsilon}{N}$ is the density of the ground state. Note that this result makes a lot of sense. It basically is Shannon's formula, $S = -\sum_i p_i \ln p_i$ with i running over the two possible states, ground and excited. We note that the fundamental equation is independent of volume and depends only on E and N.

Caloric equation of state:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_N = -k_B \, N \ \left[\frac{1}{\epsilon \, N} \ \ln \frac{E}{\epsilon \, N} + \frac{1}{\epsilon \, N} - \frac{1}{\epsilon \, N} \ \ln \left(1 - \frac{E}{\epsilon \, N} \right) - \frac{1}{\epsilon \, N} \right] = \frac{k_B}{\epsilon} \ \ln \left(\frac{N \, \epsilon}{E} - 1 \right)$$

$$\Rightarrow \quad \boxed{E = \frac{N \, \epsilon}{1 + e^{\epsilon / (k_B \ T)}} = \frac{N \, \epsilon}{1 + e^{\epsilon \beta}}}$$

where we have introduced the inverse temperature $\beta = 1/(k_B \ T)$.

If we do not use E, but stick to the density $\rho_\epsilon = \frac{E}{\epsilon \, N}$, the thermal equation of state lead to

$$\frac{\rho_\epsilon}{1 - \rho_\epsilon} = e^{-\beta \epsilon} \Rightarrow \rho_\epsilon = \frac{e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} = \frac{1}{1 + e^{\beta \epsilon}}$$

which we will see in the next chapter to be the Boltzmann factor for a two-state system. The result itself is of course the same as for the energy E.

We consider the two limits:

① High T, $\beta \epsilon \ll 1$:

$$E = \frac{N \, \epsilon}{2}, \ \rho_\epsilon = \frac{E}{\epsilon \, N} = \frac{1}{2}, \ 1 - \rho_\epsilon = \frac{1}{2}$$

$\Rightarrow$ both states have equal density.

② Low T, $\beta \epsilon \gg 1$:

$$E = 0, \ \rho_\epsilon = 0, \ 1 - \rho_\epsilon = 1$$

$\Rightarrow$ only the ground state is populated.

Here we see a very general principle of statistical physics: at low T, energy wins (every particle is in the ground state), and at high T, entropy wins (all states are equally populated). As we can see when plotting the formula for density, as T increases from 0 to ∞, ρ_ϵ increase from 0 to $1/2$ in a sigmoidal fashion.

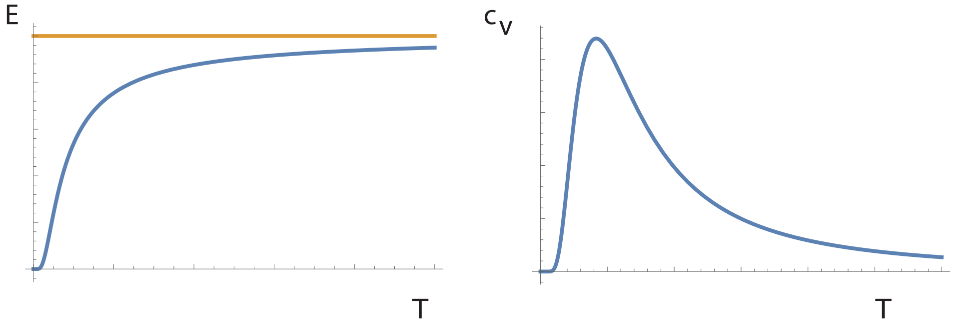

Figure 2.8: Energy and specific heat of the two-state model as a function of temperature. The energy approaches $\frac{N\,\epsilon}{2}$ asymptotically. The capacity peaks at the temperature $T_{max} = \frac{0.42\,\epsilon}{k_B}$.

Interestingly at high temperature we only reach half of the maximally possible energy. If we push more energy into the system, then formally temperature jumps from ∞ to $-\infty$ (negative temperature, see above) and then approaches 0 from below.
We finally calculate the specific heat:

$$\begin{aligned} c_v = \frac{dE}{dT} &= N\,\epsilon\,\frac{1}{(1+e^{\beta})^2}\,e^{\beta}\,\frac{\epsilon}{k_B}\,\frac{1}{T^2} \\ &= N k_B \beta^2\,\frac{1}{(e^{-\beta/2}+e^{+\beta/2})^2} \end{aligned}$$

We see that the heat capacity vanishes both for low and high T. It peaks at a finite temperature $T = \frac{0.42\,\epsilon}{k_B}$. Such a *'Schottky hump'* reveals the existence of two low lying energy states in a spectrum.

Alternative derivation

We consider an alternative way to calculate $\Omega(E)$. For each atom we introduce an occupation (or occupancy) number $n_j \in \{0,1\}$ for the external state:

$$\begin{aligned}
\Omega(E) &= \underbrace{\sum_{n_1=0}^{1} ... \sum_{n_N=0}^{1}}_{\text{sum over } 2^N \text{ states}} \delta\left(E - \epsilon \sum_{i=1}^{N} n_i\right) \\
&= \sum_{n_1=0}^{1} ... \sum_{n_N=0}^{1} \int \frac{\mathrm{d}k}{2\pi}\, e^{ik(E-\epsilon\sum_{i=1}^{N} n_i)} \\
&= \int \frac{\mathrm{d}k}{2\pi}\, e^{ikE} \left(\underbrace{\sum_{n=0}^{1} e^{-ik\epsilon n}}_{=1+e^{-ik\epsilon}} \right)^{N} \\
&= \int \frac{\mathrm{d}k}{2\pi}\, \exp\left\{ N \left[\underbrace{ik\frac{E}{N} + \ln\left(1 + e^{-i\epsilon k}\right)}_{:=f(k)} \right] \right\}
\end{aligned}$$

In the limit $N \to \infty$, we solve this integral with the *'saddle point approximation'*:

$$\Omega(E) = \frac{1}{2\pi} \int \mathrm{d}k\, e^{N\,f(k)} = \frac{1}{2\pi}\, e^{N\,f(k_0)} \int \mathrm{d}k\, e^{N\frac{1}{2}f''(k_o)(k-k_0)^2} = \frac{1}{2\pi}\, e^{N\,f(k_0)} \left(\frac{2\pi}{N\,|f''(k_0)|} \right)^{1/2}$$

where $f(k)$ has a maximum at k_0.

We define $\gamma = \frac{E}{N}$:

$$\begin{aligned}
f'(k) &= i\gamma + \frac{e^{-ik\epsilon}(-i\epsilon)}{1+e^{-ik\epsilon}} = i\gamma - \frac{i\epsilon}{e^{ik\epsilon}+1} \overset{!}{=} 0 \\
\Rightarrow \quad k_0 &= \frac{1}{i\epsilon}\ \ln\frac{\epsilon-\gamma}{\gamma} \\
\Rightarrow \quad f(k_0) &= \frac{\gamma}{\epsilon}\ \ln\frac{\epsilon-\gamma}{\gamma} - \ln\frac{\epsilon-\gamma}{\epsilon} \\
f''(k_0) &= \frac{(i\epsilon)^2\ e^{ik_0\epsilon}}{\left(e^{ik_0\epsilon}+1\right)^2} = (\gamma-\epsilon)\ \gamma \\
\Rightarrow\ \Omega(E) &= e^{Nf(k_0)}\ \left(\frac{1}{2\pi\ N\ f''(k_0)}\right)^{\frac{1}{2}} \\
&= \exp\left\{\underbrace{Nf(k_0)}_{\propto N} - \underbrace{\frac{1}{2}\ \ln\left(2\pi N f''(k_0)\right)}_{\propto \ln N\ \to\ \text{neglect in limit } N\to\infty}\right\} \\
\Rightarrow\ \Omega(E) &= e^{Nf(k_0)}
\end{aligned}$$

$$\Rightarrow\ S = k_B\ N\ f(k_0) = k_B\ N\ \left\{\rho_\epsilon\ \underbrace{\ln\left(\frac{1}{\rho_\epsilon}-1\right)}_{=-\ln\rho_\epsilon+\ln(1-\rho_\epsilon)}\ - \ln(1-\rho_\epsilon)\right\}$$

because $\rho_\epsilon = \gamma/\epsilon$.

$$\Rightarrow \quad \boxed{s = -k_B\left[\rho_\epsilon\ \ln\rho_\epsilon + (1-\rho_\epsilon)\ \ln(1-\rho_\epsilon)\right]}$$

Thus we recovered the same result as above with the combinatorial approach.

2.8 Einstein model for specific heat of a solid

In 1907, Einstein introduced a simple model for the vibrational modes in a crystal. From this, we can calculate the specific heat of a solid, which is one of the central questions in solid state physics. Later this model was improved by Debye as we will see later. The Einstein model is also of fundamental importance because it is the microcanoncial treatment of a collection of harmonic oscillators. It assumes that each atom in the crystal vibrates with a natural frequency ω around its equilibrium position. Thus for N_A atoms in three dimensions, we have a collection of $N = 3N_A$ harmonic oscillators, each with two degrees of freedom. In quantum mechanics, each harmonic oscillator is quantized and can carry an integral number of quanta $\hbar\omega$. Together with the zero energy we have

$$E = \hbar\omega(\frac{N}{2} + Q)$$

where Q is the overall number of quanta. Therefore

$$Q = (\frac{E}{\hbar\omega} - \frac{N}{2})$$

These Q quanta are distributed over N states. That's the same as distributing Q marbles over N boxes or as placing Q marbles and $N-1 \approx N$ match sticks in a row:

$$\Omega(E,N) = \frac{(Q+N)!}{Q!N!}$$

There is no role for V in this model.
We calculate the entropy using Stirling's formula:

$$\begin{aligned} S = k_B\Omega &= k_B\left[(Q+N)(\ln(Q+N)-1) - Q(\ln Q - 1) - N(\ln N - 1)\right] \\ &= k_B\left[Q\ln\frac{Q+N}{Q} + N\ln\frac{Q+N}{N}\right] \\ &= k_B N\left[(e+\frac{1}{2})\ln(e+\frac{1}{2}) - (e-\frac{1}{2})\ln(e-\frac{1}{2})\right] \end{aligned}$$

with $e = E/E_0$ and $E_0 = N\hbar\omega$. We next calculate temperature:

$$\frac{1}{T} = \frac{\partial S}{\partial E} = \frac{1}{E_0}\frac{\partial S}{\partial e} = \frac{k_B N}{E_0}\ln\left(\frac{e+\frac{1}{2}}{e-\frac{1}{2}}\right)$$

We define $\beta = \hbar\omega/k_BT$ and invert this equation:

$$2e = \frac{e^\beta+1}{e^\beta-1} \Rightarrow E = \frac{N\hbar\omega}{2}\frac{e^\beta+1}{e^\beta-1} = N\hbar\omega(\frac{1}{2}+\frac{1}{e^\beta-1})$$

For $T\to 0$, $E\to N\hbar\omega/2$, the zero energy. For $T\to\infty$, $E\to Nk_BT = 6N_A(k_BT/2)$, the classical limit for $6N_A$ degrees of freedom. We also calculate the specific heat:

$$c_v = \frac{dE}{dT} = Nk_B\beta^2\frac{1}{(e^{\beta/2}-e^{-\beta/2})^2}$$

For $T\to 0$, c_v vanishes. It then rises exponentially and plateaus at $3N_Ak_B$ for $T\to\infty$. This classical limit is also known as the law of Dulong-Petit. The crossover takes place at $T = \hbar\omega/k_B$.
It is very instructive to compare this results with the two-state system, for which we found

$$c_v = \frac{dE}{dT} = Nk_B\beta^2\frac{1}{(e^{\beta/2}+e^{-\beta/2})^2}$$

where we now define dimensionless inverse temperature as $\beta = \epsilon/k_BT$. Thus the result is essentially the same, except for the minus sign, which however makes a huge difference. Later we will see that this is typical for the difference between fermionic

systems (each state can be occupied either once or not at all) and bosonic systems (each state can be filled with arbitrarily many quanta, like the harmonic oscillator).
The theoretical curve fits the experimental data qualitatively, but if one zooms in around $T = 0$, one finds that the experimental curve rises as T^3 rather than exponentially. The reason for this discrepancy is that atoms do not vibrate by themselves at a fixed frequency, but in groups of variable frequencies. We will consider this important fact later in the Debye model.

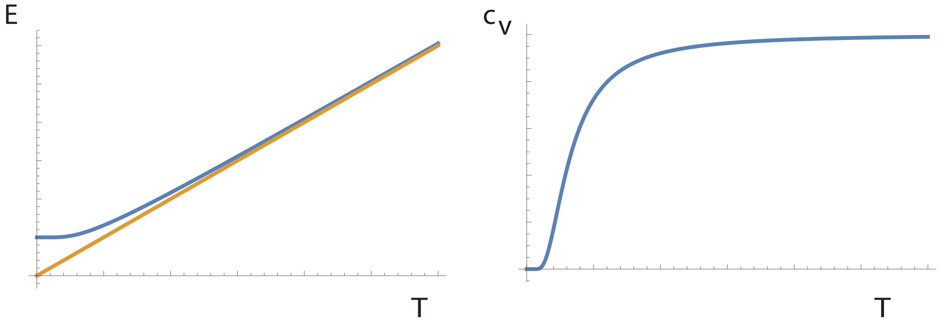

Figure 2.9: Energy and specific heat of the Einstein model as a function of temperature.

Alternative derivation

Again we consider an alternative way to calculate $\Omega(E)$ by using occupation numbers:

$$\begin{aligned}
\Omega(E) &= \sum_{n_1=0}^{\infty} \dots \sum_{n_N=0}^{\infty} \delta\left(E - \hbar\omega \sum_{i=1}^{N} (n_i + \frac{1}{2})\right) \\
&= \sum_{n_1=0}^{\infty} \dots \sum_{n_N=0}^{\infty} \int \frac{\mathrm{d}k}{2\pi} e^{ik(E - \hbar\omega \sum_{i=1}^{N} (n_i + \frac{1}{2}))} \\
&= \int \frac{\mathrm{d}k}{2\pi} e^{ikE} \sum_{n_1=0}^{\infty} \dots \sum_{n_N=0}^{\infty} e^{-ik\hbar\omega \sum_{i=1}^{N} (n_i + \frac{1}{2})} \\
&= \int \frac{\mathrm{d}k}{2\pi} e^{ikE} \left(e^{-ik\hbar\omega/2} \sum_{n=0}^{\infty} e^{-ik\hbar\omega n} \right)^N \\
&= \int \frac{\mathrm{d}k}{2\pi}\, e^{ikE} \left(\frac{e^{-ik\hbar\omega/2}}{1 - e^{-ik\hbar\omega}} \right)^N \\
&= \int \frac{\mathrm{d}k}{2\pi}\, e^{N[ik\frac{E}{N} - \ln(2i\sin(k\hbar\omega/2))]}
\end{aligned}$$

Note that this time the occupation number is not restricted, but runs over all possible numbers. Therefore in the last step we have used the geometrical sum to achieve a closed formula. It is a non-trivial question why this series converges in our case. Because the Delta-function is a distribution, in principle one has to apply the formula

above to a test function. If this test function is sufficiently well-behaved, then one can regularize the exponent by adding a small and real ϵ to the exponent. Then the series converges and one can send ϵ to zero in the end.
The remaining integral can again be solved using the method of steepest decent with

$$f(k) = ikE/N - \ln(2i\sin(k\hbar\omega/2))$$

Again the Gaussian integral can be neglected and only the first term in the Taylor expansion matters, with

$$k_0 = \frac{1}{2i\hbar\omega/2}\ln\frac{E/N+\hbar\omega/2}{E/N-\hbar\omega/2}$$

After a lengthy calculation, one then gets the same result as above.

2.9 Entropic elasticity of polymers

We consider a chain-like polymer molecule consisting of N monomers of size a in one dimension.

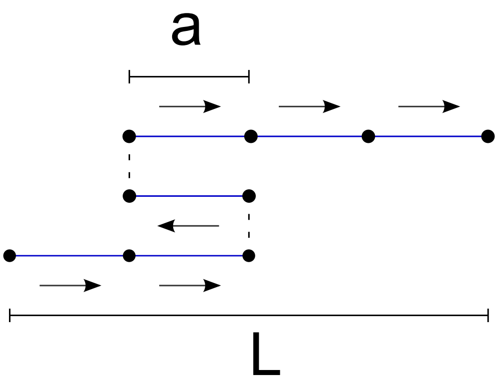

Figure 2.10: A one-dimensional polymer consisting out of monomers of size a. The sequences to the right and left are indicated by arrows. L is the effective or projected length and $L_0 = Na$ the contour length.

$$\begin{aligned} N_+ &: \quad \text{\#segments to right} \\ N_- &: \quad \text{\#segments to left} \end{aligned}$$

$$\begin{aligned} \Rightarrow\ N_+ - N_- &= \frac{L}{a} \\ N_+ + N_- &= N \\ \Rightarrow\ N_+ &= \frac{1}{2}\left(N + \frac{L}{a}\right) \\ N_- &= \frac{1}{2}\left(N - \frac{L}{a}\right) \\ \Rightarrow\ \Omega &= \frac{N!}{N_+!\, N_-!} \\ \Rightarrow\ S &= -k_B\, \left(N_+\ \ln\frac{N_+}{N} + N_-\ \ln\frac{N_-}{N}\right) \\ \Rightarrow\ S &= -k_B\, N\ (\rho_+\ \ln\rho_+ + \rho_-\ \ln\rho_-) \end{aligned}$$

Here we defined $\rho_\pm = \frac{N_\pm}{N} = \frac{1}{2}\left(1 \pm \frac{L}{L_0}\right)$ with the *'contour length'* $L_0 = Na$ in the last step.

We have a look at two limits:

① $\rho_+ \to 1,\ \rho_- \to 0\ \Rightarrow\ S \to 0$
There is only one possible configuration for the polymer.

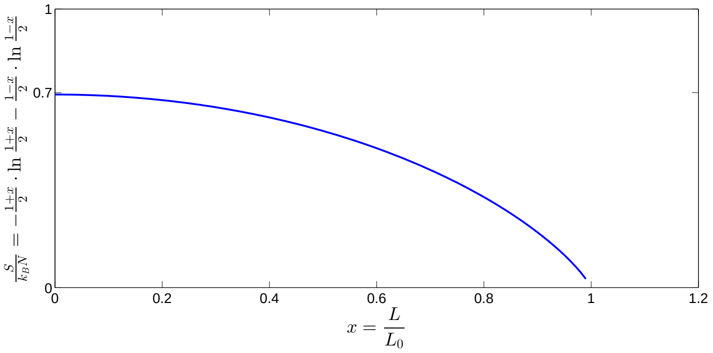

Figure 2.11: $S/(k_B N)$ as a function of x. For $x_0 = 0$ the function is maximal with $f(x_0) = \ln 2 \approx 0.7$.

② $L \to 0,\ \rho_+ \to \frac{1}{2},\ \rho_- \to \frac{1}{2}\ \Rightarrow\ S \to N\,k_B\ \ln 2$
We see that the entropy is maximal for $L = 0$, hence the polymer coils up (2.3). Stretching the polymer decreases its entropy. Therefore an entropic restoring force exists which pulls against an external stretching force.

Note that we cannot define temperature in the usual way, $1/T = \partial S/\partial E$, because $S(N, L)$ does not depend on E. However, there should be a conjugate quantity to length

L, and this should be a force F and at the same time defines temperature T:

$$
\begin{aligned}
& -\frac{F}{T} = \frac{\partial S}{\partial L} \\
\overset{x=L/L_0}{\Rightarrow} \quad & \frac{1}{L_0}\,\frac{\partial S}{\partial x} = -\frac{N\,k_B}{L_0}\left\{\frac{1}{2}\ln\left(\frac{1+x}{2}\right) + \frac{1}{2} - \frac{1}{2}\,\ln\left(\frac{1-x}{2}\right) - \frac{1}{2}\right\} \\
& = -\frac{N\,k_B}{2L_0}\,\underbrace{\ln\left(\frac{1+x}{1-x}\right)}_{\approx 2x} \approx -\frac{N\,k_B\,L}{L_0^2} \\
\overset{L_0=N\,a}{\Rightarrow} \quad & \boxed{F = \frac{k_B\,T\,L}{N\,a^2}}
\end{aligned}
$$

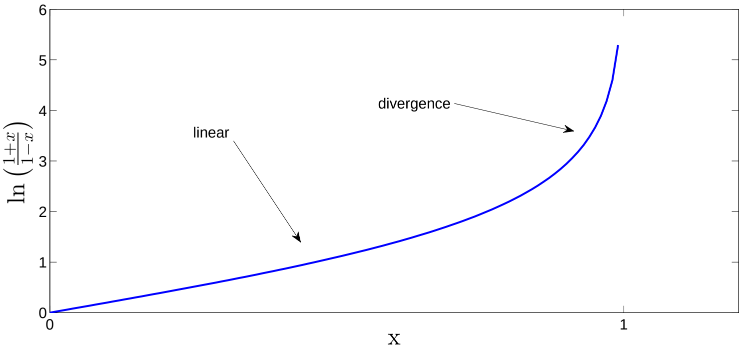

Figure 2.12: $f(x) = \ln\left(\frac{1+x}{1-x}\right)$ as a function of x.

The entropic restoring force corresponds at small extension to a harmonic potential:

$$U = \frac{k_B\,T\,L^2}{2\,N\,a^2}$$

Most materials expand with temperature (eg ideal gas: $pV = Nk_BT$). For the polymer, in contrast, increasing temperature causes a raise in restoring force and thus the system contracts. Note that this is a purely entropic effect (our polymer has no energy). In analogy with mechanics, $k = \frac{k_B\,T}{N\,a^2}$ is called the *'entropic spring constant'* and it increases with T.

2.10 Statistical deviation from average

For two systems in thermal contact, our equilibrium condition $T_1 = T_2$ followed from the maximal entropy principle. We now investigate how sharp this maximum is. As an instructive example, we consider two ideal gases in thermal contact:

$$S_i = \frac{3}{2}\,k_B\,N_i\,\ln E_i + terms\ independent\ of\ E_i\ (i \in \{1,2\})$$

$$S = S_1 + S_2$$

Maximal entropy:

$$\mathrm{d}S = 0 \quad \Rightarrow \quad \frac{\partial S_1}{\partial E_1} = \frac{\partial S_2}{\partial E_2}$$

$$\Rightarrow \frac{3}{2}\, k_B\, \frac{N_1}{\bar{E_1}} = \frac{3}{2}\, k_B\, \frac{N_2}{\bar{E_2}} \quad \text{corresponding to } T_1 = T_2$$

$$\Rightarrow \bar{E_1} = \frac{N_1}{N_2}\, \bar{E_2} = \frac{N_1}{N - N_1}\, (E - \bar{E_1}) \quad \Rightarrow \quad \bar{E_1} = \frac{N_1}{N}\, E$$

We check for maximum:

$$\frac{\partial^2 S}{\partial E_1^2} = -\frac{3}{2}\, k_B \left(\frac{N_1}{\bar{E_1}^2} + \frac{N_2}{\bar{E_2}^2} \right) < 0$$

We now consider a small deviation:

$$E_1 = \bar{E_1} + \Delta E, \quad E_2 = \bar{E_2} - \Delta E$$

$$\Rightarrow \; S(\bar{E_1} + \Delta E) = \frac{3}{2}\, k_B \left[N_1 \, \ln\left(\bar{E_1} \left(1 + \frac{\Delta E}{\bar{E_1}} \right) \right) + N_2 \, \ln\left(\bar{E_2} \left(1 - \frac{\Delta E}{\bar{E_2}} \right) \right) \right]$$

$$\underbrace{\approx}_{\frac{\Delta E}{\bar{E}_i} \ll 1} \frac{3}{2}\, k_B \left[N_1 \ln \bar{E_1} + N_2 \, \ln \bar{E_2} + \left(\underbrace{\frac{N_1}{E_1} - \frac{N_2}{E_2}}_{=0 \text{ at equil.}} \right) \Delta E - \frac{N_1}{2} \left(\frac{\Delta E}{\bar{E_1}} \right)^2 - \frac{N_2}{2} \left(\frac{\Delta E}{\bar{E_2}} \right)^2 \right]$$

$$E_i = \frac{N_i}{N}\, E$$

$$\Rightarrow \quad \boxed{\Omega = \bar{\Omega}\, \exp\left[-\frac{3}{4} \left(\frac{\Delta E}{E} \right)^2 N^2 \left(\frac{1}{N_1} + \frac{1}{N_2} \right) \right]}$$

Plugging in a typical number $N_i = 10^{27}$:

$$\Rightarrow \quad \Omega \text{ drops dramatically away from the maximum } \bar{E_1}$$

$\frac{\Delta E}{E}$	$e^{-\frac{3}{4}\left(\frac{\Delta E}{E}\right)^2 N^2 \left(\frac{1}{N_1}+\frac{1}{N_2}\right)}$
10^{-10}	$e^{-600} \approx 10^{-260}$
10^{-11}	e^{-6}
10^{-12}	$e^{-0.06} \approx 1$

We see that macrostates with $\bar{E_1}(1 \pm 10^{-12})$ and $\bar{E_2}(1 \mp 10^{-12})$ are equally likely. Fluctuations in energy are of the order 10^{-12}, hence tiny.

2.11 Foundation of the fundamental postulate

There is a long and twisted history of justifying the fundamental postulate. The main problem here is that on a microscopic basis, dynamics is governed by time-reversible equations (Newton's equation or Schrödinger equation), while macroscopic systems decay irreversibly into a unique equilibrium state. A rigorous derivation of the drive towards equilibrium from microscopic principles does still not exist (Boltzmann suggested the famous H-theorem for this, but it turns out that his assumptions includes something that has an irreversible nature) and several mechanisms are discussed that somehow smooth out phase space density (e.g. deterministic chaos or quantum mechanical mixing with the states of the environment), such that a homogeneous distribution over the energy shell is achieved, but there is no general consensus what a rigorous proof would be. Thus one has to state clearly that statistical mechanics works very well for most physical systems of interest, but it is notoriously difficult to establish a microscopic derivation. This is why we still have to speak of a fundamental *postulate*.

A very instructive approach to this question is to consider Hamiltonian dynamics. In classical systems with N particles, each microstate is a point in 6N-dimensional *'phase space'* $(\underbrace{\vec{r}_1, ..., \vec{r}_n}_{\text{positions}}, \underbrace{\vec{p}_1, ..., \vec{p}_n}_{\text{momenta}}) = (\vec{q}, \vec{p})$. In statistical mechanics, we consider many particle systems and therefore cannot say in which state the system is exactly located, but rather use a statistical ensemble of states. The probability distribution is continuous and the probability that the system is in state $(\vec{q}, \vec{p})$ is

$$\rho(\vec{q}(t), \vec{p}(t), t)\, \mathrm{d}\vec{q}\, \mathrm{d}\vec{p}$$

where ρ is the phase space probability density. For initial conditions $(\vec{q}(0), \vec{p}(0))$ the system evolves according to Hamilton's equations:

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

For an isolated system at rest, energy is conserved due to the time invariance (Noether theorem):

$$\mathcal{H} = \text{const} = E$$

The solutions to the system of ordinary differential equations are unique and do not intersect. Energy conservation reduces phase space to a $(6N-1)$-dimensional hypersurface, the energy surface or energy shell.

We now define a phase space velocity

$$\vec{v} := (\dot{\vec{q}}, \dot{\vec{p}})$$

and the corresponding current

$$\vec{j} = \rho\, \vec{v}$$

For an arbitrary region of phase space, we have a balance equation:

$$\int_{\partial V} \mathrm{d}\vec{A} \cdot \vec{j} = -\frac{\partial}{\partial t} \int_V \mathrm{d}\vec{q}\, \mathrm{d}\vec{p}\, \rho(\vec{q}, \vec{p}, t)$$

$$\underbrace{\Rightarrow}_{\text{Gauss theorem}} \quad \int_V \mathrm{d}\vec{q}\, \mathrm{d}\vec{p} \left[\frac{\partial \rho}{\partial t} + \nabla \cdot \vec{j}\right] = 0$$

$$\underbrace{\Rightarrow}_{\text{V arbitrary}} \quad \boxed{\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \vec{v}) = 0} \quad \text{continuity equation}$$

Thus the system evolves like a hydrodynamic system, with a probability fluid streaming through state space. We now use Hamilton's equation to show that

$$\begin{aligned} 0 &= \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial}{\partial q_i}(\rho \dot{q}_i) + \frac{\partial}{\partial p_i}(\rho \dot{p}_i)\right) \\ &= \frac{\partial \rho}{\partial t} + \sum_{i=1}^{3N} \left(\frac{\partial \rho}{\partial q_i}\, \dot{q}_i + \frac{\partial \rho}{\partial p_i}\, \dot{p}_i\right) + \rho \sum_{i=1}^{3N} \underbrace{\left(\frac{\partial \dot{q}_i}{\partial q_i} + \frac{\partial \dot{p}_i}{\partial p_i}\right)}_{\frac{\partial^2 \mathcal{H}}{\partial q_i\, \partial p_i} - \frac{\partial^2 \mathcal{H}}{\partial p_i\, \partial q_i} = 0} \\ &= \boxed{\frac{d\rho}{dt} = 0} \quad \text{Liouville's theorem} \end{aligned}$$

The total derivative of the probability density vanishes. The probability fluid is 'incompressible' and $\rho(\vec{q}(t), \vec{p}(t), t) = \rho(\vec{q}(0), \vec{p}(0), 0)$.
Let V_0 be the volume of some initial region R_0 of phase space. At some time t afterwards, this region can have evolved to some region R_t with complicated shape, but its volume is unchanged: $V_t = V_0$ (Figure 2.13).

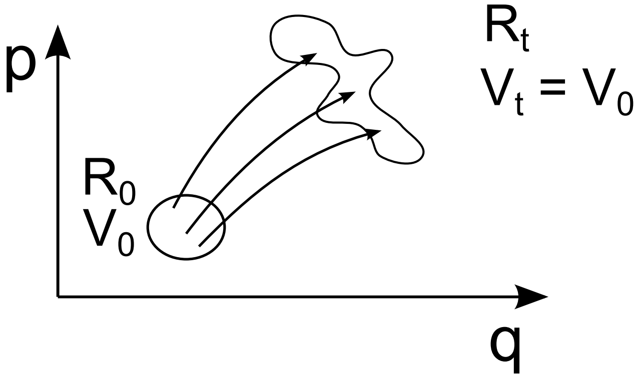

Figure 2.13: The phase space volume stays constant, although its shape may change.

At this point, we can draw two important conclusions. First the number of occupied microstates does not change, because phase space volume is conserved. More general, entropy does not change, because the system evolves deterministically and thus information content is not changed. This seems to speak against the fundamental postulate, which requires some kind of dispersion and increase in entropy. However, you can interpret this results also in another manner. Especially because the Hamiltonian system does not relax into some subset of phase space, it keeps continuing to explore all of phase space in a similar manner and this contributes to the fact that all states are equally likely to be visited.

Secondly, the fact that the phase space volume is conserved does not mean that its shape does not change. In fact for many systems of interest, one finds that a well-defined region of phase space quickly distorts into a very complex shape, especially for chaotic systems. When viewed from a coarse-grained perspective (like in a real world experiment with limited resolution), we will see a smooth distribution. This has been compared with the stirring of oil and water, which keep distinct domains on the microscale (oil and water do not mix), but appear to be uniform on a macroscopic scale. A more rigorous way to deal with the coarse-graining issue in a classical framework is the BBGKY-hierarchy (after Bogoliubov, Born, Green, Kirkwood, Yvon). Mixing in phase space is possible even in completely classical systems as proven by deterministic chaos. On the quantum level, one also could argue that completely isolated systems never exist and that coupling to the environment, even if very weak, will eventually lead to smoothing in state space. This aspect seems to suggest how the fundamental postulate might arise.

In order to learn more about the equilibrium state, we next rewrite Liouville's theorem:

$$\begin{aligned}\frac{\partial \rho}{\partial t} &= -\sum_{i=1}^{3N}\left[\dot{q}_i\,\frac{\partial \rho}{\partial q_i} + \dot{p}_i\,\frac{\partial \rho}{\partial p_i}\right]\\ &= -\sum_{i=1}^{3N}\left[\frac{\partial \mathcal{H}}{\partial p_i}\,\frac{\partial \rho}{\partial q_i} - \frac{\partial \mathcal{H}}{\partial q_i}\,\frac{\partial \rho}{\partial p_i}\right]\\ &= -\{\mathcal{H},\rho\} \quad \text{Liouville's equation}\end{aligned}$$

Here we used the notation of *'Poisson brackets'* in the last step. Liouville's equation is also known as the collisionless Boltzmann equation because it describes the streaming part for the probability fluid in phase space.

Let us now assume that $\rho(\vec{q},\vec{p},t)$ only depends on the conserved value of energy E.

$$\rho(\vec{q},\vec{p},t) = \Phi(E), \quad \frac{dE}{dt} = 0$$

We now get

$$\begin{aligned}
\frac{\partial\rho}{\partial t} &= -\{\mathcal{H},\rho\} \\
&= -\sum_{i=1}^{3N}\left[\frac{\partial\mathcal{H}}{\partial p_i}\frac{\partial\rho}{\partial q_i} - \frac{\partial\mathcal{H}}{\partial q_i}\frac{\partial\rho}{\partial p_i}\right] \\
&= \underbrace{-\sum_{i=1}^{3N}\left[\frac{\partial\mathcal{H}}{\partial p_i}\frac{\partial E}{\partial q_i} - \frac{\partial\mathcal{H}}{\partial q_i}\frac{\partial E}{\partial p_i}\right]}_{\{H,H\}=\frac{dE}{dt}=0}\frac{\partial\Phi}{\partial E} = 0
\end{aligned}$$

This result is also known as *Jean's theorem*. Thus in this case the state space probability density is constant and has the same value for a given energy. We conclude that once the microcanonical equilibrium state is reached, the system will stay there for ever.
We next consider an observable A which depends on time only through phase space:

$$A(t) = A(q(\vec{t}), p(\vec{t}))$$

$$\Rightarrow \quad \frac{dA}{dt} = \sum_{i=1}^{3N}\left[\frac{\partial A}{\partial q_i}\,\dot{q}_i + \frac{\partial A}{\partial p_i}\,\dot{p}_i\right] = \{\mathcal{H}, A\}$$

In equilibrium the ensemble average of an observable

$$\langle A\rangle = \int \mathrm{d}\vec{q}\,\mathrm{d}\vec{p}\,\rho(\vec{q},\vec{p})\,A(\vec{q},\vec{p})$$

will be time-independent. In particular we expect this to apply to all state variables.
Equivalent results can be derived for quantum mechanical systems. However, in this case we cannot use a scalar probability density, because phase space coordinates do not commute. Instead we need to introduce a density operator.
For a given state $|\Psi\rangle$, the observable has the average (projection):

$$\langle A\rangle = \langle\Psi|\,A\,|\Psi\rangle$$

We define the density operator or density matrix:

$$\rho = |\Psi\rangle\langle\Psi|$$

This then yields

$$\begin{aligned}
\langle A\rangle &= \langle\Psi|\,A\,|\Psi\rangle = \sum_n \langle\Psi|\,A\,|n\rangle\langle n|\Psi\rangle \\
&= \sum_n \langle n|\Psi\rangle\langle\Psi|\,A\,|n\rangle = \sum_n \langle n|\,\rho A\,|n\rangle \\
&= tr\{\rho A\} \quad \text{average over quantum mechanical distribution of states}
\end{aligned}$$

If we now turn to statistical mechanics, we superimpose a second layer of probability over the quantum mechanical probabilities. We call the states that follow the Schrödinger

equation *pure* states and then consider *mixed* states by adding up several of the pure states in an incoherent manner (no superposition, so probability is not tranfered from one qm state to the other and the weights stay constant).
Using an extended definition of the density matrix:

$$\rho = \sum_i p_i \; |\Psi_i\rangle \langle\Psi_i|$$

$$\begin{aligned}
\Rightarrow \; \langle A \rangle &= \sum_i p_i \; \langle\Psi_i| A |\Psi_i\rangle \\
&= \sum_n \sum_i p_i \; \langle\Psi_i| A |n\rangle \langle n|\Psi_i\rangle \\
&= \sum_n \sum_i p_i \; \langle n|\Psi_i\rangle \langle\Psi_i| A |n\rangle \\
&= \sum_n \langle n| \rho A |n\rangle = tr(\rho A)
\end{aligned}$$

Again stationary distributions result if ρ is a function of the stationary ensemble in energy representation:

$$\begin{gathered}
\mathcal{H} |n\rangle = E_n |n\rangle \\
0 = \partial_t \rho = [\mathcal{H}, \rho] \\
\Rightarrow \quad 0 = \langle m| \mathcal{H}\rho - \rho\mathcal{H} |n\rangle = (E_m - E_n)\, \rho_{mn} \\
\Rightarrow \quad \rho_{mn} = 0 \; for \; E_m \neq E_n \\
\text{without degeneracy:} \quad \rho = \sum_n \rho(E_n) \, |n\rangle \langle n|
\end{gathered}$$

We now derive the quantum mechanical analogue of Liouville's equation:

$$\begin{aligned}
\text{Schrödinger equation:} \quad & i\hbar \, \partial_t |\Psi\rangle = \mathcal{H} |\Psi\rangle \\
\text{adjoint Schrödinger equation:} \quad & -i\hbar \, \partial_t \langle\Psi| = \langle\Psi| \mathcal{H}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \quad i\hbar \, \partial_t \rho &= i\hbar \sum_i p_i \; \left(|\dot{\Psi}_i\rangle \langle\Psi_i| + |\Psi_i\rangle \langle\dot{\Psi}_i| \right) \\
&= \sum_i p_i \; (\mathcal{H} |\Psi_i\rangle \langle\Psi_i| - |\Psi_i\rangle \langle\Psi_i| \mathcal{H}) \\
&= [\mathcal{H}, \rho] \quad \text{commutator} \\
\partial_t \rho &= -\frac{i}{\hbar} \; [\mathcal{H}, \rho] \quad \text{von Neumann equation}
\end{aligned}$$

Like Liouville's equation, von Neumann's equation (also called the quantum Liouville equation) suggests that in equilibrium, probability distributions and state variables are constant.

In summary, the fundamental postulate cannot be proven rigorously and thus stays a postulate. Generations of physicists and mathematicians have tried to improve on the conceptual basis of statistical physics, e.g. trying to prove that certain model systems are ergodic, but usually this created only more riddles, e.g. in dynamical systems theory. The general idea however is clear: microscopic dynamics leads to a smearing out of correlations and if one does not resolve all of these details, entropy increases. In this sense, the increase of entropy is in fact a matter of the observer. Despite this somehow shaky foundation, the fundamental postulate has proven itself beyond doubt due to its success in explaining the physics of many particle systems.

3 The canonical ensemble

3.1 Boltzmann distribution

We consider a system in contact with a *'thermal reservoir'* or *'heat bath'*. Then it is temperature T rather than energy E that is fixed. A simple example would be a bottle of beer in a large lake, which eventually will cool down to the temperature of the lake. To qualify as a thermal reservoir, the surrounding system has to be much larger such that its temperature does not change as energy is exchanged with the system of interest. Together the two systems form an isolated system for which energy is fixed at E_{tot}.

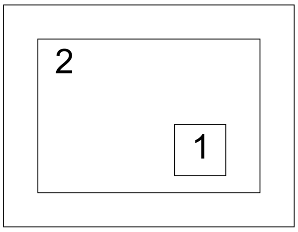

Figure 3.1: Two systems in thermal contact. System 2 is considerably larger than system 1 and serves as a *'thermal reservoir'*. Together the two systems are again a microcanoncial ensemble.

We now consider one specific microstate i in system 1. This microstate comes with an energy E_i. Its probability to occur is

$$p_i = \frac{\text{\# favorable outcomes}}{\text{\# possible outcomes}} = \frac{\Omega_{res}(E_{tot} - E_i)}{\Omega_{tot}(E_{tot})} = \frac{e^{S_{res}(E_{tot}-E_i)/k_B}}{e^{S_{tot}(E_{tot})/k_B}}$$

Here we used that the composite system is microcanonical and that we have fixed the microstate in system 1; then the number of accessible microstates is determined by system 2 (the reservoir) only.
We next introduce the average energy of system 1 as the reference energy:

$$U = \langle E \rangle = \sum_i p_i E_i$$

We now Taylor-expand the entropy of the heat bath:

$$S_{res}(E_{tot} - E_i) = S_{res}(E_{tot} - U + U - E_i) = S_{res}(E_{tot} - U) + \frac{U - E_i}{T}$$

Here we have used $\partial S/\partial E = 1/T$. Note that higher order terms do not appear because a reservoir has a constant temperature (first derivative constant, thus the second and the higher order derivatives vanish). We also can use additivity of the entropy to write

$$S_{tot}(E_{tot}) = S(U) + S_{res}(E_{tot} - U)$$

Defining the inverse temperature

$$\beta := \frac{1}{k_B\, T}$$

we thus get

$$p_i = \frac{e^{\beta U} e^{-\beta E_i} e^{S_{res}(E_{tot}-U)/k_B}}{e^{S(U)/k_B} e^{S_{res}(E_{tot}-U)/k_B}} = e^{\beta F} e^{-\beta E_i}$$

where $F = U - TS$ and where the terms with S_{res} have canceled out. We note that $p_i \sim e^{-\beta E_i}$ and that the prefactor $e^{\beta F}$ has the role of a normalization factor. In order to normalize, we use $\sum_i p_i = 1$ to write

$$\boxed{p_i = \frac{1}{Z} e^{-\beta E_i}} \quad \text{Boltzmann distribution}$$

with

$$\boxed{Z = \sum_i e^{-\beta E_i}} \quad \text{partition sum}$$

We conclude that the probability for a microstate decreases exponentially with its energy. The newly defined quantity Z is the central concept of the canonical ensemble and plays a similar role as the phase space volume Ω in the microcanonical ensemble.

Comments:

① We note that the expansion around $E_{tot} - U$ is not required to get the Boltzmann factor $e^{-\beta E_i}$. We would have obtained this result also by expanding simply around E_{tot}, because the derivative would also have given $1/T$. The normalization is ensured anyway by the new quantity Z. The expansion used here becomes important later because only in this way we get the prefactor $e^{\beta F}$. As we will discuss below in more detail, this leads to the important relation $F = -k_B T \ln Z$ connecting thermodynamics (F) and statistics (Z).

② For classical Hamiltonian systems we have

$$p(\vec{q}, \vec{p}) = \frac{1}{Z\, N!\, h^{3N}}\, e^{-\beta\, \mathcal{H}(\vec{q},\vec{p})}$$

with the Hamiltonian $\mathcal{H}$ and the partition sum (or, better, the partition function) is

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

③ From Liouville's theorem it follows that the Boltzmann distribution is a stationary distribution as $\rho = \rho\left(\mathcal{H}\right)$. Like for the microcanonical distribution, it is reasonable to associate it with equilibrium.

④ The Boltzmann distribution can also be motivated by information theory. In chapter 1 we showed that it maximizes entropy under the condition that:

$$U = \langle E \rangle = \text{const}$$

3.2 Free energy

The Boltzmann factor $e^{-\beta E_i}$ means that a microstate is the less likely the higher its energy. However, if we ask for the probability that the system has energy E, we have to consider all microstates with energy E:

$$\begin{aligned} p(E) &= \frac{1}{Z}\,\Omega(E)\,e^{-\beta E} = \frac{1}{Z}\,e^{-\beta E + S(E)/k_B} \\ &= \frac{1}{Z}\,e^{-\frac{E-TS}{k_B T}} = \frac{1}{Z}\,e^{-\beta F} \end{aligned}$$

This probability is maximal, if the quantity:

$$F(E,V,N) = E - T\,S(E,V,N)$$

has a minimum with respect to E. This is the case when:

$$0 = \frac{\partial F}{\partial E} = 1 - T\,\underbrace{\frac{\partial S}{\partial E}}_{=\frac{1}{T_1}}$$

that is when the system is at the temperature of the heat bath.
We conclude that not the smallest energy (like for vanishing T) or the highest entropy (like for very high T) is achieved in equilibrium, but the smallest F, which is a combination of energy and entropy with the relative importance determined by temperature. The smaller weight of the Boltzmann factor for higher energy is compensated by the fact that the number of microstates usually increases with energy. For example, for the ideal gas we have $\Omega \sim E^{3N/2}$ and therefore $S \sim \ln E$. Therefore $F = E - TS(E)$ diverges logarithmically at low E and linearly at high E, with a minimum somewhere inbetween (note that for this argument, we have to keep T fixed and cannot replace it be E, because this is the temperature of the heat bath given from outside).
Because in the canonical ensemble we fix T rather than E, we actually should write $F(T,V,N)$ using the caloric equation of state to convert E into T. This quantity is called the *'Helmholtz free energy'* or simply the *'free energy'*. Its appearance in the canonical ensemble provides a direct link to thermodynamics, which derives the extremum principle for the free energy soley from macroscopic arguments. We conclude:

In the canonical ensemble, equilibrium corresponds to the minimum of the free energy $F(T, V, N)$.

We can costruct $F(T, V, N)$ from $S(E, V, N)$ (entropy representation of the microcanonical ensemble) in the following way:

① introduce a new variable $\frac{1}{T} = \frac{\partial S(E,V,N)}{\partial E}$

② eliminate E in favor of T by solving for $E = E(T, V, N)$

③ construct $F(T, V, N) = E(T, V, N) - T\ S(E(T, V, N), V, N)$

Alternatively we can start from the energy representation of the microcanonical ensemble:

① $T(S, V, N) = \frac{\partial E}{\partial S}$

② $S = S(T, V, N)$

③ $F(T, V, N) = E(S(T, V, N), V, N) - T\ S(T, V, N)$

Mathematically, such procedures are known as *'Legendre transformations'*, when a function is rewritten such that it becomes a unique function of its derivative. Legendre transforms also occur in analytical mechanics:

$$\underbrace{\mathcal{L} = \mathcal{L}(q, \dot{q}, t)}_{\text{Lagrangian mechanics}} \quad \rightarrow \quad \underbrace{\mathcal{H} = \mathcal{H}(q, p, t)}_{\text{Hamiltonian mechanics}} = -(\mathcal{L} - \dot{q}\ p) \quad with\ p = \frac{\partial \mathcal{L}}{\partial \dot{q}}$$

where the Legendre transform of the (negative) Lagrange function gives the Hamiltonian function. The free energy F is in fact the Legendre transform of energy E with the argument changing from entropy S to temperature T, which is defined as a derivative of E.

We now consider the total differential of $F(T, V, N)$:

$$dF = dE + d(TS) = TdS - pdV + \mu dN - TdS - SdT = -SdT - pdV + \mu dN$$

This is the fundamental equation of the canonical ensemble and leads to three equations of state:

$$S = -\frac{\partial F}{\partial T}$$
$$p = -\frac{\partial F}{\partial V}$$
$$\mu = \frac{\partial F}{\partial N}$$

How do we calculate F? In principle, we can start from the microcanonical ensemble and calculate it as explained above. However, this will be quite tedious or even impossible for many cases. There is a much simplier way to do it. We recall that

$$p_i = e^{\beta F} e^{-\beta E_i} = \frac{1}{Z} e^{-\beta E_i}$$

and therefore we simply have

$$\Rightarrow \quad \boxed{F(T,V,N) = -k_B\, T\, \ln Z(T,V,N)}$$

Thus the free energy follows directly from the partition sum and there is no need to go through the microcanonical ensemble.
A similar comment applies to the average energy $U = \langle E \rangle$. In principle, it follows from the microcanoncial ensemble. However, if we work in the canonical ensemble, we can calculate it directly from the partition sum:

$$\begin{aligned}\langle E \rangle &= \sum_i p_i\, E_i = \frac{1}{Z} \sum_i E_i\, e^{-\beta E_i} \\ &= \frac{1}{Z} (-)\, \partial_\beta \sum_i e^{-\beta E_i} = \boxed{-\partial_\beta \ln Z(\beta) = \langle E \rangle}\end{aligned}$$

We can check consistency with the definition of the free energy:

$$\langle E \rangle = -\partial_\beta \ln Z(\beta) = \partial_\beta(\beta F) = F + \beta \frac{\partial F}{\partial \beta} = F - \frac{k_B T^2}{k_B T} \frac{\partial F}{\partial T} = F + TS$$

as it should.
For the variance of the energy we find:

$$\begin{aligned}\sigma_E^2 &= \left\langle (E - \langle E \rangle)^2 \right\rangle = \left\langle E^2 \right\rangle - \langle E \rangle^2 \\ &= \frac{1}{Z}\, \partial_\beta^2 Z - \left(\frac{1}{Z}\, \partial_\beta Z \right)^2 \\ &= \partial_\beta \left(\frac{1}{Z}\, \partial_\beta Z \right) = \partial_\beta^2\, \ln Z \\ &= -\partial_\beta \langle E \rangle = k_B\, T^2\, \partial_T \langle E \rangle \\ &= k_B\, T^2 \underbrace{c_v(T)}_{\text{specific heat}}\end{aligned}$$

Thus the second moment is related to a material property. Note that this implies $c_v(T) \geq 0$. We note that this equation connects to physical quantities that naively are unrelated: a measure for fluctuations on the left hand side and a material property or response function on the right hand side. Such surprising relations constitute the big success of thermodynamics. In the context of statistical physics, they are examples of the *fluctuation-dissipation theorem*.

We finally comment on the sharpness of the energy distribution. For the microcanonical ensemble, we showed that in the macroscopic limit $N \to \infty$ the distribution becomes very sharp around the value E. For the canonical ensemble, the same holds true for $\langle E \rangle$. We note that $c_v \approx N k_B$ and $\langle E \rangle \approx k_B T N$ and therefore

$$\frac{\sigma_E}{\langle E \rangle} = \frac{1}{\sqrt{N}} = 10^{-10}$$

for a macroscopic system with 10^{20} particles. In contrast, the single particle with $N = 1$ experiences very strong fluctuations.

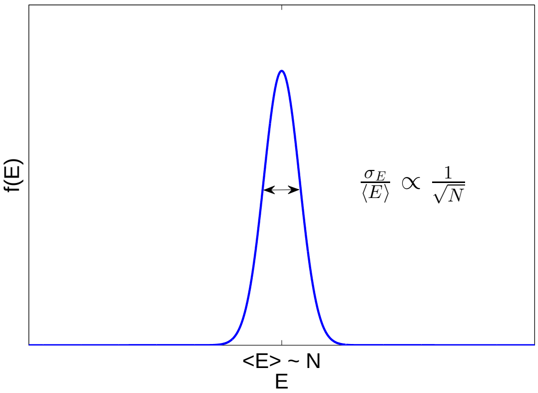

Figure 3.2: The energy distribution of a canonical system. $\langle E \rangle$ scales with N while $\sigma_E / \langle E \rangle$ is proportional to $1/\sqrt{N}$.

3.3 Non-interacting systems

The canonical formalism is especially convenient if the energy of the system is a sum over the energies of N independent elements (eg atoms in a gas or solid). We denote by ϵ_{ij} the j^{th} energy state of the i^{th} element:

$$\begin{aligned} Z &= \underbrace{\sum_{j_i} \sum_{j_2} \dots \sum_{j_N}}_{\text{microstate}} e^{-\beta \sum_{i=1}^{N} \epsilon_{i j_i}} \\ &= \left(\sum_{j_1} e^{-\beta \epsilon_{1 j_1}} \right) \left(\sum_{j_2} e^{-\beta \epsilon_{2 j_2}} \right) \dots \left(\sum_{j_N} e^{-\beta \epsilon_{N j_N}} \right) \\ &= z_1 \cdot z_2 \dots \cdot z_N = \prod_{i=1}^{N} z_i \\ \Rightarrow \quad F &= -k_B\, T \sum_{i=1}^{N} \ln z_i = -k_B\, T\, \ln Z \end{aligned}$$

Thus the partition sum factorises and the free energy is additive over the elements. For N identical elements we simply have:

$$Z = z^N, \quad F = -k_B \, T \, N \, \ln z$$

where z is the *'single particle partition sum'* .

Example:

① Two-state system

We consider a two state system as discussed above. As all particles are identical, we only have to calculate the single particle partition sum.

$$z = 1 + e^{-\beta\epsilon} \quad \Rightarrow \quad E = -\partial_\beta \ln z^N = \frac{N \, e^{-\beta\epsilon} \, \epsilon}{1 + e^{-\beta\epsilon}} = \frac{N \, \epsilon}{1 + e^{\beta\epsilon}}$$

This is exactly the result we obtained in two complicated calculations from the microcanonical ensemble.
Remember: $c = \frac{dE}{dT}$ gives the Schottky-hump.

Consider again the two limits:

① $T \to \infty \ (\beta\epsilon \ll 1) : z \to 2$

$$\left.\begin{aligned} p_0 &= \frac{1}{z} = \frac{1}{2} \\ p_\epsilon &= \frac{e^{-\beta\epsilon}}{z} = \frac{1}{2} \end{aligned}\right\} \text{Both states are equally likely.}$$

② $T \to 0 \ (\beta\epsilon \gg 1) : z \to 1$

$$\left.\begin{aligned} p_0 &= \frac{1}{z} = 1 \\ p_\epsilon &= \frac{e^{-\beta\epsilon}}{z} = 0 \end{aligned}\right\} \text{Only the ground state is occupied.}$$

② Harmonic oscillator (Einstein model)

$$E_n = \hbar \, \omega \, (n + \frac{1}{2}), \ \ n = 0, 1, ...$$

Again we only need the single particle partition sum:

$$z = \sum_{n=0}^{\infty} e^{-\beta E_n} = e^{-\frac{1}{2}\beta\hbar\omega} \left(1 + e^{-\beta\hbar\omega} + e^{-2\beta\hbar\omega} + ...\right)$$

$$= e^{-\frac{1}{2}\beta\hbar\omega} \frac{1}{1 - e^{-\beta\hbar\omega}} \quad \text{using the geometrical sum}$$

$$\Rightarrow \quad E = -\partial_\beta \ln z^N = N \left\{ \frac{\hbar\omega}{2} - \frac{e^{-\beta\hbar\omega}\,\hbar\omega}{1 - e^{-\beta\hbar\omega}} \right\}$$

$$= N\,\hbar\,\omega \left\{ \frac{1}{e^{\beta\hbar\omega} - 1} + \frac{1}{2} \right\}$$

Again this is the same result as the one we obtained from microcanonical calculations.

$$T \to 0 \; (\beta\hbar\omega \to \infty): \quad E \to \frac{N\hbar\omega}{2} \quad \text{Each oscillator is in its ground state.}$$

$$T \to \infty \; (\beta\hbar\omega \to 0): \quad E \to N k_B T \to \infty \quad \text{Each of the two modes carries energy } k_B T/2.$$

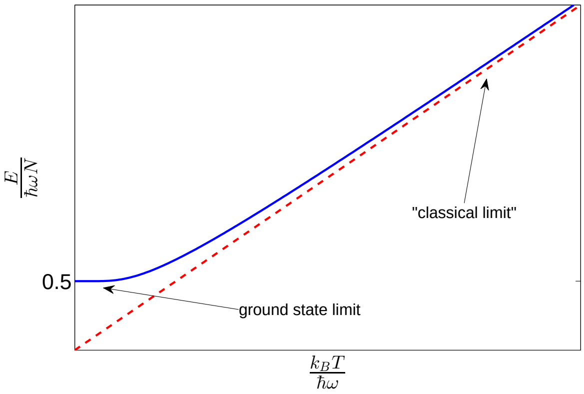

Figure 3.3: Harmonic oscillator: For T large, the energy of an harmonic oscillator (blue curve) increases linearly with temperature as in the classical case.

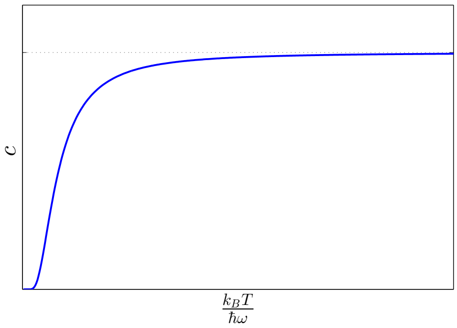

Figure 3.4: Harmonic oscillator: The heat capacity c asymptotically approaches a constant value at high temperature.

③ Ideal gas

$$
\begin{aligned}
Z &= \frac{1}{N!\,h^{3N}} \int \mathrm{d}\vec{q}\,\mathrm{d}\vec{p}\; e^{-\beta p^2/(2m)} \\
&= \frac{1}{N!\,h^{3N}}\, V^N \left(\frac{2\pi\, m}{\beta}\right)^{\frac{3N}{2}} \\
\Rightarrow \quad F &= -k_B\, T\, \ln Z \\
&= -k_B\, T\, N \left[\ln V + \frac{3}{2}\,\ln\left(\frac{2\pi m k_B T}{h^2}\right) - \frac{\ln N!}{N}\right] \\
&\underbrace{\approx}_{\text{Stirling}} \boxed{-k_B\, T\, N \left[\ln\left(\frac{V}{N}\right) + \frac{3}{2}\,\ln\left(\frac{2\pi m k_B T}{h^2}\right) + 1\right] = F(T,V,N)}
\end{aligned}
$$

$$
\langle E\rangle = -\partial_\beta\, \ln Z = \frac{3N}{2}\,\frac{1}{\beta} = \boxed{\frac{3}{2}\,N\,k_B\,T = E} \quad \text{caloric equation of state}
$$

We see that we get the same result as for the microcanonical ensemble.

$$
p = -\frac{\partial F}{\partial V} = k_B\, T\, N\, \frac{1}{V} \quad \Rightarrow \quad \boxed{p\, V = N\, k_B\, T} \quad \text{thermal equation of state (ideal gas law)}
$$

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

Replacing $T = \frac{2E}{3Nk_B}$ and using the thermal wavelength $\lambda = \left(\frac{3h^2 N}{4\pi m E}\right)^{\frac{1}{2}}$, we recover the result form the microcanonical ensemble:

$$
S = k_B\, N\, \left\{\ln\left(\frac{V}{N\lambda^3}\right) + \frac{5}{2}\right\}
$$

One can easily check that $F(T,V,N)$ also follows as Legendre transform from this expression for $S(E,V,N)$.

④ Maxwell distribution

For a classical gas with $\mathcal{H} = \frac{p^2}{2m} + V(\vec{q})$, we ask for the probability $\mathrm{d}W$ to find some particle, eg $i = 1$, in the momentum interval

$$p \leq |p_1| \leq p + \mathrm{d}p.$$

Thus we simply integrate out all other degress of freedom: $\vec{p}_2, \ldots, \vec{p}_N, \vec{q}_1, \ldots, \vec{q}_N$.

$$\Rightarrow \quad \mathrm{d}W = \frac{1}{\widetilde{Z}} \, 4\pi \, e^{-\beta p^2/(2m)} \, p^2 \, \mathrm{d}p$$

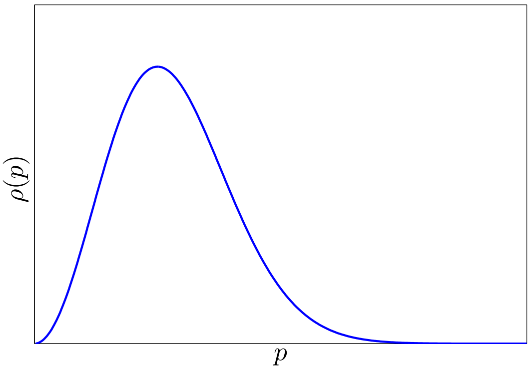

Figure 3.5: Maxwell distribution: The probability density ρ as a function of p.

The normalization constant is

$$\widetilde{Z} = (2\pi \, m \, k_B \, T)^{\frac{3}{2}}$$

such that $\int \mathrm{d}W = 1$. This result is valid for any type of gas (arbitrary interaction potential V). For an ideal gas, the average energy is the average kinetic energy:

$$E = N \left\langle \frac{p^2}{2m} \right\rangle = \frac{3}{2} \, N \, k_B \, T$$

as expected.

One can easily calculate that the most likely velocity is $v_{max} = \sqrt{2k_B T/m}$. For air, we use the mass of an oxygen molecule, 32 g / N_A with the Avogadro number $N_A = 6 \cdot 10^{23}$ (nitrogen or a mixture of oxygen and nitrogen would give us the same order of magnitude, because their molecular weights are so similar). With $k_B T = 4 \cdot 10^{-21} J$, we then get $v_{max} = \sqrt{2k_B T/m} = 400 m/s$. However, the mean free path length is only 10^{-7} m and the mean collision time $2 \cdot 10^{-10}$ s, so the molecules do not get far. One also can calculate that the coefficient of variation $\sigma_v / \langle v \rangle \approx 0.67$, which shows that the distribution is very broad.

⑤ Barometric formula

For an ideal gas in the gravitational field of the earth we have

$$\mathcal{H} = \sum_{i=1}^{N} \left\{ \frac{p_i^2}{2m} + m\, g\, h_i \right\}$$

where h_i is the height of particle i. To get the probability that a particle is at height h, we arbitrarily choose one particle and integrate out all other degress of freedom:

$$\Rightarrow \quad \mathrm{d}W = \frac{e^{-mgh/(k_B T)}}{\widetilde{Z}}\, \mathrm{d}h$$

Thus the density will decrease exponentially with distance. The length scale for this will be $(k_B T)/(mg)$. Using the mass of oxygen molecules, $32g/N_A$ with the Avogadro number $N_A = 6\ 10^{23}$, gravitational acceleration $g = 9.8 m/s^2$, and $k_B T = 4\ 10^{-21} J$, we get 7.5 km, in good agreement with the experimental scale. Because of the ideal gas law, the pressure will also decay exponentially on the same scale. Note that this model neglects the fact that temperature decreases with height (typically one degree by 100 m).

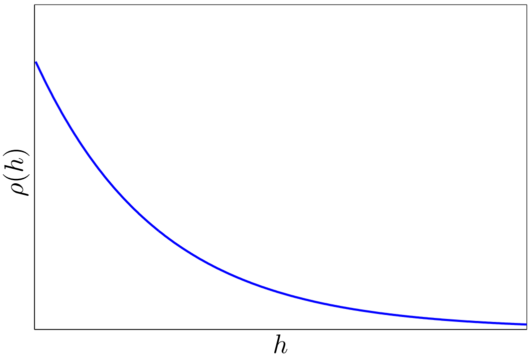

Figure 3.6: Barometric formula: The probability density ρ as a function of h. Assuming constant T, the density decays exponentially. For $T \downarrow$ the particles come down, while for $T \uparrow$ they go up.

3.4 Equipartition theorem

We define f to be the degrees of freedom (DOFs) that can be excited by thermal energy. We have a look at the heat capacities for the harmonic oscillators and the ideal gas:

① harmonic oscillator: $f = 2N$, each oscillator has kinetic and potential energy

$$T \to \infty: \quad c_v = N\, k_B = f\, \frac{1}{2}\, k_B$$

② ideal gas: $f = 3N$, each particles has three degrees of translation, but there is not interaction potential

$$c_v = \frac{3}{2}\, N\, k_B = f\, \frac{1}{2}\, k_B$$

Such a result was not obtained for the two-state system. Obviously it is related to harmonic terms in the Hamiltonian.
Consider a harmonic Hamiltonian with $f = 2$:

$$\mathcal{H} = A\, q^2 + B\, p^2$$

$$\begin{aligned} z &\propto \int \mathrm{d}q\, \mathrm{d}p\, e^{-\beta\mathcal{H}} \\ &= \left(\frac{\pi}{A\beta}\right)^{\frac{1}{2}} \cdot \left(\frac{\pi}{B\beta}\right)^{\frac{1}{2}} \\ &\propto \left(T^{\frac{1}{2}}\right)^{f} \end{aligned}$$

We conclude:

> *For sufficiently high temperature (classical limit), each quadratic term in the Hamiltonian contributes a factor $T^{\frac{1}{2}}$ to the partition sum* ***('equipartition theorem')****.*

This then leads to the following terms in other quantities:

$$\left.\begin{aligned} \text{free energy}: F &= -k_B T \ln z = -\frac{f}{2}\, k_B\, T\, \ln T \\ \text{entropy}: S &= -\frac{\partial F}{\partial T} = \frac{f}{2}\, k_B\, (\ln T + 1) \\ \text{internal energy}: U &= -\partial_\beta \ln z = \frac{f}{2}\, k_B\, T \\ \text{heat capacity}: c_v &= \frac{dU}{dT} = \frac{f}{2}\, k_B \end{aligned}\right\} \text{all extensive, scale with } f$$

Examples

① monoatomic ideal gas A monoatomic ideal gas has three translational (kinetic) degrees of freedom.

$$\Rightarrow \quad c_v = \frac{3}{2}\, N\, k_B$$

The positional degrees of freedom do not have quadratic energy. However, special wall potentials can give similar contributions. For the ideal gas, the wall only leads to the volume term.

② diatomic ideal classical gas (eg H-H)
The molecule has 6 modes, but different modes have different degrees of freedom. We consider a classical Hamiltonian in the center of mass system and look at relative rotational motions:

$$\mathcal{H} = \underbrace{\frac{1}{2M}\left(p_x^2 + p_y^2 + p_z^2\right)}_{\text{translation: } f_t=3} + \underbrace{\frac{1}{2I}\left(\frac{p_\Phi^2}{\sin^2\Theta} + p_\Theta^2\right)}_{\text{rotation: } f_r=2} + \underbrace{\left(\frac{p_r^2}{2\mu} + \frac{\mu\omega^2 r^2}{2}\right)}_{\text{vibration: } f_v=2}$$

Here I denotes moment of inertia and μ the reduced mass. The total number of degrees of freedom then is

$$f = f_t + f_r + f_v = 7$$

$$c_v = \frac{7}{2}\, N\, k_B$$

By only measuring the heat capacity c_v, we can already infer some far-reaching statements about the architecture of the molecule. Note that this measurement does not work at room temperature, when the vibrational modes are not populated yet. At room temperature, one only measures $c_v = (5/2)Nk_B$.

③ Triatomic molecules
Here one has to differ between triangular and linear shaped molecules. Each have $3 \times 3 = 9$ modes, but the degrees of freedom f are different. For the triangular shaped molecules, we have three rotational modes and three vibrations, giving $f = 3 + 3 + 2x3 = 12$. For the linear molecule, we have only two rotational modes but four vibrational ones, giving $f = 3 + 2 + 2x4 = 13$.

Virial theorem

We consider a classical Hamiltonian system. Let $x_i = q_i, p_i$ be position or momentum.

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = \frac{1}{Z} \int \mathrm{d}\Gamma\, \underbrace{\left(x_i \frac{\partial \mathcal{H}}{\partial x_j}\right) e^{-\beta\mathcal{H}}}_{=x_i \frac{\partial}{\partial x_j}\left(e^{-\beta\mathcal{H}}\right) \frac{1}{-\beta}}$$

$$= k_B\, T\, \frac{1}{Z} \int \mathrm{d}\Gamma\, \underbrace{\frac{\partial x_i}{\partial x_j}}_{\delta_{ij}}\, e^{-\beta\mathcal{H}} = \boxed{k_B\, T\, \delta_{ij} = \left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle}$$

Here we used partial integration and considered no contributions at infinity.

Specify to position and harmonic potential:

$$V_i = \frac{1}{2}\, m\, \omega^2\, q_i^2$$

$$\Rightarrow \quad \left\langle q_i \frac{\partial V}{\partial q_i} \right\rangle = \left\langle q_i \left(m\omega^2 q_i \right) \right\rangle = 2 \left\langle V_i \right\rangle$$

$$\Rightarrow \quad \boxed{\langle V_i \rangle = \frac{k_B T}{2}}$$

Application to momentum and kinetic energy instead yields:

$$T_i = \frac{p_i^2}{2m}$$

$$\Rightarrow \quad \left\langle p_i \frac{\partial T_i}{\partial p_i} \right\rangle = \left\langle p_i \frac{p_i}{m} \right\rangle = 2 \, \langle T_i \rangle = k_B \, T$$

$$\Rightarrow \quad \boxed{\langle T_i \rangle = \frac{k_B T}{2}}$$

Thus we obtain the same results as from the canonical ensembles. In classical mechanics, we derived a virial theorem for time rather than ensemble average:

$$\langle T \rangle = \frac{1}{2} \langle \vec{q} \cdot \nabla V \rangle = \langle V \rangle$$

where the last step only works for harmonic systems and the average denote time averages.

3.5 Molecular gases

Each molecule can be thermally excited in many ways. Apart from the three translational modes, there are vibrational, rotational, electronic and nuclear ones. Often one can assume that they are independent:

$$Z = Z_{trans} \cdot Z_{vib} \cdot Z_{rot} \cdot Z_{elec} \cdot Z_{nuc}$$

With respect to the N molecules in a gas we have:

$$Z_x = z_x^N$$

where x stands for the different mode types. Z_{trans} is special, because it includes the effect of interaction potential, but we will postpone this for now and treat it later in relation to phase transitions. For a vanishing interaction potential or dilute gas, Z_{trans} is the one of the ideal gas, with a factor $1/N!$ accounting for exchange of particle identities. Only Z_{trans} depends on volume and thus contributes to pressure. In the following we mainly discuss diatomic molecules like molecular hydrogen H_2. Important applications include plasma physics (fusion) and astrophysics (star formation from interstellar clouds).

Vibrational modes

Intermolecular vibrations are often described by the *'Morse potential'*:

$$V(r) = E_0 \left(1 - e^{-\alpha\,(r-r_0)}\right)^2$$

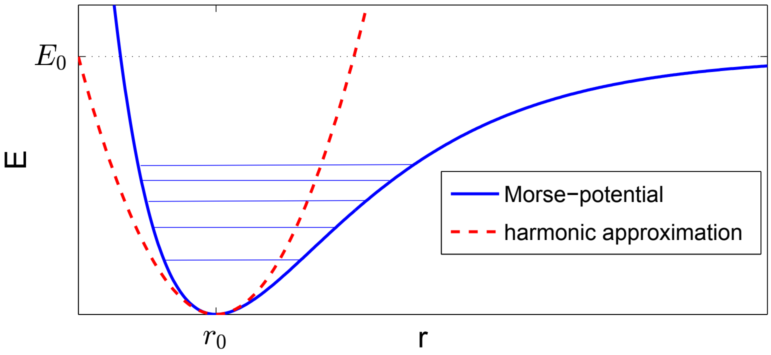

Figure 3.7: The Morse potential (blue) and its harmonic approximation (red) as a function of the interatomic distance r.

An exact solution of the Schrödinger equation gives

$$E_n = \hbar\,\omega_0 \left(n + \frac{1}{2}\right) - \frac{\hbar^2 \omega_0^2}{4E_0} \left(n + \frac{1}{2}\right)^2$$

$$\omega_0 = \frac{\alpha}{2\pi} \sqrt{\frac{2E_0}{\mu}}, \quad \mu = \frac{m}{2} \text{ reduced mass}$$

For $\hbar\omega_0 \ll E_0$ (small anharmonic contribution, valid for E_0 large or α small) we can use the harmonic approximation (anharmonic corrections can be calculated with perturbation theory):

$$z_{vib} = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega_0}}$$

Excitation of this mode occurs at:

$$T_{vib} \approx \frac{\hbar\omega_0}{k_B} \approx 6.140\ K \quad \text{for } H_2$$

Rotational modes

The standard approximation is the one of a rigid rotator. The moment of inertia for a rotator is given as

$$I = \mu\, r_0^2$$

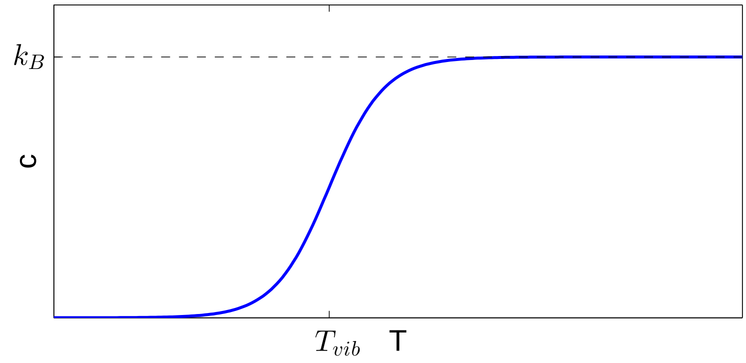

Figure 3.8: The heat capacity c_v as a function of temperature T. Around $T = T_{vib}$ the capacity increases and then levels off at k_B.

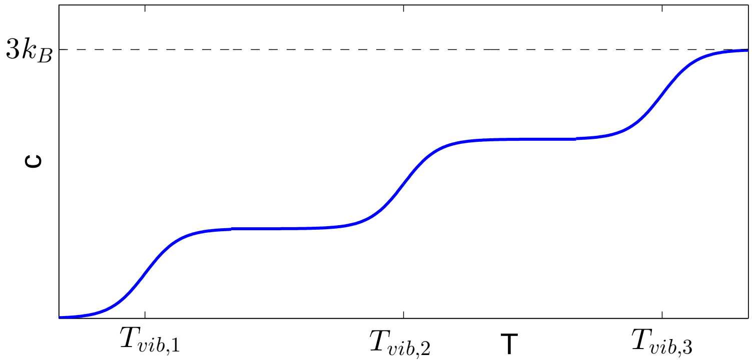

Figure 3.9: For N_{vib} vibrational modes, the heat capacity raises in a step-wise manner. Here the heat capacity c_v is plotted as a function of T for $N_{vib} = 3$.

The rotation around the horizontal axis does not matter, because the corresponding quantum mechanical wave function is rotationally symmetric. The eigenfunctions obtained from the Schrödinger equation are the spherical harmonics Y_{ml}.

$$\Rightarrow \quad E_l = \frac{\hbar^2}{2I}\, l(l+1)$$

with degeneracy $g_l = 2l + 1$.

The internal rotation contributes to the heat capacity of the diatomic molecule as shown in Figure 3.10 (small bump before the plateau). For H_2, the rise occurs for 85.4 K. Together with the vibrational contribution, this yields a specific heat with two rising parts as schematically shown in Figure 3.11. Below 20 K, the gas condensates into a liquid. Around 10.000 K, it starts to dissociate into atoms.

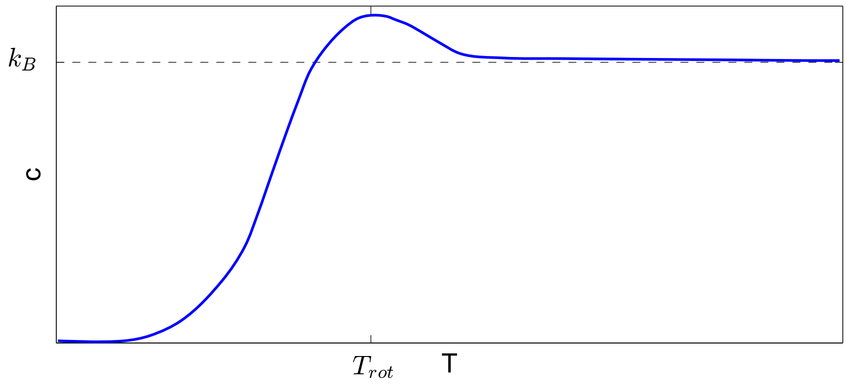

Figure 3.10: Rigid rotator: Sketch showing the contribution of internal rotation to the heat capacity of a diatomic molecule. For eg H_2: $T_{rot} = \hbar^2/(I\,k_B) \approx 85.4\,K$.

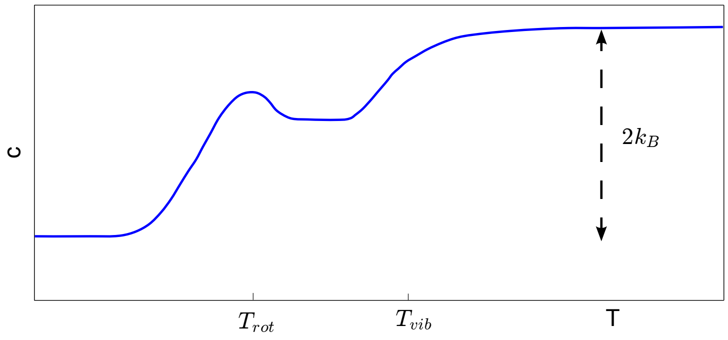

Figure 3.11: Sketch showing the specific heat of a diluted gas of H_2 as a function of temperature.

Nuclear contributions: ortho- and parahydrogen

Due to the existence of a covalent bound, the spins of the electrons are coupled to $S_{elec} = 0$. However, the spins of the protons can have variable coupling. The two protons are fermions (spin-$\frac{1}{2}$ particles). For spin coupling $S = 1$ (parallel spins), there are three possible spin states:

$$|\uparrow\uparrow\rangle\,,\ \frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle\right),\ |\downarrow\downarrow\rangle \qquad \text{triplet state with degeneracy 3:}\textit{'orthohydrogen'}$$

For spin coupling $S = 0$ (anti-parallel spins), there is only one possible spin state:

$$\frac{1}{\sqrt{2}}\left(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle\right) \qquad \text{singlet state, no degeneracy:}\textit{'parahydrogen'}$$

Under exchange of the two protons, the total wave function has to be anti-symmetric:

$$\Psi(1,2) = \Psi_0(r)\, Y_{lm}(\Theta,\Phi)\, |SS_z\rangle$$

$$\Rightarrow \quad P_{12}\,\Psi(1,2) = (-1)^{l+S+1}\,\Psi(1,2) \overset{!}{=} -\Psi(1,2)$$

$$\Rightarrow \quad \begin{array}{lll} S = 1\ (ortho): & l = 1,3,5... & \text{only odd } l \text{ are allowed} \\ S = 0\ (para): & l = 0,2,4... & \text{only even } l \text{ are allowed} \end{array}$$

As for low T $(T \ll T_{rot})$ only the ground state $(l = 0)$ is occupied, we conclude that only parahydrogen can exist at low temperatures.
For high temperatures $(T \gg T_{rot})$, all four spin states are equally likely and ortho- and parahydrogen hence exist in a ratio of 3 : 1.
Importantly, the relaxation time to equilibrium is one year due to a very weak interaction energy. Below this time scale, the system has a memory and the specific heat depends on the preparation history.

Restricted partition sums:

$$z_{ortho} = \sum_{l=1,3,5,...} (2l+1)\, e^{-\frac{l(l+1)T_{rot}}{T}}$$

$$z_{para} = \sum_{l=0,2,4,...} (2l+1)\, e^{-\frac{l(l+1)T_{rot}}{T}}$$

After a long time, we have $z_{rot} = 3z_{ortho} + z_{para}$. Figure 3.12 shows schematically how the two different contributions to c_v and their equilibrium mixture behave as a function of temperature.

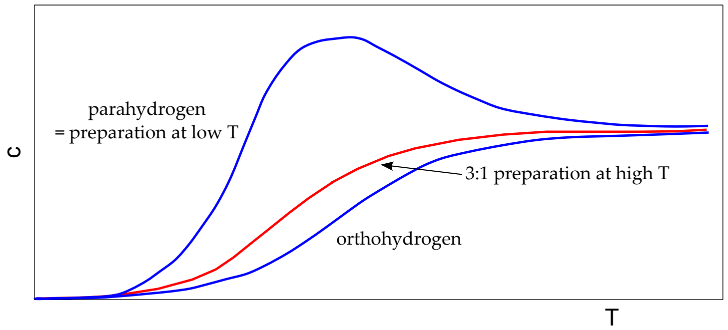

Figure 3.12: Heat capacity as a function of temperature for ortho- and parahydrogen (blue). The red curve indicates the equilibrium ratio as obtained at high temperature.

3.6 Specific heat of a solid

We consider a crystal to be a collection of masses and springs (*harmonic crystal*). The springs represent the harmonic approximations to more complicated interactions potentials. We only focus on the $3N$ vibrational modes and disregard electronic or nuclear modes (see discussion below on the electronic degrees of freedom). Translational and rotational modes of the overall object are disregarded because the total mass is much larger than the mass of its constituents.
As we discussed before, the simplest approach is the so-called *'Einstein solid'*, consisting of $3N$ oscillators with frequency ω_0 vibrating around their equilibrium positions. However, this model does not predict the experimentally measured low temperature dependence correctly ($c \propto T^3$). What is missing here is the fact that the atoms do not oscillate by themselves, but collectively. The corresponding model was suggested by Debye.

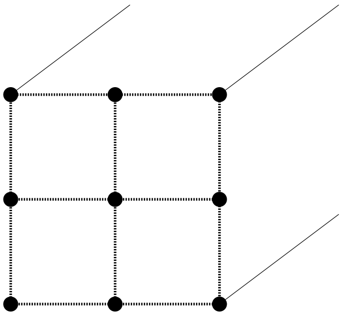

Figure 3.13: Simple model of a crystal: The regularly distributed atoms (here a simple cubic lattice) are connected by springs forming a large coupled system.

Debye model (1912)

We first consider a one-dimensional crystal (linear chain).

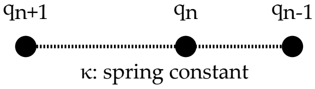

Figure 3.14: Mechanical waves in one dimension: the atoms are arranged in a linear chain and coupled with springs. q_i denotes the elongation away from equilibrium for the i-th atom.

Each atom is deplaced by $q_n(t)$ from its equilibrium position $x_n = n\ a$. We write down the equation of motion:

$$\begin{aligned} m\,\ddot{q_n} &= F_n = \kappa\,(q_{n+1}-q_n) - \kappa\,(q_n - q_{n-1}) \\ &= \kappa\ \underbrace{(q_{n+1}+q_{n-1}-2q_n)}_{\text{discretized second derivative}} \end{aligned}$$

Here κ is the spring constant (we cannot use k because this will be the wavenumber later). In the continuum limit, this gives a wave equation $\ddot{q} - c_s^2 q'' = 0$ with the wave velocity $c_s = (\kappa a^2/m)^{1/2}$ (velocity of sound). We stay in the discrete picture but use the fact that we expect wave-like solutions:

$$q_n(t) = A\ e^{i(\omega t + kna)}$$

$$\omega = \frac{2\pi}{T}\ \text{frequency}, \quad k = \frac{2\pi}{\lambda}\ \text{wave number}$$

$$\Rightarrow \quad -m\omega^2 = \kappa\left[e^{ika}+e^{-ika}-2\right] = -2\kappa\,[1-\cos(ka)] = -4\kappa\,\sin^2\left(\frac{ka}{2}\right)$$

$$\Rightarrow \quad \boxed{\omega(k) = \left(\frac{4\kappa}{m}\right)^{\frac{1}{2}}\ |\sin\left(\frac{ka}{2}\right)|}$$

The frequency is not constant, but depends on the wave number. Since the solution is 2π-periodic, we restrict it to the interval $-\pi/a < k < +\pi/a$ (*first Brillouin zone*). We note that for $k = 0$ (center of Brillouin zone), the phase difference between two neighbors is $e^{ka} = 1$ and they oscillate in phase. For $k = \pi/a$ (boundary of Brillouin zone), we have a phase difference of $e^{\pi} = -1$ and neighbors oscillate against each other. The oscillation at small frequency defines the same velocity of sound as identified above because both cases correspond to the continuum limit:

$$c_s = \left.\frac{\mathrm{d}\omega}{\mathrm{d}k}\right|_{k=0} = \sqrt{\frac{\kappa}{m}}\ a$$

We next consider a finite system size, that is N atoms and a chain length $L = N\ a$. We choose periodic boundary conditions (*Born-von Karman boundary conditions*):

$$q_n = q_{n+N}$$

$$\Rightarrow \quad e^{ikNa} = 1 \quad \Rightarrow \quad k_j = \frac{2\pi}{aN}\,j, \quad j = 0, \pm 1, ..., \pm\frac{N}{2}$$

Only discrete $k-$values are allowed in the first Brillouin zone. As a result we have N degrees of freedom, thus only N oscillators are possible. Their $k-$values are separated by:

$$\Delta k = \frac{2\pi}{L}$$

which is the density of states in $k-$space.

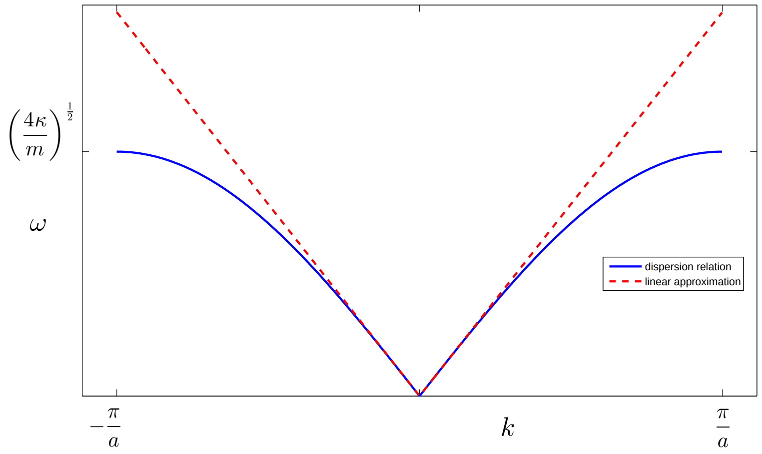

Figure 3.15: The dispersion relation of a linear chain (blue) and its linear approximation for small values of k (red). The Debye model uses the linear dispersion relation.

Going to a three-dimensional system, waves are now written as:

$$\vec{A}e^{i(\vec{k}\cdot\vec{r}+\omega t)}$$

with wave vector $\vec{k}$. In addition to the longitudinal wave discussed in one dimension, we also have two transversal waves. The number of modes then becomes:

$$\begin{aligned}\#\text{modes} &= \sum_{\vec{k}} 3 = 3 \prod_{i=x,y,z} \int_{-\pi/a}^{\pi/a} \frac{\mathrm{d}k_i}{\Delta k} \\ &= 3\left(\frac{2\pi}{a}\,\frac{L}{2\pi}\right)^3 = \frac{3V}{a^3} = 3N\end{aligned}$$

Here we introduced the volume $V = L^3$ and then used $V/a^3 = N$. This result makes a lot of sense: the overall number of modes has to be the same both in real and in wave space.

We now transfer the summation over modes from $\vec{k}$-space to ω-space. This implies that we now switch from Cartesian to spherical coordinates. As an approximation we assume a linear dispersion relation:

$$\omega = c_s\,|k|$$

We also assume that longitudinal and transversal modes have the same velocity and

and that all directions are equally probable.

$$\Rightarrow \quad \#\text{modes} = 3 \int \frac{\mathrm{d}^3 k}{(\Delta k)^3} = \frac{3V}{(2\pi)^3} \int_0^{k_{max}} 4\pi\, k^2 \, \mathrm{d}k$$
$$= \frac{3V}{2\pi^2\, c_s^3} \int_0^{\omega_D} \mathrm{d}\omega\, \omega^2 = \frac{V}{2\pi^2 c_s^3}\, \omega_D^3 \overset{!}{=} 3N$$

The maximal or *Debye frequency* follows as:

$$\omega_D = c_s \left(\frac{6\pi^2 N}{V} \right)^{\frac{1}{3}} \approx 3.9\, \frac{c_s}{a}$$

where we used $V = N\, a^3$. This agrees approximately with our earlier statement:

$$k_{max} = \frac{\omega_D}{c_s} = \frac{\pi}{a}$$

The Debye frequency as frequency cutoff makes sure that we count the number of modes in the correct manner and it reflects the fact that in a periodic crystal, waves cannot have a wavelength smaller than the lattice constant.
In summary, the Debye model assumes the following (normalized) density of states in $\omega-$space:

$$D(\omega) = \begin{cases} 3\dfrac{\omega^2}{\omega_D^3} & \text{for } \omega \le \omega_D \\ 0 & \text{for } \omega > \omega_D \end{cases}$$

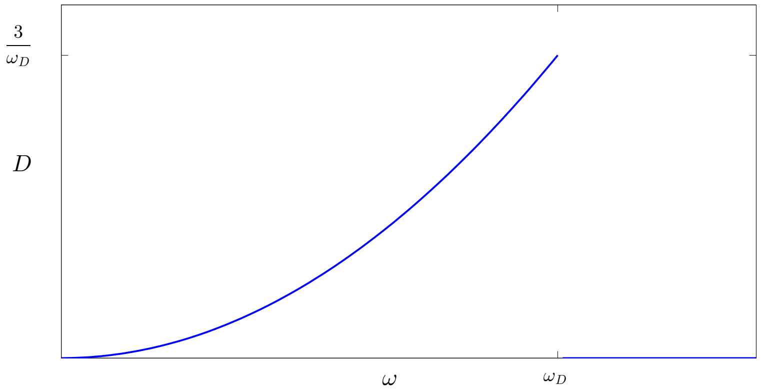

Figure 3.16: The density of states D for the Debye model as a function of ω.

We now know how to count modes in frequency-space:

$$\sum_{modes} (...) = 3 \sum_k (...) = 3N \int_0^{\omega_D} \mathrm{d}\omega\, D(\omega)\, (...)$$

The essential assumption of the Debye model is that each allowed k-mode is one harmonic oscillator that is quantized according to the rules of quantum mechanics (which means that the Schrödinger equation is implicitly used here). The ultimate justification for this choice is that it correctly predicts the experimental results for the specific heat. The Debye model is so popular that even a temperature-dependent Debye frequency $\omega_D(T)$ is used to present experimental results as deviations from the ideal Debye behaviour, although in reality these deviations show that the Debye model (which assumes constant ω_D) is not good enough to describe the details of the experimental data; for this one has to use more complicated densities of state.
We now use the single particle partition sum for each k-mode:

$$z(\omega) = \frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}}$$

$$\begin{aligned}
\Rightarrow \quad Z &= \prod_{\text{modes}} z(\omega) \\
\Rightarrow \quad E &= -\partial_\beta \, \ln Z = \sum_{\text{modes}} \hbar\omega \left(\frac{1}{e^{\beta\hbar\omega} - 1} + \frac{1}{2} \right) \\
&= E_0 + 3N \int_0^{\omega_D} \mathrm{d}\omega \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \frac{3\omega^2}{\omega_D^3} \\
c_v(T) &= \frac{\partial E}{\partial T} = \frac{(-1)}{k_B T^2} \frac{\partial E}{\partial \beta} \\
&= \frac{\hbar^2}{k_B T^2} \sum_{\text{modes}} \frac{e^{\beta\hbar\omega} \, \omega^2}{\left(e^{\beta\hbar\omega} - 1\right)^2} \\
&= \frac{3\hbar^2 N}{k_B T^2} \int_0^{\omega_D} \mathrm{d}\omega \, \frac{3\omega^2}{\omega_D^3} \, \frac{e^{\beta\hbar\omega} \, \omega^2}{\left(e^{\beta\hbar\omega} - 1\right)^2}
\end{aligned}$$

The constant contribution E_0 from the rest energy drops out for the specific heat. We change the integration variable to $u = \beta \, \hbar \, \omega$:

$$\Rightarrow \quad \boxed{c_v(T) = \frac{9Nk_B}{u_m^3} \int_0^{u_m} \frac{e^u \, u^4}{\left(e^u - 1\right)^2} \, \mathrm{d}u}$$

Unfortunately this integral cannot be solved analytically.

We consider two limits:

① $k_B T \gg \hbar \, \omega_D$, $u \ll 1$:

$$\Rightarrow \quad \frac{u^2 \, e^u}{\left(e^u - 1\right)^2} \approx 1 \Rightarrow \quad \boxed{c_v(T) = 3 \, N \, k_B} \quad \text{classical limit for 6N degrees of freedom}$$

② $k_B T \ll \hbar\, \omega_D,\ u \gg 1$:

$$\Rightarrow \qquad u_m \to \infty$$

The integral becomes a constant and the only temperature dependence comes from the factor $1/u_m^3$.

$$\Rightarrow \quad \boxed{c_v(T) = \frac{12\pi^4}{5}\, N\, k_B \left(\frac{T}{T_D}\right)^3}$$

Here we defined the *'Debye temperature'* $T_D = \hbar\omega_D/k_B$. Note that if you replace T_D by ω_D and then ω_D by the formula derived for it above, then N drops out and c_v scales simply as V, as it should as an extensive quantity.

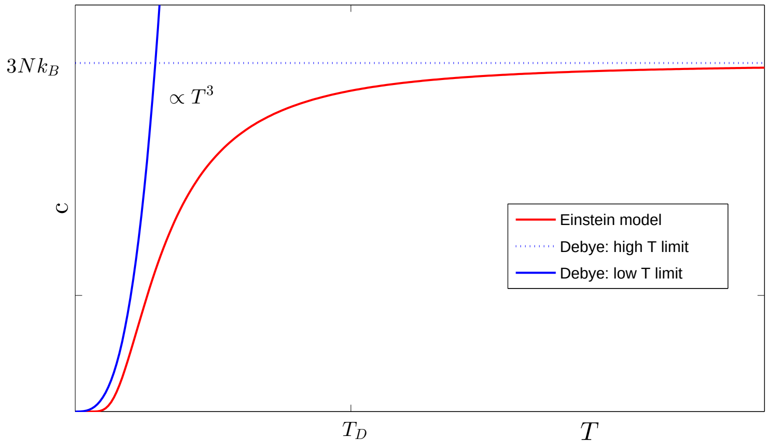

Figure 3.17: The heat capacity as a function of temperature for two limits of the Debye model (blue) compared with the Einstein model (red), which increases only exponentially at low T. For the Debye model $c_v(T) \propto T^3$ at low temperatures in agreement with experimental results.

Typical experimental values are:

$$c_s = 5.000\ \frac{m}{s}$$

$$a = 2\ \mathring{A}$$

$$\omega_D \approx \frac{4c_s}{a} = 10^{14}\ Hz$$

$$T_D = \frac{\hbar\omega_D}{k_B} =\approx \frac{10^{-34}\ Js\ 10^{14}\ Hz}{\frac{3}{2}\ 10^{-23}\frac{J}{K}} \approx 700\ K$$

As a matter of fact, most metals have Debye temperature in the range of hundreds of degrees of Kelvin. This shows that metals are in the quantum regime even at room temperature when it comes to their specific heats.

Summary

① At high T, we have the classical limit:

$$3N\,oscillators\ \times\ 2\,degrees\,of\,freedom\,each\ \times\ \frac{k_BT}{2}\,energy$$

② At low T, all modes with $\hbar\omega \leq k_BT$ are excited.

$$\Rightarrow\quad \#modes \propto \int \mathrm{d}\vec{k} \propto k^3 \propto \omega^3 \propto T^3$$

Final comments on solid state physics

① The Debye model belongs to the field of solid state physics. Classical textbooks are *Solid state physics* by Neil Ashcroft and David Mermin (2nd edition 1976) and *Introduction to Solid State Physics* by Charles Kittel (8th edition 2011). Yes, these books are old, but they are simply the best.

② The excitations of the lattice can be considered as quasi-particles with energy $\hbar\omega$ *('phonons')* in analogy to photons, which have similar statistics (*Bose statistics*, see black body radiation below). The three polarities of a wave with wave vector $\vec{k}$ can be interpreted as spins $(0, \pm 1)$.

③ The excitations discussed here are known as *acoustic* modes. The two transversal and the one longitudenal modes typically have different dispersion relations, so there are three branches for a monoatomic unit cell. All of them have vanishing frequency at $k = 0$. They are called acoustic because the define the velocity of sound. Moreover there can be also *optical* modes which have a finite frequency value at $k = 0$. These modes arise if the unit cell of the crystal contains several atoms. They are called optical because they can interact with light. For a diatomic unit cell, there can be three acoustic and three optical modes, but often some of them are degenerate.

④ The Debye model works so well because at a given temperature the average energy is fixed and much higher energies (which correspond to higher frequencies) do not matter. Thus the optical modes usually can be neglected and the linear approximation is good.

⑤ The standard way to measure the dispersion relation (or phonon spectrum) of a crystal is neutron scattering. X-rays also work, but this is much harder. One typically works around $k = 0$ in the Brillouin zone, which also satisfies the Bragg criterion for a strong reflection. One then looks for small side peaks (Stokes and anti-Stokes lines) that either come from acoustic phonons (Brillouin scattering) or from optical phonons (Raman scattering).

⑥ Experiments show that at very low temperatures, there is also a linear contribution to the specific heat of crystals. This comes from the Fermi fluid formed by the delocalized electrons and we will deal with this in the chapter on quantum fluids.

⑦ The "band" structure of the dispersion relations with their k-dependance and the different branches should not be mixed up with the band structure of the electronic degrees of freedom. The delocalized electrons have their own Schrödinger equation and due to the Bloch theorem wave vectors appear there as parameters. The eigenvalue problem leads to discrete energy levels n and the parameter $\vec{k}$ spreads them out into bands $E_n(\vec{k})$. If the different bands do not overlap, a band gap emerges. Depending on whether this gap can be bridged by thermal excitations or not, the crystal is either a semi-conductor or an insulator, respectively.

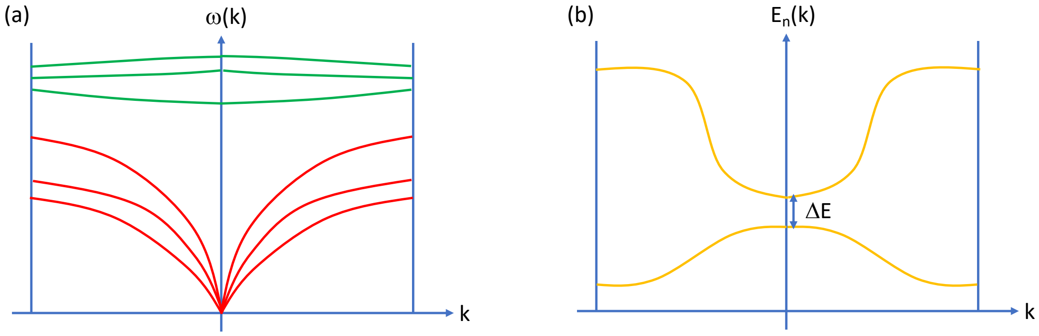

Figure 3.18: (a) Dispersion relation of the photons for a diatomic basis. The frequency ω depends on the wave number k. There can be three acoustic (red) and three optical (green) branches, if there is no degeneracy. (b)Band structure arising from the Schrödinger equation for the electrons. The band gap ΔE describes the energy difference that electrons need to get to hop to a higher energy level. In both cases, there is a dependance on wavenumber k running over the first Brillouin zone. Otherwise however the physics behind the branches in the dispersion relation and the bands in the electronic structure comes from very different physics and should not be mixed up with each other.

3.7 Black body radiation

The black body is a cavity filled with electromagnetic radiation in equilibrium with a thermal reservoir of temperature T. In the particle picture of quantum mechanics, we deal with a gas of photons. The dispersion relation for photons is

$$\epsilon(k) = \hbar\omega(k) = \hbar c k$$

with the velocity of light $c \approx 3 \cdot 10^8\ m/s$. This relation is linear in k as for the Debye model. In contrast to the Debye model, however, there is no cutoff in the frequency spectrum. Moreover there are only two possible polarizations (dictated by Maxwell equations - in contrast to lattice vibrations there is no longitudinal mode).
Like for the Debye model, each k-mode is a quantum mechanical oscillator:

$$\epsilon_{\vec{k},m} = \hbar\omega(k)\ \left(n_{\vec{k},m} + \frac{1}{2}\right)$$

$$\Rightarrow \quad E = \sum_{\vec{k},m} \epsilon_{\vec{k},m} = 2\sum_{\vec{k}} \epsilon(k) n_{\vec{k}} + E_0$$

While for the Debye model the ground state contribution E_0 is finite, this is not the case for photons. We disregard it as it is an unobservable constant.

$$\begin{aligned} \Rightarrow \quad E &= 2\left(\frac{1}{\Delta k}\right)^3 \int \mathrm{d}\vec{k}\ \frac{\epsilon(k)}{e^{\beta\epsilon(k)} - 1} \\ &= \frac{2V}{(2\pi)^3} 4\pi \int k^2\ \mathrm{d}k\ \frac{\hbar c k}{e^{\beta\hbar c k} - 1} \\ &= \frac{V}{\pi^2 \beta^4 \hbar^3 c^3} \qquad \underbrace{\int_0^\infty \mathrm{d}u\ \frac{u^3}{e^u - 1}}_{=\sum_{n=1}^\infty \int_0^\infty \mathrm{d}u\ u^3 e^{-nu} = 6\sum_{n=1}^\infty \frac{1}{n^4} = \frac{\pi^4}{15}} \end{aligned}$$

$$\Rightarrow \quad \boxed{E = \frac{4\sigma}{c} V T^4} \quad \text{with Stefan-Boltzmann constant } \sigma = \frac{\pi^2 k_B^4}{60\hbar^3 c^2}$$

For the heat capacity of a black body this implies:

$$\boxed{c_v = \frac{16\sigma}{c} V T^3}$$

The $c_V \propto T^3$ scaling is the same as for the low temperature limit of the Debye model. However, for the photon gas this result is rigorous as no linearization of the dispersion relation is required. Furthermore it is valid for all temperatures.
Looking back at our calculation for the photon gas, we see that it can be rewritten as

$$E = \int_{-\infty}^{+\infty} d\epsilon D(\epsilon) \epsilon n(\epsilon)$$

where we have a defined a density of state as

$$D(\epsilon) = \begin{cases} \frac{V}{\pi^2 (\hbar c)^3} \epsilon^2 & \text{for } \epsilon \geq 0 \\ 0 & \text{for } \epsilon < 0 \end{cases}$$

and where the distribution function per mode is

$$n(\epsilon) = \frac{1}{e^{\beta\epsilon} - 1} \, .$$

In the chapter on quantum fluids, we will see that this is the Bose distribution for bosons in the case $\mu = 0$ (no chemical potential for photons). The density of states is the same as for the Debye model, but without frequency cutoff, because in both cases we sum over all allowed wavevectors with a linear dispersion relation.
In analogy with the calculation of the mean energy, we now can also write the mean photon number, simply by leaving out the energy as the quantity to be averaged:

$$N = \int_{-\infty}^{+\infty} d\epsilon D(\epsilon) n(\epsilon) = \frac{V}{\pi^2} (\frac{k_B T}{\hbar c})^3 \int_0^\infty dx \frac{x^2}{e^x - 1} = \frac{V}{\pi^2} (\frac{k_B T}{\hbar c})^3 \Gamma(3) \zeta(3)$$

where the last integral just gives a numerical factor. If we evaluate these numbers, the final result is

$$N = 2.032 \; 10^7 \; V T^3$$

where length is measured in meter and temperature in K, respectively. Thus the number of photons increases linearly with volume (as to be expected for a non-interacting system) and with the third power of absolute temperature. Thus at room temperature, there are around $5 \; 10^{14}$ photons in one cubic meter.

Stefan-Boltzmann law

How much power P (energy/time) is emitted by a black body cavity with an opening of size A (Figure 3.19)? The energy flux J $\left(\frac{energy}{time \cdot area}\right)$ in direction of a photon is

$$J = \frac{E}{A \; \Delta t} = \frac{E \; \Delta x}{(A \; \Delta x) \; \Delta t} = \frac{Ec}{V}$$

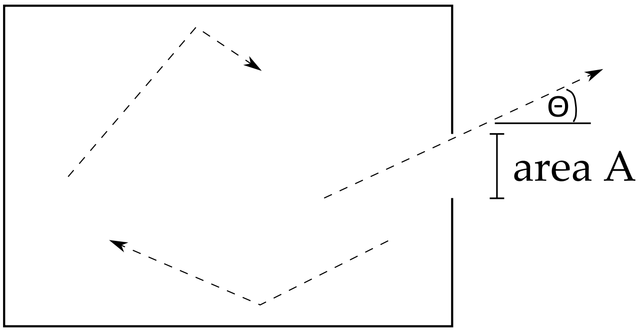

Figure 3.19: A block body cavity with opening of size A.

Photons move in all directions, but only the component perpendicular and outward to the opening contributes.

$$\Rightarrow \quad P = \frac{Ec}{V} A \frac{1}{4\pi} \underbrace{\int_0^{2\pi} \mathrm{d}\Phi}_{=2\pi} \underbrace{\int_0^{1} \mathrm{d}(\cos\Theta)\ \cos\Theta}_{=\frac{1}{2}}$$

$$\Rightarrow \quad \boxed{J = \frac{P}{A} = \sigma T^4} \quad \text{Stefan-Boltzmann law}$$

Planck spectrum

We return to the energy integral and note that it can be written as an integral over frequencies:

$$\frac{E}{V} = \int_0^\infty \mathrm{d}\omega\ u(\omega)$$

Here we defined:

$$\boxed{u(\omega) := \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\hbar\omega/(k_B T)} - 1}} \quad \text{Planck's law for black body radiation}$$

Figures 3.20 and 3.21 show the spectral radiance as a function of frequency and wavelength respectively.

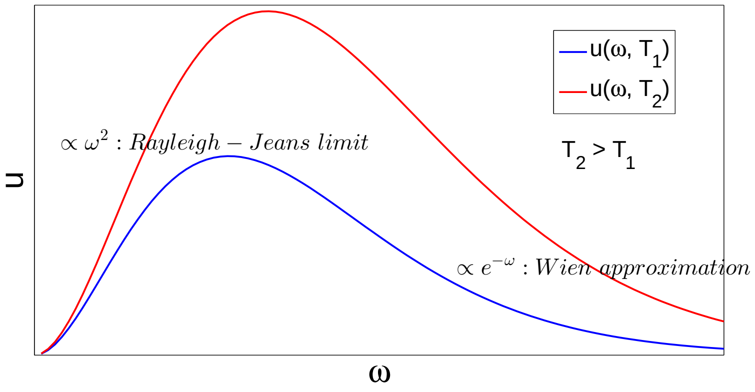

Figure 3.20: The spectral radiance u as a function of ω. For small or large ω Planck's law can be approximated by the Rayleigh-Jeans or Wien limits, respectively. Both limits were known to Max Planck when he derived his formula in 1900. For higher temperatures the spectral maximum is shifted to higher frequencies.

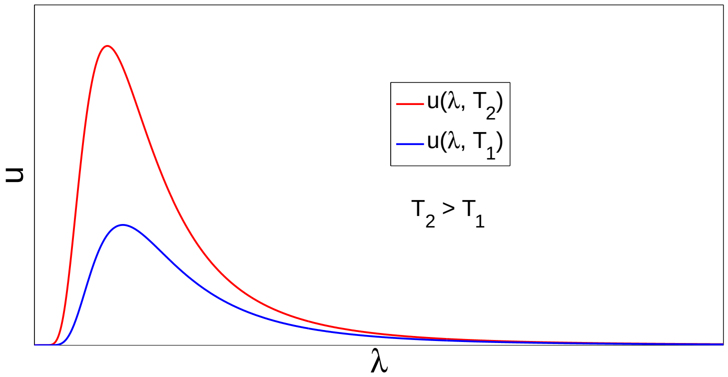

Figure 3.21: The spectral radiance u as a function of wavelength λ. For higher temperatures the spectral maximum is shifted to shorter wavelengths.

The Planck distribution has a maximum at

$$\left.\frac{\mathrm{d}u}{\mathrm{d}\omega}\right|_{\omega=\omega_m} \overset{!}{=} 0 \quad \Rightarrow \quad \beta\hbar\omega_m = 3\left(1-e^{-\beta\hbar\omega_m}\right)$$

$$\Rightarrow \quad \boxed{\hbar\omega_m = 2.82\, k_B T} \quad \text{Wien's displacement law}$$

For example, heating an iron from 1.000 to 2.000 K shifts its spectrum from red to white due to a larger contribution in the blue range.
The Planck distribution and the Stefan-Boltzmann law are very general and usually apply if matter and radiation are in thermodynamical equilibrium. The Planck distribution and Wien's displacement law can be used to extract the temperature of such a system:

① sun: $T = 5.800\ K \ \Rightarrow\ \lambda_m$ in the visible und UV range

② earth or human body: $T = 300\ K \ \Rightarrow\ \lambda_m$ in the infrared range; this is why we can see people and heated houses so well with infrared cameras.

③ cosmic background radiation: $T = 2,73\ K \ \Rightarrow\ \lambda_m \approx cm$ (microwaves) discovered by Penzias and Wilson 1964 (Nobel Prize in physics 1978)

The ideal photon gas depends on volume only through its dispersion relation (no particle interactions):

$$\epsilon(k) = \hbar c k \propto \frac{1}{L} = \frac{1}{V^{\frac{1}{3}}}$$

$$E = 2\sum_{\vec{k}} \epsilon(k) n_{\vec{k}}$$

$$p = -\frac{\partial E}{\partial V} = \frac{1}{3}\frac{E}{V} = \frac{1}{3}\frac{1}{c}\frac{cE}{V} = \frac{J}{3c}$$

$$J \approx 1 \frac{kW}{m^2} \quad \Rightarrow \quad p \approx 10^{-6} \frac{N}{m^2} = 10^{-11}\, bar$$

We note that the sun's radiation pressure on earth is much smaller than the atmospheric pressure.

Solar energy and the greenhouse effect

Our earth is heated by the sun. Because we know its temperature, radius and distance, we can calculate how much energy should arrive here and how much it should warm up the earth. The sun radiates an energy flux $J = \sigma T^4$ due to its temperature $T = 5.800\ K$. We have to multiply this by $4\pi R_s^2$ with $R_s = 696.000$ km being the radius of the sun to get the overall output. Because of energy conservation, at the distance $r = 1.5\ 10^8$ km where the earth is located, we have to divide by $4\pi r^2$ to get the flux onto the earth:

$$J_{in} = \sigma T_s^4 (\frac{R_s}{r})^2 = 1.37\ kW/m^2$$

which is known as the solar constant. Note that not all of this reaches the surface of the earth, approximately 30 percent of the sun's radiation is actually reflected by the atmosphere (*albedo*), so only 70 percent reach the surface.

In thermal equilibrium, the earth should emit as much as it receives:

$$J_{in} = J_{out} = 4\sigma T_e^4$$

where the factor 4 represents the fact that while the sun is so far away that it appears to be an emitter with parallel rays to us, the earth radiates in all directions, see the calculation above between the factor 4 difference between radiation in all directions and in one direction only. We conclude

$$T_e = T_s (\frac{R_s}{2r})^{1/2} = 280K = 7°C$$

If we take the albedo into account, this value gets even worse:

$$T_e = T_s 0.7^{1/4} (\frac{R_s}{2r})^{1/2} = 256K = -17°C$$

Thus everything should be frozen and life was not possible on earth.

The explanation for the discrepancy to the observed temperature is the famous *greenhouse effect*. Because the earth has a temperature of the order of 300 K, it emits in the infrared according to Planck (see above). While being relatively transparent in the visible and UV range, water vapor, CO_2 and CH_4 in the atmosphere strongly absorb in the infrared. Therefore they let radiation from the sun in but absorb radiation from the earth (ozone O_3 absorbs in the UV and therefore protects us from excessive DNA-damage). 40% of the outgoing radiation is absorbed, so only 60% is radiated away. A better estimate therefore is

$$T_e = T_s \frac{0.7^{1/4}}{0.6^{1/4}} (\frac{R_s}{2r})^{1/2} = 290K = 17°C$$

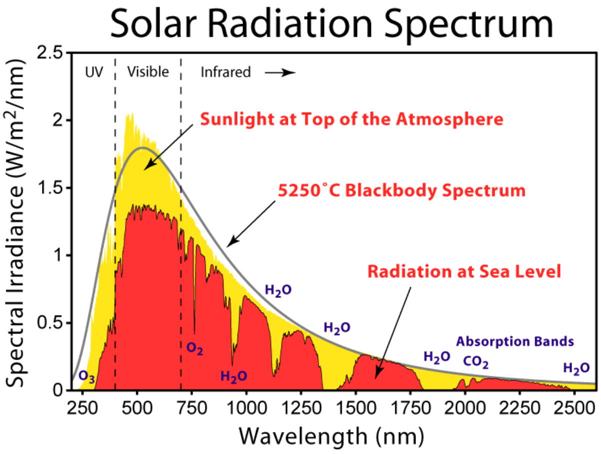

Figure 3.22: The effect of different atmospheric gases on the radiation coming from the sun, which has a perfect Planck spectrum before it hits the earth (source: wikipedia).

which agrees well with our observations.

Without any atmosphere, the earth would be around 30 degrees Celsius colder. Without clouds, it would be colder by 10-15 degrees. Interestingly, the power radiated by the sun was 25% weaker when life started on earth 4.5 billion years ago. Thus it is hard to understand why the earth was sufficiently warm to allow for life to evolve (*faint young sun problem*). However, while the sun got stronger, the amount of CO_2 in the atmosphere decreased because life started to generate oxygen (especially during the *great oxidation event* around 2 billion years ago) and luckily the two effects balanced each other such that a relatively constant average temperature of 15 degrees Celsius and a pressure of 1 atm was maintained on earth.

Today, the amount of CO_2 increases again because we burn so much coal and gas, thus freeing carbon stored in the earth. Other reasons for the increasing CO_2 levels are the destruction of forest and increased land use. In 2016 average atmospheric CO_2 concentrations have surpassed 400 parts per million, a level that has been last reached 3 million years ago, when temperatures were $2-3°C$ warmer and sea levels 10-20 meters higher; even more recently this concentration has in fact shown in increasing rate of change, which is very alarming.

On Venus, which has a similar size and composition as the earth, but is closer to the sun and therefore always had higher temperatures, no ocean formed and the greenhouse effect with too much CO_2 in the atmosphere led to very high temperatures. Life is not possible there with 470 degrees Celsius and a pressure of 90 atm. This scenario can not happen on earth, where we started with lower temperatures and developed oceans that now protect us from a run-away climate: if it becomes hotter, water evaporates, acts as green house gas and cools, thus counteracting the increase in temperature (negative feedback). However, many of our subearth systems (including the ice sheets on Green-land and West Antarctica, the Atlantic thermohaline circulation, the Arctic sea ice, the alpine glaciers and the Arctic ozone depletion) are currently in the process of passing tipping points into instability[1]. For example, the snow of ice sheets and glaciers melts with increasing temperature, the brown earth below absorbs more radiation energy and local temperature rises even more (positive feedback).

Eruptions of volcanos release sulphur dioxide into the atmosphere that absorb sun light and therefore cool the earth. 74.000 years ago the eruption of Toba in Indonesia possibly triggered a new ice age. Only thousands of humans survived as evidenced from their small genetic diversity at that time. When Tambora erupted in Indonesia in 1815, the following year was without summer in Europe and the population suffered from hunger. In 1991, Pinatubo erupted in the Philippines and temperature went down by half a degree for two years. In principle these effects could be used for climate engineering, by spraying vapor or sulphur dioxide into the atmosphere, but this might be very difficult to control, as we all know from the netflix series *Snowpiercer*.

Finally some history on climate research. Joseph Fourier in 1824 published his "Remarques Generales Sur Les Temperatures Du Globe Terrestre Et Des Espaces Planetaires", in which he concluded that the atmosphere must somehow isolate the earth. In 1859 John Tyndall found that water vapor and CO_2, but not oxygen or nitrogen in the atmosphere absorb and gives away heat. Climate research finally started fully in 1896 when Svante Arrhenius (Nobel Price 1903 for Physical Chemistry) published his treatise "On the Influence of Carbonic Acid in the Air Upon the Temperature of the Ground". He even calculated that temperature increases logarithmically with CO_2 concentration in the atmosphere, correctly predicting that a doubling leads to in increase of temperature of a few degrees. He already concluded that us burning fossil fuel will lead to a warmer earth, but thought that this will be favorable for agriculture, because he could not anticipate the massive scale on which we are doing this today.

---

[1] For a discussion of tipping points in climate compare Lenton, Timothy M., et al. "Tipping elements in the Earth's climate system." Proceedings of the national Academy of Sciences 105.6 (2008): 1786-1793; and Levermann, Anders, et al. "Potential climatic transitions with profound impact on Europe." Climatic Change 110.3 (2012): 845-878.

4 The grandcanonical ensemble

4.1 Probability distribution

We now consider a reservoir for both energy and particle number. In this *'grandcanonical ensemble'* the equilibrium conditions are $T_1 = T_2$ and $\mu_1 = \mu_2$. Thus now temperature T and chemical potential μ are fixed. An example would be a open bottle lying on the ground of a lake. Now not only heat, but also particle number can be exchanged with the environment.

The derivation of the corresponding probability distribution is similar as for the canonical ensemble. We consider one specific microstate i that then comes with an energy E_i and a particle number N_i. Then its probability to occur is

$$p_i = \frac{\text{\# favorable outcomes}}{\text{\# possible outcomes}} = \frac{\Omega_{res}(E_{tot} - E_i, N_{tot} - N_i)}{\Omega_{tot}(E_{tot}, N_{tot})} = \frac{e^{S_{res}(E_{tot}-E_i,N_{tot}-N_i)/k_B}}{e^{S_{tot}(E_{tot},N_{tot})/k_B}}$$

We next introduce the average energy E and the average particle number N as reference values and Taylor-expand the entropy of the reservoir:

$$S_{res}(E_{tot} - E + E - E_i, N_{tot} - N + N - N_i) = S_{res}(E_{tot} - E, N_{tot} - N) + \frac{E - E_i}{T} - \frac{\mu(N - N_i)}{T}$$

Here we have used $\partial S/\partial E = 1/T$ and $\partial S/\partial N = -\mu/T$. Note that higher order terms do not appear for a reservoir. We also can use additivity of the entropy to write

$$S_{tot}(E_{tot}, N_{tot}) = S(E, N) + S_{res}(E_{tot} - E, N_{tot} - N)$$

Together we therefore get

$$p_i = e^{\beta\Psi} e^{-\beta E_i + \beta\mu N_i}$$

where

$$\Psi = E - TS - \mu N$$

Thus the overall structure of the result is similar as for the canonical ensemble: there is a Boltzmann-factor $e^{-\beta E_i + \beta\mu N_i}$ for the weight of the microstate as a function of its energy E_i and its particle number N_i, and there is a normalization factor $e^{\beta\Psi}$.

We can make this derivation shorter if we make the normalization only at the end. In this case we do not even have to specify around which reference values we expand. We simply write

$$p_i \propto e^{S_{res}(E_{tot}-E_i,N_{tot}-N_i)/k_B} \propto e^{-\beta E_i + \beta\mu N_i}$$

where now the expansion could also be around E_{tot}, N_{tot} for simplicity. We therefore get

$$\boxed{p_i = \frac{1}{Z_G}\, e^{-\beta(E_i - \mu N_i)}} \quad \text{grandcanonical prob. distribution}$$

where

$$\boxed{Z_G = \sum_i e^{-\beta(E_i - \mu N_i)}} \quad \text{grandcanonical partition sum}$$

Note that we could have derived the same result through information theory by maximizing Shannon entropy with the two constraints of given average energy and given average particle number. From above we also can conclude that

$$\boxed{\Psi = -k_B T \ln Z_G}$$

must be the relevant thermodynamic potential.
Obviously our result is closely related to the canonical result. To make this clearer, we write $i = (N, j)$, where j are all quantum numbers besides N:

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} \left(\sum_j e^{-\beta E_j} \right) e^{\beta \mu N} \\ &= \sum_{N=0}^{\infty} Z(T,V,N)\, e^{\beta \mu N} \end{aligned}$$

Thus Z_G follows by another summation from the canonical partition sum Z.

4.2 Grandcanonical potential

The probability to have a macroscopic value (E, N) is:

$$\begin{aligned} p(E,N) &= \frac{1}{Z_G}\, \Omega(E,N)\, e^{-\beta(E-\mu N)} \\ &= \frac{1}{Z_G}\, e^{-\beta(E - TS - \mu N)} = \frac{1}{Z_G}\, e^{-\beta\, \Psi(T,V,\mu)} \end{aligned}$$

Here we defined the *'grandcanonical potential'* that also appeared above: $\Psi(T,V,\mu) := E - TS - \mu N$.

p is maximal, if Ψ is minimal. We note:

microcanonical ensemble: S *maximal*

canonical ensemble: F *minimal*

grandcanonical ensemble: Ψ *minimal*

$\Psi = E[T,\mu]$ is the two-dimensional Legendre transform of energy from the extensive variables (S, N) to the intensive variables (T,μ).

Total differential:

$$\begin{aligned}
\mathrm{d}\Psi &= \mathrm{d}(E - TS - \mu N) \\
&= T\mathrm{d}S - p\mathrm{d}V + \mu\mathrm{d}N - \mathrm{d}(TS + \mu N) \\
&= -S\mathrm{d}T - p\mathrm{d}V - N\mathrm{d}\mu
\end{aligned}$$

Equations of state:

$$S = -\frac{\partial\Psi}{\partial T},\ p = -\frac{\partial\Psi}{\partial V},\ N = -\frac{\partial\Psi}{\partial\mu}$$

4.3 Fluctuations

Average and variance of the internal energy E can be calculated as before for the canonical ensemble. We now calculate average and variance of the particle number N:

$$\begin{aligned}
\langle N\rangle &= \sum_i p_i\, N_i = \frac{1}{Z_G}\ \sum_i N_i\, e^{-\beta(E_i - \mu N_i)} \\
&= \frac{1}{Z_G}\ \frac{1}{\beta}\ \partial_\mu \sum_i\ e^{-\beta(E_i - \mu N_i)} \\
&= \frac{1}{\beta}\ \partial_\mu \ln Z_G
\end{aligned}$$

$$\begin{aligned}
\sigma_N^2 &= \langle N^2\rangle - \langle N\rangle^2 \\
&= \frac{1}{\beta^2}\ \frac{1}{Z_G}\ \partial_\mu^2 Z_G - \left(\frac{1}{\beta}\frac{1}{Z_G}\ \partial_\mu Z_G\right)^2 \\
&= \frac{1}{\beta}\ \partial_\mu \langle N\rangle = \frac{1}{\beta^2}\ \partial_\mu^2 \ln Z_G
\end{aligned}$$

One can show with thermodynamic considerations (Maxwell relations):

$$\begin{aligned}
\left.\frac{\partial N}{\partial\mu}\right|_{T,V} &= -\frac{N^3}{V^2}\ \left.\frac{\partial v}{\partial p}\right|_{N,T} \quad \left(v = \frac{V}{N}\right) \\
\Rightarrow \quad \kappa_T &:= -\frac{N}{V}\ \left.\frac{\partial v}{\partial p}\right|_{N,T} \\
&= \frac{V}{N^2 k_B T}\ \sigma_N^2 \geq 0 \quad \text{as } \sigma^2 \geq 0
\end{aligned}$$

Here we defined the *'isothermal compressibility'* κ_T.

Like for the canonical ensemble, where we found

$$c_v = \frac{\sigma_E^2}{k_B T^2} \geq 0$$

again the variance is related to a material property:

$$\sigma_N^2 = \frac{N^2 k_B T \kappa_T}{V} \quad \overset{\kappa_T \text{ intensive}}{\rightarrow} \quad \sigma_N^2 \propto N$$

$$\Rightarrow \quad \boxed{\frac{\sigma_N}{\langle N \rangle} \propto \frac{1}{N^{\frac{1}{2}}}}$$

Particle fluctuations are small for large systems as expected based on the law of large numbers.
Note that calculation of the average enery is a bit more tricky now than in the canonical ensemble because there are two terms related to β. If we want to get only the energy part, we have to keep the chemical potential part constant:

$$\langle E \rangle = - \left. \frac{\partial \ln Z_G}{\partial \beta} \right|_{\beta\mu}$$

A simpler solution is to do the derivative for β without constraint and then to subtract the part which is not wanted:

$$\langle E \rangle = -\partial_\beta \ln Z_G + \mu \langle N \rangle = -\partial_\beta \ln Z_G + \frac{\mu}{\beta} \partial_\mu \ln Z_G$$

4.4 Ideal gas

For the canonical partition sum we had:

$$Z(T,V,N) = \frac{1}{N!} \left(\frac{V}{\lambda^3} \right)^N \quad \text{with } \lambda = \frac{h}{(2\pi m k_B T)^{\frac{1}{2}}}$$

$$\begin{aligned} Z_G &= \sum_{N=0}^{\infty} Z(T,V,N)\, e^{\beta\mu N} = \sum_{N=0}^{\infty} \frac{1}{N!} \left(e^{\beta\mu} \frac{V}{\lambda^3} \right)^N \\ &= e^{z \frac{V}{\lambda^3}} \quad \text{with } \textit{fugacity } z := e^{\beta\mu} \\ \Rightarrow \quad \langle N \rangle &= \frac{1}{\beta}\, \partial_\mu \ln Z_G \\ &= \frac{1}{\beta}\, \partial_\mu \left(\frac{zV}{\lambda^3} \right) = \frac{V}{\lambda^3}\, e^{\beta\mu} \\ \mu &= k_B T \, \ln \left(\frac{N\lambda^3}{V} \right) \end{aligned}$$

This result is identical with the one we derived from the microcanonical ensemble. The variance then becomes:

$$\begin{aligned}
\sigma_N^2 &= \frac{1}{\beta}\,\partial_\mu \langle N \rangle = \frac{1}{\beta}\,\partial_\mu \left(\frac{e^{\beta\mu}V}{\lambda^3}\right) = N \\
\kappa_T &= \frac{VN}{N^2 k_B T} = \frac{1}{p} \quad \text{which also follows from } pV = Nk_BT
\end{aligned}$$

4.5 Molecular adsorption onto a surface

We consider a gas in contact with a solid surface (e.g. argon on graphene or molecular nitrogen on iron, as in the Haber-Bosch synthesis). The gas molecules can be adsorbed at N specific adsorption sites while one site can only bind one molecule. The energies of the bound and unbound state are ϵ and 0, respectively. ϵ can be negative or positive. The gas acts as a reservoir fixing T and μ.
The partition sum is:

$$Z_G = (z_G)^N$$

where we considered a non-interacting system with single particle partition sum:

$$z_G = 1 + e^{-\beta(\epsilon-\mu)}$$

The mean number of absorbed particles per site follows as:

$$\langle n \rangle = \frac{0 + 1 \cdot e^{-\beta(\epsilon-\mu)}}{z_G} = \frac{1}{e^{-\beta(\mu-\epsilon)} + 1}$$

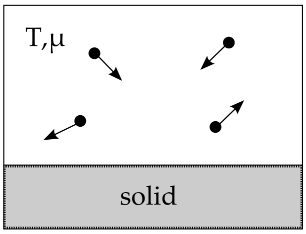

Figure 4.1: A gas at (T,μ) in contact with a solid surface. The solid contains N specific adsorption sites each able to bind one or none gas molecule.

The mean energy per site is:

$$\langle \epsilon \rangle = \frac{0 + \epsilon \cdot e^{-\beta(\epsilon-\mu)}}{z_G} = \epsilon\ \langle n \rangle$$

This model can be easily generalized to more complicated situations, e.g. if more than one molecule can be adsorbed.

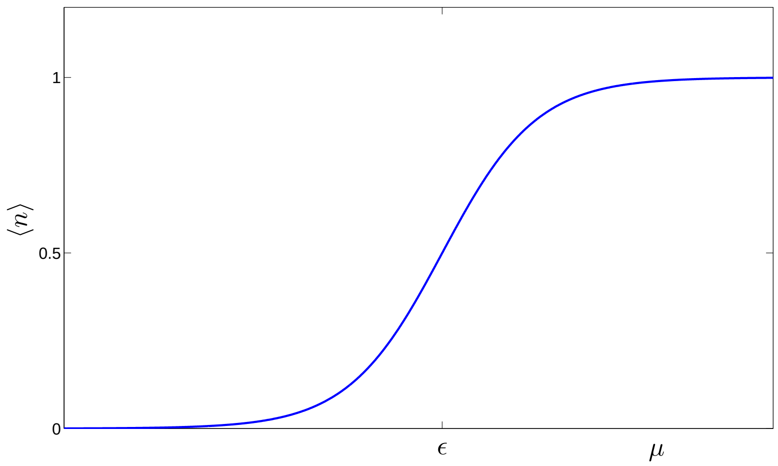

Figure 4.2: The mean number of absorbed particles $\langle n \rangle$ as a function of the chemical potential μ. Half-occupancy occurs at $\mu = \epsilon$.

If the adsorbate is an ideal gas:

$$\mu = k_B T \ln \frac{p}{p_0}$$

$$\Rightarrow \quad e^{\beta\mu} = \frac{p}{p_0}$$

$$\Rightarrow \quad \langle n \rangle = \frac{p}{p_0 e^{\beta\epsilon} + p}$$

The number of adsorbed particles first rises linearly with pressure (or, equivalently, with concentration) and then plateaus. This is known as the *Langmuir isotherm*.

4.6 Chemical reactions

The grandcanonical ensemble is also the appropriate framework to deal with chemical reactions: because particles number can change, the chemical potential becomes the central quantity. Chemical reactions are the basis of (bio)chemistry and the law of mass action is at the heart of any chemical reaction. An important example is the decomposition of water

$$H_2O \rightleftharpoons H_2 + \frac{1}{2}O_2$$

which is at the heart of a potential hydrogen economy, when this process might be fueled by regenerative energy and then the hydrogen might be stored in tanks or some type of porous material (to avoid explosions). The reverse process would release lots of

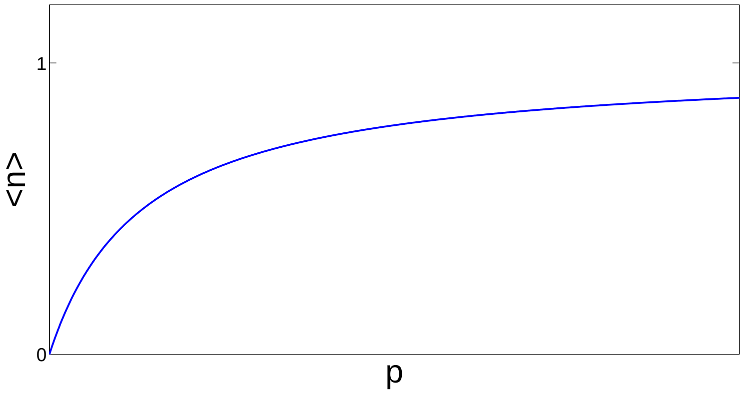

Figure 4.3: The mean number of absorbed (ideal gas) particles $\langle n \rangle$ as a function of pressure p.

energy, but only produce water, no carbon dioxide. Another important example is the production of ammonia (NH_3)

$$3\ H_2 + N_2 \rightleftharpoons 2\ NH_3$$

whose industrial importance cannot be underestimated, e.g. for the production of fertilizer or explosives. Actually the invention of the industrial process to produce ammonia by Fritz Haber and Robert Bosch earned both of them a Nobel Prize in chemistry (in 1918 and 1931, respectively); moreover in 2007 Gerhard Ertl earned the Nobel Prize in chemistry because he explained the underlying molecular processes, in particular the role of iron as catalyst. Each year 180 million tons of ammonia are produced with the help of the Haber-Bosch process, which uses pressures of 200 bar and temperatures of 500 degrees Celsius. Today you can learn more about this history by visiting the Bosch museum at Heidelberg, close to the villa Bosch (now used by the Klaus Tschira Foundation), where Robert Bosch lived while running the chemical company BASF at Ludwigshafen. In front of the museum you can see the high pressure pipes invented by Robert Bosch. While N_2 is taken from the air, H_2 is obtained from natural gas (CH_4). This step leads to a lot of CO_2-production (1.5 tons of CO_2 for one ton of ammonia); as a matter of fact ammonia production contributes one percent to our greenhouse gas emissions.

Usually chemical reactions are carried out at constant temperature T and constant pressure p. Therefore the relevant TD-potential is the Gibbs free energy $G(T, p, N_i) = E - TS + pV$. The index $1 \leq i \leq r$ is used to number the r different components of the reaction mixture. The internal energy $E = TS - pV + \sum \mu_i N_i$. We therefore have the fundamental equation

$$dG = -SdT + Vdp + \sum \mu_i dN_i$$

At equilibrium, $dG = 0 = \sum \mu_i dN_i$ at constant T and p. Each chemical reaction of interest has a fixed ratio of how the mole numbers change (e.g. $-1 : +1 : +1/2$ for water decomposition and $-3 : -1 : +2$ for ammonia production). We therefore write

$$0 \rightleftharpoons \sum \nu_i A_i$$

where the v_i are the stoichiometric coefficients and the A_i are the symbols for the chemical components. We now note that

$$\frac{dN_i}{\nu_i} = const = dN$$

because the relative changes are fixed by the stoichiometric coefficients and therefore the requirement $dG = 0$ becomes

$$\boxed{\sum_i \mu_i \nu_i = 0}$$

The essential aspect here is that we have effectively only *one* reaction coordinate over which we can minimize G. Our result is completely general and allows to determine the mole numbers at equilibrium given the initial mole numbers, the relations $\mu_i(T, p, N_i)$, as well as T and p.

We now apply this result to a mixture of ideal gases. Before we do this, however, we make some general comments on these systems. Because the different species do not interact, for each of them we have for the internal energy

$$E_i = \frac{3}{2} N_i k_B T \Rightarrow E = \sum E_i = \frac{3}{2} (\sum N_i) k_B T = \frac{3}{2} N k_B T$$

where for simplicity we have used the factor $3/2$ for a monoatomic ideal gas. We conclude that the internal energy is the same as the one for an ideal gas of N molecules.

This however is different for entropy. We next turn to entropy which for one species can be written as

$$\begin{aligned} S_i &= N_i s_{i0} + N_i k_B \ln \left[\left(\frac{T}{T_0} \right)^{3/2} \frac{V}{V_0} \frac{N_0}{N_i} \right] \\ \Rightarrow S = \sum S_i &= \sum N_i s_{i0} + (\sum N_i) \frac{3}{2} k_B \ln \left(\frac{T}{T_0} \right) + \sum N_i k_B \ln \left(\frac{V}{V_0} \frac{N_0}{N_i} \frac{N}{N} \right) \\ &= \sum N_i s_{i0} + \frac{3}{2} N k_B \ln \left(\frac{T}{T_0} \right) + N k_B \ln \left(\frac{V}{V_0} \frac{N_0}{N} \right) - N k_B \sum x_i \ln x_i \end{aligned}$$

where $x_i = N_i / N$ is the mole fraction of species i. The last term is the *entropy of mixing*. If a collection of separate gases each at the same temperature T and the same density $\rho = N/V$ (or, equivalently, at the same pressue P) is allowed to mix, then it will gain entropy because the mole fractions obey $0 \leq x_i \leq 1$ and therefore the entropy gain will be positive.

We are now ready to discuss the chemical reactions. From the discussion of mixtures, we understand that the chemical potentials are simply the ones of the separate gases:

$$\mu_i = k_B T \ln \left(\frac{N_i \lambda^3}{V} \right)$$

where the thermal wavelength $\lambda \sim T^{-1/2}$. We rewrite this equation as

$$\mu_i = k_B T \ln \left(\frac{N_i \lambda^3}{V} \frac{N}{N} \right) = k_B T \left[\phi_i(T) + \ln p + \ln x_i \right]$$

using the ideal gas law. Here $\phi_i(T)$ is a function of T only. We insert these relations into the general result $\sum \mu_i \nu_i = 0$:

$$\sum \nu_i \ln x_i = -(\sum \nu_i) \ln p - \sum \nu_i \phi_i(T)$$

We define the last term as $\ln K(T)$ where $K(T)$ is called the *equilibrium constant*. After exponentiation we then have

$$\boxed{\prod x_i^{\nu_i} = p^{-(\Sigma \nu_i)} K(T)} = K(p, T)$$

This is the famous *law of mass action* (*Massenwirkungsgesetz* in German). On the left we have a product of mole fractions and on the right we have a constant that depends on T and p. If one wants to suppress the pressure-dependance, one can write $K(T, p)$ for the complete right hand side.

Examples

① Dissociation of water: as initial conditions we take 2 moles of water. We take $p = 1MPa$ and $T = 2000K$. Then the equilibrium constant $K(T) = 0.088 Pa^{1/2}$. The law of mass action now reads

$$\frac{x_{H_2} x_{O_2}^{1/2}}{x_{H_2O}} = p^{-1/2} K(T)$$

To solve this scalar equation, we have to introduce a reaction coordinate ΔN. Then

$$N_{H_2O} = 2 - \Delta N, N_{H_2} = \Delta N, N_{O_2} = \Delta N/2$$

and thus

$$x_{H_2O} = \frac{2 - \Delta N}{2 + \Delta N/2}, x_{H_2} = \frac{\Delta N}{2 + \Delta N/2}, x_{O_2} = \frac{\Delta N/2}{2 + \Delta N/2}$$

If we insert this into the law of mass action, we get a polynomial equation that we cannot solve easily:

$$\frac{(\Delta N)^{3/2}}{\sqrt{2}(2 - \Delta N)(2 + \Delta N/2)^{1/2}} = p^{-1/2} K(T)$$

A numerical solution however is easy to get:

$$x_{H_2O} = 0.996, x_{H_2} = 0.003, x_{O_2} = 0.001$$

Thus the reaction is very much on the side of the water. Decreasing pressure pushes it slightly away from there, but not much. In fact this reaction has $\Delta G_0 = +237 kJ/mol$ and therefore does not occur spontaneously.

② Ammonia production: this reaction is in principle favorable and should work well ($\Delta G_0 = -10.9 kJ/mol$ at ambient conditions). High pressure favors the reaction because it has a negative volume balance. In practise, there is a high kinetic reaction barrier (not described by statistical physics, this needs a quantum mechanical calculation, the main problem is the high stability of N_2), therefore iron catalysts are required. Note that nature takes a similar route. Many soil-dwelling bacteria use the metalloenzyme nitrogenase containing iron atoms to convert N_2 from air into ammonia.

Mass action kinetics

In (bio)chemistry, one often wants to understand also the time-dependance of the chemical reaction. If the reaction is determined mainly by collisions, then this is easy. Consider the bimolecular reaction

$$A + B \rightleftharpoons C$$

with a forward rate constant k_+ (also called on-rate or association rate) and a backward rate constant k_- (also called off-rate or dissociation rate). Assuming a homogeneous mixture (no spatial effects), we write the ordinary differential equation

$$\frac{dA}{dt} = k_- C - k_+ A\ B$$

The gain term results from a simple "radioactive" decay and the loss term results from a collision between one A- and one B-molecule. At equilibrium, $dA/dt = 0$ and therefore

$$\frac{C_{eq}}{A_{eq} B_{eq}} = \frac{k_+}{k_-} = K_A = \frac{1}{K_D}$$

where we have defined the association constant K_A and its inverse, the dissociation constant K_D. Obviously we have recovered the law of mass action and $K_A = K(T,p)$ (except for the different dimensions, here we use concentrations and above we used mole fractions for the left hand side). Note that the dimensions of k_- and k_+ are $1/s$ and $1/(s\ mol)$, respectively, such that the dimension of K_D is mol.
We next note that $A + C = const = A_0$ and therefore we have

$$K_D = \frac{(A_0 - C_{eq}) B_{eq}}{C_{eq}} \Rightarrow C_{eq} = A_0 \frac{B_{eq}}{K_D + B_{eq}}$$

Interestingly, this law has the same form as the Langmuir isotherm discussed above (number of adsorbed particles as a function of pressure). We conclude that we reach half binding ($C_{eq} = A_0/2$) if $B_{eq} = K_D$. Therefore K_D is a measure for how strong the two partners react: the smaller K_D, the weaker dissociation compared to association and the less B_{eq} is required to achieve the same amount of binding (high affinity).
It is important to note that the law of mass action only makes a equilibrium statement and that mass action kinetics is only valid if the reaction is determined mainly by collisions. In fact this argument might not work at all if the reaction in addition has to

cross some transition state barrier which slows down the time for reaction. The law of mass action connecting the initial and final states would still be valid, but the time for this process to occur might be very large. In such cases, one typically looks for catalysts to speed up the reaction (e.g. iron in the case of the famous Haber-Bosch-synthesis of ammonia, which in addition uses very high pressure; living systems have evolved enzymes for this purpose).

A small value of K_D also means that there is a large gain in Gibbs free energy G during the course of the reaction. For one mole of ideal gas, we have

$$\begin{aligned} G_i &= G_{i0}(T,P) + RT\ln x_i \\ \Rightarrow \Delta G &= \sum \nu_i G_i = \Delta G_0 + RT\ln\left(\prod x_i^{\nu_i}\right) = \Delta G_0 + RT\ln K \end{aligned}$$

At equilibrium, $\Delta G = 0$ and therefore

$$\Delta G_0 = RT\ln K_D \, , K_D = e^{\Delta G_0/RT}$$

The more negative ΔG_0, the smaller K_D and the stronger the reaction is driven to the right.

Living systems have singled out a few central biochemical reactions that have a very high gain in Gibbs free energy and therefore those are used to drive other reactions. The two most important examples are:

① oxidation of food (glucose):

$$C_6H_{12}O_6 + 6O_2 \rightleftharpoons 6CO_2 + 6H_2O$$

has $\Delta G_0 = -2890 kJ/mol$. This is the gradient which essentially drives our metabolism. If you use divide by Avogadro's number and the usual value for k_BT, you see that this free energy gain is around 1000 k_BT per molecule, which is huge and only possible because glycose is a complicated molecule with many bonds. In fact metabolism uses many enzymes to completely digest glycose (part of this is the famous Krebs cycle) and to get as much free energy out of it as possible.

② ATP-hydrolysis:

$$ATP \rightleftharpoons ADP + P_i$$

has $\Delta G_0 = -35 kJ/mol$. Per molecule, this is around 10 k_BT, which is a large amount for such a small molecule in which basically only one bond is cleaved. As this was not enough, nature makes sure that this reaction is very favorable by additionally keeping the relevant concentrations out of equilibrium: with $ATP = P_i = mM$ and $ADP = 10\mu M$, we have

$$\Delta G = \Delta G_0 + RT\ln\frac{ADP\ P_i}{ATP\ 1M} = -60 kJ/mol$$

thus the reaction becomes even more favorable (here $1M$ is used as reference concentration). ATP-hydrolysis drives many processes in our cells, including the movement of our muscle.

5 Quantum fluids

5.1 Fermions versus bosons

For multi-particle systems, experiments have shown that the rules of single-particle quantum mechanics have to be extended by one more principle (generalized Pauli principle):

> *Particles with half-integer (integer) spin are called fermions (bosons). Their total wave function (space and spin) must be antisymmetric (symmetric) under the exchange of any pair of identical particles.*

Comments

① For electrons (spin $= \frac{1}{2}$, fermions), we get the Pauli principle (antisymmetric wave function) and the Pauli exclusion principle (no two electrons can occupy the same state). If two electrons have opposite spins (singlet), their spin wave function is antisymmetric ($\frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle - |\downarrow\uparrow\rangle)$) and the spatial wave function can be symmetric. If two electrons have the same spin (triplet), their spin wave function is symmetric ($|\uparrow\uparrow\rangle, |\downarrow\downarrow\rangle, \frac{1}{\sqrt{2}}(|\uparrow\downarrow\rangle + |\downarrow\uparrow\rangle)$) and the spatial wave function hence has to be antisymmetric:

$$\Phi(1,2) = a(1)b(2) - a(2)b(1)$$

Since $a = b$ for identical particles and thus $\Phi(1,2) = 0$, two electrons can not be at the same position. This is the basis of atomic stability. Note that this argument does not really need positions, but just position-dependent wave functions, so it is not contradicted by Heisenberg's uncertainty principle that we cannot exactly know the positions of the particles.

② Fermions are the quantum mechanical analogues of classical particles (matter can not penetrate itself). Bosons, in contrast, are the quantum mechanical analogue of classical waves, which can be superimposed freely and thus transfer lots of information. The standard example are photons which can be freely created and annihilated and which are the exchange particles of the electromagnetic interaction. Thus the dichotomy between fermions and bosons is not really new in quantum physics, but also present in classical physics.

③ The fact that there are two possible kinds of particles comes from the fact that two permutations lead back to the initial situation, so there can be a sign in the wavefunction ($-1 \times -1 = 1$) or not ($1 \times 1 = 1$). The relation between this property and the spin is described by the spin-statistics theorem from relativistic quantum theory, thus it is a genuinely relativistic effect.

Consequences for counting

We consider a system of two identical particles, each of which can exist in either of two states.

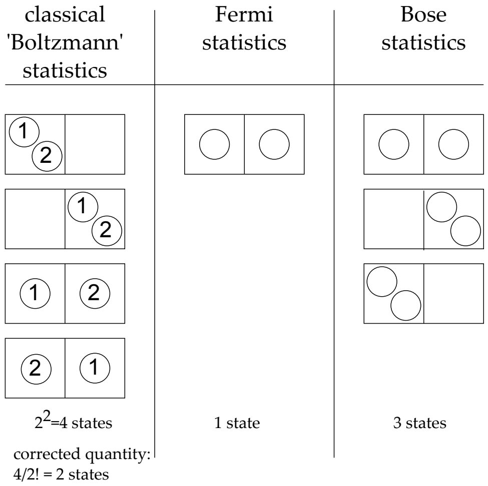

Figure 5.1: Table of states for a system of two identical particles from a point of view of different statistics. Each particle can be either in state one or two.

We see from Figure 5.1 that all cases differ in their outcome for the number of states: four for classical counting, two for corrected classical counting, one for fermions and three for bosons. Correct are the results for fermions and bosons. Both ‘classical’ and ‘corrected classical’ counting are incorrect. The rules of quantum mechanics imply more than simply saying that identical particles cannot be distinguished. They actually mean that we have to do the calculation separately for fermions and bosons. However, at high temperature the particles are distributed over all states due to entropy and the probability of two particles being in one state becomes very small. As one can see from Figure 5.1, in this limit, fermionic, bosonic and corrected classical counting all give exactly one state. All quantum fluids become ideal gases at high temperature.

Canonical ensemble

We now formalize the differences in terms of partition sums. In the canoncial ensemble, we consider two particles that are distributed over two states with energies 0 and ϵ (two-state system) as shown in Figure 5.1. For Fermi-Dirac statistics we have only one state and the partition sum is

$$Z_F = e^{-\beta\epsilon}$$

For Bose-Einstein statistics, we have three different states:

$$Z_B = 1 + e^{-\beta\epsilon} + e^{-2\beta\epsilon}$$

For classical Maxwell-Boltzmann statistics, we have four states, where the two mixed ones have the same energies. In addition we have the factor 2! for corrected counting:

$$Z_M = \frac{1 + 2e^{-\beta\epsilon} + e^{-2\beta\epsilon}}{2}$$

One sees that each kind of statistics leads to a different result. However, at high temperatures only the mixed case is relevant and then all three statistics give the same result.

Grand canonical ensemble

Quantum fluids are best treated in the grand canonical formalism. Thus we fix T and μ. For a non-interacting system, the partition sum for a single energy state for a fermion (only two states in regard two occupancy) is:

$$z_F = 1 + e^{-\beta(\epsilon-\mu)}$$

We have called this quantity the *single particle partition sum* before, but note that in the grandcanoncial ensemble, the number of particles is not fixed, so we might also call it the *single unit partition sum* or *single state partition sum*.
The average occupation number n_F becomes:

$$\Rightarrow \quad n_F = \frac{0 \times 1 + 1 \times e^{-\beta(\epsilon-\mu)}}{1 + e^{-\beta(\epsilon-\mu)}} = \frac{1}{\beta}\,\partial_\mu \ln z_F = \frac{e^{-\beta(\epsilon-\mu)}}{1 + e^{-\beta(\epsilon-\mu)}}$$

$$\Rightarrow \boxed{n_F = \frac{1}{e^{\beta(\epsilon-\mu)} + 1}} \quad \text{Fermi function}$$

$$n_F(\epsilon = \mu, T) = \frac{1}{2}$$

$n_F(\epsilon, T)$ (compare Figure 5.2) is symmetric under inversion through the point $\epsilon = \mu\,, n_F = \frac{1}{2}$. For $T \to 0$, the Fermi function approaches a step function:

$$n_F = \begin{Bmatrix} 1 & \epsilon < \mu \\ 0 & \epsilon \geq \mu \end{Bmatrix} = \Theta(\mu - \epsilon)$$

Here Θ is the Heaviside step function. For vanishing temperature, the fermions fill up all energy states up to the chemical potential μ.
For bosons we have:

$$\begin{aligned} z_B &= 1 + e^{-\beta(\epsilon-\mu)} + e^{-\beta(2\epsilon-2\mu)} + \ ... \\ &= \sum_{n=0}^{\infty} \left(e^{-\beta(\epsilon-\mu)}\right)^n = \frac{1}{1 - e^{-\beta(\epsilon-\mu)}} \end{aligned}$$

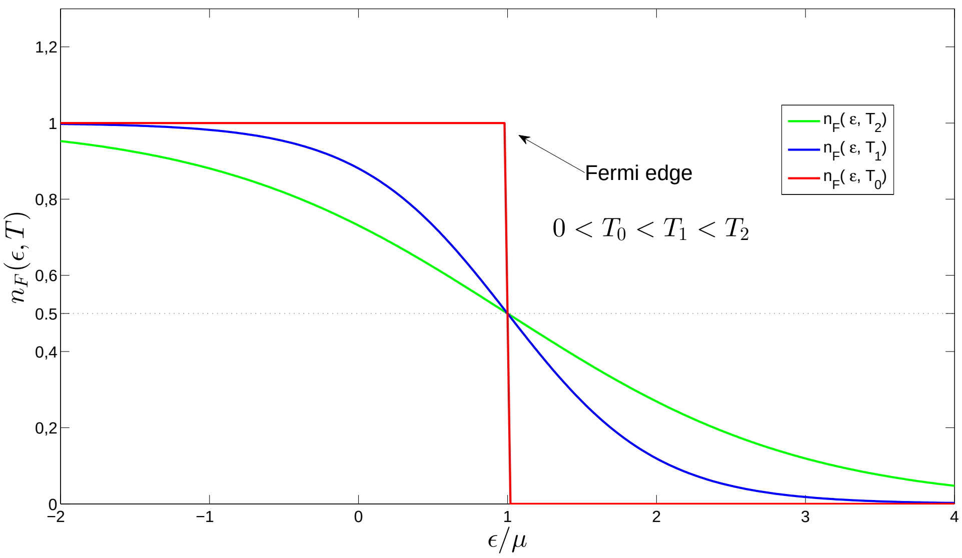

Figure 5.2: The Fermi function as a function of ϵ/μ for different temperature values. For $T_0 \to 0$ the function becomes step-like with the *'Fermi edge'* at $\epsilon = \mu$. At this point the probability for one state to be occupied is $\frac{1}{2}$. μ at $T = 0$ is called the *'Fermi energy'*.

The average occupation number is:

$$n_B = \frac{e^{-\beta(\epsilon-\mu)} + 2e^{-\beta(2\epsilon-2\mu)} + \dots}{z_B} = \frac{1}{\beta}\,\partial_\mu \ln z_B = \frac{e^{-\beta(\epsilon-\mu)}}{1 - e^{-\beta(\epsilon-\mu)}}$$

$$\Rightarrow \boxed{n_B = \frac{1}{e^{\beta(\epsilon-\mu)} - 1}} \quad \text{Bose distribution}$$

The result differs only by a minus sign from the fermion result, but this difference has dramatic consequences because this distribution diverges at $\epsilon = \mu$ (Figure 5.3). For $T \to 0$, the Bose-Einstein function vanishes everywhere except for the lowest energy state.

In summary we have found a dramatic difference between fermions and bosons at low temperature:

- ⇒ Fermions tend to fill up energy states one after the other (horizontal stacking).
- ⇒ Bosons tend to condense all into the same low energy state (vertical stacking).

We finally compare the three statistics in form of the occupation number (Figure 5.4). If we define $x := \beta(\epsilon - \mu)$, then Fermi, Bose and classical statistics correspond to distri-

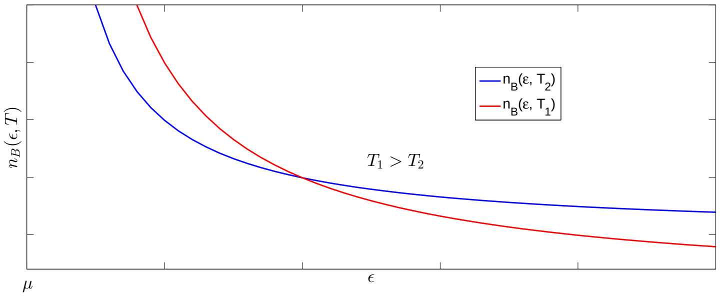

Figure 5.3: These curves are schematic with $T_1 > T_2$ and $\mu_1 < \mu_2$. They diverge as $\epsilon \to \mu$. For a system with a finite number of particles, we therefore must have $\mu \leq \epsilon$. At low temperature, all particles accumulate in the state of lowest energy.

bution functions $n_F = 1/(e^x + 1)$, $n_B = 1/(e^x - 1)$ and $n_M = 1/e^x$, respectively. Obviously they all agree with each other for large x, which is the classical limit in which fugacity $z = e^{\beta\mu} = \lambda^3/v \ll 1$ (using the result for the ideal gas) and the chemical potential becomes very negative, meaning that density is low (specific volume is large), the wavefunctions do not overlap and there are no quantum effects.

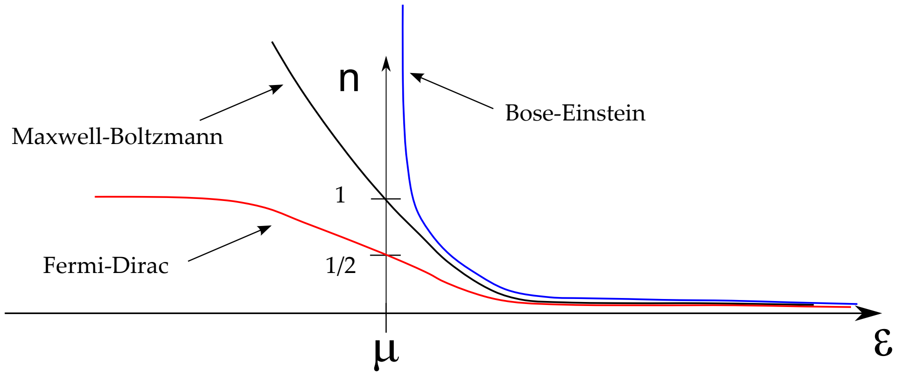

Figure 5.4: A schematic sketch of the occupation number n for the three statistics: Bose-Einstein (blue), Fermi-Dirac (red) and Maxwell-Boltzmann (black).

Examples of physical fluids

① Fermi fluids

- a) electrons in a metal or in a white dwarf
- b) neutrons in the atomic nuclei or in neutron stars

c) helium three (^3He: 2 protons, 1 neutron, 2 electrons, half-integer spin)

② Bose fluids

a) photons (black body radiation)

b) quasi-particles like phonons (lattice vibrations), polarons (polarization waves), magnons (spin waves)

c) helium-four (^4He: 2 protons, 2 neutrons, 2 electrons, integer spin)
^4He becomes *'superfluid'* below the critical temperature $T_{crit} = 2.17\ K$.

d) BCS-theory: Two electrons can form a bosonic pair in certain materials due to the interaction with the crystal lattice of ions. These become *'superconducting'* below T_{crit}.

e) ^{87}Rb-atoms are bosons. In 1995 rubidium atoms have been used to realize a *'Bose-Einstein condensate'* in atomic traps at very low temperatures ($T_{crit} \approx 10^{-7}\ K$).

5.2 Calculating with occupation numbers

Above we have presented simple arguments that counting has to be very different for fermions versus bosons. If we now look into the differences in the formalism, we will see that it is mainly in an innocently looking and strategically placed minus sign, but that this minus sign has dramatic consequences. For non-interacting, non-relativistic particles of mass m in a volume V, we only have kinetic energy

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m}$$

From solving Schrödinger's equation, the momentum vector has the shape

$$\vec{p} = \hbar\vec{k} = \frac{2\pi\hbar}{L}(n_1, n_2, n_3)$$

where $n_i \in \mathbb{Z}$. In addition each particle can have a spin quantum number m_S. For spin S and without an external magnetic field, this number leads to a degeneracy $g_S = 2S+1$. For example, for an electron we have $S = 1/2$, $g_S = 2$ and $m_S = \pm 1/2$. The wavefunction for the overall system now follows as

$$|\Psi\rangle = N \sum_P (\pm 1)^{|P|} P \prod_{i=1}^{N} |\vec{p}\rangle\, |m_{S,i}\rangle$$

where P are the permutations allowed for the system (symmetric for bosons, anti-symmetric for fermions) and N is some normalization factor.

In the next and decisive step, we switch from this wavefunction picture to a picture in which we only consider occupation numbers $n_{\vec{p}s}$. Then the equations for particle number and energy look very simple:

$$N = \sum_{\vec{p}} \sum_{m_S} n_{\vec{p}s} \,, \quad E = \sum_{\vec{p}} \sum_{m_S} n_{\vec{p}s} \epsilon_{\vec{p}}$$

with $\epsilon_{\vec{p}} = p^2/(2m)$. The grandcanonical partition sum follows as

$$Z_G = \sum_{0}^{\infty} \sum_{\{n_{\vec{p}s}\}, N\text{fixed}} e^{-\beta E + \mu N} = \sum_{\{n_{\vec{p}s}\}} e^{-\beta \sum_{\vec{p},m_s} (\epsilon_p - \mu) n_{\vec{p}s}} = \prod_{\vec{p},m_s} \sum_{n_{\vec{p}s}} e^{-\beta(\epsilon_p - \mu) n_{\vec{p}s}}$$

For fermions, we have only two possible occuptation numbers and thus

$$Z_G = \prod_{\vec{p},m_s} (1 + e^{-\beta(\epsilon_p - \mu)})$$

For bosons, we get the geometrical sum as above

$$Z_G = \prod_{\vec{p},m_s} \frac{1}{(1 - e^{-\beta(\epsilon_p - \mu)})}$$

The essence of these manipulations is that particle number N drops out. Later we can use μ to calculate back to a desired N. The grandcanonical potential now reads

$$\Psi = -k_B T \ln Z_G = \mp k_B T \sum_{\vec{p},m_s} \ln(1 \pm e^{-\beta(\epsilon_p - \mu)})$$

From here we get the average particle number N and the internal energy E by partial derivatives. For the particle number in particular we have

$$N = \sum_{\vec{p},m_s} \frac{1}{e^{\beta(\epsilon_p - \mu)} \pm 1}$$

Thus the essential difference here is indeed a minus sign: plus for fermioins and minus for bosons.

5.3 The photon gas

Our first application for quantum fluids is the photon gas, which we already discussed as black body radiation in the chaper on the canonical ensemble. In this case we have a linear dispersion relation: $\epsilon = \hbar\omega = cp = \hbar ck$ and the chemical potential $\mu = 0$ because photons can be created and destroyed. The grand canonical potential can be calculated to be

$$\Psi = -k_B T \ln Z_G = -\frac{\pi^2 V}{45 (\hbar c)^3} (k_B T)^4 \,.$$

To calculate the average photon number, one has to do this calculation with a chemical potential, then get the desired result as

$$\langle N\rangle = \frac{1}{\beta}\,\partial_\mu \ln Z_G$$

and finally set $\mu = 0$. Alternatively, one can do the calculation as done before with the density of state. The result should be the same as the one obtained before:

$$\langle N\rangle = \frac{V}{\pi^2}(\frac{k_B T}{\hbar c})^3\Gamma(3)\zeta(3)$$

which also used $\mu = 0$ in the Bose function.

5.4 The ideal Fermi fluid

We consider N particles with mass m, that is we now go to the canonical ensemble, but our starting point are the results from above for the grandcanonical ensemble, because this is mathematically much more simple and equivalent. The spatial states are characterized by the wave vector $\vec{k}$ of the wave function. The spin of each particle can be up or down ($m_S = \pm\frac{1}{2}$). Considering no additional interaction apart from the intrinsic one due to the particles being fermions, this implies:

$$\begin{aligned} Z_G &= \prod_{\vec{k},m_S} z_{\vec{k},m_S} \\ z_{\vec{k},m_S} &= 1 + e^{-\beta(\epsilon_{\vec{k},m_S}-\mu)} \\ n_{\vec{k},m_S} &= \frac{1}{e^{\beta(\epsilon_{\vec{k},m_S}-\mu)}+1} \end{aligned}$$

For the dispersion relation, we have the classical relation for a massive particle combined with the de Broglie relation:

$$\epsilon_{\vec{k},m_S} = \frac{p^2}{2m} = \frac{\hbar^2 k^2}{2m}$$

In the following we will use a factor 2 for the spin degrees and replace the sums by integrals:

$$\begin{aligned} \sum_{\vec{k},m_S} \,... &= 2\,\frac{V}{h^3}\int \mathrm{d}\vec{p}\,... = 2\,\frac{V}{h^3}\int \mathrm{d}p p^2 4\pi\,... \\ &= 2\,\frac{V}{h^3}\int_0^\infty \mathrm{d}\epsilon\, 4\pi\, m^{\frac{3}{2}}\ (2\epsilon)^{\frac{1}{2}}\ ... \\ &= N\int_0^\infty \mathrm{d}\epsilon\ \underbrace{\frac{V}{2\pi^2 N}\left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}}\sqrt{\epsilon}}_{\text{density of states } D(\epsilon)}\ ... \end{aligned}$$

The concepts used here are the same ones as used before for the Debye solid and the black body radiation. While here $D \propto \sqrt{\epsilon}$, for the phonons and photons we had $D \propto \epsilon^2$ due to the linear dispersion relation.

Fermi energy

For given particle number N, the chemical potential μ has to be determined from:

$$N = \sum_{\vec{k},m_S} n_{\vec{k},m_S} = N \int_0^\infty \mathrm{d}\epsilon \; D(\epsilon) \; \underbrace{\frac{1}{e^{\beta(\epsilon-\mu)}+1}}_{n(\epsilon)}$$

We first consider the limit $T \to 0$:

$$n(\epsilon) \to 1 - \Theta(\epsilon - \mu) = \Theta(\mu - \epsilon)$$

The value of μ at $T = 0$ is called *'Fermi energy'*:

$$\epsilon_F = \frac{p_F^2}{2m} = \mu\left(T = 0, v = \frac{V}{N}\right)$$

$$\Rightarrow \quad N = \sum_{\vec{k},m_S \; for \; p \le p_F} 1 = 2\frac{V}{h^3} \int_{p \le p_F} \mathrm{d}\vec{p} = \frac{2V}{h^3}\frac{4\pi}{3}\, p_F^3$$

Here we integrated over the 'Fermi sphere'.

$$\Rightarrow \quad \boxed{\epsilon_F = \left(3\pi^2\right)^{\frac{2}{3}} \; \frac{\hbar^2}{2mv^{\frac{2}{3}}} = \left(3\pi^2\right)^{\frac{2}{3}} \frac{\hbar^2 \rho^{\frac{2}{3}}}{2m}}$$

Typical values for the Fermi energy:

$$\epsilon_F = \begin{cases} 10^{-4} \; eV & {}^3He \\ 10 \; eV & \text{electrons in metal} \\ 1 \; MeV & \text{electrons in white dwarf} \\ 35 \; MeV & \text{neutrons in atomic nucleus} \end{cases}$$

Since $k_B T_R = \frac{eV}{40}$, for electrons in metals at room temperature T_R we typically have $\epsilon_F \gg k_B T_R$. We evaluate the occupancy around the Fermi-edge (compare Figure 5.5):

$$n(\epsilon) = 0.5 \pm 0.23 \quad \text{for } \epsilon = \mu \mp k_B T$$

$$n(\epsilon) = 0.5 \pm 0.45 \quad \text{for } \epsilon = \mu \mp 3k_B T$$

We see that the width of the step at finite T is only a few $k_B T$. Therefore only a few of the N electrons in the *'Fermi sea'* are thermally excited above ϵ_F.

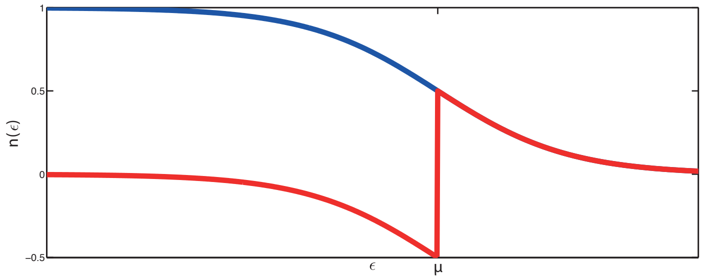

Figure 5.5: The occupation number n as a function of ϵ at finite temperature (blue) and the difference of the curve with respect to the one at $T = 0$ (red and red-dashed above blue).

Specific heat

We use this result to calculate the specific heat based on the *'Sommerfeld method'*. We consider an arbitrary function $f(\epsilon)$ (eg $f(\epsilon) = \epsilon^{\frac{1}{2}}$):

$$I = \int_0^\infty \mathrm{d}\epsilon\, f(\epsilon)\, n(\epsilon) = \int_0^\mu \mathrm{d}\epsilon\, f(\epsilon) + \int_0^\infty \mathrm{d}\epsilon\, f(\epsilon) \quad \underbrace{[n(\epsilon) - \Theta(\mu - \epsilon)]}_{\neq 0 \text{ only in small region around } \mu}$$

Expansion of $f(\epsilon)$ around the Fermi edge:

$$f(\epsilon) = f(\mu) + f'(\mu)(\epsilon - \mu) + \frac{1}{2} f''(\mu)(\epsilon - \mu)^2 + \ldots$$

We introduce $x = \beta(\epsilon - \mu)$:

$$\begin{aligned} \Rightarrow \quad \eta(x) &= n(\epsilon) - \Theta(\mu - \epsilon) = \frac{1}{e^x + 1} - \Theta(-x) \\ &= \left(\frac{1}{e^x + 1} - (1 - \Theta(x)) \right) = -\left(\frac{1}{e^{-x} + 1} - \Theta(x) \right) = -\eta(-x) \end{aligned}$$

$\eta(x)$ being odd in x implies that all even terms on the Taylor expansion vanish.

$$\Rightarrow \quad I = \int_0^\mu \mathrm{d}\epsilon\, f(\epsilon) + \frac{1}{\beta} \int_{-\beta\mu}^\infty \mathrm{d}x \left[f(\mu) + f'(\mu) \frac{x}{\beta} + \ldots \right] \eta(x)$$

For low temperatures: $\beta\mu \to \infty$:

$$= \int_0^\mu \mathrm{d}\epsilon\, f(\epsilon) + \frac{f'(\mu)}{\beta^2} \qquad \underbrace{\int_{-\infty}^\infty \mathrm{d}x\; x\, \eta(x)}_{=2\int_0^\infty \mathrm{d}x\; x\eta(x) = 2\int_0^\infty \mathrm{d}x\; \frac{x}{e^x + 1} = \frac{\pi^2}{6}}$$

$$= \int_0^\mu \mathrm{d}\epsilon\, f(\epsilon) + \frac{\pi^2}{6\beta^2}\, f'(\mu)$$

We now apply this result to our normalization condition:

$$1 = \int_0^\infty \mathrm{d}\epsilon\, D(\epsilon) n(\epsilon)$$

with $D(\epsilon) \propto \epsilon^{\frac{1}{2}}$ (compare Figure 5.6) to determine the chemical potential $\mu(T, v)$.

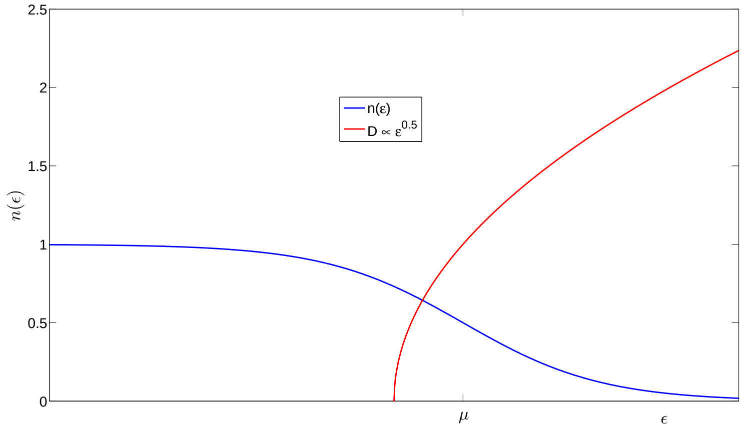

Figure 5.6: The average occupancy n (blue) as a function of ϵ and the density of states $D \propto \epsilon^{\frac{1}{2}}$ (red).

$$\begin{aligned} 1 = & \underbrace{\int_0^\mu \mathrm{d}\epsilon\, D(\epsilon)} + \frac{\pi^2}{6\beta^2} D'(\mu) \\ & = \int_0^{\epsilon_F} \mathrm{d}\epsilon\, D(\epsilon) + \int_{\epsilon_F}^{\mu} \mathrm{d}\epsilon\, D(\epsilon) \\ & = 1 + (\mu - \epsilon_F)\, D(\widetilde{\epsilon}) \end{aligned}$$

Here 1 is the result for $T = 0$ and $\widetilde{\epsilon}$ some value between μ and ϵ_F according to the mean value theorem.

$$\Rightarrow \quad \mu - \epsilon_F = -\frac{\pi^2}{6\beta^2} \frac{D'(\mu)}{D(\widetilde{\epsilon})} \approx -\frac{\pi^2}{6\beta^2} \underbrace{\frac{D'(\epsilon_F)}{D(\epsilon_F)}}_{=\frac{1}{2}\frac{1}{\epsilon_F}} \quad \text{using } \mu - \epsilon_F \propto T^2$$

$$\Rightarrow \quad \boxed{\mu = \epsilon_F \left[1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F}\right)^2\right]} \quad \text{for } T \ll \frac{\epsilon_F}{k_B} \text{ (Figure 5.7)}$$

We note that in general the chemical potential μ has to go down with temperature because for fixed particle number the joint integral with the density of states has to

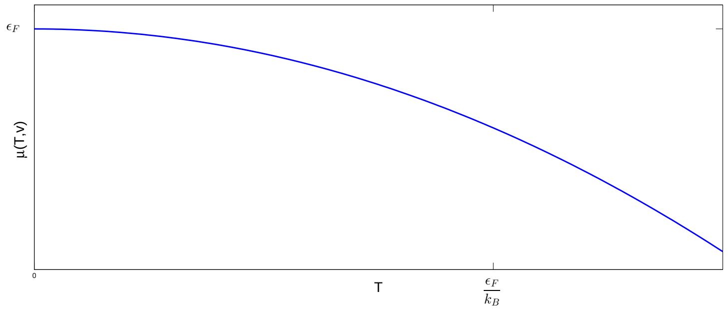

Figure 5.7: The chemical potential $\mu(T, v)$ decreases with increasing temperature. For $T \ll \frac{\epsilon_F}{k_B}$ it can be taken to be $\mu = \text{const} = \epsilon_F$.

stay constant, compare Figure 5.8. Therefore higher order term in this expansion are not expected to change the general picture.

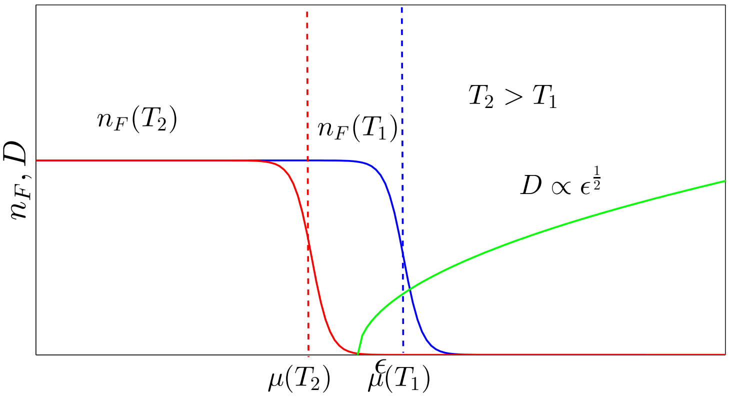

Figure 5.8: The fermionic occupation number n_F as a function of ϵ for different temperatures (blue and red curves) and the density of states D (green). Because increasing T moves more particles into the high energy tail, one has to decrease chemical potential μ to keep the overlap integral determining N constant. Thus the curves shift to the left with increasing T.

We next evaluate the energy:

$$\begin{aligned}
\frac{E}{N} &= \int_0^\infty \mathrm{d}\epsilon\ \underbrace{D(\epsilon)\ \epsilon}_{f(\epsilon)\propto\epsilon^{\frac{3}{2}}}\ n(\epsilon) \\
&= \int_0^{\epsilon_F} \mathrm{d}\epsilon\ D(\epsilon)\epsilon + (\mu-\epsilon_F)\widetilde{\epsilon}D(\widetilde{\epsilon}) + \frac{\pi^2}{6\beta^2}\left[\mu D'(\mu) + D(\mu)\right] \\
&\approx \frac{E_0}{N} + \underbrace{(\mu-\epsilon_F)}_{-\frac{\pi^2}{6\beta^2}\frac{D'(\epsilon_F)}{D(\epsilon_F)}}\epsilon_F D(\epsilon_F) + \frac{\pi^2}{6\beta^2}\left[\epsilon_F D'(\epsilon_F) + D(\epsilon_F)\right]
\end{aligned}$$

$$\Rightarrow \quad E = E_0 + N\frac{\pi^2}{6}D(\epsilon_F)\left(k_B T\right)^2$$

$$\Rightarrow \quad \boxed{c_V = \left.\frac{\partial E}{\partial T}\right|_V = N\frac{\pi^2}{3}k_B^2 D(\epsilon_F)T}$$

The specific heat of an electron gas at $T \ll \frac{\epsilon_F}{k_B}$ is linear in temperature T.

We use $D(\epsilon) = A\ \epsilon^{\frac{1}{2}}$ to write

$$\begin{aligned}
1 &= \int_0^{\epsilon_F} \mathrm{d}\epsilon\ D(\epsilon) = A\int_0^{\epsilon_F} \mathrm{d}\epsilon\ \sqrt{\epsilon} \\
&= \frac{2}{3}A\epsilon_F^{\frac{3}{2}} = \frac{2}{3}D(\epsilon_F)\epsilon_F
\end{aligned}$$

$$\Rightarrow \quad D(\epsilon_F) = \frac{3}{2\epsilon_F}$$

$$\Rightarrow \quad \boxed{c_V = N\frac{\pi^2}{2}\frac{k_B T}{\epsilon_F}k_B}$$

Disregarding the numerical prefactor, this result is easy to understand: a fraction $\frac{k_B T}{\epsilon_F}$ of the electrons from the Fermi sea is thermally excited, each contributing around k_B.
Our calculation is only valid for $T \ll \frac{\epsilon_F}{k_B T}$. At high temperature, we have to recover the classical limit:

$$c_V = \frac{3}{2}Nk_B$$

Therefore the complete result schematically has to look like shown in Figure 5.9.
We also comment on the role of lattice vibrations. From the Debye model we know that lattice vibrations contribute a term $\propto T^3$.

$$\Rightarrow \quad c_V = a\ T + b\ T^3$$

One can measure a and b experimentally and thus extract the Fermi energy ϵ_F and the Debye frequency ω_D. With these two values, we know the most important numbers for a given solid.

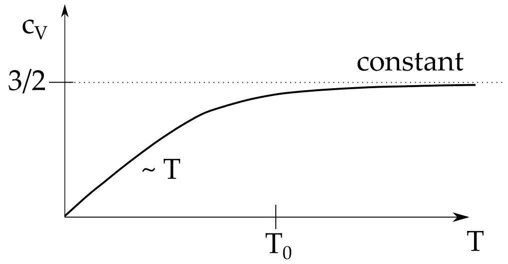

Figure 5.9: Two regime behaviour of the specific heat at constant volume: While for $T \ll T_0 = \frac{\epsilon_F}{k_B}$ $c_V \propto T$, c_V is approximately constant for $T \gg T_0$.

Full solution

Until now we have worked in an expansion around the $T = 0$-case. We can also write the full solution for arbitrary T, however we will end up with integrals that cannot be solved but rather lead to definitions of new functions.

We start with the grandcanonical potential and use the same concepts as above:

$$\begin{aligned}\Psi(T, V, \mu) &= -k_B T \ln Z_G \\ &= \frac{-k_B T}{h^3} 2V(4\pi) \int_0^\infty p^2 dp \ln\left(1 + e^{\beta(\epsilon-\mu)}\right) \\ &= \frac{-2k_B T V}{\lambda^3} f_{5/2}(z)\end{aligned}$$

where we have defined a new function

$$f_{5/2}(z) := \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \ln\left(1 + z e^{-x^2}\right) = \sum_{\alpha=1}^{\infty} (-1)^{\alpha+1} \frac{z^\alpha}{\alpha^{5/2}}$$

and where we have used the dimensionless momentum x defined by $x^2 = \beta p^2/2m$ and fugacity $z = e^{\beta\mu}$.

Particle number can be written in a similar manner:

$$\begin{aligned}N &= \frac{2V(4\pi)}{h^3} \int_0^\infty p^2 dp \left(\frac{1}{e^{\beta(\epsilon-\mu)} + 1}\right) \\ &= \frac{2V}{\lambda^3} \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \left(\frac{z}{e^{x^2} + z}\right) = \frac{2V}{\lambda^3} f_{3/2}(z)\end{aligned}$$

with another new function

$$f_{3/2}(z) := \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \left(\frac{z}{e^{x^2} + z}\right) = \sum_{\alpha=1}^{\infty} (-1)^{\alpha+1} \frac{z^\alpha}{\alpha^{3/2}}$$

As a function of z, both functions increase monotonously from 0 with a decreasing slope.

One can easily check that the two formula are consistent:

$$N = \frac{1}{\beta}\partial_\mu \ln Z_G = \frac{1}{\beta}(\beta z)\partial_z \ln Z_G = \frac{2V}{\lambda^3}(z\partial_z) f_{5/2}(z) = \frac{2V}{\lambda^3} f_{3/2}(z)$$

One can also calculate the variance as $\sigma_N^2 = (1/\beta)\partial N/\partial\mu$. For low temperature, we would get the same results as above.

Fermi pressure

We consider the definition of pressure:

$$p = -\left.\frac{\partial E}{\partial V}\right|_{T,N} = -\sum_{\vec{k},m_S} \frac{\partial \epsilon_{\vec{k},m_S}}{\partial V}\, n_{\vec{k},m_S}$$

where in the last step we have neglected any temperature-dependent change in the occupation level (second order effect, a more rigorous treatment would again start from the grandcanoncial ensemble). Since $\epsilon_{\vec{k}} = \frac{(\hbar k)^2}{2m}$ and $k_i \propto \frac{1}{V^{\frac{1}{3}}}$, we have

$$\epsilon_{\vec{k}} \propto \frac{1}{V^{\frac{2}{3}}} \;\Rightarrow\; \frac{\partial \epsilon_{\vec{k},m_S}}{\partial V} = -\frac{2}{3}\frac{\epsilon_{\vec{k},m_S}}{V}$$

$$\Rightarrow \quad p = \frac{2}{3}\frac{E}{V} = \frac{2E_0}{3V} + \underbrace{\frac{\pi^2}{6}\frac{N k_B T}{V}\frac{k_B T}{\epsilon_F}}_{\to 0 \text{ for } T\to 0 \text{ like for the ideal gas}}$$

Interestingly, this contribution to the pressure is always positive, showing that the Fermi gas is effectively like a gas with repulsive interactions. There is also a temperature-independent term:

$$\frac{E_0}{N} = \int_0^{\epsilon_F} \mathrm{d}\epsilon \; D(\epsilon)\epsilon = \frac{3}{5}\epsilon_F$$

$$\Rightarrow \quad \boxed{p \overset{T\to 0}{\to} \frac{2}{5}\frac{N}{V}\epsilon_F = \frac{\left(3\pi^2\right)^{\frac{2}{3}}}{5}\frac{\hbar^2}{m v^{\frac{5}{3}}}}$$

The *'Fermi pressure'* in a Fermi fluid at very low temperature accounts for the incompressibility of matter and essentially results from the Pauli principle. For example, it prevents that the earth collapses under gravitation. This is also true for white dwarfs (electrons) or neutron stars, but not for the sun. In the latter case classical ideal gas pressure at $T = 5 \cdot 10^7\ K$ (temperature in the center of the sun) balances gravitational attraction.

5.5 The ideal Bose fluid

We now turn to Bose fluids with

$$\epsilon = \frac{\hbar^2 k^2}{2m}$$

and conserved particle number N.

Example:

A fluid of ^4He, which in contrast to ^3He is a boson. Both types of helium have the same weak van der Waals interaction and the same chemistry, but one is a bosonic, the other a fermionic fluid. We demonstrate now that the bosonic system undergoes a *'Bose-Einstein condensation'* at a critical temperature T_c.

$\epsilon = \frac{p^2}{2m}$ has two main consequences for Bose fluids:

① $D(\epsilon) \propto \epsilon^{\frac{1}{2}}$: like for Fermi fluids

② $\epsilon \geq 0 \;\Rightarrow\; \mu \leq 0$: otherwise the mode with $\epsilon = \mu$ would have an infinite occupation number

$\mu = 0$ is allowed as the contribution to n_B at $\epsilon = \mu$ vanishes in a continuum framework:

$$n_B = \frac{1}{e^{\beta(\epsilon-\mu)} - 1} \overset{\mu=0}{\rightarrow} \frac{1}{e^{\frac{\beta p^2}{2m}} - 1} \overset{p\rightarrow 0}{\rightarrow} \frac{2m}{\beta p^2} \Rightarrow \mathrm{d}\vec{p}\, n_B \propto p^2\, \mathrm{d}p\, \frac{1}{p^2} \propto \mathrm{d}p \rightarrow 0$$

For $\mu > 0$, one can use a Taylor expansion to see that now the divergence is not canceled by the Jacobian and that the weight would diverge.
We consider the particle number N for spin-0 bosons (degeneracy $g = 1$):

$$N = \frac{V}{h^3} \int \mathrm{d}\vec{p}\, \frac{1}{e^{\beta(\epsilon-\mu)} - 1} = \frac{V}{(2\pi)^3} \left(\frac{2m}{\hbar^2}\right)^{\frac{3}{2}} \int_0^\infty \mathrm{d}\epsilon\, \frac{\epsilon^{\frac{1}{2}}}{e^{\beta(\epsilon-\mu)} - 1}$$

For fixed μ, increasing T shifts n_B to higher values. In order to keep N constant, μ must decrease (as for the Fermi fluid).
For $T \rightarrow \infty$, $\mu \rightarrow -\infty$ we recover the classical limit. Compare with the ideal gas result:

$$\mu = k_B T \ln \rho\lambda^3 = k_B T \ln \frac{p}{p_0} \text{ with } p_0 = \frac{k_B T}{\lambda^3} \approx GPa, \text{ atmospheric pressure } 0.1\, MPa \rightarrow \mu < 0$$

We now consider the integral for N:

$$N = \frac{V}{h^3} \int \mathrm{d}\vec{p}\, \underbrace{\frac{1}{e^{\beta(\epsilon-\mu)} - 1}}_{=\sum_{l=1}^\infty \left(e^{-\beta(\epsilon-\mu)}\right)^l} = \frac{V}{h^3} \sum_{l=1}^\infty e^{\beta\mu l} \int \mathrm{d}\vec{p}\, e^{-\beta \frac{p^2}{2m} l}$$

$$= V \sum_{l=1}^\infty \frac{e^{\beta\mu l}}{l^{\frac{3}{2}}} \underbrace{\frac{1}{h^3} \int \mathrm{d}\vec{x}\, e^{-\beta \frac{x^2}{2m}}}_{=\frac{1}{\lambda^3}} = \frac{V}{\lambda^3} \underbrace{\sum_{l=1}^\infty \frac{z^l}{l^{\frac{3}{2}}}}_{=g_{\frac{3}{2}}(z)}$$

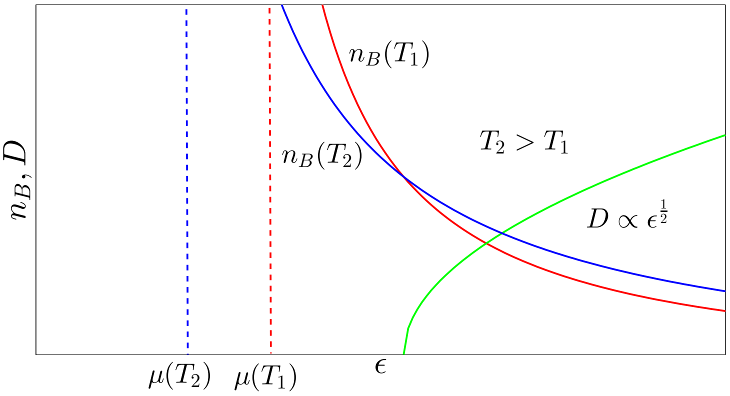

Figure 5.10: The bosonic occupation number n_B as a function of ϵ for different temperatures (blue and red curves) next to the density of states D (green).

$$\Rightarrow \quad \boxed{N = \frac{V}{\lambda^3} g_{\frac{3}{2}}(z)}$$

Here we used the substitution $x^2 = p^2 l$, identified the fugacity z and the thermal wavelength λ

$$z = e^{\beta\mu}, \quad \lambda = \frac{h}{(2\pi m k_B T)^{\frac{1}{2}}}$$

and the generalized Riemann Zeta function

$$g_\nu(z) = \sum_{l=1}^{\infty} \frac{z^l}{l^\nu}$$
$$g_{\frac{3}{2}}(z) = z + \frac{z^2}{2\sqrt{2}} + \frac{z^3}{3\sqrt{3}} + ...$$

For high temperatures:

$$\lambda \to 0, g_{\frac{3}{2}}(z) \to 0, z \to 0, \mu \to -\infty$$

$$\Rightarrow N = \frac{V}{\lambda^3} e^{\beta\mu} \quad \text{classical result (with corrected counting)}$$

Upon lowering temperature, z approaches 1, where

$$g_{\frac{3}{2}}(1) = \zeta\left(\frac{3}{2}\right) = 2.612$$

with the Riemann Zeta function $\zeta(\nu) = g_\nu(1) = \sum_{l=1}^{\infty} \frac{1}{l^\nu}$.

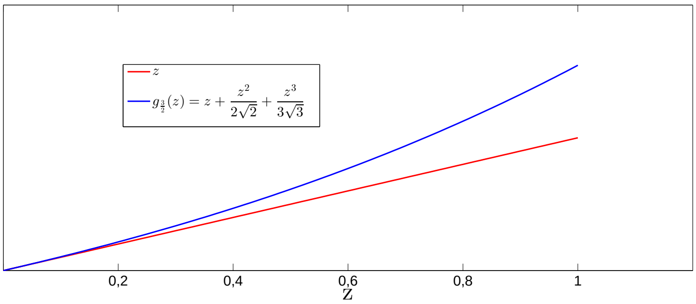

Figure 5.11: The generalized Riemann Zeta function $g_{\frac{3}{2}}$ plotted against fugacity z. In contrast to the function $f_{\frac{3}{2}}$ required for fermions, it is only defined up to $z = 1$, reflecting that the chemical potential for bosons cannot become positive.

The maximal value $\mu = 0$ is reached when

$$\lambda^3 \rho = \zeta\left(\frac{3}{2}\right) = 2.612 \quad \text{where } \rho = \frac{N}{V}$$

surprisingly, this happens at a finite temperature T_c:

$$\boxed{T_c = \frac{2\pi}{\left(\zeta\left(\frac{3}{2}\right)\right)^{\frac{3}{2}}} \frac{\hbar^2 \rho^{\frac{2}{3}}}{k_B m}} \quad \text{Einstein 1924}$$

T_c is the critical temperature for the *'phase transition'*. At this point the system changes its state characteristics.

In the case of ^4He we estimate

$$v = \frac{V}{N} \approx 46\ \mathring{A}^3,\ m = 4\ u\ \Rightarrow\ T_c = 3.13\ K$$

Experimentally one finds $T_c = 2.17\ K$ for the so-called λ-transition to superfluidity. The difference comes from direct interactions which we neglected here (ideal gas of point particles).

Below T_c, the chemical potential μ must be identical zero and we get a plateau in the value for μ. The problem here is that as we lower T below T_c, we cannot satisfy anymore the condition to keep N constant by increasing μ. This might indicate that the theory fails, but in fact it is the other way around, the theory immediately suggests a way out: we can rescue the theory by decreasing the number of particles while lowering temperature. The main idea is that we remove particles from the system and put them into another phase, namely a condensate of particles with cero energy (vanishing momentum). Because for a finite sized system the momenta are discretized anyway, we

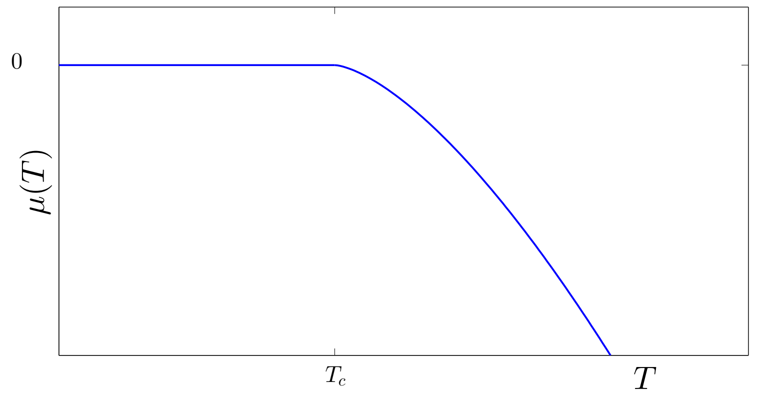

Figure 5.12: The chemical potential μ as a function of temperature. Below $T = T_c$, μ equals zero.

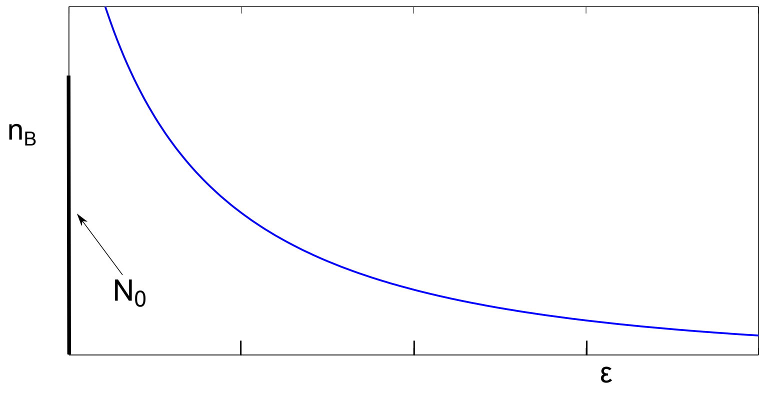

Figure 5.13: The occupation number n_B as a function of ϵ. In the continuum description the ground state does not contribute. Therefore we introduce it here as an additional degree of freedom (black box at $\epsilon = 0$) into which particles from the excited states can disappear, thus allowing us to satisfy the equation for particle number while still lowering temperature and keeping chemical potential constant.

also can consider this ground state to be a discrete state into which the excited states can disappear. The range of excited states is treated as continuum as before.
At T below T_c, some particles condense into the ground state for which we introduce an extra particle number:

N_0: number of atoms in the ground state, N_e: number of atoms in the excited state

$$\Rightarrow \quad N_e = \frac{V}{\lambda^3}\zeta\left(\frac{3}{2}\right) = \frac{\lambda_c^3}{\lambda^3}\underbrace{\frac{V}{\lambda_c^3}\zeta\left(\frac{3}{2}\right)}_{=N} = N\left(\frac{T}{T_c}\right)^{\frac{3}{2}} \quad \text{for } T < T_c$$

$$N = N_0 + N_e \quad \Rightarrow \quad N_0 = N\Big(1 - \left(\frac{T}{T_c}\right)^{\frac{3}{2}}\Big)$$

This result is plotted in Figure 5.14.

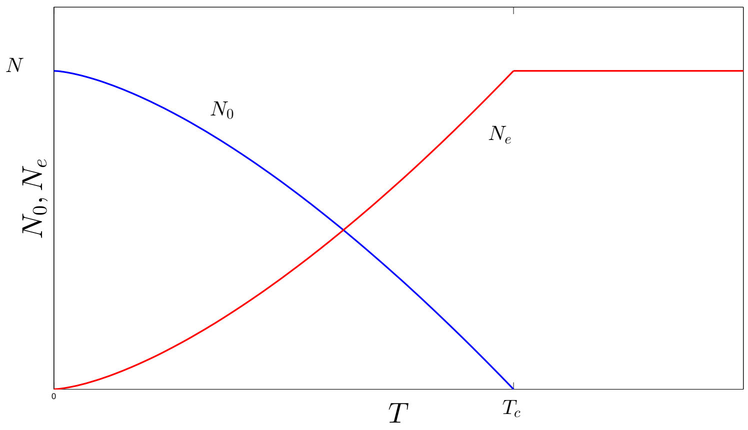

Figure 5.14: The number of atoms in ground (N_0) and excited state (N_e) as a function of temperature. Above $T = T_c$ all atoms are in the excited state. The more one goes below T_c, the more particles condense into the groundstate of the condensate.

At $T = 0$, all N particles have condensed into the ground state. For $0 \leq T \leq T_c$, a finite fraction is condensed. For $T > T_c$, all particles are excited.

Next we calculate the energy of the ideal Bose fluid:

$$E = \underbrace{N_0\epsilon_0}_{=0 \text{ since } \epsilon_0 = 0} + \frac{V}{h^3}\sum_{l=1}^{\infty} e^{\beta\mu l}\underbrace{\int \mathrm{d}\vec{p}\,\frac{p^2}{2m}e^{-\beta\frac{p^2}{2m}l}}_{=-\frac{1}{\beta}\,\partial_l \int \mathrm{d}\vec{p}\; e^{-\beta\frac{p^2}{2m}l}}$$

$$= V\sum_{l=1}^{\infty} e^{\beta\mu l}\left(-\frac{1}{\beta}\,\partial_l\right)\frac{1}{\lambda^3 l^{\frac{3}{2}}}$$

$$\boxed{E = \frac{3}{2}k_B T\frac{V}{\lambda^3}g_{\frac{5}{2}}(z) = \frac{3}{2}k_B T N_e\frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)}}$$

This result (Figure 5.16)is valid for all temperatures.
Next we calculate the specific heat for $T \leq T_c$:

$$\Rightarrow \quad \mu = 0,\ z = 1,\ \lambda \propto T^{-\frac{1}{2}},\ E \propto T^{\frac{5}{2}}$$

$$\begin{aligned}
\Rightarrow \quad & c_V = \frac{dE}{dT} = \frac{3}{2} k_B N \frac{d}{dT} T \left(\frac{T}{T_c}\right)^{\frac{3}{2}} \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} \\
\Rightarrow \quad & \boxed{c_V = \frac{15}{4} k_B N \left(\frac{T}{T_c}\right)^{\frac{3}{2}} \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)}} \quad T \leq T_c
\end{aligned}$$

For $T > T_c$ we get

$$c_v = \frac{dE}{dT} = \frac{15}{4} k_B \frac{V}{\lambda^3} g_{\frac{5}{2}}(z) + \frac{3}{2} k_B T \frac{V}{\lambda^3} g'_{\frac{5}{2}}(z) \frac{dz}{dT}$$

The chemical potential μ is determined by

$$\begin{aligned}
& N = \frac{V}{\lambda^3} g_{\frac{3}{2}}(z) \\
\Rightarrow \quad & 0 = \frac{3}{2} \frac{V}{T\lambda^3} g_{\frac{3}{2}}(z) + \frac{V}{\lambda^3} g'_{\frac{3}{2}}(z) \frac{dz}{dT} \\
\Rightarrow \quad & \frac{dz}{dT} = -\frac{3}{2} \frac{g_{\frac{3}{2}}(z)}{T g'_{\frac{3}{2}}(z)} \\
\Rightarrow \quad & c_V = \frac{15}{4} k_B N \frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{9}{4} k_B N \frac{g'_{\frac{5}{2}}(z)}{g'_{\frac{3}{2}}(z)} \quad T > T_c \\
\Rightarrow \quad & \boxed{c_V = \frac{15}{4} k_B N \frac{g_{\frac{5}{2}}(z)}{g_{\frac{3}{2}}(z)} - \frac{9}{4} k_B N \frac{g_{\frac{3}{2}}(z)}{g_{\frac{1}{2}}(z)}} \quad T > T_c
\end{aligned}$$

Here we used

$$\frac{dg_\nu(z)}{dz} = \sum_{l=1}^{\infty} \frac{l z^{l-1}}{l^\nu} = \frac{1}{z} \sum_{l=1}^{\infty} \frac{z}{l^{\nu-1}} = \frac{1}{z} g_{\nu-1}(z)$$

in the last step.

$$\text{High temperature limit :} \quad z \to 0,\ g_\nu(z) \to z$$

$$\Rightarrow \quad c_V = \left(\frac{15}{4} - \frac{9}{4}\right) k_B N = \frac{3}{2} k_B N \quad \text{classical limit}$$

$$T = T_c: \quad z \to 1,\ g_{\frac{1}{2}}(1) = \infty$$

$$\Rightarrow \quad c_V = \frac{15}{4} k_B N \frac{\zeta\left(\frac{5}{2}\right)}{\zeta\left(\frac{3}{2}\right)} = 1.925\, k_B N$$

The specific heat has a unique cup at $T = T_c$ (Figure 5.15). A similar behaviour has been experimentally observed for ${}^4He$. The energy E and its first derivative c_V are continuous; only the second derivative has a jump.

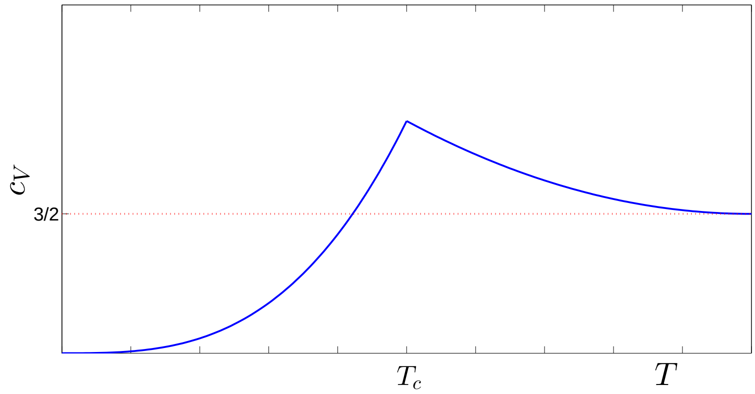

Figure 5.15: The specific heat c_V as a function of temperature has a unique cup at $T = T_c$.

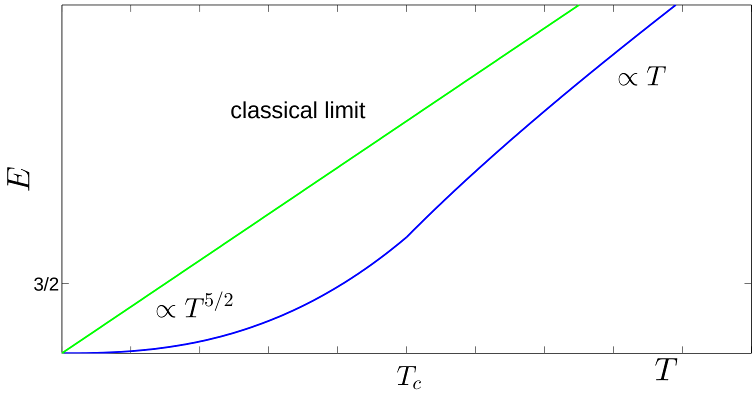

Figure 5.16: The energy E of the ideal Bose fluid as a function of temperature. For low temperatures (quantum regime), $E \propto T^{\frac{5}{2}}$ while $E \propto T$ for high temperatures (classical limit).

Ultracold atoms

Bosonic atoms like ^{87}Rb or ^{23}Na can be captured in magnetic traps because they have unpaired electrons that interact with magnetic fields. To first order, this is a harmonic oscillator with frequency ω.

$$\epsilon_{n_x,n_y,n_z} = \hbar\omega \left(n_x + n_y + n_z + \frac{3}{2} \right)$$

For $\mu = \mu_c = \epsilon_{000} = \frac{3}{2}\hbar\omega$, N_0 atoms condensed into a BEC:

$$N = N_0 + \sum_{n_x=0}^{\infty} \sum_{n_y=0}^{\infty} \sum_{n_z=0}^{\infty} \frac{1}{e^{\beta\hbar\omega(n_x+n_y+n_z)} - 1}$$

Here the contribution $\epsilon_{000} - \mu = 0$ vanishes from the exponential. Converting the three sums into integrals, using the geometrical sum and performing the integrals gives

$$N = N_0 + \zeta(3) \left(\frac{k_B T}{\hbar\omega} \right)^3 \quad T \leq T_c$$

The critical temperature follows from $N_0 \to 0$:

$$T_c = \frac{\hbar\omega}{k_B} \left(\frac{N}{\zeta(3)} \right)^{\frac{1}{3}}$$

For $\hbar\omega \approx 10^{-8}\ k_B K$ and $N = 4 \cdot 10^4$ this gives

$$T_c \approx 3 \cdot 10^{-7}\ K$$

In 1995 such a Bose-Einstein condensate was achieved for the first time (Nobel Prize 2001 for Ketterle, Cornell and Wieman). Usually they are demonstrated by the expansion following shutoff of an atomic trap. In 2010 a BEC was achieved for photons (Weitz group, Bonn).

5.6 Classical limit

As we have seen above, both for Fermi and Bose fluids the classical limit emerges as $\mu \to -\infty$. Then the two grandcanonical distribution functions become the same classical Maxwell-Boltzmann distribution:

$$n_{F/B} = \frac{1}{e^{\beta(\epsilon-\mu)} \pm 1} \to e^{\beta\mu} e^{-\beta\epsilon}$$

In this limit, all occupied states are in the tail of the distributions, which is the same for both. The normalization condition now becomes, using $\epsilon = p^2/2m$, which is valid for both cases and therefore give the same density of states:

$$N = g \frac{V}{(2\pi)^2} (\frac{2m}{\hbar^2})^{3/2} e^{\beta\mu} \int d\epsilon \epsilon^{1/2} e^{-\beta\epsilon} = g \frac{V}{\lambda^3} e^{\beta\mu}$$

where degeneracy $g = 2$ for electrons and $g = 1$ for spin-0 particles. For the average energy we get

$$E = g\frac{V}{(2\pi)^2}(\frac{2m}{\hbar^2})^{3/2}e^{\beta\mu}\int d\epsilon\epsilon^{3/2}e^{-\beta\epsilon} = \frac{3}{2}k_BTg\frac{V}{\lambda^3}e^{\beta\mu} = \frac{3}{2}k_BTN$$

which is simply the caloric equation of state of the classical ideal gas. The boundary to the quantum regime occurs at fugacity $e^{\beta\mu} = z \approx 1$, which from the first equation corresponds to $v \approx \lambda^3$. When the density is sufficiently high or the temperature is sufficiently low that the specific volume approaches the volume defined by the thermal wavelength, then quantum effects will dominate. Note that the classical limit corresponds to the calculation with corrected counting. Without this factor $N!$, we would not have achieved agreement with the full quantum calculations (and also not with the thermodynamic result).

6 Ising model

6.1 History and definition

Until now we have only discussed so-called non-interacting systems. Of course fermions and bosons feel each other through the exchange rules, but we did not consider yet the possibility of a direct interaction between particles. As we will see, such interactions typically lead to phase transitions. We first discuss this for lattice models and then for classical fluids. While lattice models are abstract, they allow us to introduce the relevant concepts and methods. Moreover some lattice models can be solved exactly. The classical fluids are more directly related to experiments, but here exact solutions are more challenging.

The most famous lattice model is the Ising model. In fact it is the most important model in statistical physics anyway, both historically and conceptually. It is one of the few analytically solvable models with interactions and a phase transition. In 1920 it was invented by Wilhelm Lenz as an exercise in ferromagnetism and given to his PhD student Ernst Ising, who solved the one-dimensional case (publication Z. Phys. 31, 253-258, 1925), which has a phase transition only at $T = 0$ (some textbooks like Landau and Lifshitz therefore state that it has no phase transition at all, which is true as a statement on finite temperature). In 1933 Rudolf Peierls used indirect arguments to show that the 2D version must have a phase transition at finite temperature. In 1944 Lars Onsager solved the two-dimensional Ising model in vanishing magnetic field by mapping it onto a fermionic problem using a transfer matrix method (publication Phys. Rev. 65, 117, 1944). Later several other and less difficult proofs have been reported, often focusing more on the geometrical aspects of random walks, as also suggested by the Peierls argument. One of these proofs is by Vdovichenko (publication N. V. Vdovichenko, J. Exptl. Theor. Phys. (USSR) 47, 715 (1964);English translation: Soviet Physics JETP 20, 477 (1965)) and has become very popular through the textbook by Landau and Lifshitz. Generations of theoretical physicists worked on the Ising model and related models (including the Heisenberg, Potts, n-vector and XY-models), leading to the developments of concepts such as spontaneous magnetization, symmetry breaking and universality, and methods such as perturbation theory, transfer matrix, renormalization group (RG) theory and Monte Carlo computer simulations. The two-dimensional Ising model with magnetic field and the three-dimensional Ising model are still not solved today and this is one of the greatest challenges in theoretical physics (like solving the Navier-Stokes equation). However, it has been extensively studied with numerical methods, so that we can say that in principle, we know everything about it. Overall the Ising model is for statistical physics what the harmonic oscillator is for mechanics, the hydrogen atom for quantum physics and the fruit fly for biology: the most important reference system both in terms of concepts and methods.

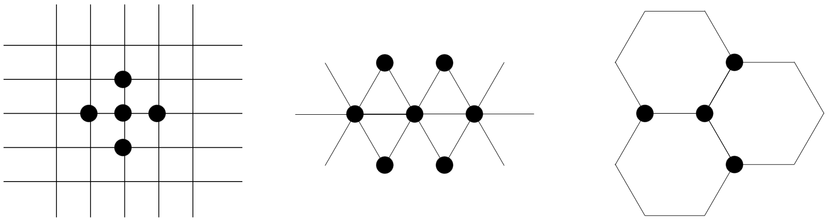

Figure 6.1: Ising lattice examples in two dimensions: cubic (# neighbours $z = 4$, left), triangular ($z = 6$, center) and hexagonal (honeycomb) ($z = 3$, right).

The Ising model is defined on a lattice of given connectivity and dimension (e.g. 1D Ising chain, 2D square lattice Ising model, Ising model on a Cayley tree, Ising model on a small network, etc), thus in principle, there are infinitely many versions of the Ising models. Figure 6.1 shows some examples in two dimensions. Usually however one studies it on a simple cubic lattice.

The Ising model is an example for a *lattice spin model*. In each lattice site i, we consider a spin S_i with two possible states: $S_i = \pm 1$ (corresponding to $|\uparrow\rangle$ and $|\downarrow\rangle$). Nearest neighbours interact with an energy scale J. In addition there might be an external magnetic field B giving a preference for one direction. Note that although we talk about spins and the cartoon show vectors, we really have only a scalar degree of freedom, which is binary; this makes it the simplest model possible, like the flip of a coin in probability theory. More advanced models like the XY- or Heisenberg models considers spins as vectors with a scalar product between them.

The Hamiltonian of the Ising model reads

$$\boxed{\mathcal{H} = -J \sum_{\langle ij \rangle} S_i \, S_j - B\mu \sum_i S_i}$$

Here $\langle ij \rangle$ indicates summation over nearest neighbours and μ is the magnetic moment of a spin. In non-dimensional units we write

$$\boxed{\beta\mathcal{H} = -K \sum_{\langle ij \rangle} S_i \, S_j - H \sum_i S_i}$$

where now both the coupling constant $K = \beta J$ and the external field $H = \beta B \mu$ depend on temperature. The H used here should not be confused with the H of the magnetic field in the macroscopic Maxwell equations, which sometimes is used in the same sense as we use B here. Many books also use h for the H we use here.

For $J > 0$ the symmetric configurations $\uparrow\uparrow$ and $\downarrow\downarrow$ are favorable and $\uparrow\downarrow$ and $\downarrow\uparrow$ are unfavorable. Thus the system wants to avoid grain boundaries between regions with up and down spins, at least at low temperature. At high temperature, grain boundaries will proliferate because they correspond to a lot of entropy.

For $B = 0$ the system is invariant under $S_i \to -S_i$. If $B > 0$, $\uparrow$-spins are favored, and for $B < 0$, $\downarrow$-spins are favored. Using the canonical formalism the partition sum for N spins reads:

$$\boxed{Z_N(K,H) = \underbrace{\sum_{S_1=\pm1}\sum_{S_2=\pm1} \dots \sum_{S_N=\pm1}}_{2^N \text{ states}} e^{-\beta\mathcal{H}} = \sum_{\{S_i\}} e^{-\beta\mathcal{H}}}$$

In practice one often uses periodic boundary conditions for lattice models to avoid boundary effects, or finite size scaling to get rid of boundary effects by making the system larger and larger.

Due to its history, usually the Ising model is treated in magnetic language. However, today it is used in many other ways. Here are a few examples of important applications of the Ising model:

① ferromagnetism:
The Ising model is the scalar version of the three-dimensional *'Heisenberg model'* for ferromagnetic systems:

$$\mathcal{H} = -J\sum_{\langle i,j\rangle}\vec{J}_i\vec{J}_j - \mu\vec{B}\sum_i\vec{J}_i$$

② binary alloys and lattice gases:
Each lattice site is occupied either by an atom A or B. Nearest neighbour interactions are t_{AA}, t_{BB} and t_{AB}. We identify A with $S_i = 1$ and B with $S_i = -1$. The Hamiltonian then is:

$$\mathcal{H} = -\sum_{\langle i,j\rangle} J_{ij}S_iS_j$$

with $J_{ij} = t_{AB} - \frac{1}{2}(t_{AA} + t_{BB})$. Thus the Ising model describes order-disorder transitions in regard to composition.

③ spin glasses:
now each bond is assigned an individual coupling constant J_{ij} and they are drawn from a random distribution. E.g. one can mix ferromagnetic and anti-ferromagnetic couplings. This is an example for a structurally disordered system, on top of which we can have a thermal order-disorder transition.

④ conformations in biomolecules:
a famous example is the helix-coil transition from biophysics. $S_i = 1$ a hydrogen bond in a DNA-molecules is closed; $S_i = -1$ the bond is open. The phase transition is between a straight DNA-molecule (helix) and a coiled DNA-molecule. Other examples are the oxygen-binding sites in hemoglobin, chemotactic receptors in the receptor fields of bacteria, or the molecules building the bacterial flagellum, which undergos a conformational switch if the flagellum is rotated in the other direction (switch from run to tumble phases).

⑤ neural networks representing the brain:
$S_i = 1$ a synapse is firing, $S_i = -1$ it is resting. The Hopfield model for neural networks is a dynamic version of the Ising model (the coupling constants are learned, with or without supervision) and Boltzmann machines recognise handwriting by using the Ising model. Neural networks have recently become very important again due to the huge success of deep learning (layered neural networks with an intermediate number of layers) and artificial intelligence.

⑥ spread of opinions or diseases:
Spread of opinions, rumours or diseases in a society; these kinds of models are used in socioeconomic physics and epidemiology. If nearest neighbour coupling is sufficiently strong, the system gets *'infected'*. A lot of the spatial model efforts in the context of COVID-19 are based on the Ising model.

In order to decide if the microscopic rules lead to a macroscopic change, one has to introduce an *order parameter*. For a magnetic model like the Ising model, the natural choice is the *magnetisation*:

$$M(K, H) = \left\langle \mu \sum_{i=1}^{N} S_i \right\rangle$$

which is a measure for the averaged spin orientation. $B > 0$ will lead to $M > 0$. If for $B = 0$ we find $M \neq 0$, then the system has spontaneously polarized itself (an example of *spontaneous symmetry breaking*). In the following we will discuss two important results:

① The one-dimensional Ising model shows a phase transition only at $T = 0$ (Figure 6.2).

② The two-dimensional Ising model shows a phase transition at finite temperature T_c (Figure 6.3).

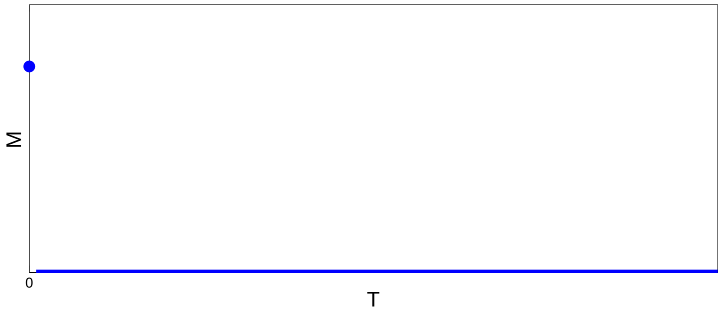

Figure 6.2: 1D Ising model: Magnetisation M as a function of temperature T. For $T \neq 0$ M vanishes and shows a jumping behaviour at $T = 0$.

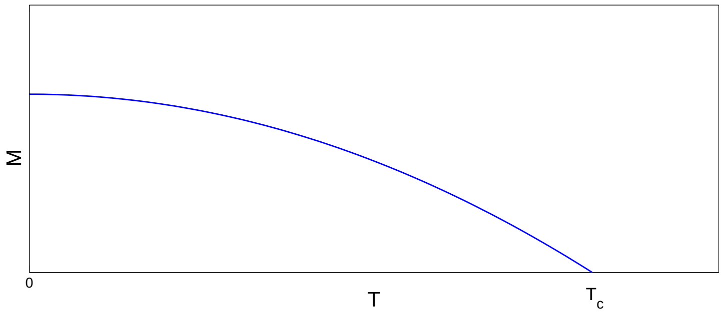

Figure 6.3: 2D Ising model: Magnetisation M as a function of temperature T. For values $T \leq T_c$ M has a finite value.

If M changes in a smooth way at the transition (no jumps), we talk about a phase transition of second order or continuous phase transition. The 2D Ising model is the paradigmatic case for such a transition at the critical temperature T_c. In the region around the critical point, the system has very unusual properties (large fluctuations, universality, scale invariance, critical slowing down). The Ising model below the critical temperature has a phase transition of first order as a function of the magnetic field (with jump in M). The terms first and second order phase transitions go back to Paul Ehrenfest, who suggested a classification of phase transitions in terms of jumps in the derivatives of the chemical potentials; this idea has been discarded, but the two terms first and second order phase transitions have persisted.
Although the 1D Ising model has a jump at $T = 0$. one should not consider it as an example for a discontinuous or first order phase transition; rather it is a version of the 2D Ising model that has been squished into one point. Because it is somehow pathological, some authors prefer to say that it does not have a phase transition at all, as mentioned above, but here we like to talk about one, because many of the relevant signatures are here, like divergent suspectibilities (see below). Although $T = 0$ cannot be reached in experiments, one therefore should see a signature of this change already at finite T.

6.2 The 1D Ising model

In one dimension the Ising model is an *'Ising chain'* of spins (Figure 6.4). With periodic boundary conditions this chain becomes a ring.
Without periodic boundary conditions and considering the external field to vanish, hence $H = 0$, Z_N becomes:

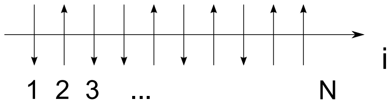

Figure 6.4: 1D Ising model: Ising chain of N spins.

$$\begin{aligned}
Z_N &= \sum_{S_1=\pm1}\sum_{S_2=\pm1}\cdots\sum_{S_N=\pm1} e^{K(S_1S_2+S_2S_3+\ldots+S_{N-1}S_N)} \\
&= \sum_{S_1=\pm1}\cdots\sum_{S_{N-1}=\pm1} e^{K(S_1S_2+\ldots+S_{N-2}S_{N-1})} \underbrace{\sum_{S_N=\pm1} e^{KS_{N-1}S_N}}_{=e^K+e^{-K}=2\cosh K} \\
&= Z_{N-1}\, 2\ \cosh K \\
&= (2\ \cosh K)^{N-1} \underbrace{Z_1}_{=2} \overset{N\gg1}{\approx} \boxed{(2\ \cosh K)^N = Z_N}
\end{aligned}$$

Another way to get rid of the boundary conditions would be to put the Ising chain on a ring, thus we would have N sites and N bonds and the extra 2 from the last site would not appear anyway. At any rate, the free energy expression in the TD-limit becomes, remembering $K = \beta J$:

$$\boxed{F = -k_BTN\ln\left(2\cosh\frac{J}{k_BT}\right)}$$

Because this is an analytical function for finite temperature, one already expects that no phase transition takes place at finite T. We show this by considering spin correlations: For each spin pair we introduce a different coupling constant K_i :

$$\beta\mathcal{H} = -K\sum_{i=1}^{N-1} S_iS_{i+1} \ \rightarrow\ -\sum_{i=1}^{N-1} K_iS_iS_{i+1}$$

$$\begin{aligned}
\Rightarrow\quad \left\langle S_iS_{i+j}\right\rangle &= \frac{1}{Z_N}\sum_{\{S_i\}}(S_iS_{i+j})e^{-\beta\mathcal{H}} \\
&= \frac{1}{Z_N}\sum_{\{S_i\}} \underbrace{(S_iS_{i+j})} \qquad e^{-\sum_{i=1}^{N-1}K_iS_iS_{i+1}} \\
&\qquad\qquad = (S_i\ S_{i+1})(S_{i+1}\ S_{i+2})\ \ldots\ (S_{i+j-1}\ S_{i+j}) \\
&\qquad\qquad = S_i\ \underbrace{S_{i+1}S_{i+1}}_{=1}\ \underbrace{S_{i+2}\,.}_{=1}\ .\ \underbrace{.\,S_{i+j-1}}_{=1}\,S_{i+j} \\
&\qquad\qquad = \partial_{K_i}\,\partial_{K_{i+1}}\ \ldots\ \partial_{K_{i+j-1}}
\end{aligned}$$

Z_N then can be calculated iteratively as above:

$$Z_N = 2^N \prod_{i=1}^{N-1} \cosh K_i$$

$$\Rightarrow \quad \left\langle S_i S_{i+j} \right\rangle = \frac{\cosh K_1 \, ... \, \sinh K_i \, ... \, \sinh K_{i+j-1} \, ... \, \cosh K_{N-1}}{\cosh K_1 \, ... \, \cosh K_i \, ... \, \cosh K_{i+j-1} \, ... \, \cosh K_{N-1}} = \prod_{k=1}^{j} \tanh K_{i+k-1}$$

$$\forall i: \; K_i = K \quad \Rightarrow \quad \left\langle S_i S_{i+j} \right\rangle = (\tanh K)^j$$

The resulting spin correlations are shown in Figure 6.6. Despite the short-ranged interaction - we only consider nearest neighbour interactions - a longer ranged correlation emerge from the statistical average, which decays exponentially with distance:

$$\boxed{\left\langle S_i S_{i+j} \right\rangle = (\tanh K)^j = \left(e^{\ln \tanh K}\right)^j = e^{j \ln \tanh K} = e^{-j/\xi}}$$

where we have defined the *correlation length*

$$\boxed{\xi = -(\ln(\tanh(K)))^{-1}}$$

Because $\tanh(K) \leq 1$, the correlation length is positive and finite, except at $K = \infty$ ($T = 0$), where it diverges (in an exponential fashion with T in 1D, as a power law in 2D). This is one of the most important signatures of a critical point: the system becomes correlated over its whole size.

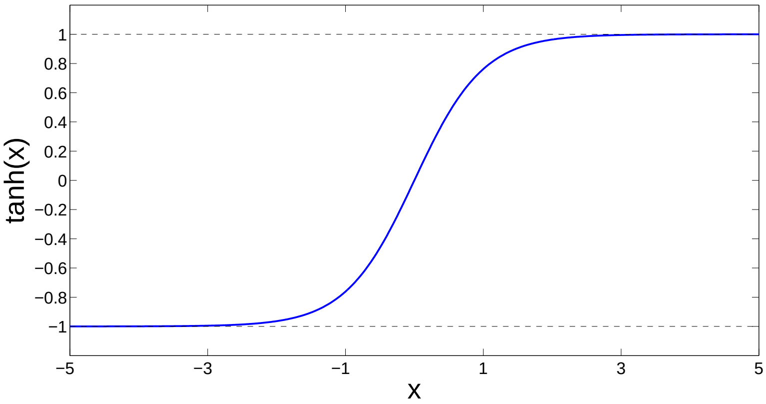

Figure 6.5: $\tanh(x)$ as a function of x. For $x > 0$ $\tanh(x) > 0$; $\tanh(x) \overset{x \to \pm\infty}{\to} \pm 1$.

Because the system is homogeneous:

$$\forall i: \quad \langle S_i \rangle = \langle S \rangle$$

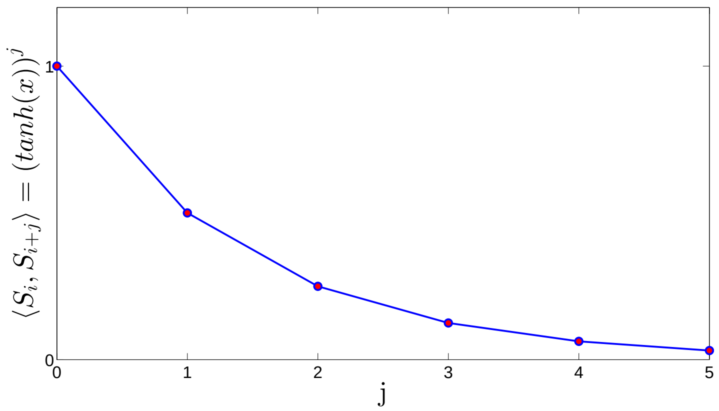

Figure 6.6: $\langle S_i, S_{i+j}\rangle = (\tanh(x))^j$ as a function of j. As can be seen in Figure 6.5 , $\tanh(x) > 0$ for $x > 0$. For the plot $\tanh(x)$ was taken to be 0.5. For $T = 0$, this curve would not decay.

$$\Rightarrow \quad M = \mu N \langle S\rangle$$

$$\langle S_i S_{i+j}\rangle \overset{j\to\infty}{\to} \langle S_i\rangle \langle S_{i+j}\rangle = \langle S\rangle^2$$

$$M^2 = \mu^2 N^2 \lim_{j\to\infty} \langle S_i S_{i+j}\rangle = \begin{cases} \mu^2 N^2 & T = 0 \\ 0 & T > 0 \end{cases}$$

At finite T no spontaneous magnetisation occurs. At $T = 0$ we have a phase transition (compare Figure 6.2). For $T \to 0$ we have first made this limit and then the thermodynamic limit $N \to \infty$.
In summary, the exact solution of the Ising chain at cero magnetic field shows that it has a critical phase transition at $T = 0$. At finite T, spontaneous magnetization does not occur and correlations decay exponentially with distance. In the 2D Ising model, the critical point will be at a finite temperature.

6.3 Transfer matrix

Next we investigate the case $H \neq 0$ using the transfer matrix method introduced by Onsager for the 2D Ising model:

$$\beta\mathcal{H} = -K \sum_{\langle ij\rangle} S_i S_j - H \sum_i S_i$$

We now use periodic boundary conditions: $S_{N+1} = S_1$ (compare Figure 6.7). In the thermodynamic limit $N \to \infty$, boundaries become irrelevant.

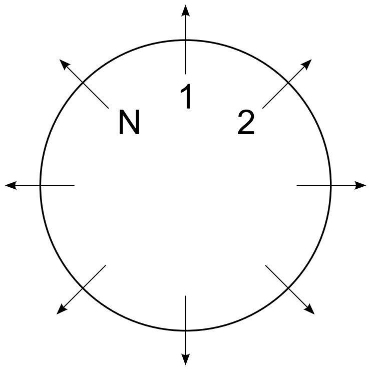

Figure 6.7: With periodic boundary conditions the one-dimensional Ising chain becomes a ring.

We define a *'transfer function'* :

$$
\begin{aligned}
T_{i,i+1} &:= e^{KS_iS_{i+1}+\frac{1}{2}H(S_i+S_{i+1})} \\
\Rightarrow \quad e^{-\beta\mathcal{H}} &= T_{1,2}\ T_{2,3}\ ...\ T_{N,1}
\end{aligned}
$$

Each transfer function has four possible values which define a symmetric *'transfer matrix'*:

$$
T = \begin{pmatrix} e^{K+H} & e^{-K} \\ e^{-K} & e^{K-H} \end{pmatrix}
$$

In quantum mechanical notation:

$$
|S_i = +1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad |S_i = -1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}
$$

$$
\Rightarrow \quad T_{i,i+1} = \langle S_i|\, T\, |S_{i+1}\rangle
$$

$$
\begin{aligned}
\Rightarrow \quad Z_N &= \sum_{\{S_i\}} e^{-\beta\mathcal{H}} \\
&= \sum_{\{S_i\}} \langle S_1|\, T\, |S_2\rangle\, \langle S_2|\, T\, |S_3\rangle\ ...\ \langle S_N|\, T\, |S_1\rangle \\
&= \sum_{S_1=\pm 1} \langle S_1|\, T^N\, |S_1\rangle \\
&= \left(T^N\right)_{11} + \left(T^N\right)_{22} = tr\left(T^N\right) \\
&= \lambda_1^N + \lambda_2^N
\end{aligned}
$$

We note that solving the Ising model amounts to an eigenvalue problem with λ_i being the eigenvalues of T. This implies:

$$\det\begin{pmatrix} e^{K+H}-\lambda & e^{-K} \\ e^{-K} & e^{K-H}-\lambda \end{pmatrix} = 0$$
$$\left(e^{K+H}-\lambda\right)\left(e^{K-H}-\lambda\right) - e^{-2K} = 0$$
$$\lambda^2 - 2e^K \cosh H\,\lambda + e^{2K} - e^{-2K} = 0$$

$$\begin{aligned} \Rightarrow \quad \lambda_{1,2} &= e^K \cosh H \pm \sqrt{e^{2K}\cosh^2 H - 2\sinh 2K} \\ &= e^K \left[\cosh H \pm \sqrt{\cosh^2 H - 2e^{-2K}\sinh 2K}\right] \end{aligned}$$

Thus we have arrived at an exact solution for the one dimension Ising model with external field:

$$\boxed{Z_N = \lambda_1^N + \lambda_2^N}$$

In the thermodynamic limit, only the larger eigenvalue λ_1 is relevant:

$$Z_N = \lambda_1^N \left(1 + \left(\frac{\lambda_2}{\lambda_1}\right)^N\right) \overset{N\to\infty}{\to} \lambda_1^N$$

For $H = 0$ we get:

$$\lambda_1 = e^K + \sqrt{e^{2K} - (e^{2K} - e^{-2K})} = e^K + e^{-K} = 2\cosh K$$

$$\Rightarrow \quad Z_N = (2\cosh K)^N \quad \text{for } N \gg 1$$

like before from the solution by direct summation (but different boundary conditions). With the full solution we now can calculate any thermodynamic quantity of interest. The thermal equation of state describes the magnetisation:

$$\begin{aligned} M(T,B) &= \frac{1}{Z}\sum_{\{S_i\}} \left(\mu \sum_i S_i\right) e^{-\beta\mathcal{H}} \\ &= \mu\partial_H \ln Z_N = \frac{\mu N}{\lambda_1}\partial_H \lambda_1 \\ &= \frac{\mu N \sinh H}{\sqrt{\cosh^2 H - 2e^{-2K}\sinh 2K}} \end{aligned}$$

We note that $M(T \neq 0, B = 0) = 0$ and no spontaneous magnetisation at finite T occurs. For strong fields, hence

$$M(T \neq 0, B \to \pm\infty) \to \pm\mu N$$

magnetisation saturates. For $T \to 0$ M turns into a step function (compare Figure 6.8).

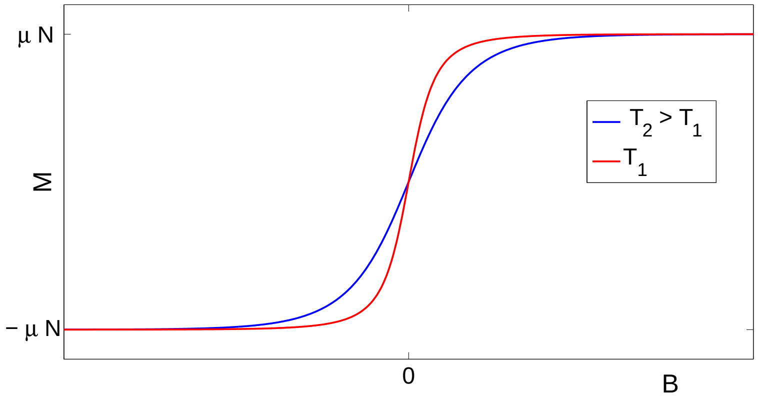

Figure 6.8: The magnetisation M as a function of magnetic field B plotted for different temperatures.

Next we calculate the entropy for $B = 0$:

$$F = -Nk_B T \ln(2\cosh K)$$

$$\Rightarrow \quad S = -\frac{\partial F}{\partial T} = Nk_B \left[\ln(2\cosh K) - K\tanh K\right] \quad \text{(Figure 6.9)}$$

Considering the low and high temperature limits:

$$S \overset{T\to\infty,\, K\to 0}{\longrightarrow} Nk_B \ln 2$$

$$S \overset{T\to 0,\, K\to\infty}{\longrightarrow} Nk_B(K-K) = 0$$

where we recovered the third law of thermodynamics.

From this we calculate the heat capacity in absence of a field:

$$c_B = T\left.\frac{\partial S}{\partial T}\right|_{B=0} = k_B \frac{K^2}{\cosh^2 K} \quad \text{(Figure 6.10)}$$

Finally we consider the *isothermal susceptibility* (needs H):

$$\chi_T = \frac{1}{N}\left.\frac{\partial M}{\partial B}\right|_T = \frac{\beta\mu}{N}\partial_H M$$
$$= \frac{\beta\mu^2}{N}\partial_H^2 \ln Z_N = \beta\mu^2\partial_H^2 \ln\lambda_1$$

In the special case $B = 0$, χ_T becomes:

$$\boxed{\chi_T = \frac{\beta\mu^2}{(1-\tanh K)}}$$

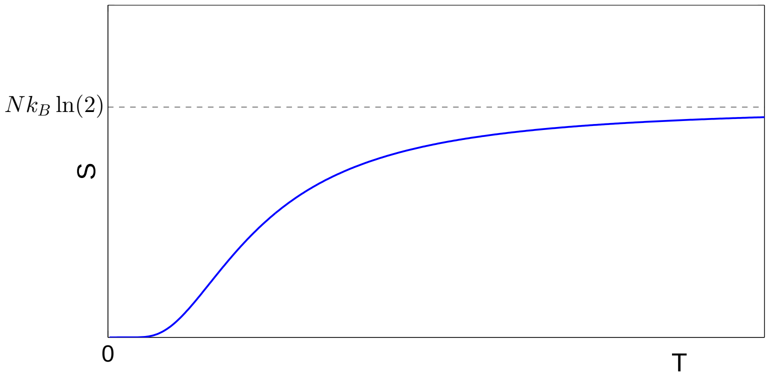

Figure 6.9: The entropy S as a function of temperature T. For high temperatures S approaches $S_0 = Nk_B \ln 2$ asymptotically.

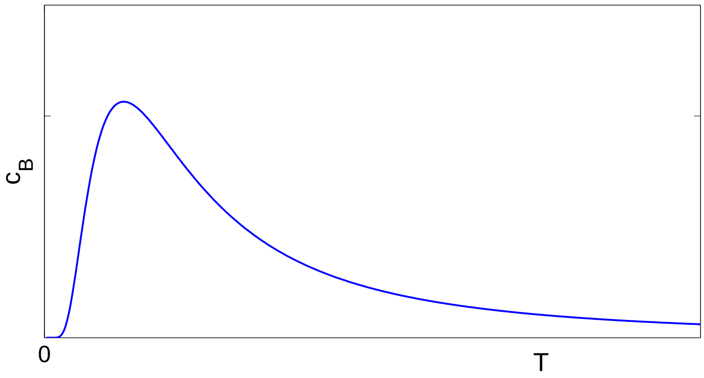

Figure 6.10: The heat capacity c_B as a function of temperature shows a similar shape as the one for the two state model.

$$\chi_T \overset{T\to\infty}{\to} \frac{1}{T} \quad \text{law of Curie}$$

In Figure 6.11 χ_T is plotted as a function of temperature.
We also note an interesting relation between susceptibility (a response function like viscosity) and spin correlations (describing thermally activated *'fluctuations'*):

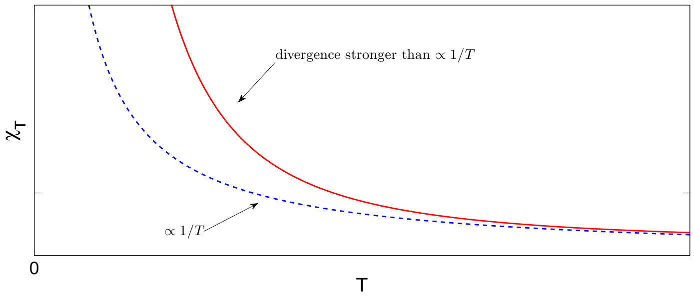

Figure 6.11: The susceptibility χ_T as a function of temperature. χ_T diverges for $T \to 0$, hence for T approaching the phase transition. In general this is a typical signature of a phase transition. For large temperatures: $\chi_T \propto 1/T$ (law of Curie).

$$\chi_T = \frac{1}{N} \left.\frac{\partial M}{\partial B}\right|_T = \frac{\beta\mu}{N} \partial_H \left\{ \frac{1}{Z} \sum_{\{S_i\}} \left(\mu \sum_i S_i \right) e^{K \sum S_i S_j + H \sum S_i} \right\}$$

$$= \frac{\beta\mu^2}{N} \frac{1}{Z} \sum_{\{S_i\}} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} S_i S_j \right) e^{-\beta\mathcal{H}}$$

$$= \boxed{\frac{\beta\mu^2}{N} \left(\sum_{i=1}^{N} \sum_{j=1}^{N} \langle S_i S_j \rangle \right) = \chi_T} \quad \text{'fluctuation dissipation theorem' for the Ising model}$$

For the one-dimensional Ising model and the limit $N \to \infty$ the result becomes:

$$\chi_T = \frac{\beta\mu^2}{N} N \sum_{j=0}^{\infty} (\tanh K)^j = \beta\mu^2 \frac{1}{1 - \tanh K}$$

which is the same as above.

As $-1 \leq \langle S_i S_j \rangle \leq 1$ this implies χ_T can only diverge when

① $N \to \infty$ (thermodynamical limit)

② The range of the correlations must diverge, such that infinitively many terms give a non-finite contribution. Therefore phase transitions are related to a divergence of the *'correlation length'*. This implies that microscopic details become irrelevant because the system becomes correlated on a macroscopic scale (*'critical fluctuations'*).

6.4 Renormalization of the Ising chain

Above we have seen that the correlation length ξ diverges at the critical point at $T = 0$. We turn this observation around and ask if we can identify a critical point by the fact that there the correlation length has to diverge. This feature is also known as *scale invariance*: the system should look the same irrespective of on which scale we investigate it. We therefore perform the following procedure: we coarse-grain the Ising model step-by-step by decimating half of its spins in each step and adjusting the model and its parameters (*renormalization*), as shown in Figure 6.12. We can do this infinitely many times in the thermodynamic limits, because the number of spins stays infinite. We call the set of all these scale transformations the *renormalization group* (RG), because it has the group property of associativity (because strictly speaking there is no inverse element, RG is actually only a semi-group and the name is a misnomer). Obviously, in each RG-step the correlation length should decrease because we move spins closer to each other, except at the critical point, where correlations should persist. Thus the critical point should be a fixed point under the RG-transformation. In detail, it should be unstable or repulsive, because for all other cases, correlation should decay. These ideas of real space renormalization goes back to Leo Kadanoff and later was refined by Ken Wilson, who won the Nobel Prize for RG. Alternatively one can also perform renormalization in momentum space by integrating out a momentum shell of small wavelength modes and then follow the same arguments. This however first requires a Fourier transform (either on the lattice or in a continuum version, like the ϕ^4 field theory), so for time reasons we do not discuss this version of RG here.

The RG of the Ising chain can be performed analytically and gives very good results. We do this by separating spins with even and odd indices and integrating out the odd ones:

$$Z_N = \sum_{S_2=\pm 1} \sum_{S_4=\pm 1} \cdots \sum_{S_N=\pm 1} \left(\sum_{S_1=\pm 1} \sum_{S_3=\pm 1} \cdots \sum_{S_{N-1}=\pm 1} e^{-\beta \mathcal{H}} \right)$$

where we assume N to be even, although this does not matter in the thermodynamic limit. We put the Ising chain onto a ring, so that S_1 has nearest neighbors S_N and S_2, and integrate out this spin:

$$\sum_{S_1=\pm 1} e^{K S_1 (S_N + S_2) + H S_1} = 2\cosh(K(S_N + S_2) + H)$$

If we now consider the remaining spins, we see that this term can have three different outcomes, for the combination $(S_N, S_2) = (+1, +1), (-1, +1)/(+1, -1)$ (the two mixed cases give the same result) and $(-1, -1)$. We now guess a new Hamiltonian which should be as close as possible to the old one, but also has to accomodate these three possible outcomes. The following three-parameter model works:

$$2\cosh(K(S_N + S_2) + H) = \zeta_0 e^{K' S_N S_2 + \frac{\delta H}{2}(S_N + S_2)}$$

Here K' is a renormalized coupling constant and δH is a new contribution to the magnetic field (the factor of 2 arises because the same contribution will also arise from the

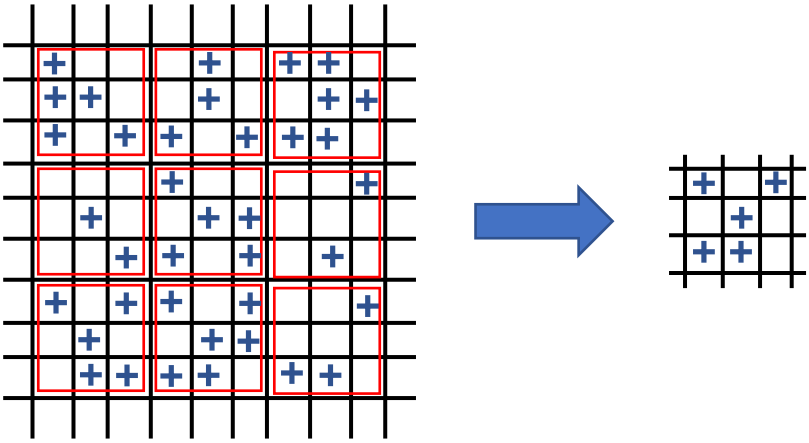

Figure 6.12: Kadanoff real space renormalization: b times b spins (here $b = 3$) are grouped into a block and the new block is assigned a new spin, e.g. by the majority rule. This coarse-graining procedure should decrease correlation except at the critical point.

other side, that is when summing over S_{N-1} and S_3, until now we only discuss S_1). Considering the three possible outcomes we get three equations:

$$\begin{aligned} 2\cosh(2K+H) &= \zeta_0 e^{K'+\delta H} \\ 2\cosh(H) &= \zeta_0 e^{-K'} \\ 2\cosh(-2K+H) &= \zeta_0 e^{K'-\delta H} \end{aligned}$$

We rearrange each equation to

$$\begin{aligned} e^{\delta H} &= 2\cosh(2K+H)\frac{e^{-K'}}{\zeta_0} \\ \zeta_0 &= 2\cosh(H)e^{K'} \\ e^{-\delta H} &= 2\cosh(2K-H)\frac{e^{-K'}}{\zeta_0} \end{aligned}$$

By taking product and quotient of the first and third equations, one can eliminate δH and K', respectively, and solve for the other (replacing ζ_0 by the second equation). We also replace ζ_0 by a new definition through $(\zeta_0)^{1/2} = e^g$. We thus find:

$$\begin{aligned} K' &= \frac{1}{4} \ln \frac{\cosh(2K+H)\cosh(2K-H)}{\cosh^2(H)} \\ H' &= H + \delta H = H + \frac{1}{2} \ln \frac{\cosh(2K+H)}{\cosh(2K-H)} \\ g &= \frac{1}{2} \ln \zeta_0 = \frac{1}{8} \ln(16 \cosh(2K+H)\cosh(2K-H)\cosh^2(H)) \end{aligned}$$

If we perform the same procedure with every odd spin, then we finally have

$$Z_N = \sum_{S_2=\pm1} \sum_{S_4=\pm1} \dots \sum_{S_N=\pm1} e^{Ng+K'\sum_i S_{2i}S_{2i+2}+H'\sum_i S_{2i}}$$

By our definition of g, it now comes with a prefactor N and therefore has the meaning of an absolute free energy gain per spin per transformation. The second part on the RHS is simply the partition sum for a system in which have of the spins have been eliminated, but with renormalized parameters. We thus have

$$Z_N(K,H) = e^{Ng(K,H)} Z_{N/2}(K',H')$$

Obviously this procedure can now be iterated infinitely many times, defining a RG-flow of the model parameters, e.g. K', K'', etc (or $K_0 = K$, $K_1 = K'$, $K_2 = K''$, etc) for the coupling constant. For $H = 0$, we see that $H' = 0$, that is a cero field stays cero. The coupling constant then flows as

$$\boxed{K' = \frac{1}{2} \ln \cosh(2K)}$$

In Figure 6.13(a) we plot the function $K'(K)$. Because this function is always smaller than K (it starts $\sim K^2$ and asymptotically approaches K), the flow (shown by the stair function) goes towards $K = 0$, which is a stable fixed point under the RG-transformation. $K = 0$ corresponds to $T = \infty$ and therefore corresponds to an interaction-free system. Importantly, there is another fixed point, namely at $K = \infty$. Because this corresponds to $T = 0$, it is the critical point we already know from the exact solution. This fixed point is repulsive as expected: in the moment we go away from $T = 0$, coarse-graining will decrease correlations until they are gone. Thus the RG correctly predicts the existence of a critical point at $T = 0$. In Figure 6.13(b) we also show the flow with non-cero field. We see that the RG-flow increases H, which makes sense, because at the same time K decreases, showing again that it goes away from the strongly interacting (critical) system at $K = \infty$ to a non-interacting situation at $K = 0$.

The RG-procedure also gives as a neat way to calculate the free energy. We non-dimensionalise and normalize for a single spin:

$$f = \frac{-\beta F}{N} = \frac{1}{N} \ln Z_N = g(K,H) + \frac{1}{N} \ln Z_{N/2}(K',H') = \sum_{i=0}^{\infty} (\frac{1}{2})^i g(K_i,H_i)$$

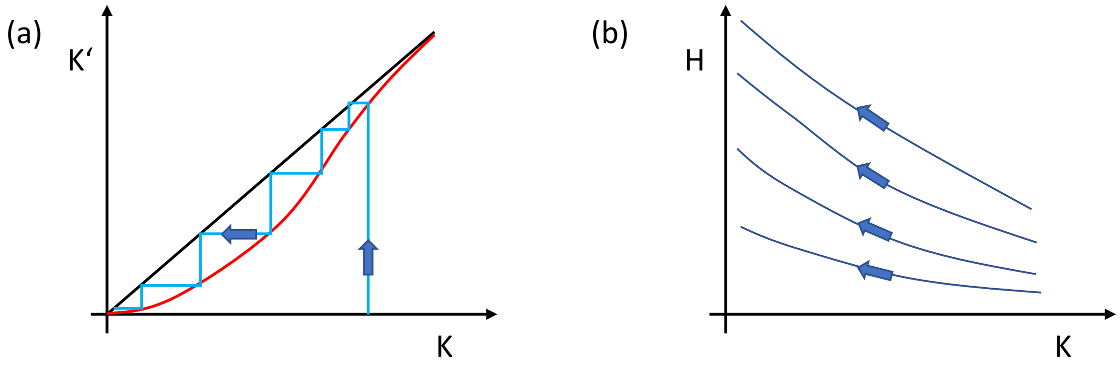

Figure 6.13: (a) Iteration of the RG-transformation at $H = 0$ decreases K towards the stable (attractive) $K = 0$ ($T = \infty$) fixed point without interactions. The other fixed point at $K = \infty$ ($T = 0$) is the unstable (repulsive) critical point. (b) For finite field, H increases under RG, again flowing away from the critical point.

where the last expression is obtained by iteration. For $H = 0$, we have

$$g = \frac{1}{2}\ln 2 + \frac{1}{4}\ln\cosh(2K)$$

where the second term can be neglected once K gets close to 0. The remainder of the series is like a geometrical series and converges quickly. The overall result can be compared with the exact value $f = \ln(2\cosh(K))$ obtained above from the analytical solution of the Ising chain.

6.5 Renormalization of the 2D Ising model

For the 2D Ising model, many different RG-schemes have been invented. Here we describe the most simplest one. The 2D square lattice is bipartite and we divide it into two sublattices, each with a new lattice constant that is $b = \sqrt{2}$ times larger than the old one, as shown in Figure 6.14(a). Spin S_0 has four nearest neighbors S_1, S_2, S_3 and S_4. Motivated by our procedure in 1D, where we were able to keep the same form of the Hamiltonian, we would like to write

$$\sum_{S_0=\pm 1} e^{kS_0(S_1+S_2+S_3+S_4)} = 2\cosh(K(S_1+S_2+S_3+S_4)) = e^{(K'/2)(S_1S_2+S_2S_3+S_3S_4+S_4S_1)}$$

where the factor of 2 takes into account that the same contribution will come from the spin on the other side of these bonds. Unfortunately, this is impossible, because now there are too many possibilities for the outcomes. The only way to proceed is to consider

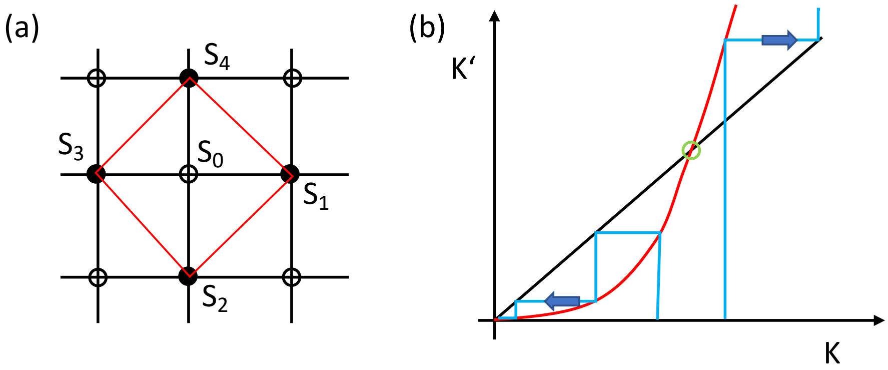

Figure 6.14: (a) The 2D square lattice can be separated into two sublattice. We integrate out S_0 which has nearest neighbors only in the other sublattice. (b) RG-flow of the coupling constant K. This time we have a (unstable, repulsive) fixed point at finite K. Smaller and larger K-values flow away from this fixed point. Thus RG predicts a critical point at finite temperature, in agreement with the exact solution by Onsager.

new types of interactions, that is to make the space of the Hamiltonian larger. We now write

$$2\cosh(K(S_1+S_2+S_3+S_4)) = \zeta_0 e^{(K'/2)(S_1S_2+S_2S_3+S_3S_4+S_4S_1)} e^{K_1(S_1S_3+S_2S_4)} e^{L(S_1S_2S_3S_4)}$$

where the K_1-interaction is a next nearest neighbor interaction and the L-interaction is a four-spin interaction. Like in the 1D case, we consider all possible outcomes. There are four cases one has to distinguish (e.g. (+,+,+,+), (+,+,+,-), (+,+,-,-) and (+,-,+,-)), justifying the four model parameters and leading to four equations:

$$\begin{aligned} 2\cosh(4K) &= \zeta_0 e^{2K'} e^{2K_1} e^{L} \\ 2\cosh(2K) &= \zeta_0 e^{-L} \\ 2 &= \zeta_0 e^{-2K_1} e^{L} \\ 2 &= \zeta_0 e^{-2K'} e^{2K_1} e^{L} \end{aligned} \tag{6.1}$$

By combining these equations we find the flow equations:

$$\begin{aligned}
K' &= \frac{1}{4}\ln\cosh(4K) \\
K_1 &= \frac{1}{8}\ln\cosh(4K) \\
L &= \frac{1}{8}\ln\cosh(4K) - \frac{1}{2}\ln\cosh(2K) \\
\zeta_0 &= 2(\cosh(2K))^{1/2}(\cosh(4K))^{1/8}
\end{aligned} \tag{6.2}$$

The problem is that we cannot iterate these equations because our starting point was more simple. We therefore neglect the four-spin interaction and project the next nearest neighbor interactions on the nearest neighbor interactions by assuming the case that all spins are parallel:

$$K'\sum_{nn} S_i S_j + K_1 \sum_{nnn} S_i S_j = \bar{K}' \sum_{nn} S_i S_j$$

Note that this step is not rigorous, but it reflects the fact that there will be an extra term that will increase the value for K', even if the projection would be more accurate. Now we can iterate and get a full RG. Considering all spins to be parallel gives us

$$\boxed{\bar{K}' = K' + K_1 = \frac{3}{8}\ln\cosh(4K)}$$

As shown in Figure 6.14(b), this equation does give a non-trivial fixed point at finite $K_c = 0.50689$. Thus RG predicts a phase transition for the 2D Ising model (exact value $K_c = 0.440687$, see below). It also gives a good prediction for the critical exponent $\nu = 1.07$ (exact value $\nu = 1$).

6.6 The Peierls argument

Starting around 1933, Peierls published scaling arguments why a phase transition should occur in 2D as opposed to 1D. Here we report a few of these kinds of arguments to demonstrate their spirit. Note that their validity also comes from the fact that thanks to Onsager, we have an exact solution and thus can check back if they describe the core of the problem or not.

Simple argument for 1D

We consider an Ising chain with all spins up and then select a few neighboring spins and flip the whole island over. This creates two *domain walls* (also called *grain boundaries* or *defects*) in the chain. The change in energy is

$$\Delta E = 2 \cdot 2J$$

because there are two defects, each with an energy penalty $2J$. The change in entropy corresponds to the number of ways to choose the positions of the two defects:

$$\Delta S = k_B \ln \frac{N(N-1)}{2} \approx 2k_B \ln N$$

where we assume the number of lattice sites $N \gg 1$ in the thermodynamic limit. Thus the change in free energy reads

$$\Delta F = 4J - 2k_B T \ln N < 0$$

for any temperature T in the thermodynamic limit. This means that it is always favorable to create grain boundaries due to entropic reasons and a phase transition to order cannot occur at finite temperature.

More complex argument for 1D

We now look at an arbitrary number of domain walls, not only at one island with two of them. We introduce the number of such domain walls M and write the free energy in the domain wall picture:

$$F = 2JM - k_B T \ln \binom{N}{M}$$

In the thermodynamic limit and with the Stirling formula we get

$$\frac{F}{N} = 2Jx + k_B T(x \ln x + (1-x)\ln(1-x))$$

where $x = M/N$ is the domain wall density. If we minimize F for x we get

$$x_{eq} = \frac{1}{e^{2J/k_B T} + 1}$$

thus at finite T there is always a finite domain wall density and correlations decay over a finite distance. Moreover the system will not feel the effect of the boundary conditions. Only at $T = 0$ we have $x_{eq} = 0$, because then entropy does not matter.

Simple argument for 2D

We now want to make the simple argument for 2D rather than for 1D. We immediately encounter the problem that now there are two processes we have to account for: where to place the domain walls, and which shape to assign to them. With some intuition, we anticipate that shape fluctuations are now more important than where the islands are located. Thus we consider one island of down spins in a sea of up spins. The change in energy is

$$\Delta E = L \cdot 2J$$

where L is the contour length of the domain. A crude estimate for the number of possible shapes is 3^L, assuming a random walk on a 2D cubic lattice and neglecting intersections and the fact that it has to close onto itself (at each lattice site, there are three possibilities to proceed). Thus for entropy we have

$$\Delta S = k_B \ln 3^L \, .$$

Together we get

$$\Delta F = L(2J - k_B T \ln 3)$$

and thus $\Delta F < 0$ only for $T > T_c = 2J/(\ln 3 k_B)$ even in the thermodynamic limit $L \to \infty$. Thus this simple argument predicts that in 2D a phase transition can take place at finite T, and the reason is a feature that is only present in two and higher dimensions, namely shape fluctuations of the domain walls.

More complex argument for 2D

Another way to identify a phase transition is to investigate the effects of boundaries. We consider a quadratic field of spins and fix all the ones at the boundary to point up. We then consider the spin in the middle and ask if it keeps the up-preference of the boundary in the TD-limit ($p_+ > 1/2$?). One can show that for sufficiently low but finite T, indeed this happens. This means that correlations do not decay completely and that spontaneous magnetisation can emerge, indicating a phase transition.
We consider the quantity $m = p_+ - p_- = 2p_+ - 1$, which will be finite if spontaneous magnetization exists and vanish otherwise. We can write

$$m = \frac{1}{Z}\sum_{\Sigma_+} e^{-\beta\mathcal{H}} - \frac{1}{Z}\sum_{\Sigma_-} e^{-\beta\mathcal{H}} = \frac{1}{Z}\sum_{\Sigma_+} e^{-\beta\mathcal{H}}(1-\Sigma)$$

The first and second terms are sums over all configurations with a positive and negative central spin, respectively. The basic idea of the newly defined quantity Σ is that each configuration with a positive central spin can be turned into one with a negative central spin by flipping all spins in the surrounding positive domain. Importantly, the difference in energy is simply $2Jl$, where l is the length of the domain wall surrounding this domain. Therefore one can write

$$\Sigma = \sum e^{-2J\beta l} = \sum_{l=4}^{\infty} g(l) e^{-2J\beta l}$$

where the sum is now over all configurations which have been obtained by flipping. In the second step we have rewritten the sum in terms of the length of the boundary. Here $g(l)$ is the number of domains with length l. We note that the minimum l is 4 (one spin flipped) and that one only will have even values ($l = 4, 6, \ldots$), because adding spins one by one to the domain increases l by 2.

In order to prove the polarization, we have to show that Σ can be smaller than 1. We do this by establishing an upper bound for $g(l)$:

$$g(l) < (\frac{l}{4})^2 \cdot 4 \cdot 3^{l-1} \cdot \frac{1}{2l} = \frac{l}{24} 3^l$$

The first term is the maximal area corresponding to the contour length l. The second term is the number of possible paths starting from each point within this area: 4 for the first step and 3 for each additional step (on a 2D simple cubic lattice). The last term corrects for the fact that a path can go in two directions and can start at any point along the contour of a boundary. We now transfer this into an upper bound for Σ:

$$\Sigma < \sum_{l=4}^{\infty} \frac{l}{24} w^l = \frac{1}{24} \sum_{n=2}^{\infty} (2n) w^{(2n)} = \frac{w^4(2-w^2)}{12(1-w^2)^2}$$

where $w = 3e^{-2\beta J}$. We thus obtain $\Sigma < 1$ for $w < w_c = 0.87$. This in turn translates into a critical temperature

$$T_c = \frac{2J}{k_B \ln(3/w_c)} = 1.6 J/k_B$$

The exact result for the 2D Ising model is $T_c = 2.269 J/k_B$ (see below). Thus the Peierls argument does not only prove the transition, but even gives a reasonable first estimate for its value. Note that here we have established only an upper bond for Σ. This does not mean that Σ will be different from 1 above the critical temperature, we only showed that it will certainly become smaller than this value at sufficiently low temperature. Our argument is obviously very crude because we neglect interactions between boundary loops, which will strongly bring down the number of possible paths.

6.7 The 2D Ising model

Several methods of solution have been reported since Onsager's 1944 proof. The main insight of Onsager was that the problem can be mapped onto a fermionic one. Moreover he used the idea of the transfer matrix that we have introduced for the 1D Ising model in a magnetic field. The main idea for 2D is that one writes

$$\beta\mathcal{H} = -K \sum_{r,c} S_{r,c} S_{r+1,c} - K \sum_{r,c} S_{r,c} S_{r,c+1}$$

which means that one differs between interactions in columns c and interactions in rows r. The first can be treated like Ising chains, and the second give non-trivial couplings in the transfer matrix. Using a Jordan-Wigner transformation from Pauli spins to fermionic creation / destruction operators gives a fermionic problem (this was to be expected because we deal with spins $S_i = \pm 1$). A second (canonical) transformation leads to Fourier series, which in the thermodynamic limit lead to integrals (see below). Inbetween one has to calculate the eigenvalues of a 4x4 matrix, as was to be expected from a transfer matrix approach. The fact that the final integrals cannot be solved is actually good, because it means that one can obtain the non-analytical behaviour required for a phase transition to occur.

As mentioned above, other proofs have been reported that exploit the role of random walks, most prominently the one by Vdovichenko which appears in the popular textbook by Landau and Lifshitz. The main idea here is to start from

$$Z = \sum_{\{S_i\}} e^{K \sum_{\langle ij \rangle} S_i \, S_j} = \sum_{\{S_i\}} \prod_{\langle ij \rangle} e^{K S_i \, S_j}$$

so we have a large product of exponential factor. We then use

$$e^{K S_i S_j} = \cosh K + S_i S_j \sinh K = \cosh K (1 + x S_i S_j)$$

with $x := \tanh K$. In this way, the spins in the exponential factors come down and together form a huge polynomial. We write

$$Z = \left(\frac{1}{1 - x^2} \right)^N \sum_{\{S_i\}} \prod_{\langle ij \rangle} (1 + x S_i S_j)$$

Next we realize that the single terms appearing in these large products are all possible combinations of bonds on the lattice and that single spins S_i can occur only to powers of 0, 2 or 4: odd powers drop out because we eventually sum over $S_i = \pm 1$, and there are only up to four bonds connected to one lattice site. So if a spin appears, than it can only come with two or four bonds. This means that we deal with all *closed loops* (or graphs or random walks) on the lattice. Formally one can thus write

$$Z = \left(\frac{2}{1 - x^2} \right)^N \sum_r x^r g_r$$

where the sum if now over all closed graphs of length r and g_r is the multiplicity of this length. From here one needs to do graph theory and make sure that the counting is right. Surprisingly, this can be done analytically and boils down to calculating the eigenvalues of a 4x4 matrix, similar as in Onsager's solution, where the 4 comes from the 4 directions on the lattice.

Irrespective of the method of solution, the end result for the free energy is always the same. Here we report it using $x = \tanh K$ introduced for the second solution:

$$\frac{-\beta F}{N} = \ln 2 - \ln(1 - x^2) + \frac{1}{2} \int_{-\pi}^{\pi} \frac{dq dp}{(2\pi)^2} \ln \left[(1 + x^2)^2 - 2x(1 - x^2)(\cos p + \cos q) \right]$$

We note that the integral has a singularity when the argument of the logarithm becomes 0, which is $x_c = \sqrt{2} - 1$. The corresponding critical temperature with $x_c = \tanh K_c = J / k_B T_c$ then is

$$\boxed{K_c = \frac{1}{2} \ln(1 + \sqrt{2}) \approx 0.4407}$$

$$\Rightarrow \quad \boxed{T_c = 2J / \ln(1 + \sqrt{2}) \approx 2.269 J / k_B}$$

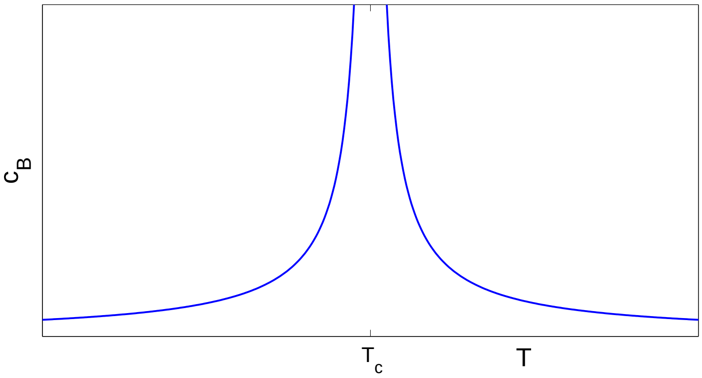

Figure 6.15: c_B as a function of temperature with a divergence at $T = T_c$.

We define the *'reduced temperature'*:

$$t := \frac{T - T_c}{T_c}$$

and *'critical exponents'* for the divergences (for $B = 0$) around T_c:

$$c_B = \begin{cases} (-t)^{-\alpha'} & T < T_c \\ t^{-\alpha} & T > T_c \end{cases}$$

$$M = \begin{cases} (-t)^{\beta} & T < T_c \\ 0 & T > T_c \end{cases}$$

From the exact solution one finds:

① c_B has a logarithmic divergence (Figure 6.15).
$\Rightarrow \quad \alpha = \alpha' = 0$ motivated by $\lim_{\alpha \to 0} \frac{1}{\alpha}(x^{-\alpha} - 1) = -\ln x$

② $M = \left(1 - \sinh^{-4} 2K\right)^{\frac{1}{8}}$ (Figure 6.16)
$\Rightarrow \quad \beta = \frac{1}{8}$
This result was announced by Onsager in 1948 at a conference, but never published by himself.

From the result for the magnetisation (which is the order parameter of the phase transition) one can construct the phase diagram. Figure 6.17 (left) shows the phase diagram in the T-M-plane. Values for the magnetisation in the grey area (two-phase region) cannot be realized in one system, because a self-polarized system jumps to the upper or lower values of M. However, such a magnetisation can be realized by two systems, so the system has to split into two. For example, M = 0 can be realized by two equally

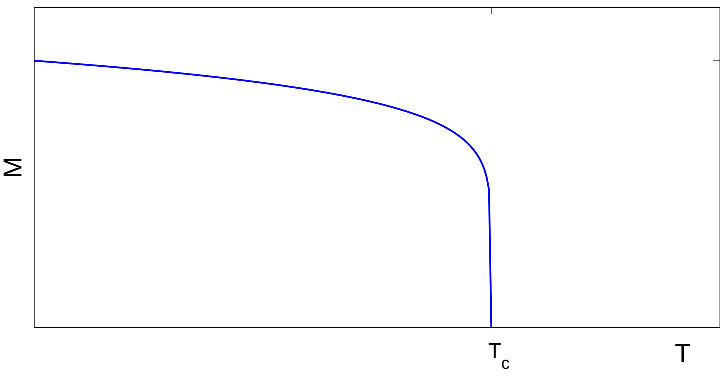

Figure 6.16: M as a function of temperature.

large systems with up and down magnetisation, respectively. Using the lever rule, each desired value of M can be realized. Figure 6.17 (right) shows the phase diagram in the T-B-plane. Now the two-phase region reduce to a line because any small external field will immediately bias the system to up or down. Only for B = 0 phase coexistence can occur.

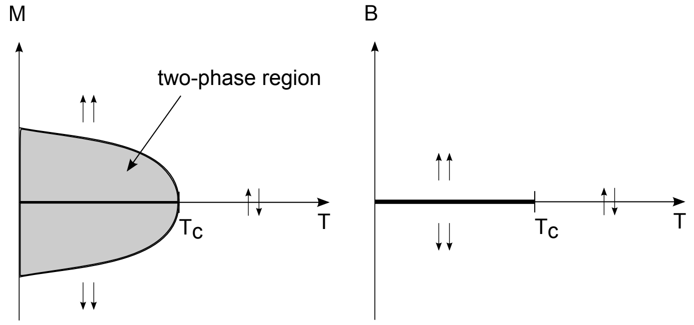

Figure 6.17: Left: Phase diagram with excluded *'two-phase region'* where the system splits into two parts. Right: The two-phase region becomes a line in the $B(T)$ diagram.

6.8 Perturbation theory

In order to understand the mechanisms underlying this phase transition, we now consider the *'mean field theory'* for the Ising model. This theory approximates a system of interacting particles by a system of non-interacting particles. It can be made rigorous by the *'Gibbs-Bogoliubov-Feynman inequality'* and as such is a *'perturbation theory'* (simi-

lar to the *'Hartree-Fock approximation'* in quantum mechanics). In general, it is important to have as many exactly solvable models in Statistical Physics as possible, even if they might be physically not so realistic because they are built around some mathematical trick to solve them. Nevertheless they can be very useful as starting points for perturbative analyses.
We start from a model Hamiltonian $\mathcal{H}_0$ for which an exact solution is known:

$$\mathcal{H}(\lambda) = \mathcal{H}_0 + \lambda \mathcal{H}_1$$

① $\mathcal{H}(\lambda = 0) = \mathcal{H}_0$ reference case

② $\mathcal{H}(\lambda = 1) = \mathcal{H}$ case of interest

③ $\mathcal{H}_1 = \mathcal{H} - \mathcal{H}_0$

$$\Rightarrow -\beta F(\lambda) = \ln \sum_j e^{-\beta E_j(\lambda)} = \ln \left(tr \left\{ e^{-\beta \mathcal{H}(\lambda)} \right\} \right)$$

where $\sum_j = tr$ is the sum over all states.

$$F(0) = F_0$$

$$F(1) = F \quad \text{result of interest}$$

$$\Rightarrow \quad \frac{dF}{d\lambda} = \frac{tr \left\{ \mathcal{H}_1 e^{-\beta(\mathcal{H}_0 + \lambda \mathcal{H}_1)} \right\}}{tr \left\{ e^{-\beta(\mathcal{H}_0 + \lambda \mathcal{H}_1)} \right\}} = \langle \mathcal{H}_1 \rangle (\lambda)$$

$$\begin{aligned} \frac{d^2F}{d\lambda^2} &= -\beta \left\{ \frac{tr \left\{ \mathcal{H}_1^2 e^{-\beta(\mathcal{H}_0 + \lambda \mathcal{H}_1)} \right\}}{tr \left\{ e^{-\beta(\mathcal{H}_0 + \lambda \mathcal{H}_1)} \right\}} - \left(\frac{tr \left\{ \mathcal{H}_1 e^{-\beta(H_0 + \lambda H_1)} \right\}}{tr \left\{ e^{-\beta(H_0 + \lambda H_1)} \right\}} \right)^2 \right\} \\ &= -\beta \left(\langle \mathcal{H}_1^2 \rangle - \langle \mathcal{H}_1 \rangle^2 \right) = -\beta \left\langle \left(\mathcal{H}_1 - \langle \mathcal{H}_1 \rangle \right)^2 \right\rangle \leq 0 \end{aligned}$$

$$\Rightarrow \quad F(\lambda) \leq F(0) + \lambda \left. \frac{\mathrm{d}F}{\mathrm{d}\lambda} \right|_{\lambda=0}$$

$$\overset{\lambda=1}{\Rightarrow} \quad \boxed{F \leq F_u = F_0 + \langle \mathcal{H}_1 \rangle_0} \quad \text{Bogoliubov inequality}$$

A visualisation of the Bogoliubov inequality is sketched in Figure 6.18. Note that the real F is everywhere concave, not only at $\lambda = 0$, so we can use $\lambda = 1$ without problems. In order to optimize the approximation one minimizes the upper bound with respect to the free model parameters. The modern master of this type of perturbation theory was Richard Feynman.

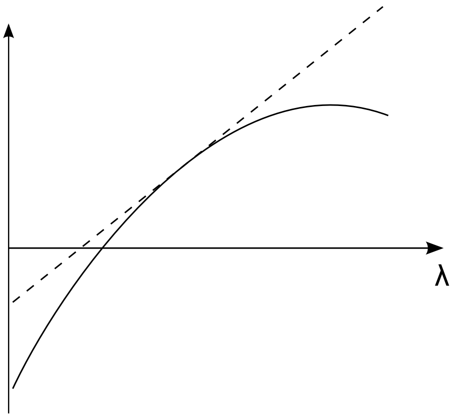

Figure 6.18: Sketch visualising the Bogoliubov inequality: $F(\lambda)$ (solid line) $\leq F(0) + \lambda \frac{dF}{d\lambda}(0)$ (dashed line).

6.9 Mean field theory for the Ising model

We consider no external field:

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} S_i S_j$$

However, we note that a spontaneous magnetization looks like there was an effective magnetic field. We therefore choose as our unperturbed reference Hamiltonian

$$\mathcal{H}_0 = -B \sum_i S_i$$

where we set $\mu = 1$ for convenience and have introduced an effective magnetic field B. For $\mathcal{H} = \mathcal{H}_0$ we know the free energy expression:

$$F_0 = -Nk_BT \ln \left(\underbrace{e^{\beta B} + e^{-\beta B}}_{=2\cosh \beta B} \right)$$

The Bogoliubov inequality then states

$$\begin{aligned} F &\leq F_0 + \langle \mathcal{H} - \mathcal{H}_0 \rangle_0 \\ &= -Nk_BT \ln\left(2\cosh(\beta B)\right) - J \underbrace{\sum_{\langle i,j \rangle} \left\langle S_i S_j \right\rangle_0}_{N(z/2)\langle S \rangle_0^2} + B \underbrace{\sum_i \langle S_i \rangle_0}_{N\langle S \rangle_0} = F_u \end{aligned}$$

Here z is the number of nearest neighbours and we have to correct with a factor of 2 so that we count each bond only once (compare Figure 6.1).

$$\langle S \rangle_0 = \frac{e^{\beta B} - e^{-\beta B}}{e^{\beta B} + e^{-\beta B}} = \tanh \beta B$$

We now fix B such that the upper bound F_u becomes minimal:

$$0 = \frac{1}{N}\frac{dF_u}{dB} = -\langle S\rangle_0 - Jz\langle S\rangle_0 \frac{d\langle S\rangle_0}{dB} + \langle S\rangle_0 + B\frac{d\langle S\rangle_0}{dB}$$
$$\Rightarrow \quad \boxed{B = Jz\langle S\rangle_0 = Jz\tanh\beta B}$$

Note that a factor of 2 has canceled here. We note that our central result is a self-consistent relation for the effective field B. We could have obtained this result directly from a mean field reasoning, but it is more rigorous to derive it from the Bogoliubov inequality.

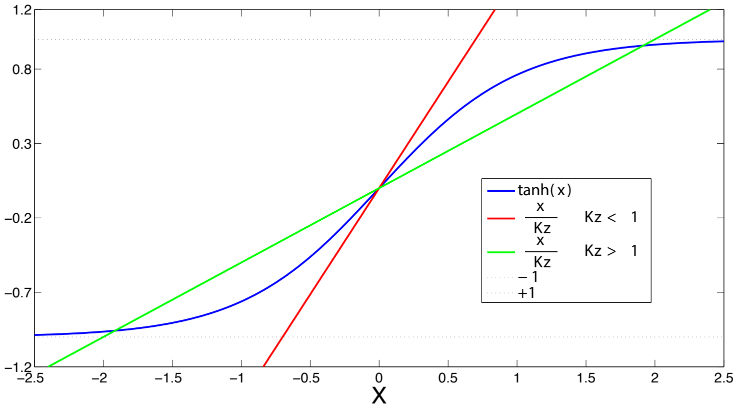

Figure 6.19: $M(x) = \tanh(x)$ (blue) as a function of x. For $Kz < 1$ there is only one intersection with $g(x) = \frac{x}{Kz}$ (red) at $x = 0$. For $Kz > 1$ there is also an intersection with $g(x) = \frac{x}{Kz}$ (green) at finite x.

We define $x = \beta B$ and have a look at the intersection of $f(x) = \tanh(x)$ and $g(x) = \frac{x}{Kz}$ (Figure 6.19). We note:

① $Kz < 1 \quad \Rightarrow$ only intersection at $x = 0$

② $Kz > 1 \quad \Rightarrow$ also two interactions at finite $x \neq 0$

$$\Rightarrow \quad K_c = \frac{1}{z} \quad \Rightarrow \quad \boxed{T_c = \frac{zJ}{k_B}}$$

For the two-dimensional Ising model with cubic arrangement:

$$z = 4 \quad \Rightarrow \quad K_c = \frac{1}{4} = 0.25$$

Compare exact solution: $K_c = 0.4407$. Obviously the mean field theory is just a crude approach because it predicts a phase transition in $\underline{\text{any}}$ dimension d. It becomes exact for $d \to \infty$.
How does magnetisation behave below T_c? Assuming a small magnetisation $m = \langle S \rangle$ just below T_c, we can perform a Taylor expansion:

$$m = \tanh \beta B \approx \beta B - \frac{1}{3} (\beta B)^3$$

where $\beta B = zKm$. Therefore one power of m cancels and we have

$$m^2 = 3\frac{(zK-1)}{(zK)^3} = 3\frac{(\frac{zJ}{k_B T} - 1)}{(\frac{zJ}{k_B T})^3} = 3T^3 \frac{(\frac{zJ}{k_B T} - 1)}{(\frac{zJ}{k_B})^3}$$

As above, m vanishes at $T_c = zJ/k_B$ and we can write

$$m^2 = 3\left(\frac{T}{T_c}\right)^3 \left(\frac{T_c}{T} - 1\right) = 3\left(\frac{T}{T_c}\right)^2 \left(\frac{T_c - T}{T_c}\right)$$

Taking the positive square root finally gives

$$m = \sqrt{3}\left(\frac{T}{T_c}\right)\left(\frac{T_c - T}{T_c}\right)^{1/2} \approx \sqrt{3}\left(\frac{T_c - T}{T_c}\right)^{1/2}$$

to lowest order in the expansion (note that $T = T_c - (T_c - T)$). We see that our approximative calculation yields a critical exponent $\beta = \frac{1}{2}$ (compare exact solution $\beta = \frac{1}{8}$).

6.10 Monte Carlo computer simulations of the Ising model

Exact enumerations are not feasible for large systems. Consider that for a 50 x 50 2D Ising system we would have already $2^{2500} = 10^{753}$ states. Therefore it is better to use *importance sampling*. Here only the relevant configurations are sampled. For the canonical ensemble this means that we want to sample according to the Boltzmann distribution

$$p_i = \frac{1}{Z} e^{-\beta E_i}$$

We now look for a procedure which effectively generates this distribution.
The standard procedure for this purpose is a *Monte Carlo simulations* (the name results from the use of random numbers, which are also at the heart of gambling at Monte Carlo). In a Monte Carlo simulation, we generate a series of configurations (*Markov chain*, that is a memory-less process) such that:

1. Any configuration can be reached in principle.
2. Averaging over all configurations in the Markov chain amount to doing the average with exact enumeration.

For the Ising model, the simplest procedure is to flip single spin at random (if in average we have tried to flip every spin once, we call this a Monte Carlo step). We compare two such configurations i and j with:

$$\frac{p_i}{p_j} = e^{-\beta(E_i - E_j)}$$

We define $p_{i\to j}$ to be the *transition probability* for one spin to go from state i to j.

$$\Rightarrow \quad \sum_j p_{i\to j} = 1$$

We now require that locally we have *detailed balance* (follows from time reversal invariance):

$$\frac{p_{i\to j}}{p_{j\to i}} = \frac{p_j}{p_i} = e^{-\beta(E_j - E_i)}$$

$$\Rightarrow \quad p_i = \left(\sum_j p_{i\to j}\right) p_i = \sum_j p_{j\to i} p_j$$

We note that p_i is an eigenvector of the transition probability matrix and thus corresponds to a steady state. Thus a rule that obeys detailed balance should bring us to a steady state distribution $\{p_i\}$.
The simplest implementation of this is the *Metropolis algorithm*:

① Pick a spin i by random.

② Calculate the energy change ΔE upon flipping the spin.

③ If $\Delta E < 0$, accept the spin flip.

④ If $\Delta E > 0$, accept the flip with probability $e^{-\beta \Delta E}$.

This is the simplest version because on the one hand it implements the condition on the ratio of the transition probabilities resulting from detailed balance and on the other hand, it uses the remaining freedom by simply setting the transition probability for one direction to 1. By going downhill, the algorithm samples regions of high importance. By hopping over barriers (compare 6.20), it allows to escape from metastable states.
There are two examples for more sophisticated MC-algorithms. First an algorithm in which one does not decide wheather to flip or not, but places a new spin (*heat-bath algorithm*):

① Pick a spin i by random.

② Calculate the effective magnetization m_i resulting from the neighboring spins.

③ Calculate the energies of placing up or down spins in this magnetic environment: $E_\pm = \mp J m_i \mp B$

④ Place a new spin with the corresponding probabilties: $p_{\pm} = e^{-\beta E_{\pm}}/(e^{-\beta E_{+}} + e^{-\beta E_{-}})$

Monte Carlo dynamics of course is not physical dynamics and was only invented to do importance sampling. Yet often one also uses it to explore dynamical phenomena. Then a good choice is *Glauber dynamics*, which in some sense combines the first two algorithms:

① Pick a spin i by random.

② Calculate the effective magnetization m_i resulting from the neighboring spins.

③ Calculate the energy change $\Delta E = 2 s_i m_i$ upon flipping the spin.

④ Accept the flip with probability $p = e^{-\beta \Delta E}/(1 + e^{-\beta \Delta E})$

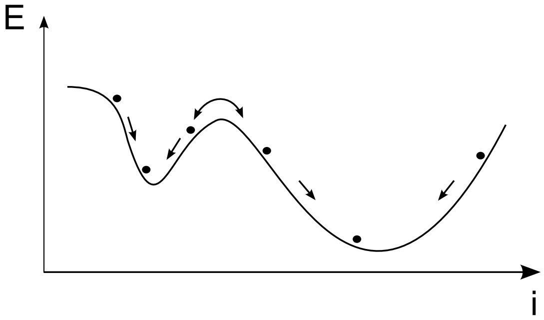

Figure 6.20: Sketch visualising the Metropolis algorithm and how it recovers from local minima.

One can come up with more complicated MC-moves than simply flipping single spins. In MC-simulations of the Ising model, it is very common to flip whole clusters (e.g. the Swendsen-Wang or Wolff algorithms). In MC-simulations of polymers, it is common to flip large parts of the polymer in a way that never would be possible in real systems due to self-intersection, or to grow polymers in the computer such that they give the correct importance sampling. Sometimes one also changes temperature, e.g. to solve optimization problems like the traveling salesman by *simulated annealing* (lowering temperature by optimal protocols to fall into the global minimum) or replica MC (when one runs several simulations in parallel and at different temperatures). The latest development in this field is the use of machine learning, e.g. Boltzmann generators, which are invertible neural networks that learn transformations into latent space where sampling is much easier (Noe, F., Olsson, S., Köhler, J., and Wu, H. (2019). Boltzmann generators: Sampling equilibrium states of many-body systems with deep learning. Science, 365(6457), eaaw1147).
With schemes like these, one can get very good results, e.g. $K_c = 0.2216544$ for the critical coupling constant of the 3D Ising model (that is T_c around 4.5), which is known

to be correct to all given digits due to a detailed analysis of its uncertainty. In particular, this allows one to conclude that the analytical suggestion $K_c = \tanh^{-1}((\sqrt{5} - 2)\cos(\pi/8)) = 0.2216586$ by Rosengreen 1986 cannot be correct, despite being so close (for a discussion on MC-simulations of the 3D Ising model, check the work by Martin Hasenbusch, e.g. his paper in International Journal of Modern Physics C, Vol. 12, No. 7 (2001) 911–1009). For more details on Monte Carlo simulations, check the book *Monte Carlo simulation in statistical physics* by Kurt Binder and Dieter Heermann (6th edition Springer 2019).

7 Classical fluids

7.1 Virial expansion

We now turn to classical fluids as the second example for directly interacting systems, after lattice models like the Ising model. Our reference system here is the ideal gas. We recollect that for the ideal gas, the Hamiltonian for N particle reads:

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p_i^2}{2m}$$

The canonical partition sum in this case reads:

$$Z_{id} = \frac{1}{N!} \left(\underbrace{\frac{1}{h^3} \int \mathrm{d}\vec{p}\; e^{-\beta p^2/(2m)}}_{=1/\lambda^3} \right)^N \left(\underbrace{\int \mathrm{d}\vec{q}}_{=V} \right)^N = \frac{1}{N!} \frac{V^N}{\lambda^{3N}}$$

with the thermal wavelength $\lambda = \frac{h}{(2\pi m k_B T)^{\frac{1}{2}}}$. Based on the partition sum we can calculate the free energy:

$$\begin{aligned} \Rightarrow \quad F_{id} &= -k_B T \ln Z_{id} \\ \Rightarrow \quad p_{id} &= -\left.\frac{\partial F_{id}}{\partial V}\right|_{T,N} = \frac{N k_B T}{V} = \rho k_B T \end{aligned}$$

The ideal gas is an appropriate description for diluted gases (small density ρ). At higher densities, direct interactions become important.
We consider a pairwise additive and isotropic interaction potential U as the simplest case:

$$\mathcal{H} = \sum_{i=1}^{N} \frac{p^2}{2m} + \sum_{j<i} U \left(\underbrace{|\vec{r}_i - \vec{r}_j|}_{:=r_{ij}} \right)$$

Here $j < i$ means a summation over all interaction pairs (or *bonds*) between different particles; this is similar to the situation with the Ising model, but because the interaction potential can be long-ranged, there is no restriction to nearest neighbors here. In the potential, there is no dependence on

1. momenta (only positions)

2. relative orientations

of the particles. An example for which the second assumption does not hold are *liquid crystals* (Figure 7.1).

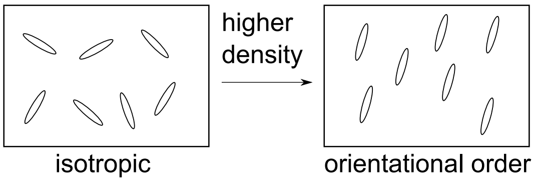

Figure 7.1: Liquid crystals: For increased density orientational, but not positional order is established. This is the *'isotropic-nematic transition'* of liquid crystals that has been calculated by Lars Onsager in 1949. The full phase diagram for hard spherocylinders also includes a smetic phase (in addition to the orientational order, there is positional ordering in one dimension) and a solid phase (here positional order is established in all three dimensions). Because in hard systems there is no attractive energy, these effects are all driven by entropy: by becoming globally ordered, the system generates more entropy for the single particles (larger configurational phase space volume).

An example for an isotropic potential is the *'Lennard-Jones potential'* introduced by John Lennard-Jones in 1924:

$$U(r) = 4\epsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$$

The potential consists of two elements:

① a universal attraction between neutral atoms and molecules (*'van der Waals interaction'*) proportional to $1/r^6$

② stability is provided by short-ranged *'Born repulsion'* ($\propto 1/r^{12}$).

For computer simulations one typically shifts and truncates the potential to achieve a finite interaction range (this also allows the use of neighbor lists). These simulations can be done based on *'Monte Carlo'* (MC) or *'Molecular Dynamics'* (MD) procedures; the resulting phase diagrams will be the same. Figure 7.3 shows a phase diagram which is typical for a simple one-component system (as for example described by the Lenard-Jones potential; experimentally a good example would be CO_2).
We now return to the analytical description:

$$Z = \underbrace{\frac{1}{N!} \left(\frac{1}{h^3} \int \mathrm{d}\vec{p}\, e^{-\beta p^2/(2m)} \right)^N V^N}_{=Z_{id}} \cdot \underbrace{\frac{1}{V^N} \int \mathrm{d}^N\vec{q}\, e^{-\beta \sum_{i<j} U(r_{ij})}}_{:=Z_{int}}$$

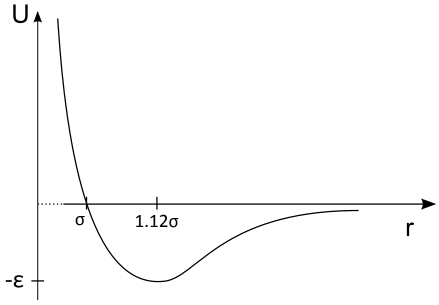

Figure 7.2: The Lennard-Jones potential U as a function of inter-particle distance r.

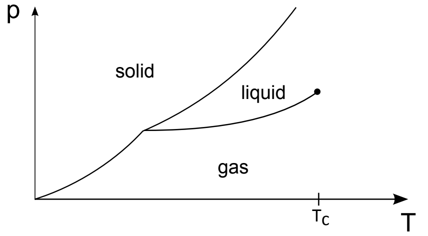

Figure 7.3: A generic phase diagram typical for a simple one-component system. T_c indicates the temperature of the *'critical point'* where phase boundaries cease to exist.

$$\Rightarrow \quad F = -k_B T \ln Z = F_{id} + F_{int}$$
$$p = -\left.\frac{\partial F}{\partial V}\right|_{T,N} = p_{id} + p_{int}$$

The interaction part does not factorise into single particle properties. Hence one needs approximations. Because we understand the dilute case, we now introduce the *'virial expansion'*, which is an expansion in low density around the ideal gas as a reference system. We note that corrections to the ideal case pressure have to be of order ρ^2 or higher, because they arise if two particles or more collide.

$$\Rightarrow \quad p_{int} = k_B T \sum_{i=2}^{\infty} B_i(T)\rho^i = k_B T B_2(T)\rho^2 + O(\rho^3)$$

Here the $B_i(T)$ are termed *'virial coefficients'*. In lowest order of the correction we thus have

$$\Rightarrow \quad F = N k_B T \left[\ln\left(\rho\lambda^3\right) - 1 + B_2\rho\right]$$
$$\boxed{p = \rho k_B T \left[1 + B_2\rho\right]}$$

7.2 Second virial coefficient

In the following we use the grand canonical formalism to calculate $B_2(T)$ from $U(r)$:

$$Z_G(T,V,\mu) = \sum_{N=0}^{\infty} \underbrace{Z(T,V,N)}_{:=Z_N} \Bigg(\underbrace{e^{\beta\mu}}_{\text{fugacity z}} \Bigg)^N$$

In the high temperature limit $z \ll 1$ we expand this expression in z:

$$\Rightarrow \quad Z_G = Z_0 + Z_1 z + Z_2 z^2 + O\left(z^3\right)$$

$$Z_0 = 1,\ Z_1 = \frac{V}{\lambda^3}$$

$$Z_2 = \frac{1}{2!\lambda^6} \int \mathrm{d}\vec{r}_1 \int \mathrm{d}\vec{r}_2 \, e^{-\beta U(|\vec{r}_1 - \vec{r}_2|)} = \frac{V4\pi}{2\lambda^6} \int \mathrm{d}r \, r^2 \, e^{-\beta U(r)}$$

Next we use the Euler relation for the grand canonical potential:

$$\Psi = -k_B T \ln Z_G = -pV$$

$$\Rightarrow \quad \frac{pV}{k_B T} = \ln Z_G \overset{z \ll 1}{\approx} \ln\left(Z_0 + Z_1 z + Z_2 z^2 + O\left(z^3\right)\right)$$

$$\overset{z \ll 1}{\approx} Z_1 z + (Z_2 - \frac{Z_1^2}{2}) z^2 + O\left(z^3\right)$$

Were we used the approximation $\ln(1+x) \approx x - \frac{x^2}{2}$ for $x \ll 1$.

Obviously the virial expansion is similar to the expansion in fugacity z:

$$\frac{pV}{k_B T} = V\left[\rho + B_2 \rho^2 + O\left(\rho^3\right)\right]$$

To make a comparison of coefficients we need the relation between z and ρ.

$$\rho = \frac{\langle N \rangle}{V}, \quad z = e^{\beta\mu} \ \Rightarrow \ \mu = \frac{\ln z}{\beta}$$

$$\Rightarrow \quad \partial_\mu = \beta z \partial_z$$

$$\langle N \rangle = \frac{1}{\beta} \partial_\mu \ln Z_G = z \partial_z \ln Z_G$$

$$\approx Z_1 z + \left(2 Z_2 - Z_1^2\right) z^2 + O\left(z^3\right)$$

We note that the first order

$$z \approx \frac{\langle N \rangle}{Z_1} = \rho \lambda^3$$

in our approximation reproduces our well known ideal gas result ($p = \rho k_B T$).

We need the next higher order to calculate B_2:

$$\underbrace{\frac{\langle N \rangle}{Z_1}}_{:=c} = z + \underbrace{\left(\frac{2Z_2 - Z_1^2}{Z_1}\right)}_{:=a} z^2$$

$$\begin{aligned}
\Rightarrow \quad z &= \frac{-1+\sqrt{1+4ac}}{2a} \approx \frac{-1+1+\frac{1}{2}(4ac)-\frac{1}{8}(4ac)^2}{2a} \\
&= c(1-ac) \\
&= \frac{\langle N \rangle}{Z_1}\left[1 - \frac{\langle N \rangle}{Z_1}\frac{2Z_2 - Z_1^2}{Z_1}\right]
\end{aligned}$$

Here we used $\sqrt{1+x} \approx 1 + \frac{1}{2}x - \frac{1}{8}x^2$ for $x \ll 1$ in the first step.

$$\begin{aligned}
\Rightarrow \quad \frac{pV}{k_B T} &= V\left[\rho + B_2\rho^2 + O\left(\rho^3\right)\right] = \langle N \rangle \left[1 + B_2\rho + O\left(\rho^2\right)\right] \\
&= \ln Z_G = Z_1 z + \left(Z_2 - \frac{Z_1^2}{2}\right) z^2 + O\left(z^3\right) \\
&= Z_1 \frac{\langle N \rangle}{Z_1}\left[1 - \frac{\langle N \rangle}{Z_1}\left(\frac{2Z_2 - Z_1^2}{Z_1}\right)\right] + \left(Z_2 - \frac{Z_1^2}{2}\right)\frac{\langle N \rangle^2}{Z_1^2} + O\left(\rho^3\right) \\
&= \langle N \rangle \left[1 + \frac{\langle N \rangle}{Z_1^2}\left(-2Z_2 - Z_1^2 + Z_2 - \frac{1}{2}Z_1^2\right)\right] + O\left(\rho^2\right) \\
&= \langle N \rangle \left[1 - \left(Z_2 - \frac{Z_1^2}{2}\right)\frac{\langle N \rangle}{Z_1^2}\right]
\end{aligned}$$

We now find our final result for $B_2(T)$, which relates the microscopic potential to its macroscopic effect (change in pressure compared to ideal gas):

$$\Rightarrow \quad B_2(T) = -V\left(\frac{Z_2}{Z_1^2} - \frac{1}{2}\right) = -\frac{1}{2}\int \mathrm{d}\vec{r}\left(e^{-\beta U(r)} - 1\right)$$

$$\Rightarrow \boxed{B_2(T) = -2\pi \int r^2 \,\mathrm{d}r \,\left(e^{-\beta U(r)} - 1\right)}$$

Note that the integrand (which is known as the Mayer f-function) has very good properties for a perturbation analysis: it is close to -1 in repulsive hard core part and then decays to 0 with the attractive interation. This observation is the starting point for Ursell's cluster expansion leading to the higher virial coefficients (in a graphical manner, similar to an expansion in Feynman diagrams). However, for our purpose it is sufficient to discuss the second virial coefficient, because it already leads to a phase transition.

Examples

① hard spheres

Spheres of radius $d/2$ which cannot penetrate each other. This yields an excluded region of radius d (Figure 7.4).

$$\Rightarrow \quad B_2(T) = -2\pi \int_0^d r^2(-1)\,\mathrm{d}r = \frac{2\pi}{3}d^3 = \frac{1}{2}V_{excl} = 4V_{sphere} > 0$$

Due to B_2 being positive, a finite, excluded volume increases the pressure. B_2 does not depend on temperature, because there is no finite interaction energy.

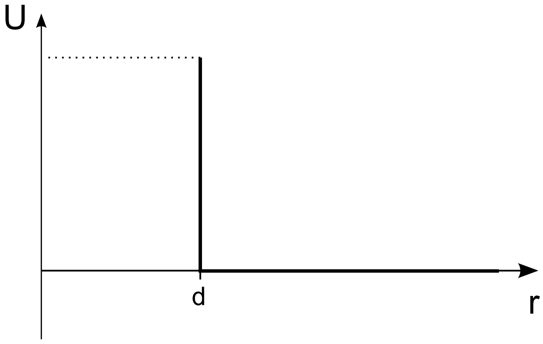

Figure 7.4: Potential for the hard spheres with excluded region $r < d$.

② attractive square well

We consider a potential well of depth ϵ between d and $d+\delta$ (Figure 7.5).

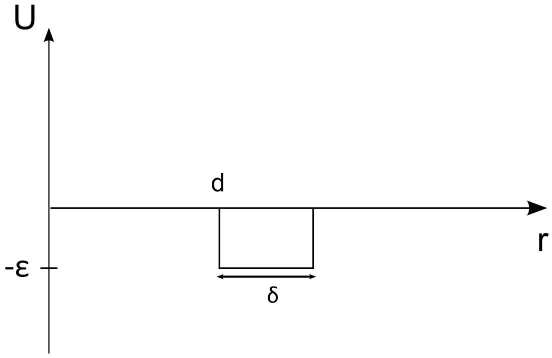

Figure 7.5: Square well potential with range δ and depth ϵ.

$$\Rightarrow \quad B_2(T) = -2\pi \int_d^{d+\delta} r^2\,\mathrm{d}r\,\left(e^{\beta\epsilon}-1\right) \overset{\delta\ll d,\ \beta\epsilon\ll 1}{\approx} -2\pi d^2\delta\beta\epsilon < 0 \quad \text{for } \epsilon \ll k_B T$$

$B_2(T)$ vanishes at high temperature. The attraction of the particles decreases the pressure. This effect increases with increasing interaction energy ϵ and range δ.

③ Hard core repulsion and attraction
As a course approximation to particle interactions with repulsive core and attractive well, we combine a hard hard sphere with an attractive well (Figure 7.6):

$$\Rightarrow B_2(T) = \frac{2\pi}{3}d^3 - 2\pi d^2 \delta \beta \epsilon = b - \frac{a}{k_B T}$$

with constants $a, b > 0$.

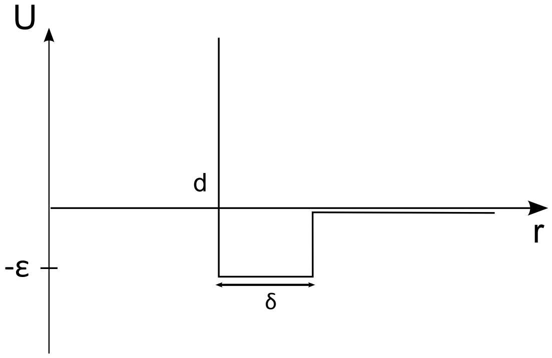

Figure 7.6: Combination of the potentials for the hard spheres and the square well. The resulting form is similar to the Lennard-Jones potential.

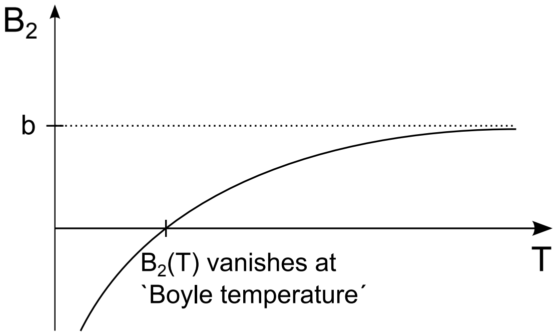

Figure 7.7: B_2 as a function of temperature. It vanishes at the *'Boyle temperature'*.

④ Lennard-Jones potential
This case cannot be solved analytically, but one can come up with a helpful approximation. We first non-dimensionalize the equation for $B_2(T)$ by measuring r in units of σ and T in units of ϵ / k_B. Next we note that one can rewrite the integral

by partial integration, because it is a product of the Mayer f-function $f(r)$ and a derivative $g'(r) = r^2$ (that is $g(r) = r^3/3$):

$$B_2(T) = -2\pi\sigma^3 \int \mathrm{d}r \, (r^2) \left(e^{-\frac{4}{T}\left[\frac{1}{r^{12}}-\frac{1}{r^6}\right]} - 1\right) = -2\pi\sigma^3 \int \mathrm{d}r \, g'(r) f(r)$$

We note that $fg = 0$ at the two boundaries and therefore

$$B_2(T) = 2\pi\sigma^3 \int \mathrm{d}r g(r) f'(r) = \frac{8\pi\sigma^3}{3T} \int \mathrm{d}r \, r^2 \left[\frac{12}{r^{12}} - \frac{6}{r^6}\right] e^{-\frac{4}{T}\left[\frac{1}{r^{12}}-\frac{1}{r^6}\right]}$$

Until here the treatment is exact. In order to proceed, we now expand $e^{4/(Tr^6)}$ for small arguments and use the definitions of the Gamma-functions, which we then evaluate numerically. The result is a power series

$$B_2(T) = \frac{2\pi}{3}\sigma^3 \left[\frac{1.73}{T^{1/4}} - \frac{2.56}{T^{3/4}} + \dots\right]$$

The resulting functional form is similar to the one for hard spheres with attraction: a divergence to negative values at $T = 0$ and a cero crossing to positive values at finite t. However, at very large T, $B_2(T)$ for the Lennard-Jones potential goes back to cero, because this potential is not infinitely hard. Such a decrease indeed can be observed experimentally for atomic and molecular systems at very high temperature.

We conclude that cases 3 and 4 above gave very similar results. For simplicity, we continue our discussion with case 3, which is sufficient to demonstrate that the combination of repulsion at short distances and attraction at larger distances leads to a phase transition. To see this, we again study pressure and use the simple result for the combination of hard spheres repulsion with the attractive square well:

$$\begin{aligned} pV &= Nk_BT \left(1 + B_2(T)\frac{N}{V}\right) \\ &= Nk_BT \left(1 + b\frac{N}{V}\right) - \frac{N^2}{V}a \approx \frac{Nk_BT}{1 - b\frac{N}{V}} - \frac{N^2}{V}a \end{aligned}$$

Introducing the specific volume $v = \frac{V}{N} = \frac{1}{\rho}$ this yields

$$\boxed{p = \frac{k_BT}{v - b} - \frac{a}{v^2}} \quad \text{van der Waals equation of state}$$

The excluded volume (b) reduces the accessible volume for the particles while an attractive interaction (a/v^2) reduces pressure. For $T < T_c = 8a/(27bk_B)$, $p(v)$ will have a minimum and maximum (see Figure 7.8).

In the region between the minimum and maximum we have $\frac{dp}{dv} > 0$. This implies that a local fluctuation to higher density (smaller v) causes a decrease in pressure, which leads

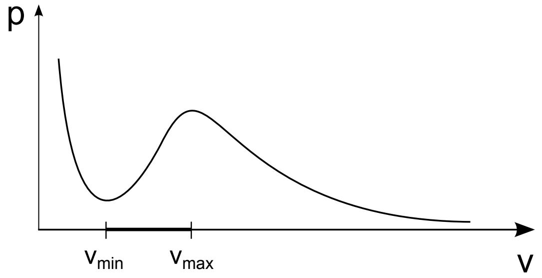

Figure 7.8: Pressure isotherm for a van der Waals gas below the critical temperature. In the region between v_{min} and v_{max} the system is unstable.

to a further increase in density (decrease in internal pressure is equivalent to increase in external pressure, so the fluid is compressed more). Likewise a fluctuation to lower density (larger v) leads to an increase in pressure, which leads to a further decrease in density (larger internal pressure is equivalent to decrease in external pressure, so the fluid expands). Thus we see that in this region, the fluid is instable to density fluctuations. The system can avoid this instability by jumping over it with a phase transition between a liquid and a gas. In the region of the phase transition, two different phases coexist. This can be understood even better by considering the free energy of the system, which becomes concave in the instable region. This means that a linear superposition of the two states obtained by a common tangent construction will have lower free energy than the system predicted by the instable free energy. We now discuss how one can calculate the region of the phase transition.

7.3 Maxwell construction

The details of the phase transition follow from the *'Maxwell construction'*. We consider the Gibbs free energy as we control temperature T and pressure p:

$$G = \underbrace{E - TS}_{:=F} + pV = \mu N$$

$$\Rightarrow \quad \mu = \underbrace{\frac{F}{N}}_{:=f} + p \underbrace{v}_{=V/N}$$

For two coexisting phases L and G in equilibrium the intensive parameters T, p and μ have to be the same:

$$\mu_L(T,p) = \mu_G(T,p)$$
$$\Rightarrow \quad f_G - f_L = p_t\,(v_L - v_G)$$

Here p_t is the transition pressure.

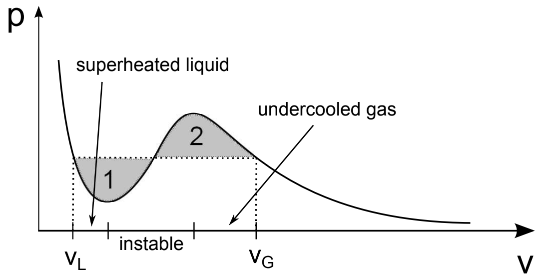

Figure 7.9: Van der Waals isotherm with Maxwell construction based on the equality of areas 1 and 2.

The left hand side can be calculated by integration along the isotherm:

$$f_G - f_L = \int_{v_L}^{v_G} \mathrm{d}v \, \left.\frac{\partial f(T,v)}{\partial v}\right|_T = -\int_{v_L}^{v_G} \mathrm{d}v \, p(T,v)$$

$$\Rightarrow \quad p_T(v_L - v_G) = \int_{v_G}^{v_L} \mathrm{d}v \, p(T,V)$$

Geometrically this means that in Figure 7.9 the dotted area has to equal the one below the solid line. Hence p_t can be determined based on the equality of areas 1 and 2.

We therefore have arrived at the following picture:
If several phases can exist at constant (T,p), the one with the lowest chemical potential μ is stable (lowest G/N). At the transition point, the chemical potentials are equal. However, their slopes can have jumps (compare Figure 7.10).

In order to bring the fluid from liquid to gas, we need the *'heat of evaporation'* or *'latent heat'* Q:

$$Q = \int_{T_{t-}}^{T_{t+}} T \, \mathrm{d}S = \int_{T_{t-}}^{T_{t+}} \mathrm{d}H = H_G - H_L$$

where we used $\mathrm{d}H = T\mathrm{d}S + V\mathrm{d}p$ and $p = p_t = \text{const.}$

$$\Rightarrow \quad h = \frac{H}{N} = \frac{E + pV}{N} \overset{vdW\ eq}{=} \underbrace{e(T)}_{\text{kinetic energy contr.}} - \frac{a}{v} + pv$$

$$q = \frac{Q}{N} = h_G - h_L = \frac{a}{v_L} - \frac{a}{v_G} + p\,(v_G - v_L) \approx \frac{a}{v_L} + pv_G \quad (v_G \gg v_L)$$

$\frac{a}{v_L}$ is the energy required to overcome attraction while pv_G is the energy required for expansion.

$$G = \mu N, \quad \mathrm{d}G = -S\mathrm{d}T + V\mathrm{d}p + \mu \mathrm{d}N$$

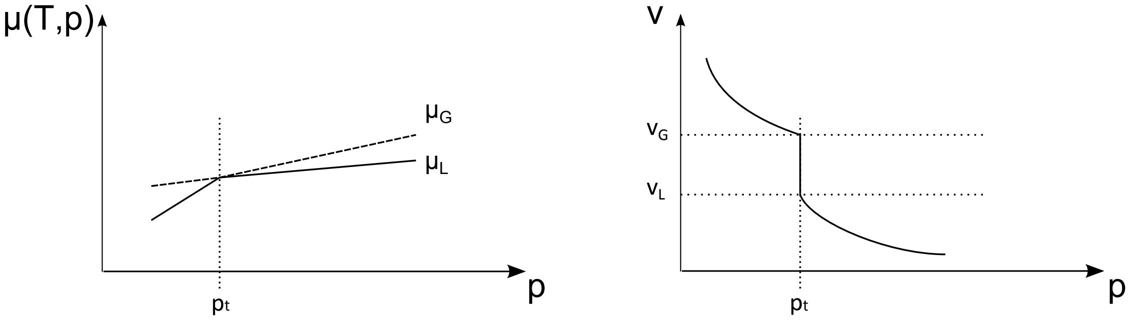

Figure 7.10: Left: The chemical potential μ as a function of pressure for phases G and L. At the transition point $\mu_G = \mu_L$, but the slopes have jumps.
Right: The specific volume $v = \frac{\partial\mu}{\partial p}\Big|_T$ as a function of pressure has a jump at the transition pressure p_t.

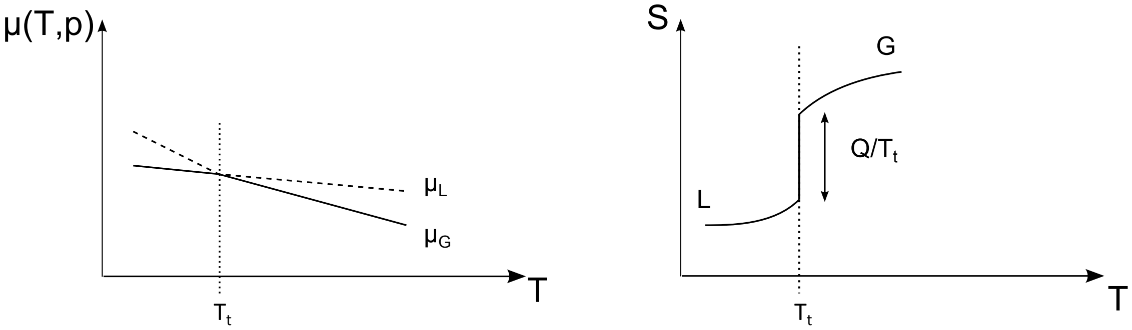

Figure 7.11: Left: The chemical potential μ as a function of temperature for phases G and L. At the transition point $\mu_G = \mu_L$, but the slopes have jumps.
Right: The entropy as a function of temperature jumps at the transition point.

$$\Rightarrow \quad \frac{\partial\mu}{\partial T}\Big|_{p,N} = \frac{1}{N}\frac{\partial G}{\partial T}\Big|_{p,N} = -\frac{1}{N}S$$

We conclude that both $v = \frac{\partial\mu}{\partial p}\Big|_T$ and $s = -\frac{\partial\mu}{\partial T}\Big|_p$ jump at the transition (compare Figures 7.10 and 7.11 respectively). Therefore this phase transition is called to be of *'first order'* or *'discontinuous'*. Both jumps disappear at the critical point, where the isotherm becomes horizontal at the transition. From

$$\frac{\partial p}{\partial v}\Big|_T = \frac{\partial^2 p}{\partial v^2}\Big|_T = 0$$

one calculates the critical values:

$$v_c = 3b, \quad T_c = \frac{8a}{27bk_B}, \quad p_c = \frac{a}{27b^2}$$

$$\text{for water:} \quad p_c = 217\ atm, \quad T_c = 647\ K$$

$$\Rightarrow \quad \frac{p_c v_c}{k_B T_c} = \frac{3}{8} = 0.375 \text{ independent of a and b}$$

Experimental values are similar, but slightly smaller (around 0.3).

If p, v and T are expressed in terms of their critical values:

$$\widetilde{p} = \frac{p}{p_c}, \quad \widetilde{v} = \frac{v}{v_c}, \quad \widetilde{T} = \frac{T}{T_c}$$

the van der Waals equation becomes

$$\boxed{\left(\widetilde{p} + \frac{3}{\widetilde{v}^2}\right)(3\widetilde{v} - 1) = 8\widetilde{T}}$$

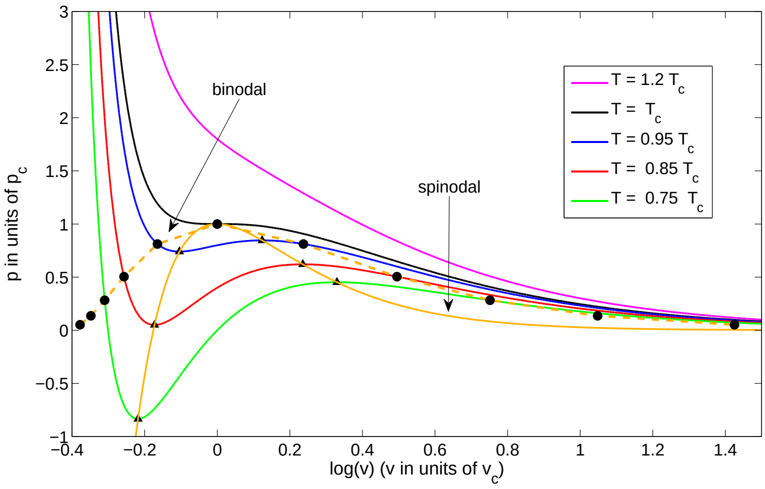

Figure 7.12: Van der Waals isotherms for different temperatures. The *'spinodal'* is the boundary between metastable and unstable states. The *'binodal'* separates metastable and absolutely stable states. The latter curve was calculated numerically based on Maxwell's construction for different temperatures.

Figure 7.12 shows van der Waals isotherms for different temperatures with respect to T_c.

This reduced equation leads to the *'law of corresponding states'*: Two fluids with the same

$\left(\widetilde{p}, \widetilde{v}, \widetilde{T}\right)$ are in equivalent states. Indeed experimental curves show surprisingly good data collapse. Even more surprisingly, their behaviour becomes almost identical at the critical point - large fluctuations render microscopic details irrelevant.

7.4 Fluid-solid phase transition

The *'van der Waals'* equation of states predicts the fluid-fluid phase transition caused by attractive interactions. The fluid-solid phase transition can be predicted by a simple entropic argument. Recall the van der Waals theory for a hard sphere fluid:

$$F = Nk_BT \left\{ \ln \left(\frac{N\lambda^3}{V - Nb} \right) - 1 \right\}$$

$$\Rightarrow \quad p = - \left. \frac{\partial F}{\partial V} \right|_{T,N} = \frac{Nk_BT}{V - Nb}$$

$$b = 4V_s \quad \Rightarrow \quad V - Nb = V(1 - \rho b) = \alpha V$$

with $\alpha = 1 - \rho/\rho_0$ and $\rho_0 = 1/b$. αV is the free volume in the fluid.
Based on Figure 7.13 and $L = V^{1/3}$ we define the free volume of a solid as:

$$\alpha V \approx \left(V^{\frac{1}{3}} - d \right)^3 = \left[1 - \left(\frac{\rho}{\rho_0} \right)^{\frac{1}{3}} \right]^3 V$$

The free volume vanishes at close packing.

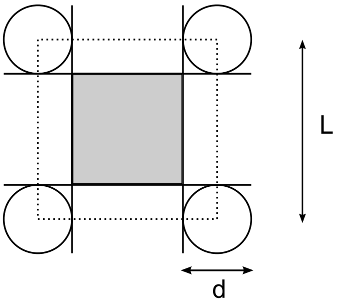

Figure 7.13: Unit cell with hard spheres of diameter d. The grey shaded region indicates the free volume.

$$\Rightarrow \quad F_1 - F_2 = Nk_BT \ln \frac{\alpha_2}{\alpha_1}$$

Hence the phase with larger α is favored. For that reason the fluid F and the solid S are stable at low and high densities, respectively. Figure 7.14 shows how the Maxwell construction looks like in this case.
We now can understand the complete phase diagram of a simple one-component fluid:

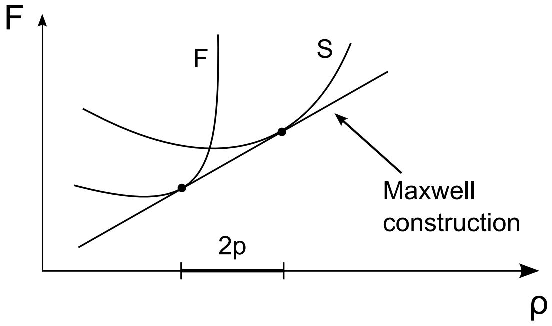

Figure 7.14: Free energy for liquid (F) and solid phase(S) as a function of density ρ. The tangent represents the Maxwell construction.

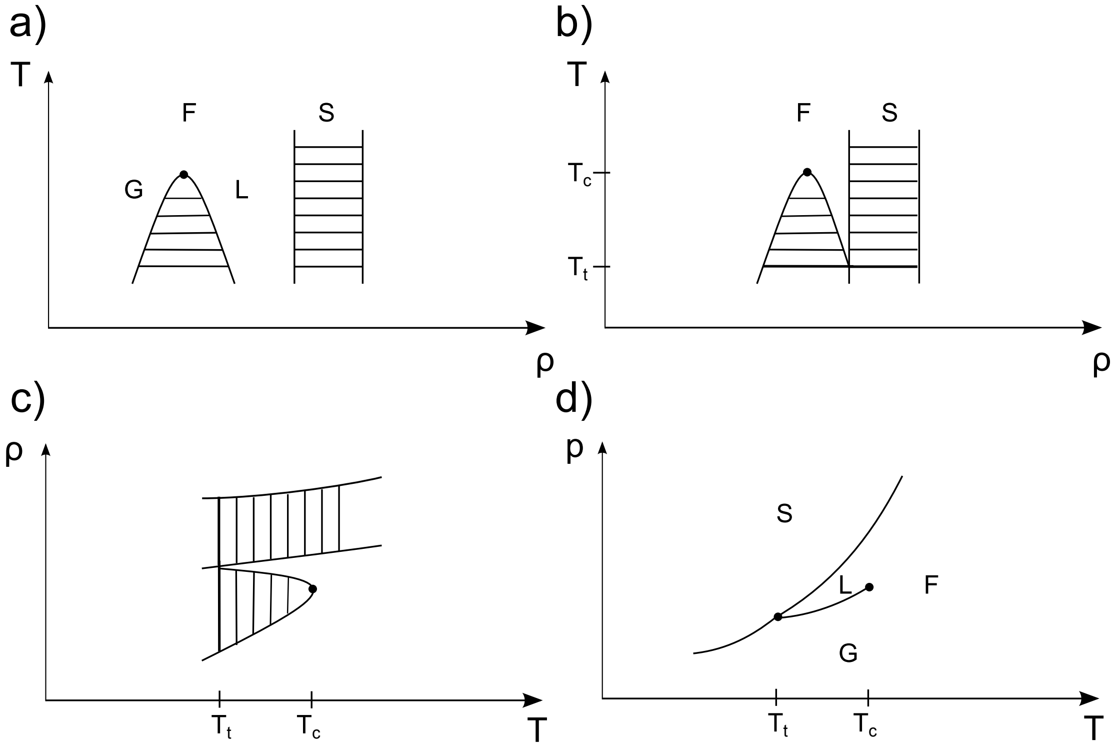

Figure 7.15: Combining the two transitions in (a), one gets the complete phase diagram in (b). In (c) we swap T and ρ axes. By replacing ρ by p, we get the final phase diagram in (d). Two-phase coexistence regions become lines in this representation.

7.5 Structure and correlation

In contrast to thermodynamics, statistical physics not only predict phase behaviour, but also the way the fluid is structured on a molecular level. The central quantity in both theory and experiments is the *pair correlation function* (also known as the *two-point correlation function* or the *pair distribution function*, and for isotropic systems as the *radial distribution function*). Below we explain that this function is essentially the Fourier transform of the *structure factor* which is measured in the detector of scattering experiments. Possible probes are neutrons, electrons or photons. Neutrons are a great choice because the scattering contrast with the nucleus makes the interaction very localized and thus uncertaintly about position is very small. However, for neutrons one needs large facilities, either reactors (like the FRM II at TU Munich or the Institut Laue-Langevin at Grenoble, France) or spallation sources (like the European Spallation Source at Lund, Sweden). Electrons, visible light or normal X-ray are much easier to handle and are available locally in many labs. For synchotron radiation one again has to use large facilities (like the Free Electron Laser at Hamburg or the European Synchrotron Radiation Facility at Grenoble, France). [1]

We start with the statistical physics side and first define the distribution function for absolute position:

$$p(\vec{r}_1, ..., \vec{r}_N) = \frac{e^{-\beta \sum_{i<j} U(r_{ij})}}{\int \mathrm{d}\vec{r}_1 \, ...\mathrm{d}\vec{r}_N \; e^{-\beta \sum_{i<j} U(r_{ij})}}$$

By defining $W := \sum_{i<j} U(r_{ij})$, the probability that any particle is at position $\vec{r}$ can be written as:

$$\begin{aligned} n_1(\vec{x}) &= \sum_{k=1}^{N} \langle \delta\left(\vec{x} - \vec{r}_k\right) \rangle \\ &= \frac{\int \mathrm{d}\vec{r}_1 ...\mathrm{d}\vec{r}_N \left(\sum_{k=1}^{N} \delta\left(\vec{x} - \vec{r}_k\right) \right) e^{-\beta W}}{\int \mathrm{d}\vec{r}_1 ...\mathrm{d}\vec{r}_N \; e^{-\beta W}} \\ &= N \frac{\int \mathrm{d}\vec{r}_2 ...\mathrm{d}\vec{r}_N \; e^{-\beta W(\vec{x}, \vec{r}_2, ..., \vec{r}_N)}}{\int \mathrm{d}\vec{r}_1 ...\mathrm{d}\vec{r}_N \; e^{-\beta W}} \end{aligned}$$

$$n_1\left(\vec{x}\right) \underbrace{\overset{\text{ideal gas}}{=}}_{W=0} \frac{N}{V} = \rho$$

The probability that some particle is at $\vec{x}_1$ and another at $\vec{x}_2$ is:

$$\begin{aligned} n_2\left(\vec{x}_1, \vec{x}_2\right) &= \sum_{i \neq j} \left\langle \delta\left(\vec{x}_1 - \vec{r}_i\right) \delta\left(\vec{x}_2 - \vec{r}_j\right) \right\rangle \\ &= N(N-1) \frac{\int \mathrm{d}\vec{r}_3 ...\mathrm{d}\vec{r}_N \; e^{-\beta W(\vec{x}_1, \vec{x}_2, ..., \vec{r}_N)}}{\int \mathrm{d}\vec{r}_1 ...\mathrm{d}\vec{r}_N \; e^{-\beta W}} \end{aligned}$$

---

[1] For more information on scattering, consult the recent proceedings of the 50th IFF spring school at Research Center Jülich (https://publications.rwth-aachen.de/record/756021).

$$n_2\left(\vec{x}_1,\vec{x}_2\right)\underbrace{\overset{\text{ideal gas}}{=}}_{W=0}\frac{N(N-1)}{V^2}\overset{N\to\infty}{\to}\rho^2$$

For the pairwise additive potential everything follows from n_1 and n_2. Eg the averaged interaction energy:

$$\begin{aligned}\langle W\rangle &= \sum_{i<j}\left\langle U(r_{ij})\right\rangle\\ &= \frac{1}{2}\sum_{i\neq j}\int \mathrm{d}\vec{x}_1\,\mathrm{d}\vec{x}_2\ \left\langle U\left(\vec{x}_1-\vec{x}_2\right)\delta\left(\vec{x}_1-\vec{r}_i\right)\delta\left(\vec{x}_2-\vec{r}_j\right)\right\rangle\\ &= \frac{1}{2}\int \mathrm{d}\vec{x}_1\,\mathrm{d}\vec{x}_2\ U\left(\vec{x}_1-\vec{x}_2\right)\ n_2\left(\vec{x}_1,\vec{x}_2\right)\end{aligned}$$

In a homogeneous system:

$$n_2\left(\vec{x}_1,\vec{x}_2\right)=n_2\left(\left|\vec{x}_1-\vec{x}_2\right|\right)$$

We define the *'radial distribution function'* g by:

$$n_2\left(\left|\vec{x}_1-\vec{x}_2\right|\right)=\rho^2 g\left(\left|\vec{x}_1-\vec{x}_2\right|\right)$$

$$\Rightarrow\qquad\boxed{\langle W\rangle=\frac{N^2}{2V}\int \mathrm{d}\vec{r}\ U(r)g(r)}$$

$\rho g(r)4\pi r^2\ \mathrm{d}r$ is the average number of particles in a spherical shell of width $\mathrm{d}r$ at a distance r from any particle.

For the ideal gas, $g(r)=1$ as shown above. For a real gas, volume exclusion means that $g(r)=0$ for small r, because no other particle can be at the same position. Above the atomic or molecular radius, $g(r)$ has to jump to values higher than 1 because of the attractive interaction. However, it then has to decay again because at large distances, fluctuations smear out the correlations and we should get the ideal gas result $g(1)=1$. Interestingly, this decay occurs with damped oscillations, indicating a shell structure: each particle is surrounded by a certain number of nearest neigbors, and this effect repeates over and over again. Unfortunately, it is not easy to calculate pair correlation functions directly from the interaction potential. Analytically, this is similar to a virial expansion, but involves hierarchies with integrals that cannot easily be closed (e.g. the BBGKY-hierarchy or the hypernetted chain). Therefore it is usually easier to simulate this function with molecular dynamics or Monte Carlo approaches.

We finally discuss how the correlation function is measured in scattering experiments. Scattering is the deviation of radiation from its straight path and thus you want to measure how many particles are scattered away from the incoming axis. For isotropic systems like fluids only the scattering angle matters, so in principle one only has to move the detector on a cirular path around the sample. The angle between the incoming wave $\vec{k}$ and the outgoing wave $\vec{k}'$ is called 2θ. Usually one has the case of elastic scattering (no energy transfer, wavelength λ constant), thus $|\vec{k}'|=|\vec{k}|$. Then the two wave vector make an equilateral triangle with the momentum transfer $q = \vec{k}' - \vec{k}$ and one has

$$\sin\theta = \frac{Q\lambda}{4\pi}$$

Therefore the scattering angle 2θ and the modulus Q of the momentum transfer vector $\vec{Q}$ are equivalent. We consider a solid angle Ω with opening $d\Omega$ and count the number of particles n which are scattered into this direction. This is quantified by the differential crosssection:

$$\frac{d\sigma}{d\Omega} = \frac{n}{jd\Omega} = b^2$$

Here j is the incoming flux (number of particles per cross-sectional area) and the result has to be the square of a length for dimensional reasons. This defines the scattering length b which describes the power of the scatting. For a detailed more description of this interaction, one actually had to use Fermi's golden rule to calculate the scattering matrix and consider the form factor $f(\vec{q})$ which describes the interaction of the probe with the scattering ostacle, but one can show that this gives a multiplicative factor that we can ignore here.

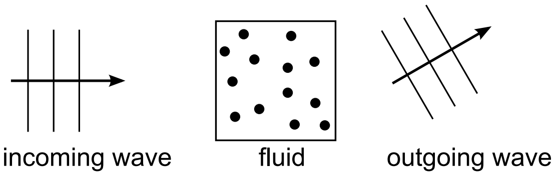

Figure 7.16: Schematic sketch of a scattering experiment. If $\vec{k}$ denotes the wave vector of the incoming wave and $\vec{k}'$ for the outgoing wave, the wave-fluid interaction results in a momentum transfer $\vec{Q} = \vec{k}' - \vec{k}$ with $|\vec{k}'| = |\vec{k}|$ (for elastic scattering, without energy transfer).

How can be now make contact to statistical physics ? The key idea is that the detector only measures the number of particles, or in other words only intensity, which is the squared modulus of the wavefunction. We assume that each scattering particle in the fluid gives rise to an outgoing spherical wave and that we sum them all up at the detector. Then we can write

$$\frac{d\sigma}{d\Omega} = b^2 N S(\vec{Q})$$

where we have introduced the *structure factor* $S(\vec{Q})$

$$S(\vec{Q}) = \frac{1}{N}\left\langle |\sum_i e^{i\vec{Q}\vec{r}_i}|^2 \right\rangle = \frac{1}{N}\left\langle \sum_i e^{-i\vec{Q}\vec{r}_i} \sum_j e^{i\vec{Q}\vec{r}_j} \right\rangle = \frac{1}{N}\left\langle \sum_{i,j} e^{i\vec{q}(\vec{r}_j - \vec{r}_i)} \right\rangle$$

Although it seems like a disadvantage that we can only measure intensity and therefore have to take the square modulus, the advantage now is that it brings in the pairs of particles and thus their correlation. We note that this result is translationally invariant and only depends on all relative distances. We also note that the sum still contains identical particles, which leads to a singularity that we have to remove. We therefore define the pair correlation function now as

$$g(\vec{r}) = \frac{n_2(\vec{r})}{\rho^2} - \frac{\delta(\vec{r})}{\rho}$$

and then get as final result:

$$S\left(\vec{Q}\right) = 1 + \rho \int \mathrm{d}\vec{r}\, e^{i\vec{Q}\cdot\vec{r}}\, (g(r) - 1)$$

The first 1 comes from the Fourier transform of the delta function and the second 1 (under the integral) removes another divergence, namely the one resulting from $\vec{Q} = 0$, that is the scatting in the forward direction, that is of no concern here (actually the detector never looks into this direction, because the beam would destroy it).

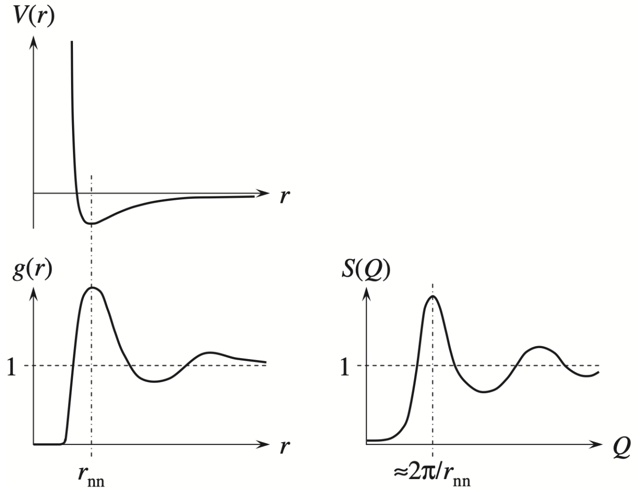

Figure 7.17: Typical shapes for interaction potential, pair correlation function and structure factor for a classical fluid. Taken from Reiner Zorn, IFF Spring School. Experimental results, e.g. for argon, give exactly these types of shape for the structure factor.

Overall, this is the promised result: the structure factor essentially is the Fourier transform of the pair correlation function. Because the situation is isotropic for fluids, we can do the angular integration and get:

$$S\left(\vec{Q}\right) = 1 + \rho \frac{4\pi}{Q} \int r dr\, e^{iQr}\, (g(r) - 1)$$

With this formula we can now convert any predicted pair correlation function into a statement on the structure factor (and vice versa).

8 Thermodynamics

8.1 Axiomatic structure

Thermodynamics is closely related to statistical physics. In fact it emerges from it in the limit of large system size, when fluctuations become irrelevant and only the mean values are relevant. Historically, thermodynamics developed in the 19th century as a theory for heat transfer and thus preceeded statistical physics, which emerged at the end of the 19th century from kinetic gas theory. In this script, however, we focus on statistical physics and therefore discuss thermodynamics only in the end.

In contrast to statistical physics, thermodynamics is a phenomenological theory and does not have a microscopic basis. One can consider this either as strength or weakness, depending on the viewpoint. Traditionally, it is explained starting from the four laws of thermodynamics. Here we choose the axiomatic viewpoint following the presentation by Callen. The following four axioms together completely determine the formal structure of classical thermodynamics:

① Simple isolated systems in equilibrium are characterised by the state variables (E, V, N). These three variables are singled out by basic symmetries: time-invariance of physical law leads to energy conservation, the existence of Goldstein modes (linear dispersion) leads to homogeneous excitations of the whole volume, and particle number is usually conserved in classical systems.

② For each equilibrium state an entropy function $S(E, V, N)$ exists. After removal of an internal constraint, the system obtains the state of maximal entropy. In thermodynamics, this is the main postulate, while in statistical physics, it follows from the fundamental postulate of a homogeneous distribution in equilibrium.

③ Entropy is additive over subsystems and increases monotonously with E. This implies that temperature T is always positive.

④ Entropy vanishes at $\partial E/\partial S = 0$ (Nernst postulate for $T = 0$).

Thus a thermodynamic system is located on the *'entropy manifold'* in a four dimensional state space (E, V, N, S). This is similar to a classical system being located on a *'energy manifold'* in $6N$ dimensional phase space $(\vec{p}, \vec{q})$. Quasi-static processes only involve equilibrium states. They can proceed only up or sidewards on the entropy manifold.

We first note that many of these elements we already encountered in the statistical physics part, in particular the maximum entropy principle. There is one interesting difference here, namely that entropy can only increase with energy, which means that temperature can only be positive. As we have seen in the statistical physics part, for systems with an upper limit of energy, we also can get negative temperature; this however is excluded in thermodynamics.

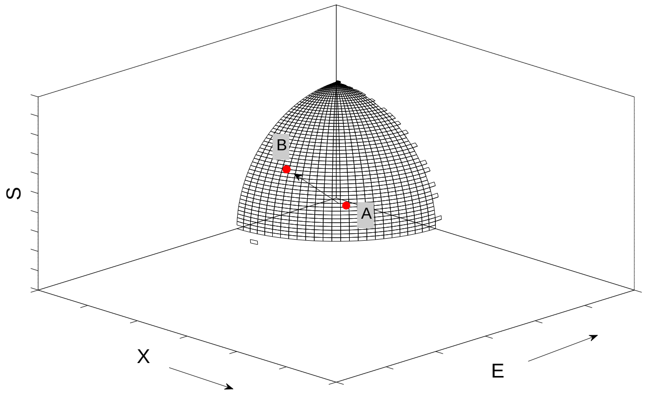

Figure 8.1: Entropy manifold with a quasi-static process from A to B. X stands for volume V or particle number N. Quasi-static processes only involve equilibrium states. Moving 'up' on the manifold indicates *'irreversible'* processes while *'reversible'* processes can be regarded as moving sidewards.

The total differential of the entropy function is:

$$\begin{aligned} \mathrm{d}S &= \left.\frac{\partial S}{\partial E}\right|_{V,N} \mathrm{d}E + \left.\frac{\partial S}{\partial V}\right|_{E,N} \mathrm{d}V + \left.\frac{\partial S}{\partial N}\right|_{V,E} \mathrm{d}N \\ &= \frac{1}{T}\,\mathrm{d}E + \frac{p}{T}\,\mathrm{d}V - \frac{\mu}{T}\,\mathrm{d}N \end{aligned}$$

Because S is an increasing function of E, one can always solve for $E = E(S, V, N)$:

$$\mathrm{d}E = \underbrace{T\,\mathrm{d}S}_{\text{heat}} - \underbrace{p\,\mathrm{d}V}_{\text{mechanical work}} + \underbrace{\mu\,\mathrm{d}N}_{\text{chemical energy}}$$

The three terms represent different ways to transfer energy.
Both $S(E, V, N)$ and $E(S, V, N)$ are fundamental equations which contain the complete thermodynamical information. Their partial derivatives can be interpreted as *'thermo-*

dynamical forces' in analogy to the mechanical forces $F = -\nabla V$:

$$T = \left.\frac{\partial E(S,V,N)}{\partial S}\right|_{V,N} \quad \text{temperature as driving force for entropy (heat) exchange}$$

$$p = -\left.\frac{\partial E(S,V,N)}{\partial V}\right|_{S,N} \quad \text{pressure as driving force for volume exchange}$$

$$\mu = \left.\frac{\partial E(S,V,N)}{\partial N}\right|_{S,V} \quad \text{chemical potential as driving force for particle exchange}$$

Each of these three equations of state contains only incomplete thermodynamical information. Knowledge of all three means that we know the tangential plane at every point and so can reconstruct the fundamental equation $E = E(S,V,N)$.

8.2 Variational principles

We next consider a composite system with two subsystems. The thermodynamical state space is now spanned by

$$(E_1, V_1, N_1, E_2, V_2, N_2, S).$$

In Figure 8.2 we schematically sketch the subsystem represented by S, $E = E_1 + E_2$ and one extensive variable of one of the two subsystems, eg $X = V_1$.

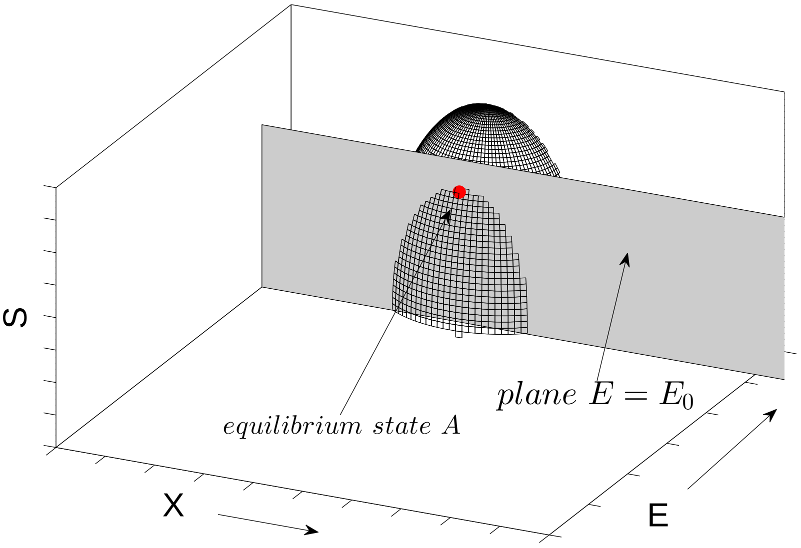

Figure 8.2: The equilibrium state A as a point of maximum S for constant $E = E_0$.

Due to the maximal entropy principle, the equilibrium state is in fact at A.

⇒ The equilibrium value of any unconstrained internal parameter X_j is such as to maximize the entropy S for a given value of the total energy E.

Note that the final state might be reached by a non-equilibrium process, but to identify the equilibrium state, one minimizes over the manifold of possible equilibrium states. We next note that the equilibrium state A can also be identified from a minimal energy principle:

⇒ The equilibrium value of any unconstrained internal parameter X_j is such as to minimize the energy E for a given value of the entropy S.

This is sketched in Figure 8.3.

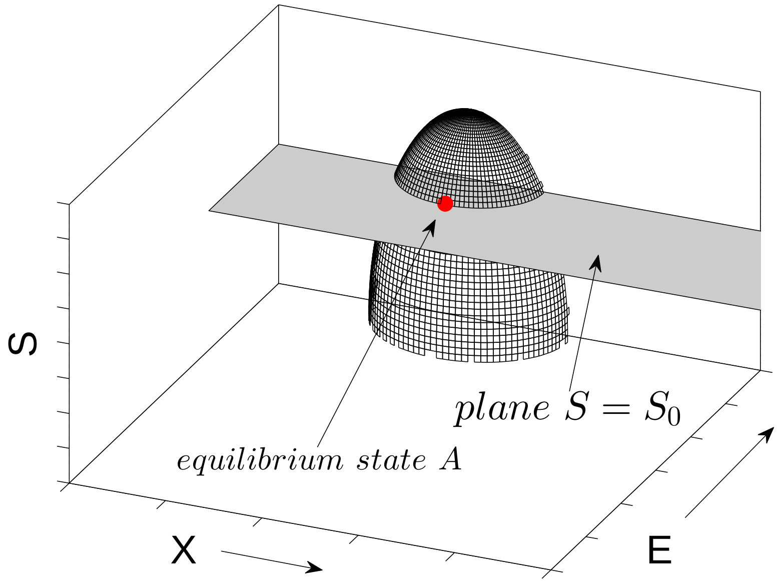

Figure 8.3: The equilibrium state A as a point of minimum E for constant $S = S_0$.

Proof:

If the energy was not minimal, we could withdraw some energy as work $-p\mathrm{d}V$ and return it as heat $T\mathrm{d}S$. Then the system would be restored to its initial energy but with increased entropy. This is a contradiction to the maximal entropy principle.

Example:

We consider the heat flow between two systems of fixed volumes and particle numbers.

① maximal entropy principle:

$$\mathrm{d}S = \frac{1}{T_1}\,\mathrm{d}E_1 + \frac{1}{T_2}\,\mathrm{d}E_2 = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)\,\mathrm{d}E_1 \overset{!}{=} 0 \;\Rightarrow\; T_1 = T_2$$

Before the equilibrium is reached: $\mathrm{d}S > 0 \quad \Rightarrow \quad \underline{\text{maximum}}$ of S.

② minimal energy principle:

$$\mathrm{d}E = T_1\,\mathrm{d}S_1 + T_2\,\mathrm{d}S_2 = (T_1 - T_2)\,\mathrm{d}S_1 \overset{!}{=} 0 \;\Rightarrow\; T_1 = T_2$$

Before the equilibrium is reached: $\mathrm{d}E < 0 \quad \Rightarrow \quad \underline{\text{minimum}}$ of E.

We again consider the two subsystems with an internal constraint X_j. We now connect them to a heat reservoir with temperature T_r.
The minimal energy principle implies:

$$\mathrm{d}\,(E + E_r) = \mathrm{d}\,(E_1 + E_2 + E_r) = 0$$

Terms related to heat exchange with the reservoir:

$$T_1\mathrm{d}S_1 + T_2\mathrm{d}S_2 + T_r\mathrm{d}S_r = (T_1 - T_r)\;\mathrm{d}S_1 + (T_2 - T_r)\;\mathrm{d}S_2 = 0$$

$\Rightarrow \quad T_1 = T_2 = T_r = T$ equilibrium condition independent of internal constraint

$$\begin{aligned} 0 &= \mathrm{d}\,(E_r + E) = T\,\mathrm{d}S_r + \mathrm{d}E = -T\mathrm{d}S + \mathrm{d}E \\ &= \mathrm{d}(E - TS) = \mathrm{d}F \end{aligned}$$

Hence the free energy F has an extremum at equilibrium. Since T is a constant and since S_r does not have a second derivative,

$$\mathrm{d}^2\,(E_r + E) > 0$$

implies $\mathrm{d}^2F > 0$ and thus F is minimal. The equilibrium value of any unconstrained internal parameter X_j in contact with a heat reservoir minimizes the free energy F over all states with $T = T_r$.

8.3 Euler and Gibbs-Duhem relations

Energy is extensive and therefore has to be a homogeneous function of order one:

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

$$\Rightarrow \frac{\partial E}{\partial(\lambda S)}\frac{\partial(\lambda S)}{\partial\lambda} + \frac{\partial E}{\partial(\lambda V)}\frac{\partial(\lambda V)}{\partial\lambda} + \frac{\partial E}{\partial(\lambda N)}\frac{\partial(\lambda N)}{\partial\lambda} = E(S,V,N)$$

$$\overset{\lambda=1}{\Rightarrow} \quad \boxed{E = TS - pV + \mu N} \quad \text{Euler relation}$$

$$\Rightarrow \quad \mathrm{d}E = T\mathrm{d}S + S\mathrm{d}T - p\mathrm{d}V - V\mathrm{d}p + \mu\mathrm{d}N + N\mathrm{d}\mu$$

$$\Rightarrow \quad S\mathrm{d}T - V\mathrm{d}p + N\mathrm{d}\mu = 0$$

$$\Rightarrow \quad \boxed{\mathrm{d}\mu = -s\mathrm{d}T + v\mathrm{d}p} \quad \text{Gibbs-Duhem relation}$$

The three intensive variables (T, p, μ) are not independent. There are only two thermodynamic degrees of freedom. If one knows the equations of state $s = s(T, p)$ and $v = v(T, p)$, one can integrate the Gibbs-Duhem relation to get $\mu = \mu(T, p)$.
If one knows the fundamental equation, the Gibbs-Duhem relation can be directly calculated in the integrated form:

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

$$\Rightarrow \quad p_i = \frac{\partial E}{\partial X_i} = p_i(S,V,N) = p_i(\frac{S}{N}, \frac{V}{N}, 1) = p_i(s,v)$$

Elimination of (s, v) from these three equations of state gives the Gibbs-Duhem relation between (T, p, μ).
The same arguments can be made in the entropy representation. The Euler-relation becomes

$$S = \frac{1}{T}E + \frac{p}{T}V - \frac{\mu}{T}N$$

and the Gibbs-Duhem relation then reads:

$$0 = d(\frac{1}{T})E + d(\frac{p}{T})V - d(\frac{\mu}{T})N$$

Example: fundamental equation for ideal gas

We now use these results to discuss the ideal gas. In thermodynamics one starts with the phenomenological observations, so we take the two well-documented gas laws

$$pV = nRT = Nk_BT, \; E = \frac{3}{2}nRT = \frac{3}{2}Nk_BT$$

Because E appears here, we choose the entropy representation and rewrite these equations as

$$\frac{p}{T} = \frac{k_B}{v}, \; \frac{1}{T} = \frac{3k_B}{2e}$$

We next integrate the Gibbs-Duhem relation

$$d(\frac{\mu}{T}) = ed(\frac{1}{T}) + vd(\frac{p}{T}) = e(\frac{-3k_B}{2e^2})de + v(\frac{-k_B}{v^2})dv = -\frac{3k_Bde}{2e} - \frac{k_Bdv}{v}$$

to give

$$\frac{\mu}{T} = (\frac{\mu}{T})_0 - \frac{3k_B}{2}\ln\frac{e}{e_0} - k_B\ln\frac{v}{v_0}$$

Inserting this result into the Euler relation gives

$$S = S_0 + k_B N \ln\left[(\frac{E}{E_0})^{3/2} (\frac{V}{V_0}) (\frac{N_0}{N})^{5/2} \right]$$

with

$$S_0 = \frac{5 k_B N}{2} - (\frac{\mu}{T})_0 N$$

Note that we get the same result as from the microcanonical or canoncial ensemble, except that we have a constant of integration that we cannot determine in thermodynamics. Here we see the essential difference to statistical physics: we cannot give a microscopic expression, but nevertheless we can get full thermodynamic information from a few experimental observations because the formal structure gives very strong constraints on the fundamental equation. We also note that in this specific and very fortunate case, we are able to integrate term-by-term, which usually is not the case. We finally note that we could have obtained the same result much easier by simply integrating the differential $ds = (1/T)de + (p/T)dv$, but here we wanted to demonstrate the use of the Gibbs-Duhem equation.
Similar procedures can be used to obtain the fundamental equations for e.g. the van der Waals gas, the photon gas, the rubber band, etc.

8.4 Thermodynamic potentials and Legendre transformations

We reconsider the situation that the temperature T is fixed by a reservoir. Now the relevant quantity is free energy F.

$$\begin{aligned} S(E) &= S_1(E_1) + S_2(E - E_1) \\ &= S_1(E_1) + S_2(E) - \left.\frac{\partial S_2}{\partial E_2}\right|_E E_1 \\ &= \text{const} + \left(S_1(E_1) - \frac{1}{T_1} E_1 \right) \end{aligned}$$

As S is maximal at equilibrium, free energy $F = E_1 - T S_1(E_1)$ is minimal.
Equilibrium is a compromise between order (E) and disorder (S). The larger T, the more disorder takes over. Free energy F is the relevant *'thermodynamic potential'* for the choice (T, V, N) as state variables.

In general one expects $2^3 = 8$ thermodynamic potentials, one for each choice of state variables:

state variables	thermodynamic potential	
S, V, N	E	internal energy
T, V, N	$F = E - TS$	Helmholtz free energy
T, p, N	$G = F + pV$	Gibbs free energy
S, p, N	$H = E + pV$	enthalpy
T, V, μ	$\Psi = F - \mu N$	grand canonical potential
S, V, μ	A_1	
S, p, μ	A_2	
T, p, μ	A_3	

E, F, G, H and Ψ are physically relevant. Each is minimal for the given variables (derivation as before). Obviously this range of potentials gives raise to many Maxwell relations. Note that S is not a thermodynamic potential. Although it also comes with a fundamental equation, it does not obey a minimization principle (rather it obeys a maximization principle).

Different thermodynamic potentials are related to each other by *Legendre transformations*. We now discuss this concept for the one-dimensional case. Consider a monotonous function $y = f(x)$ with a unique inverse $x = f^{-1}(y)$. We want to rewrite $y(x)$ as a function of its derivative $p = f'(x) = g(x)$.

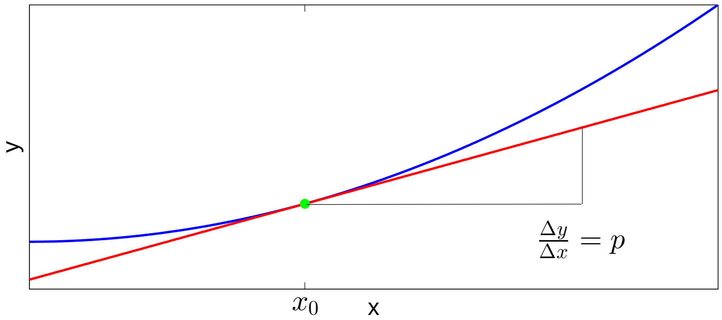

Figure 8.4: y as a function of x with tangent at $x = x_0$

$$\Rightarrow \quad x = g^{-1}(p) = x(p)$$
$$\Rightarrow \quad y(p) = y(x(p)) = f(g^{-1}(p)) = (f \circ g^{-1})(p)$$

However this procedure is not unique:
Curves shifted in x-direction have the same result $y(p)$.

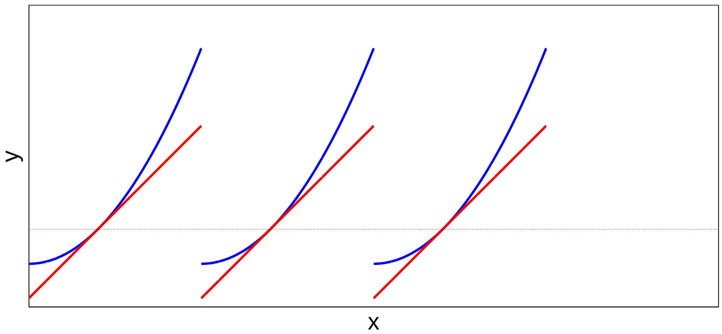

Figure 8.5: Several curves $y_i(x)$ shifted along the x-axis are shown with illustrative tangents.

The underlying reason is that we work with an ODE of first order:

$$y(x) = \left(f \circ g^{-1}\right)(y'(x)),$$

which leaves a constant of integration undetermined.

To solve this problem, we describe the curve $y(x)$ as enveloped by the family of its tangents.

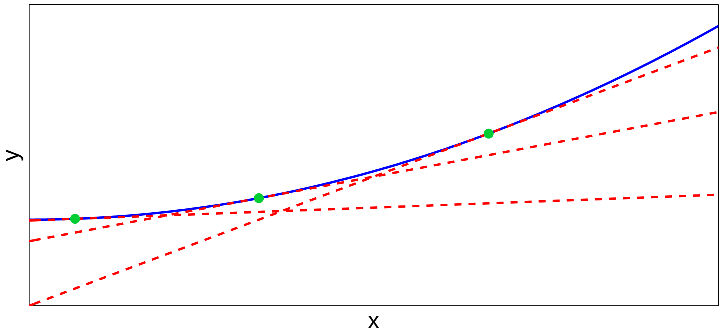

Figure 8.6: y(x) can be described as enveloped by the family of its tangents.

Each tangent is characterised by the slope $p = f'(x)$ and the intercept:

$$p(x) = \frac{y(x) - \Psi(x)}{x} \Rightarrow \Psi(x) = y(x) - p(x)\,x$$

The resulting Legendre transform

$$\Psi(p) = y(x(p)) - p\,x(p)$$

is a unique function of $y(x)$.

Back transform:

$$\begin{aligned} \mathrm{d}y = p\,\mathrm{d}x \quad &\Rightarrow \quad \mathrm{d}\Psi = \mathrm{d}y - p\,\mathrm{d}x - x\,\mathrm{d}p = -x\,\mathrm{d}p \\ &\Rightarrow \quad \frac{d\Psi}{dp} = -x \quad \Rightarrow \quad y(x) = \Psi(p(x)) + p(x)\,x \end{aligned}$$

We note that applying the Legendre transformation twice brings us back to the original function.

Example:

$$y = (x - x_0)^2 \quad \Rightarrow \quad p = 2(x - x_0) \quad \Rightarrow \quad x = \frac{p}{2} + x_0$$

$$\begin{aligned} &\Rightarrow \quad y = (\frac{p}{2})^2 \text{does not depend on } x_0 \\ &\Rightarrow \quad \Psi = y - px = \left(\frac{p}{2}\right)^2 - p\left(\frac{p}{2} + x_0\right) = -\frac{p^2}{4} - px_0 \text{depends on } x_0 \\ &\Rightarrow \quad -x = \frac{d\Psi}{dp} = -\frac{p}{2} - x_0 \quad \Rightarrow \quad p = 2(x - x_0) \\ &\Rightarrow \quad y = \Psi + px = \frac{-(2(x - x_0))^2}{4} - 2(x - x_0)\,x_0 + 2(x - x_0)\,x \\ &\qquad = -(x - x_0)^2 + 2(x - x_0)^2 = (x - x_0)^2 \end{aligned}$$

The Legendre transform of $y(x)$ in regard to $p = y'(x)$ is denoted by $y\,[p]$.

We now see that the free energy F is actually the Legendre transform $E\,[T]$ of energy E from entropy S to temperature T (alternatively it can be derived from $\beta F = S\,[\beta]$).

$$\begin{aligned} E &= E\,(S, V, N) \\ T(S, V, N) &= \frac{\partial E}{\partial S} \quad \Rightarrow \quad S = S(T, V, N) \\ F(T, V, N) &= E\,[T] = E(S(T, V, N), V, N) - T\,S(T, V, N) \end{aligned}$$

$$\Rightarrow \quad \frac{\partial F}{\partial T} = \underbrace{\frac{\partial E}{\partial S}}_{\mathrm{T}} \frac{\partial S}{\partial T} - S - T\,\frac{\partial S}{\partial T} = -S$$

$$\frac{\partial F}{\partial V} = \underbrace{\frac{\partial E}{\partial S}}_{\mathrm{T}} \frac{\partial S}{\partial V} + \frac{\partial E}{\partial V} - T\,\frac{\partial S}{\partial V} = \frac{\partial E}{\partial V} = -p$$

$$\frac{\partial F}{\partial N} = \underbrace{\frac{\partial E}{\partial S}}_{\mathrm{T}} \frac{\partial S}{\partial N} + \frac{\partial E}{\partial N} - T\,\frac{\partial S}{\partial N} = \frac{\partial E}{\partial N} = \mu$$

$$\Rightarrow \quad \boxed{\mathrm{d}F = -S\,\mathrm{d}T - p\,\mathrm{d}V + \mu\,\mathrm{d}N}$$

8.5 Maxwell relations

In the last section we introduced 8 thermodynamic potentials. Each generates $3 \cdot 2/2 = 3$ separate pairs of mixed second derivatives, giving rise to 24 Maxwell relations like this one:

$$\frac{\partial^2 E}{\partial S \partial V} = -\left.\frac{\partial p}{\partial S}\right|_{V,N} = \frac{\partial^2 E}{\partial V \partial S} = \left.\frac{\partial T}{\partial V}\right|_{S,N}$$

It is a unique strength of thermodynamics to generate such surprising relations between seemingly unrelated physical quantities. In order to memorize these relations, the *'thermodynamical square'* or *'König-Born diagram'* (compare Figure 8.7) has been introduced. We keep N fixed and consider the four most important potentials.

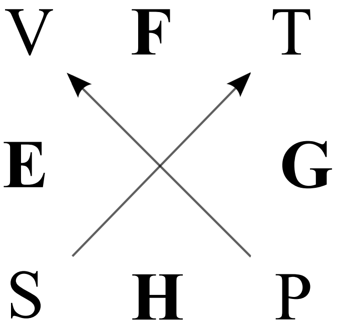

Figure 8.7: The natural variables flank the potentials while the arrows indicate signs.

$$\begin{aligned}
\mathrm{d}E &= T\mathrm{d}S - p\mathrm{d}V \\
\mathrm{d}F &= -p\mathrm{d}V - S\mathrm{d}T \\
\mathrm{d}G &= -S\mathrm{d}T + V\mathrm{d}p \\
\mathrm{d}H &= V\mathrm{d}p + T\mathrm{d}S
\end{aligned}$$

From two neighboring corners a Maxwell relation can be read off (derived from the edge in between, ie H):

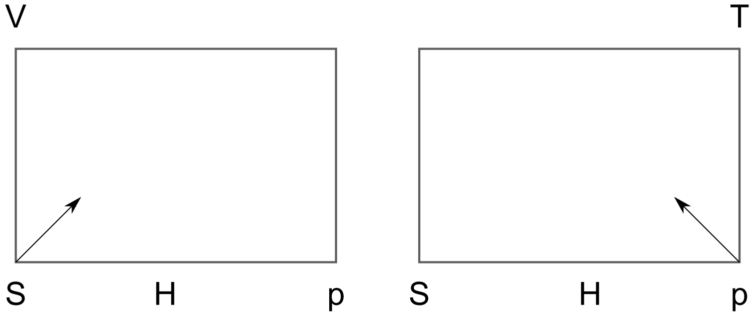

Figure 8.8: Identifying Maxwell relations using the thermodynamical square exemplified for the enthalpy H.

$$\left.\frac{\partial V}{\partial S}\right|_p = \left.\frac{\partial T}{\partial p}\right|_S$$
$$\frac{\partial^2 H}{\partial S\,\partial p} = \frac{\partial^2 H}{\partial p\,\partial S}$$

Three other Maxwell relations follow by rotation of the scheme.

Response functions

Derivatives like $\left.\frac{\partial T}{\partial p}\right|_S$ can be measured in experiments and are called *'response functions'*.

Theorem:

There exist only three independent response functions. All others can be expressed through them.

Proof:

We reduce the derivatives using the thermodynamical square, the mathematical relations between partial derivatives (see last section of this chapter) and the Gibbs-Duhem relation for the chemical potential.

Standard choice:

For constant particle number, we define the following quantities:

$$\alpha := \frac{1}{V}\left.\frac{\partial V}{\partial T}\right|_p \quad \text{coefficient of thermal expansion}$$

$$\kappa_T := -\frac{1}{V}\left.\frac{\partial V}{\partial p}\right|_T \quad \text{isothermal compressibility}$$

$$c_p := \left.\frac{đQ}{\mathrm{d}T}\right|_p = T\left.\frac{\partial S}{\partial T}\right|_p \quad \text{specific heat at constant pressure}$$

This essentially corresponds to a transformation to the Gibbs representation ($g = \frac{G}{N}$):

$$\frac{\partial^2 g}{\partial T^2} = -\frac{c_p}{T}$$

$$\frac{\partial^2 g}{\partial T\,\partial p} = v\,\alpha$$

$$\frac{\partial^2 g}{\partial p^2} = -v\,\kappa_T$$

Example

Relation between c_p and c_V ($N = \text{const}$)

$$\text{heat capacity at constant pressure:} \quad c_p = \left.\frac{đQ}{\mathrm{d}T}\right|_p = T\left.\frac{\partial S}{\partial T}\right|_p$$

$$\text{heat capacity at constant volume:} \quad c_V = \left.\frac{đQ}{\mathrm{d}T}\right|_V = T\left.\frac{\partial S}{\partial T}\right|_V$$

$$\mathrm{d}S = \underbrace{\left.\frac{\partial S}{\partial T}\right|_{V,N}}_{=\frac{c_V}{T}} \mathrm{d}T + \underbrace{\left.\frac{\partial S}{\partial V}\right|_{T,N}}_{=\left.\frac{\partial p}{\partial T}\right|_V} \mathrm{d}V$$

$$= \frac{c_V \mathrm{d}T}{T} + \left.\frac{\partial p}{\partial T}\right|_V \left[\left.\frac{\partial V}{\partial T}\right|_p \mathrm{d}T + \left.\frac{\partial V}{\partial p}\right|_T \mathrm{d}p\right]$$

$$\Rightarrow \quad T\left.\frac{\partial S}{\partial T}\right|_p = c_p = c_V + T\left.\frac{\partial p}{\partial T}\right|_V \underbrace{\left.\frac{\partial V}{\partial T}\right|_p}_{=V\alpha}$$

$$\Rightarrow \quad \left.\frac{\partial p}{\partial T}\right|_V = -\frac{\left.\frac{\partial V}{\partial T}\right|_p}{\left.\frac{\partial V}{\partial p}\right|_T} = \frac{\alpha}{\kappa_T}$$

$$\Rightarrow \quad \boxed{c_p = c_V + TV\,\frac{\alpha^2}{\kappa_T}} > c_V \quad \text{as } \kappa_T > 0,\ \frac{\alpha^2}{\kappa_T} > 0$$

For an ideal gas this yields:

$$\begin{aligned} c_V &= \left.\frac{\mathrm{d}E}{\mathrm{d}T}\right|_V = \frac{3}{2}Nk_B \\ \alpha &= \frac{1}{V}\left.\frac{\partial V}{\partial T}\right|_p = \frac{1}{T} \\ \kappa_T &= -\frac{1}{V}\left.\frac{\partial V}{\partial p}\right|_T = \frac{1}{p} \\ \Rightarrow \quad c_p &= c_V + \frac{VTp}{T^2} = c_V + Nk_B = \frac{5}{2}Nk_B \end{aligned}$$

8.6 Process-dependence of work and heat

Recall the total differential for energy:

$$\mathrm{d}E = T\,\mathrm{d}S - p\,\mathrm{d}V + \mu\,\mathrm{d}N$$

It is instructive to compare with mechanics:

$$\begin{aligned} \mathrm{d}E &= \vec{v}\cdot\mathrm{d}\vec{p} - \vec{F}\cdot\mathrm{d}\vec{r} \\ &= \frac{\vec{p}}{m}\cdot\mathrm{d}\vec{p} + \nabla V\cdot\mathrm{d}\vec{r} \\ &= \mathrm{d}\left(\frac{p^2}{2m}\right) + \mathrm{d}V \end{aligned}$$

We see that each term by itself is a total differential. This is not true in thermodynamics because $T = (S, V, N) \neq T(S)$.
Note that in both cases $E(B)$ is a state function whose value is independent of how one gets from A to B. In this sense the system is conservative. However, the way in which energy is divided between heat and mechanical work is not universal and in this sense the system is not conservative.
Despite the path-dependent weights of the different energy forms, the partial derivatives are not arbitrary because they must belong to a fundamental equation. Therefore we must have eg

$$\frac{\partial^2 E}{\partial V \partial S} = \underbrace{\left.\frac{\partial T}{\partial V}\right|_{S,N}}_{\text{change in T during adiabatic expansion}} = \frac{\partial^2 E}{\partial S \partial V} = \underbrace{-\left.\frac{\partial p}{\partial S}\right|_{V,N}}_{\text{change in p during isochoric heating}}$$

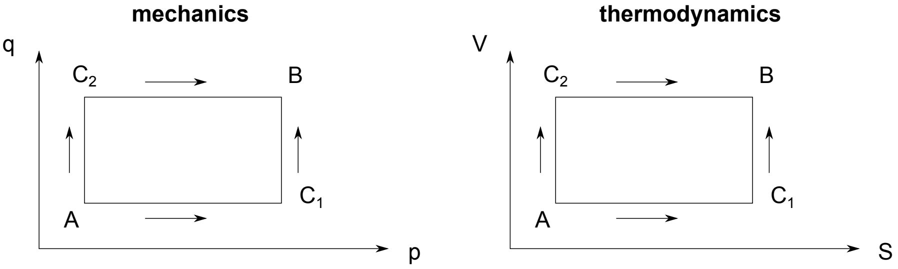

Figure 8.9: Mechanics: The contributions of the integrals $\int \frac{\partial E}{\partial p} \mathrm{d}p$ and $\int \frac{\partial E}{\partial q} \mathrm{d}q$ are the same for the different paths C_1 and C_2 going from A to B. The change in kinetic and potential energy is path-independent.
Thermodynamics: The contributions of the integrals $\int \frac{\partial E}{\partial S} \mathrm{d}S$ and $\int \frac{\partial E}{\partial V} \mathrm{d}V$ can differ for the different paths C_1 and C_2. The choice of path can determine how heat and work are distributed.

This expression relates quantities which at first seem unrelated. We will later meet the whole set of these *'Maxwell relations'*. It is the particular strength of thermodynamics to provide these surprising relations which all have been experimentally verified.
If we neglect changes in particle number, we have

$$\mathrm{d}E = \underbrace{T\mathrm{d}S}_{\text{heat}} - \underbrace{p\mathrm{d}V}_{\text{mechanical work}}$$

We have seen before that the two quantities depend on the path taken from state A to B. We therefore write:

$$\boxed{\mathrm{d}E = đQ + đW} \quad \text{first law of thermodynamics (energy conservation)}$$

Here đ indicates *'incomplete differentials'*.

sign convention: $đW > 0$ If the work increases the energy of the system.
$đQ > 0$ Heat flows into the system and increases its energy.

Very often one discusses expansion of gases. Then the mechanical work is $đW = -p\mathrm{d}V < 0$, i.e. the system does work and loses energy, because p is positive and $\mathrm{d}V > 0$.
James Joule was the first to realize that one can always measure $đQ$ and $đW$ for given states A and B:

① First make an adiabatic (isentropic) experiment (thermally isolating walls). This implies:

$$đQ = 0 \quad \Rightarrow \quad \mathrm{d}E = đW$$

If $A \to B$ does not work because $\Delta S < 0$, use $B \to A$ with $\Delta S > 0$.
Because $đW$ can be measured by mechanical means, one gets ΔE_{AB} in this way. Joule invented many devices to do exactly this.

② Now make an experiment of interest with heat exchange (diathermal walls) connecting the same two states and measuring W_{AB} (now with another result).

$$\Rightarrow \quad Q_{AB} = \Delta E_{AB} - W_{AB}$$

As ΔE_{AB} is known due to our first step, we now can calculate Q_{AB}.

The essential point here is that E is a state function, while Q and W are not.

Pond analogy

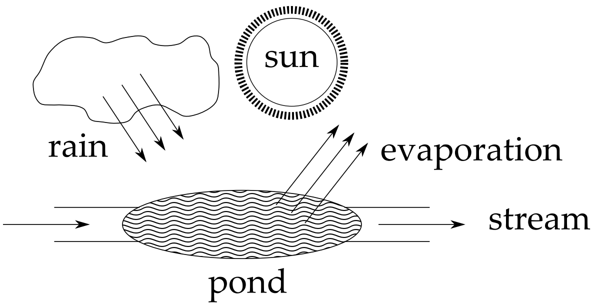

Figure 8.10: The amount of water in the pond depends on rain, evaporation and in- and outflow due to a stream.

Consider a farmer who wants to know how much water is in his pond. There are two ways its amount can change:
in- and outflow through a stream, and increase/ decrease by rain/ evaporation.
How can he control the relative importance of these two channels?
The solution is simple: First he covers the pond by a tarpaulin. He then can calibrate the water height by using flow meters in the stream. Finally removing the tarpaulin, he now can calculate back how much water results from rain/ evaporation.

Expansion of an ideal gas

We now discuss the expansion of an ideal gas as an example for the process-dependance of thermodynamic processes. We go from state A to B as shown in Figure 8.11 .

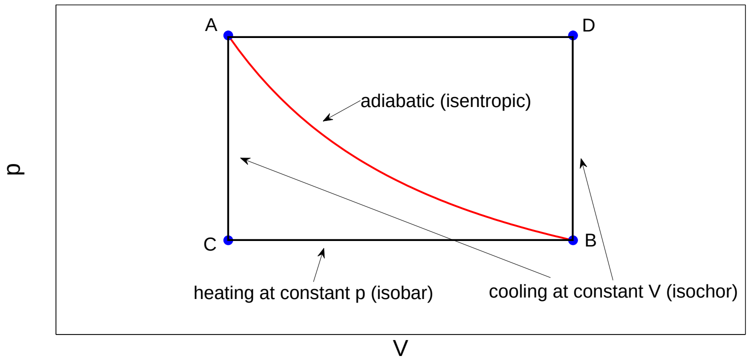

Figure 8.11: Three different paths from A to B. In each case the system cools.

Fundamental equation:

$$S = S_0 + k_B N \ \ln \left(\frac{V E^{\frac{3}{2}}}{N^{\frac{5}{2}}} \right)$$

For an adiabatic process (S=const) at N=const we therefore have

$$V \ E^{\frac{3}{2}} = \text{const}$$

$$E = \frac{3}{2} N k_B T = \frac{3}{2} p V \quad \Rightarrow \quad V^5 p^3 = \text{const}$$

$$\Rightarrow \quad \Delta E_{AB} = \int \text{đ}W = - \int_{V_A}^{V_B} p \ \mathrm{d}V = -p_A \int_{V_A}^{V_B} \left(\frac{V_A}{V} \right)^{\frac{5}{3}} \mathrm{d}V$$
$$= \frac{3}{2} p_A V_A^{\frac{5}{3}} \left(V_B^{-\frac{2}{3}} - V_A^{-\frac{2}{3}} \right) < 0$$

The gas is doing work and looses energy.

We now calculate work and heat for the paths through D and C.

$$W_{ADB} = - \int p \ \mathrm{d}V = -p_A \ (V_B - V_A) < \Delta E_{AB} < 0$$

The system is doing even more work.

$$Q_{ADB} = \Delta E_{AB} - W_{ADB} > 0$$

Heat flows into the system to compensate for this extra work.

$$W_{ACB} = - \int p \ \mathrm{d}V = -p_B \ (V_B - V_A) > \Delta E_{AB} < 0$$

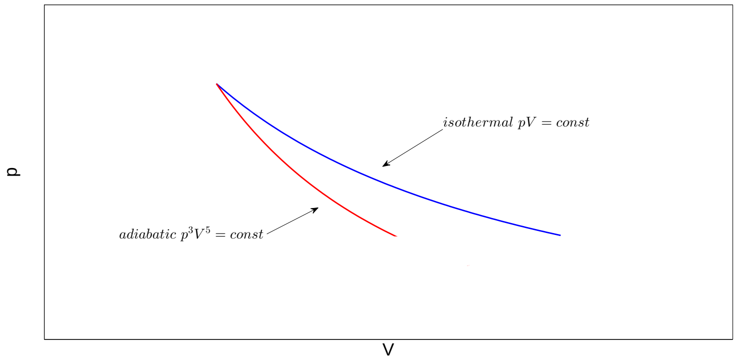

Figure 8.12: The adiabatic curve is stepper than the isothermal one. A combination can be used to get from A to D in Figure 8.11.

The system is doing work, but less compared with the two other paths.

$$Q_{ACB} = \Delta E_{AB} - W_{ACB} < 0$$

Heat flows from the system, lowering its energy.
Note that we did not need to calculate ΔE_{AD} or ΔE_{AC}. Indeed this is more complicated and requires the construction shown in Figure 8.12: by combining an adiabatic with an isothermal process, one can go from A to D.

8.7 Reversible and irreversible processes

Both statistical physics and classical thermodynamics state

$$\boxed{\Delta S \geq 0} \quad \text{second law of thermodynamics}$$

$\Delta S = 0$ reversible process: can go both ways
$\Delta S > 0$ irreversible process: other way cannot occur spontaneously

Examples

① adiabatic expansion of an ideal gas

The piston is moved out with different velocities and the complete system is thermally isolated. Therefore there is no heat flux, $\text{đ}Q = 0$. We consider the extreme cases concerning the piston's velocity:

① very fast: expansion into vacuum

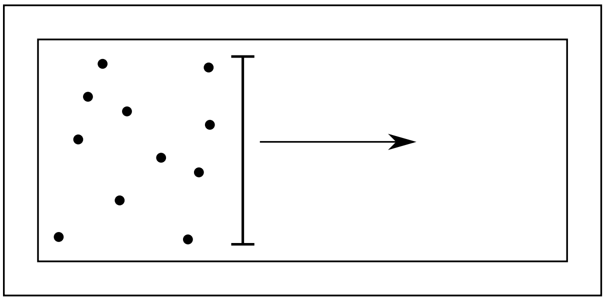

Figure 8.13: The ideal gas expands while the piston is moved out.

② very slow: quasi-static, pressure is always equilibrium pressure $p = Nk_BT/V$
Fundamental equation:

$$S = S_0 + k_B N \ln\left(\frac{VE^{\frac{3}{2}}}{N^{\frac{5}{2}}}\right)$$

<u>expansion into vacuum (case ①):</u>

$$\begin{aligned} dE &= đW = 0 \\ \mathrm{d}S &= \frac{\partial S}{\partial V} dV = k_B N \frac{1}{V}\,\mathrm{d}V > 0 \end{aligned}$$

We see that no work is being done and that the process is irreversible. The entropy has to go up because now many more microstates become available.

<u>quasi-static (case ②):</u>

$$\begin{aligned} \mathrm{d}E &= đW = -p\,\mathrm{d}V \\ &= -\frac{Nk_BT}{V}\,\mathrm{d}V = -\frac{2}{3}\frac{E}{V}\,\mathrm{d}V \\ \mathrm{d}S &= \frac{\partial S}{\partial V} dV + \frac{\partial S}{\partial E} dE = Nk_B \frac{1}{V}\,\mathrm{d}V + \frac{3}{2} Nk_B \frac{1}{E}\,\mathrm{d}E = 0 \end{aligned}$$

The gas looses energy because it does work (and hence also becomes colder). The process is reversible, because the entropy increase due to the volume increase is exactly balanced by the loss of entropy due to the decrease in energy.

This situation is easy to analyze because we consider an isolated system for which we know everything. The situation is more complex if we couple it to the environment. Consider for example isothermal quasi-static expansion, so we couple the piston to a heat bath. Then it does the same work as above. However, because now T is constant, E is also constant and heat has to flow in such that $đQ = -đW$.

The entropy of the piston goes up, but the reservoir loses exactly the same amount of entropy due to the heat flow and the overall entropy is constant. Therefore isothermal quasi-static expansion is also reversible.

② warm bottle B in cold lake L

We assume that the heat capacities of bottle (c_B) and lake (c_L) are constant.
Fundamental equation:

$$S = S(E, V, N) = S_0 + c\ \ln \frac{E}{E_0}$$

$$\frac{1}{T} = \frac{\partial S}{\partial E} = c\ \frac{1}{E} \quad \Rightarrow \quad E = c\ T$$

$$\Rightarrow c = \frac{dE}{dT}$$

We now bring the two systems in thermal contact. The bottle will be cooled from T_a to T_b. It will give away heat

$$Q_B = \Delta E_B = c_B\,(T_b - T_a) \leq 0$$

while no work is done due to the volume being constant.

$$\Delta E = \Delta E_B + \Delta E_L = Q_B + Q_L = 0$$

$$\Rightarrow \quad \Delta T_L = \frac{|Q_L|}{c_L} = \frac{|Q_B|}{c_L} = \frac{c_B}{c_L}\ \Delta T_B \approx 0$$

As $c_B \ll c_L$ the temperature change of the lake can be neglected.

Changes in entropy:

$$\Delta S_L = \int_a^b \frac{đQ_L}{T} = \frac{Q_L}{T_b} = \frac{-Q_B}{T_b} = c_B \frac{T_a - T_b}{T_b} \geq 0 \quad \text{lake gains entropy}$$

$$\Delta S_B = \int_a^b \frac{đQ_B}{T} = \int_{T_a}^{T_b} \frac{c_B\ \mathrm{d}T}{T} = c_B\ \ln\left(\frac{T_b}{T_a}\right) \leq 0 \quad \text{bottle looses entropy}$$

The overall change in entropy thus is (defining $z := T_a/T_b \geq 1$):

$$\Delta S = \Delta S_L + \Delta S_B = c_B\ \ (z - 1 - \ln z)$$

Our result is in agreement with the second law of thermodynamics:

$$\ln z \leq z - 1 \quad \Rightarrow \quad z - 1 - \ln z \geq 0 \quad \Rightarrow \quad \Delta S \geq 0$$

The equal sign is valid for $z = 1$ ($T_a = T_b$). Otherwise $\Delta S > 0$ and the process is irreversible as heat flows from the warmer to the cooler system. Note that the same conclusion holds if we place a cold bottle in a warm lake ($z < 0$).

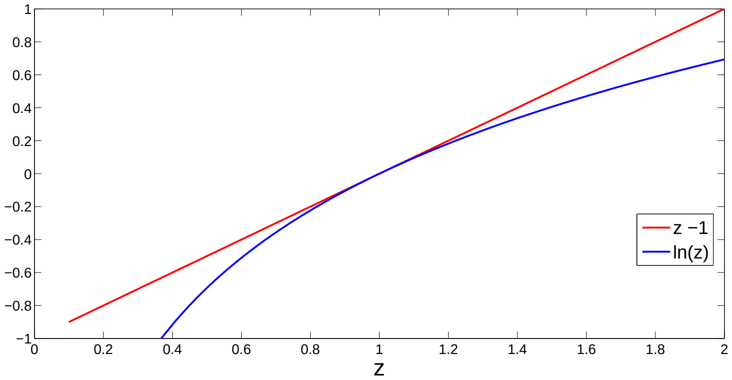

Figure 8.14: $f(z) = z - 1$ and $g(z) = \ln(z)$ as a function of z.

8.8 Thermodynamic engines

We again consider heat flow from a warm to a cold body.
We assume constant heat capacities. The change in energy of the complete system then can be expressed as:

$$\Delta E = \int_{T_{10}}^{T_f} c_1 \, \mathrm{d}T_1 + \int_{T_{20}}^{T_f} c_2 \, \mathrm{d}T_2 \overset{!}{=} 0$$

$$\Rightarrow \quad T_f = \frac{c_1 T_{10} + c_2 T_{20}}{c_1 + c_2}$$

The change in entropy is:

$$\Delta S = \int_{T_{10}}^{T_f} \frac{c_1 \mathrm{d}T_1}{T_1} + \int_{T_{20}}^{T_f} \frac{c_2 \mathrm{d}T_2}{T_2}$$

$$= c_1 \ln \frac{T_f}{T_{10}} + c_2 \ln \frac{T_f}{T_{20}}$$

Assuming equal heat capacities: $c_1 = c_2 = c$

$$\Rightarrow \quad T_f = \frac{T_{10} + T_{20}}{2}$$

$$\Delta S = 2c \ln \left(\frac{T_f}{\sqrt{T_{10} T_{20}}} \right)$$

$$= 2c \ln \left(\frac{T_{10} + T_{20}}{2 \sqrt{T_{10} T_{20}}} \right) \geq 0$$

The change in entropy is always positive as the arithmetic mean is always larger than or equal to the geometrical mean.

$$\frac{T_{10} + T_{20}}{2} \geq \sqrt{T_{10} T_{20}}$$

We can see this as follows:

$$(a-b)^2 \geq 0 \;\Rightarrow\; a^2 + 2ab + b^2 \geq 4ab \;\Rightarrow\; \frac{(a+b)^2}{4} \geq ab$$

The process would be reversible for

$$T_f = \sqrt{T_{10} T_{20}}$$

However, this would lead to an energy

$$E = 2c\,\sqrt{T_{10} T_{20}} < 2c\,\frac{T_{10} + T_{20}}{2}$$

The energy difference

$$E = 2c\,\left[\frac{T_{10} + T_{20}}{2} - \sqrt{T_{10} T_{20}}\right]$$

had to be spent as work. This can be accomplished by a thermodynamic engine (*Wärmekraftmaschine'*).

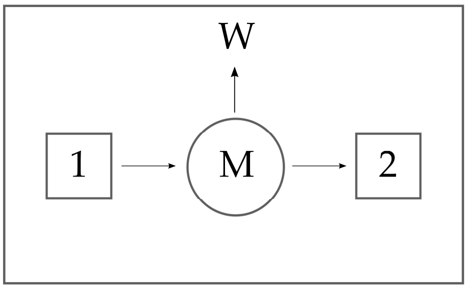

Figure 8.15: Scheme of a power plant: 1) furnace (coal, oil ...), 2.) environment (cooling tower), M) turbine, W) power line

How much work can one get out of the system?
We consider a final temperature T_f:

$$\sqrt{T_{10} T_{20}} \leq T_f \leq \frac{T_{10} + T_{20}}{2}$$

This range for the final temperature is bounded by the reversible case from below (a smaller value would correspond to negative entropy change) and the completely spontaneous process without any work being done from above (a larger value would correspond to influx of work into the system, corresponding to a heat pump but not to a heat engine).

$$\begin{aligned} Q_1 &= c\ \left(T_{10} - T_f\right) \quad \text{heat taken up by M} \\ Q_2 &= c\ \left(T_f - T_{20}\right) \quad \text{heat given away by M} \\ W &= Q_1 - Q_2 \quad \text{work done by the system} \end{aligned}$$

The lower T_f the more work we can get out of our machine. Work is produced by 'slowing down' the heat exchange.

We consider the two extreme cases:

① $T_f = \frac{1}{2}\ (T_{10} + T_{20})$
no work, maximal entropy production (compare gas expansion to vacuum)

② $T_f = \sqrt{T_{10} T_{20}}$
maximal work, reversible process (compare gas expansion under quasi-static pressure changes)

We note that producing entropy is a waste of work.

We define a thermodynamic efficiency *('Wirkungsgrad')*:

$$\eta := \frac{W}{Q_1} = 1 - \frac{Q_2}{Q_1}$$

For case 1 we get:

$$\eta = 1 - \frac{Q_2}{Q_1} = 0$$

$Q_2 = 0$ would describe a perpetuum mobile of the second kind which cannot exist. This is due to a negative entropy change for the furnace system which violates the second law of thermodynamics:

$$\Delta S = -\frac{Q_1}{T_1}$$

As it always takes two to tango: Heat flow away from system 1 requires heat uptake by a second system to ensure $\mathrm{d}S \geq 0$.

$$\begin{aligned} \Rightarrow \quad \mathrm{d}S &= -\frac{đQ_1}{T_1} + \frac{đQ_2}{T_2} \geq 0 \\ \Rightarrow \quad đQ_2 &\geq đQ_1\,\frac{T_2}{T_1} \quad \text{lower bound for } Q_2 \\ \Rightarrow \quad \eta &= 1 - \frac{Q_2}{Q_1} \leq 1 - \frac{T_2}{T_1} = \eta_{ideal} \end{aligned}$$

No thermodynamic engine can be more efficient than η_{ideal}.

Maxwell demon

The following thought experiment has been suggested for a perpetuum mobile of the second kind and goes back to Maxwell (1867).
A device *('Maxwell demon')* (see 8.16) opens a little trapdoor each time a gas molecule comes from the right (*'ratchet mechanism'*) . Then pressure raises on the left, which can be used to extract work. This also cools the system, which is compensated by heat flux from the environment. In this way, the device can produce work out of heat, seemingly violating the second law.

This and many other similar setups have been analyzed by many researchers. Feynman argued that it cannot work because the device itself will fluctuate (*'trembling demon'*). Landauer built his argumentation on the need for storing information, thereby raising entropy.

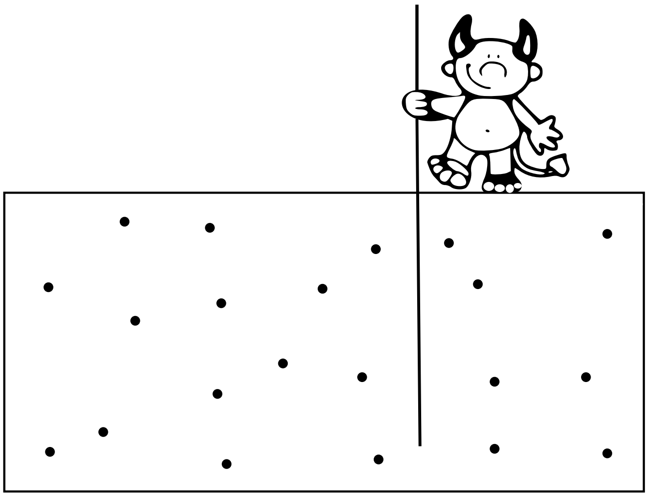

Figure 8.16: Maxwell's demon at work selectively letting gas particles pass.

When analyzing the thermodynamic engine, we assumed that M does not store heat or work. This is certainly true for a machine working in a cyclic manner.
How can one construct such a thermodynamic engine in practice?

Carnot cycle

The Carnot cycle 8.17 uses an ideal gas and four steps.

① The gas is in contact with a hot reservoir and isothermally expanded. Entropy increases while energy is constant. Heat is taken up and work is delivered.

② The gas is adiabatically expanded. Entropy is constant, more work is delivered.

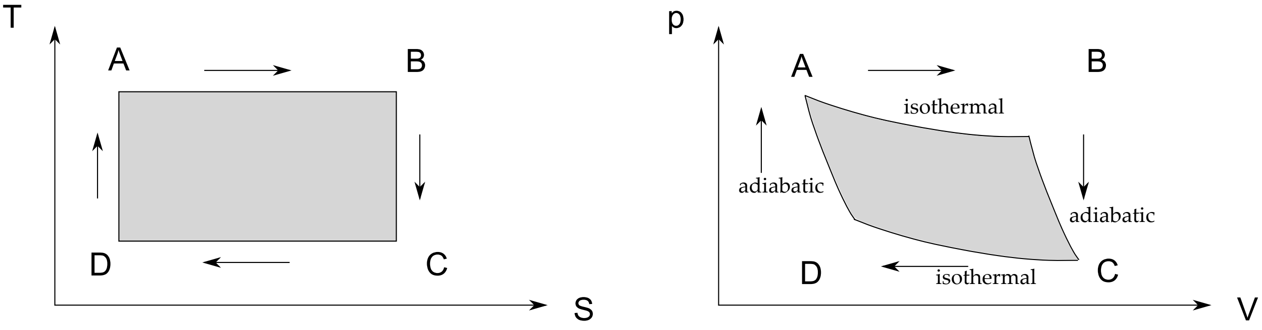

Figure 8.17: Carnot cycle

③ The gas is contacted with a cold reservoir and isothermally contracted. Work is consumed by the system, entropy decreases and heat is given to the reservoir.

④ The gas is adiabatically contracted. Entropy is constant, more work is consumed.

The grey area in the S-T-diagram is the heat delivered to the outside. The grey area in the V-p-diagram is the work delivered. The thermodynamic efficiency can be shown to be:

$$\eta = \frac{W}{Q_{AB}} = 1 - \frac{T_2}{T_1} = \eta_{ideal}$$

Because the Carnot cycle is reversible, it can also be used in the other direction. Then we use work to transfer heat from the cold to the hot reservoir (*'heat pump'* or *'refrigerator'*).

For a power plant we typically have:

$$\begin{aligned} T_1 &= 540\ ^\circ\mathrm{C} = 813\ K \\ T_2 &= 40\ ^\circ\mathrm{C} = 313\ K \\ \Rightarrow \quad \eta_{ideal} &= 62\% \end{aligned}$$

In practice one can reach around 45% which corresponds to 5 kWh output energy from 1 l of oil.
Heat production is unavoidable but can be used for heating (*'Wärmekraftkopplung'*).

9 Stochastic thermodynamics

9.1 A new research field between statistical physics and thermodynamics

Thermodynamics emerges from statistical physics in the limit of large system size, when fluctuations vanish due to the law of large numbers and only the mean values are relevant. Therefore quantities like the internal energy E, which have probability distributions in statistical physics, are deterministic in thermodynamics. The same then applies to work W and heat Q, which in thermodynamics have incomplete differentials, but have no fluctuations.

Recently this traditional viewpoint of thermodynamics has been extended to smaller ("mesoscopic")systems, and a new field has emerged that considers work and heat to be realized according to distributions that can be calculated. The main idea is to go from the ensemble picture to a description of single trajectories and to assign quantities like work, heat and entropy, that traditionally have been considered to be systems properties, to such trajectories. This approach is closely related to the method of path integrals, which in turn is related to partition sums, so it is actually a natural approach and not completely new. Also the fact that work fluctuates is clear to anyone who always has done an actual experiment. Because we deal with dynamic trajectories, this is a subject of non-equilibrium physics, but surprisingly, many exact relations have been demonstrated (for example the Jarzynski relation or the integral fluctuation theorem) which have similarities with the laws of thermodynamics and thus earlier have not been thought to be possible out of equilibrium. Stochastic thermodynamics comes in many different variants, but here we will report only on the one that builds on the overdamped Langevin equation (other interesting examples are Hamiltonian systems or master equations). This corresponds to the canonical ensemble with fixed temperature T.[1]

9.2 Langevin and Fokker-Planck equations

We first introduce the concept of stochastic differential equations (the Langevin equation, which is equivalent to the Fokker-Planck equation). Such an equation describes

---

[1] The first book on this subject is *Stochastic thermodynamics: an introduction* by Luca Peliti and and Simone Pigolotti. Excellent reviews are Felix Ritort, *Work Fluctuations, Transient Violations of the Second Law and Free-Energy Recovery Methods: Perspectives in Theory and Experiments*, Poincare Seminar 2003, and Udo Seifert, *Stochastic thermodynamics, fluctuation theorems and molecular machines*, Reports on Progress in Physics 2012.

Brownian or random walks, which we have already touched on in the introduction. We consider the simplest case for a microscopic dynamics: a spherical particle in aqueous solution (compare Fig. 9.1). Collisions with the water molecules keep it in continuous motion. Obviously this is an effect of finite temperature. How can we describe this *'Brownian motion'*?

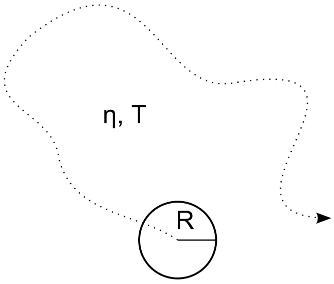

Figure 9.1: A spherical particle of radius R in a fluid with viscosity η and temperature T is performing a *'Brownian random walk'*.

For simplicity we work in one dimension. Newton's second law yields a differential equation:

$$\underbrace{m\ddot{x}}_{\text{inertial force}} = m\dot{v} = \underbrace{-\xi v}_{\text{friction force}}$$

where ξ is the friction coefficient. The solution to the equation describes an exponential decay:

$$v(t) = v_0\, e^{-t/t_0} \quad \text{with } t_0 = \frac{m}{\xi}$$

and hence over time the particle comes to rest, which is not what we observe.
By adding a random force continuously kicking the particle we arrive at the *'Langevin equation'*:

$$m\dot{v} = -\xi v + \sigma\eta(t)$$

σ is the amplitude of the thermal noise and η describes Gaussian white noise which obeys:

1. $\langle\eta(t)\rangle = 0$
2. $\langle\eta(t)\eta(t')\rangle = 2\delta\,(t-t')$

The formal solution is given by:

$$v(t) = e^{-t/t_0}\left(v_0 + \int_0^t \mathrm{d}s\; e^{s/t_0}\frac{\sigma}{m}\eta(s)\right)$$

$$\Rightarrow \quad \langle v(t)\rangle = v_0 e^{-t/t_0}$$

$$\langle v(t)v(t')\rangle = v_0^2 e^{-\frac{t+t'}{t_0}} + \left(\frac{\sigma}{m}\right)^2 e^{-\frac{t+t'}{t_0}} \underbrace{\int_0^t \mathrm{d}s \int_0^{t'} \mathrm{d}s'\, e^{\frac{s+s'}{t_0}} 2\delta\left(s-s'\right)}_{\overset{t\leq t'}{=} \int_0^t \mathrm{d}s\ 2e^{2s/t_0} = t_0\left(e^{2t/t_0}-1\right)}$$

$$= \underbrace{e^{-\frac{t+t'}{t_0}} \left(v_o^2 - \frac{\sigma^2}{m\xi}\right)}_{=0 \text{ for } t,t'\gg t_0} + \frac{\sigma^2}{m\xi} e^{(t'-t)/t_0}$$

$$\Rightarrow \quad \left\langle v(t)^2 \right\rangle = \frac{\sigma^2}{m\xi}$$

The random kicks from the environment keep the particle in motion.

Equipartition theorem:

$$\frac{1}{2} m \left\langle v^2 \right\rangle = \frac{1}{2} k_B T$$

$$\Rightarrow \quad \boxed{\sigma^2 = \xi k_B T} \quad \text{fluctuation-dissipation theorem}$$

The noise amplitude σ (fluctuations) is related to the friction coefficient (dissipation) through temperature. The higher T, the stronger the noise.

For $t \gg t_0$, we can neglect inertia:

$$\Rightarrow \quad \xi v = \sigma\eta(t) = \xi \dot{x}$$

$$\Rightarrow \quad x(t) = x_0 + \frac{1}{\xi}\int_0^t \mathrm{d}t' \sigma\eta(t')$$

$$\Rightarrow \quad \langle x(t) \rangle = x_0$$

$$\left\langle \left(x(t) - x_0\right)^2 \right\rangle = \frac{1}{\xi^2} \int_0^t \mathrm{d}t' \int_0^t \mathrm{d}t''\ 2\sigma^2 \delta\left(t'-t''\right)$$

$$= \frac{1}{\xi^2} 2\sigma^2 t \overset{!}{=} 2Dt$$

Here we identified the diffusion constant D from the one dimensional random walk.

$$\Rightarrow \quad \boxed{D = \frac{\sigma^2}{\xi^2} = \frac{k_B T}{\xi}} \quad \text{Einstein relation}$$

If we use for the friction coefficient Stoke's law from hydrodynamics, $\xi = 6\pi\eta R$ with viscosity η we get:

$$\Rightarrow \quad \boxed{D = \frac{k_B T}{6\pi\eta R}} \quad \text{Stokes-Einstein relation}$$

The Langevin equation is a *'stochastic differential equation'* and requires *'stochastic calculus'*. Alternatively one can derive an equation for the probability $p\,(x,t)$ to be at position x at time t starting from the Langevin equation:

$$\xi \dot{x} = F(x) + \sigma \eta(t)$$

$$\Rightarrow \quad \boxed{\dot{p}(x,t) = \underbrace{-\frac{1}{\xi}\partial_x \left(F(x)p(x,t)\right)}_{\text{drift}} + \underbrace{D\partial_x^2 p(x,t)}_{\text{diffusion}}}$$

This is the *Fokker-Planck* or *Smoluchovski* equation. It can be written as a continuity equation

$$\dot{p} + \partial_x J = 0$$

with probability current

$$J = \frac{1}{\xi}Fp - D\partial_x p$$

In the case of detailed balance:

$$J = 0$$

$$\begin{aligned} \Rightarrow \quad & \frac{1}{\xi}Fp = D\partial_x p \\ \Rightarrow \quad & \frac{\partial_x p}{p} = \partial_x \ \ln p = \frac{F}{\xi D} \\ \Rightarrow \quad & p = p_0 e^{\int \mathrm{d}x \ \frac{F}{\xi D}} = p_0 e^{-\frac{V}{\xi D}} \\ & = p_0 e^{-\beta V} \end{aligned}$$

Here we used the definition $F = -V'$ of the potential in the second and the Einstein relation in the last step. We note that the Boltzmann distribution arises as stationary solution of the Fokker-Planck equation.

9.3 Stochastic energetics and the first law of thermodynamics

The Langevin equation can be considered to be a balance of forces. To go from forces to energy, one has to multiply with distance and then to integrate. Let us first discuss how this works for Newton's 2nd law by multiplying it with $\dot{x}$:

$$m\ddot{x} = F = -\frac{dV}{dx} \Rightarrow \frac{d}{dt}(\frac{1}{2}m\dot{x}^2) = -\frac{dV}{dx}\dot{x} = -\frac{dV}{dt} \Rightarrow \frac{1}{2}m\dot{x}^2 + V = const$$

Bingo, we managed to show energy conservation. Easy. Surprisingly, the same idea has not been applied to the Langevin equation for a long time, until Sekimoto did so in

1998 (Sekimoto, Ken. "Langevin equation and thermodynamics." Progress of Theoretical Physics Supplement 130 (1998): 17-27). We take the Langevin equation from above. Very importantly, we consider a conservative force, $F = -dV/dx$ (in the general case, we also would add a non-conservative force here), but we now consider the potential $V(x(t), \lambda)$ to depend not only on position $x(t)$, but also on some external control parameter λ, which is our way to do work on the system. A simple example would be $\lambda = t$, if we have a time-dependent potential for example by switching on a laser potential for a colloidal bead. We now write the Langevin equation and then multiply with dx to go from forces to energies:

$$\xi \dot{x} = -\frac{\partial V(x(t), \lambda)}{\partial x} + \sigma \eta(t) \Rightarrow 0 = -(-\xi \dot{x} + \sigma \eta(t))dx + \frac{\partial V(x(t), \lambda)}{\partial x} dx$$

We first note that the term in brackets is the reaction force to the heat bath and therefore the corresponding energy should be identified with the differential heat dQ released into the environment. Another way is to argue from force balance that this term is $Fdx = -\frac{\partial V}{\partial x} dx$, and that in an overdamped system all forces are dissipated. We next note that the term on the right hand side is not a total differential, because V depends on x and λ. We therefore add 0 here:

$$0 = dQ + \frac{\partial V(x(t), \lambda)}{\partial x} dx + \frac{\partial V(x(t), \lambda)}{\partial \lambda} d\lambda - \frac{\partial V(x(t), \lambda)}{\partial \lambda} d\lambda$$

We now have a total differential dV on the right hand side and can identify $dW = \frac{dV(x(t),\lambda)}{d\lambda} d\lambda$ as the work that is being done on the system. Thus we get

$$dW = dQ + dV$$

which is our microscopic version of the first law of thermodynamics (energy conservation), $dV = dE = dW - dQ$. Note that heat here appears with a minus sign, different from the convention in thermodynamics (the convention here is that heat is positive if is released into the environment).

In order to complete this procedure, we still have to integrate over time. This converts the differentials into functionals of the trajectories $x(t)$:

$$W[x(t)] = \int dt \frac{\partial V}{\partial \lambda} \dot{\lambda}, \quad Q[x(t)] = -\int dt \frac{\partial V}{\partial x} \dot{x}$$

The work and heat functionals can be measured in experiments for colloidal beads and have been shown to add up to the change in internal energy, as expected (Blickle, Valentin, et al. "Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential." Physical review letters 96.7 (2006): 070603).

9.4 Jarzynski relation, the second law of thermodynamics and violating trajectories

Having identified the first law in the Langevin equation, we now can ask if we also can identify entropy and the second law on the level of trajectories. This can be achieved with the help of the Jarzynski equality, which has been published in 1997 in a paper that now is cited more than 5.000 times (Jarzynski, Christopher. "Nonequilibrium equality for free energy differences." Physical Review Letters 78.14 (1997): 2690):

$$\left\langle e^{-W/k_B T} \right\rangle = e^{-\Delta F/k_B T}$$

This equation has been first proven for Hamitonian systems and later also extended to Langevin and quantum systems. It is very surprising because the left is an average over all non-equilibrium trajectories connecting two states with each other, while the right is an equilibrium quantity (ΔF is the free energy difference). We can also write this equation as

$$\left\langle e^{-(W-\Delta F)/k_B T} \right\rangle = 1$$

and then consider the energy term in more detail using the first law from above and the definition of free energy:

$$W - \Delta F = Q + \Delta E - (\Delta E - T\Delta S) = T\Delta S_r + T\Delta S = T\Delta S_{tot} = W_{diss}$$

where $\Delta S_r = Q/T$ (Q is the heat dissipated into the reservoir) and ΔS are the entropies of the reservoir and the subsystem, respectively. $\Delta E = \Delta V$ has dropped out here. In the last step we have identified this expression with the dissipated work, that is the part of the work which cannot be recovered again (the thermodynamic potential ΔF can be argued to be the maximal work one can extract from the system). If we now use Jensen's inequality on Jarzynski's equation, we can write:

$$1 = \left\langle e^{-W_{diss}/k_B T} \right\rangle \geq e^{-\langle W_{diss} \rangle / k_B T}$$

Because e^{-x} decays down from 1 if x starts at 0, we conclude that

$$\langle W_{diss} \rangle = T\Delta S_{tot} \geq 0$$

which is nothing else than the second law of thermodynamics (increase in total entropy).

Up to now it looks as if we managed to identify the second law of thermodynamics on the level of single trajectories, but in fact we found something which is even more surprising. Going beyond Jensen's inequality and looking at the full average again, we see that one can average the function e^{-x} to an overall value 1 only if there is a finite probability also for negative values of W_{diss} (e^{-x} is always smaller than 1 for positive x, so we need some negative x with values larger than 1 to get an average of 1). So although for most trajectories the realization is positive, to get the positive mean, there must be some with a negative value. These are called *violating trajectories* and mean that the reservoir can help the system to perform work. Such a process is completely prohibited in thermodynamics, but can occur in mesoscopic systems. It has been proven experimentally (e.g. for colloidal beads or biomolecules) and it is ongoing speculation if this can be used to extract useful work (or information) from thermal reservoirs (with the help of *Maxwell demon*, which filters out the violating trajectories).

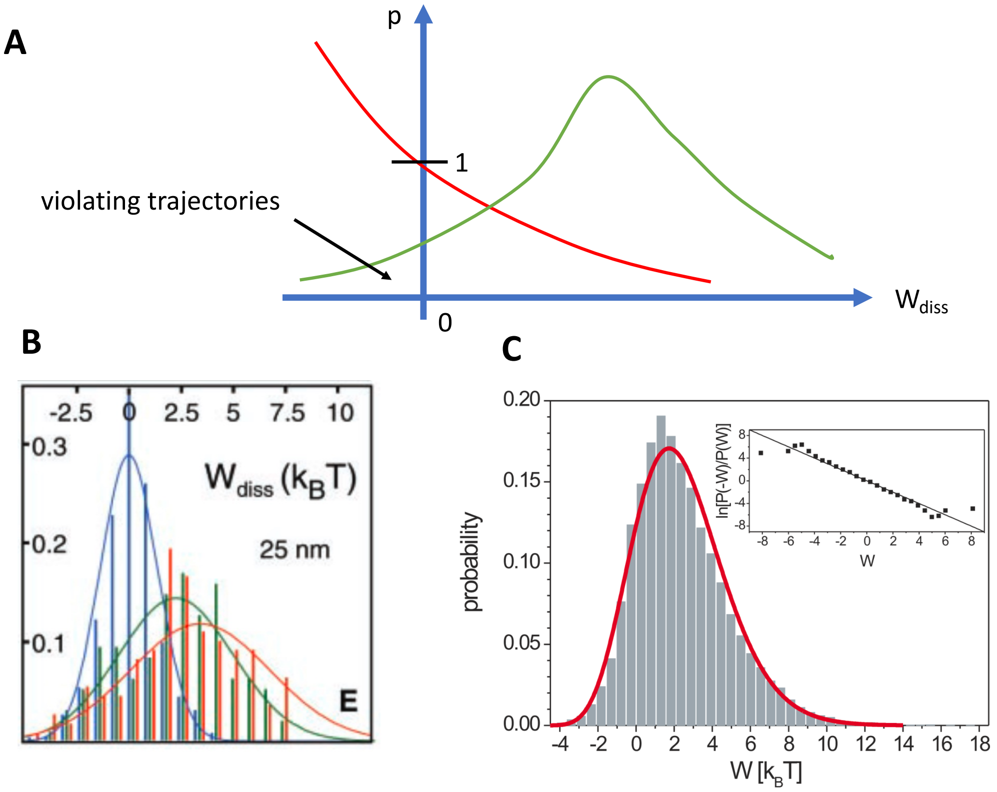

Figure 9.2: (A) The Jarzynski equation predicts that violating trajectories must exist in the distribution of dissipated work $W_{diss} = T\Delta S_{tot}$. While the average has to be positive, in agreement with classical thermodynamics, the overall averaging of e^{-x} (red line) with the distribution (green line) can only reach 1 if also negative values existed. (B) The existence of violating trajectories has been demonstrated first for biomolecules that have been pulled at different speeds (Liphardt, Jan, et al. "Equilibrium information from nonequilibrium measurements in an experimental test of Jarzynski's equality." Science 296.5574 (2002): 1832-1835). (C) It also has been demonstrated for a colloidal particle pushed against a wall by a laser (Blickle, Valentin, et al. "Thermodynamics of a colloidal particle in a time-dependent nonharmonic potential." Physical review letters 96.7 (2006): 070603).

Finally we discuss how classical thermodynamics emerges again from this system. We split the averaging into regions of positive and negative values for dissipated heat:

$$1 = \left\langle e^{-W_{diss}/k_BT} \right\rangle = p_+ \left\langle e^{-W_{diss}/k_BT} \right\rangle + p_- \left\langle e^{-W_{diss}/k_BT} \right\rangle$$

with $p_+ + p_- = 1$. We note that energy is an extensive quantity, thus $W_{diss} \sim N$. For the negative part not to diverge, we therefore must have $p_- \sim e^{-N}$. Therefore $p_+ \sim (1 - e^{-N}) \to 1$ and violating trajectories are exponentially suppressed with increasing system size. To measure these violations, one therefore needs small systems and small energies (of the order of $k_B T$).

10 Appendix: some useful relations between partial derivatives

We consider a function $f(x,y,z)$ of three variables, which is a typical situation in statistical physics and thermodynamics. We write its total differential and then keep it constant:

$$\mathrm{d}f = \left.\frac{\partial f}{\partial x}\right|_{y,z} \mathrm{d}x + \left.\frac{\partial f}{\partial y}\right|_{x,z} \mathrm{d}y + \left.\frac{\partial f}{\partial z}\right|_{x,y} \mathrm{d}z \overset{!}{=} 0$$

Now in addition we keep z constant and divide by dx to get

$$0 = \left.\frac{\partial f}{\partial x}\right|_{y,z} + \left.\frac{\partial f}{\partial y}\right|_{x,z} \left.\frac{\partial y}{\partial x}\right|_{f,z}$$

We rearrange to get

$$\boxed{\left.\frac{\partial y}{\partial x}\right|_{f,z} = -\frac{\left.\frac{\partial f}{\partial x}\right|_{y,z}}{\left.\frac{\partial f}{\partial y}\right|_{x,z}}}$$

We repeat the same procedure, but now we divide not by dx, but by dy to get

$$0 = \left.\frac{\partial f}{\partial x}\right|_{y,z} \left.\frac{\partial x}{\partial y}\right|_{f,z} + \left.\frac{\partial f}{\partial y}\right|_{x,z}$$

thus

$$\boxed{\left.\frac{\partial x}{\partial y}\right|_{f,z} = -\frac{\left.\frac{\partial f}{\partial y}\right|_{x,z}}{\left.\frac{\partial f}{\partial x}\right|_{y,z}}}$$

Comparing the two boxed results, we conclude

$$\boxed{\left.\frac{\partial x}{\partial y}\right|_{f,z} = \left(\left.\frac{\partial y}{\partial x}\right|_{f,z}\right)^{-1}}$$

Repeating the first procedure from above, we find two more analogous relations to the first boxed relation:

$$\left.\frac{\partial z}{\partial x}\right|_{f,y} = -\frac{\left.\frac{\partial f}{\partial x}\right|_{y,z}}{\left.\frac{\partial f}{\partial z}\right|_{x,y}}$$

$$\left.\frac{\partial z}{\partial y}\right|_{f,x} = -\frac{\left.\frac{\partial f}{\partial y}\right|_{x,z}}{\left.\frac{\partial f}{\partial z}\right|_{x,y}}$$

Combining these three equations and using the relation with the inverse gives

$$\boxed{\left.\frac{\partial x}{\partial y}\right|_{f,z} \left.\frac{\partial y}{\partial z}\right|_{f,x} \left.\frac{\partial z}{\partial x}\right|_{f,y} = -1}$$

All the relations given here find frequent applications in statistical physics and thermodynamics.