

Theoretical Statistical Physics

Prof. Dr. Christof Wetterich
Institute for Theoretical Physics
Heidelberg University

Last update: March 25, 2014

Script prepared by Robert Lilow,
using earlier student's script by Birgir Ásgeirsson, Denitsa Dragneva and Emil Pavlov.
For corrections and improvement suggestions please send a mail to
lilow@stud.uni-heidelberg.de.

Contents

I. Foundations of Statistical Physics	9
1. Introduction	11
1.1. From Microphysics to Macrophysics	11
1.2. Macroscopic laws and simplicity	12
1.3. Fundamental laws are of probabilistic nature	12
2. Basic concepts of statistical physics	15
2.1. Probability distribution and expectation values	15
2.1.1. States (Microstates)	15
2.1.2. Probability distribution	15
2.1.3. Observables	15
2.1.4. Expectation value	16
2.1.5. Dispersion	16
2.1.6. Expectation values and probability distribution	16
2.2. Simple examples	17
2.2.1. Two-state system	17
2.2.2. Four-state system	18
2.3. Reduced system or subsystem	18
2.4. Probabilistic observables	20
2.5. Quantum statistics and density matrix	23
2.5.1. Expectation value	23
2.5.2. Dispersion	24
2.5.3. Diagonal operators	24
2.5.4. Properties of the density matrix	25
2.5.5. Diagonal elements of ρ define the probability distribution	25
2.5.6. Positivity constraints for the elements of ρ	26
2.5.7. Non-diagonal operators	26
2.5.8. Pure quantum states	29
2.5.9. Change of the basis by unitary transformation	29
2.5.10. Pure and mixed states	30
2.6. Microstates	31
2.7. Partition function	33
2.8. Continuous distribution and continuous variables	33

2.9.	Probability density of energy	35
2.9.1.	Equal probability for all values τ	35
2.9.2.	Boltzmann distribution for microstates τ	36
2.9.3.	Number of one particle states in a momentum interval	36
2.10.	Correlation functions	38
2.10.1.	Conditional probability $p(\lambda_B \lambda_A)$	38
2.10.2.	Product observable $B \circ A$	38
2.10.3.	Correlation function	39
2.10.4.	Independent observables	39
2.10.5.	Classical correlation function	40
2.10.6.	Quantum correlations	41
3.	Statistical description of many-particles systems	43
3.1.	Chain with \mathcal{N} lattice sites	43
3.1.1.	Probability distribution	43
3.1.2.	Macroscopic probability	43
3.1.3.	Expectation value	45
3.1.4.	Dispersion	45
3.2.	Uncorrelated probability distribution	46
3.2.1.	\mathcal{N} spins on a chain	46
3.2.2.	Correlation functions	47
3.2.3.	Independent measurement	49
3.3.	Statistical Ensembles	50
3.4.	Gaussian distribution	51
3.5.	Thermodynamic limit	53
4.	Equilibrium Ensembles	55
4.1.	Microcanonical ensemble	55
4.1.1.	Fundamental statistical postulate	55
4.1.2.	Microcanonical partition function	56
4.1.3.	Entropy S	56
4.1.4.	Partition function for ideal classical gas	57
4.1.5.	Entropy of an ideal classical gas	61
4.1.6.	Thermodynamic limit	61
4.1.7.	The limit $E \rightarrow E_0$	63
4.1.8.	Derivatives of the entropy	63
4.2.	Canonical ensemble	65
4.2.1.	System in a heat bath	65
4.2.2.	Canonical partition function and Boltzmann factor	66
4.2.3.	Relation between temperature and energy	68
4.2.4.	Energy distribution	68
4.2.5.	Equivalence of microcanonical and canonical ensemble	70
4.2.6.	Temperature	70
4.2.7.	Temperature and energy distribution in subsystems	71

4.2.8.	Thermal equilibrium	71
4.3.	Grand canonical ensemble	72
4.4.	Relations for entropy	73
4.4.1.	Entropy in canonical ensemble	73
4.4.2.	Entropy in grand canonical ensemble	74
4.4.3.	General expression for entropy	74
4.4.4.	Entropy and density matrix	75
5.	Thermodynamics	77
5.1.	Thermodynamic potentials	77
5.1.1.	State variables, equations of state	77
5.1.2.	Thermodynamic potentials and partition functions	79
5.1.3.	Differential relations	80
5.1.4.	Legendre transform	81
5.1.5.	Properties of entropy	83
5.1.6.	Zero of absolute temperature	84
5.1.7.	3rd law of thermodynamics	85
5.2.	Reversible and irreversible processes	85
5.2.1.	Subsystems and constraints	85
5.2.2.	Removing the constraint	86
5.2.3.	Increase of entropy by removing the constraints	87
5.2.4.	2nd law of thermodynamics	88
5.2.5.	Reversible processes for isolated systems	89
5.2.6.	Reversible processes for open systems	89
5.2.7.	Thermodynamics and non-equilibrium states	90
II.	Statistical Systems	91
6.	Ideal quantum gas	93
6.1.	Occupation numbers and functional integral	93
6.2.	Partition function for the grand canonical ensemble	94
6.3.	Mean occupation numbers	95
6.3.1.	Bosons	95
6.3.2.	Fermions	96
6.3.3.	Mean energy	97
6.3.4.	Chemical potential	97
6.3.5.	Boltzmann statistics	98
6.4.	Photon gas	99
6.4.1.	Black body radiation	99
6.4.2.	Spectrum of black body radiation	101
6.4.3.	Energy density	103
6.4.4.	Canonical partition function	104
6.4.5.	Free energy and equation of state	105

6.5.	Degenerate Fermi gas	105
6.5.1.	Quasiparticles	105
6.5.2.	The limit $T \rightarrow 0$	106
6.5.3.	Computation of the Fermi energy	106
6.5.4.	Thermodynamics	107
6.5.5.	Ground state energy	107
6.5.6.	One-particle density of states	108
6.5.7.	$\Omega_1(\varepsilon)$ for free non-relativistic fermions	109
6.5.8.	$E(T, N)$ for arbitrary temperature	110
6.5.9.	Smoothing of Fermi surface	110
6.5.10.	Computation of $\mu(T)$ from the normalisation condition	111
6.5.11.	Occupation numbers for particles and holes	112
6.5.12.	Solve for $\mu(T)$	113
6.6.	Non-relativistic bosons	114
6.6.1.	Chemical potential	114
6.6.2.	Number of bosons in the ground state	115
6.6.3.	Qualitative behaviour of $\bar{n}(E_\nu)$ for $T \rightarrow 0$	115
6.6.4.	Gibbs potential	115
6.7.	Classical ideal gas	116
6.7.1.	Classical approximation	116
6.7.2.	Gibbs potential	116
6.7.3.	Thermal de Broglie wavelength	117
6.7.4.	Thermodynamics	118
6.7.5.	Validity of the approximation for a diluted gas (classical gas)	118
6.8.	Bose-Einstein condensation	119
6.8.1.	$n(\mu)$	119
6.8.2.	Critical temperature	119
6.8.3.	Ground state of bosonic gas	120
6.8.4.	Bose-Einstein condensation	121
7.	Gas of interacting particles	123
7.1.	Interactions and complexity	123
7.2.	Real gases	123
7.3.	Mean field approximation	124
7.3.1.	Classical statistics in phase space	124
7.3.2.	Average potential \cong mean field	125
7.3.3.	Expansion of canonical partition function	125
7.3.4.	Computation of $Z_{\text{ww}}^{(1)}$	126
7.3.5.	Excluded volume	127
7.3.6.	Mean potential	127
7.3.7.	Free energy	128
7.4.	Equation of state	128
7.5.	Specific heat of real gases	129

7.6.	Equipartition theorem	129
8.	Magnetic systems	133
8.1.	Magnetization	133
8.1.1.	Degrees of freedom	133
8.1.2.	Hamiltonian	134
8.1.3.	Magnetization	134
8.1.4.	Thermodynamic potential	135
8.1.5.	Effective potential	136
8.1.6.	Minimal principle	137
8.2.	Spontaneous symmetry breaking	137
8.2.1.	Symmetries	137
8.2.2.	Vanishing magnetic field	138
8.2.3.	Domains	138
8.2.4.	General form of U	139
8.3.	Landau theory and phase transition	139
8.3.1.	Quartic expansion of U	139
8.3.2.	Phase transition for $j = 0$	140
8.3.3.	Discontinuity in magnetization for non-zero B	142
8.4.	Magnetic susceptibility	143
9.	Phase transitions	147
9.1.	Order of phase transition	147
9.1.1.	First order phase transition	147
9.1.2.	Second order phase transition	147
9.1.3.	Probabilistic interpretation	149
9.2.	Liquid-gas phase transition	150
9.2.1.	Effective potential	150
9.2.2.	First order transition	151
9.2.3.	Compressibility	152
9.2.4.	Chemical potential	152
9.3.	Phase diagram	153
9.3.1.	Vapor pressure curve	153
9.3.2.	$p - T$ diagram	153
9.3.3.	Clausius-Clapeyron relation	154
9.3.4.	Latent heat	155
9.3.5.	Superheating, supercooling	155
10.	Critical Phenomena	157
10.1.	Critical exponents	157
10.1.1.	Specific heat	157
10.1.2.	Order parameter as function of temperature	157
10.1.3.	Susceptibility	157
10.1.4.	Order parameter at critical temperature as function of the “source”	158

10.1.5.	Mean field exponents	158
10.1.6.	Universal critical exponents	160
10.2.	Correlation functions	161
10.2.1.	Correlation length	161
10.2.2.	Correlation function and generating functional	162
10.2.3.	Correlation function and susceptibility	163
10.2.4.	Critical exponent of correlation length and anomalous dimension	164
10.2.5.	Critical opalescence	164
10.3.	Universality	164
10.3.1.	Macroscopic laws for second order phase transition	164
10.3.2.	Ising model in one dimension	165
10.3.3.	Ising model in $d > 1$ dimensions	167
10.3.4.	Renormalization flow	170
10.3.5.	Fixed points	173

Part I.

Foundations of Statistical Physics

1. Introduction

1.1. From Microphysics to Macrophysics

Fundamental laws for basic interactions do not explain everything, e. g. the description of macroscopic objects, consisting of many particles (solids, fluids, gases).

A typical number of particles is Avogadro's number: $N_A \simeq 6 \cdot 10^{23}$. It denotes the number of molecules per mol, e. g. the number of water molecules in 18 g water.

Typical questions are:

- connections between pressure (p), volume (V), and temperature (T) of gases (equation of state)
- specific heat capacity of a material
- conductivity of a solid, colour of a fluid
- connection between magnetization and magnetic field
- melting of ice
- How does a refrigerator work?
- How does a supernova explode?
- energy budget of the Earth's atmosphere
- diffusion through a membrane in a plant
- particle beam in a high energy collider

Typical area of application:

- systems with many degrees of freedom \leftrightarrow macroscopic objects

For example a water droplet:

- Consider an isolated system – a droplet of water with volume V and number of particles N . A classical mechanical description needs $6N$ initial conditions $\leftrightarrow 6N$ degrees of freedom.
- The distance between two trajectories, which initially are close to each other, grows exponentially with time. Initial state can never be prepared (known) with arbitrary precision: Limit for predictability (weather forecast)

1. Introduction

- Try deterministic classical mechanics as a description of a water droplet for one second. One needs more information than can be saved in the entire universe, even if one stores one bit in every volume with Planck size $(10^{-34} \text{ m})^3$

The goal of the statistical physics is to develop laws for the macrostates by deriving them from the microphysics with the help of statistical methods.

1.2. Macroscopic laws and simplicity

- qualitatively new laws
- new concepts, e. g. entropy
- statistic laws are very exact, relative fluctuations $\sim \frac{1}{\sqrt{N}}$
- predictive power only for systems with many degrees of freedom; uncertainty about the behaviour of single atoms (quantum physics), but precise predictions about the whole system
- thermodynamics
- material properties
- non-equilibrium processes

Scales for macroscopic systems:

- $L \gg$ mean distance between particles in fluids and solids $\simeq 10^{-7} \text{ cm}$, $N \gg 1$
- $t \gg$ timescale of the microscopic motions:
 - motion of an atom: 10^{-13} s
 - motion of an electron (in solids): 10^{-15} s

1.3. Fundamental laws are of probabilistic nature

Fundamental laws are of probabilistic nature

The role of the statistical physics in the building of physics is huge: it is the foundation of the modern physics! There are two two major strategies in physics, as illustrated in Figure 1.1:

1. Fundamental theory: systems with many (infinite) degrees of freedom; QFT, String theory (for scattering of two particles) are formulated as statistical theories.
2. The role of statistics to understand complex systems is obvious.

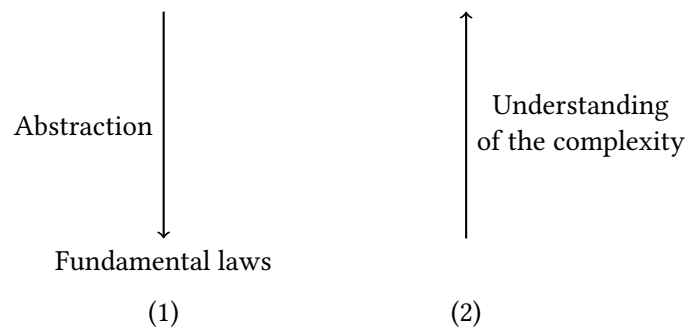


Figure 1.1.: The two major strategies in physics

To calculate the predictions, which follow from the fundamental theory, one uses methods of statistical physics just like in this lecture.

Basic laws of physics are statistical in nature!

They make statements about:

- the probability that, when measuring an observable A once, one gets the value A_i
- the conditional probability for two measurements: if one measures a value A_i for an observable A , what is the probability that one measures the value B_j for a second observable B ?

Remarks

1. From microphysics to microphysics and from Planck's scale (10^{19} GeV) to the energy scale of LHC (1000 GeV) – 16 orders of magnitude.
2. Universality: one does not consider all the details of the microscopic laws \rightarrow predictive power of the QFT!
3. Conceptual revolution due to the statistical physics.
Modern physics: Plank, Boltzmann, Einstein.
4. ~~Determinism~~ \rightarrow probabilistic theory \rightarrow predictive power

2. Basic concepts of statistical physics

2.1. Probability distribution and expectation values

2.1.1. States (Microstates)

We will denote states with τ and the number of states with Ω . In the beginning, we will examine only a finite number of states $\Omega < \infty$.

2.1.2. Probability distribution

To every state τ we associate a probability p_τ , the “probability that this state is realized”:

$$\tau \rightarrow p_\tau . \quad (2.1)$$

The probabilities p_τ have to obey two conditions

$$p_\tau \geq 0 \quad \text{positivity,} \quad (2.2)$$

$$\sum_{\tau} p_\tau = 1 \quad \text{normalization.} \quad (2.3)$$

The range of p_τ is therefore

$$0 \leq p_\tau \leq 1 . \quad (2.4)$$

The probability distribution is $\{p_\tau\} = \{p_1, p_2, p_3, \dots\}$, which is an ordered sequence of Ω numbers.

2.1.3. Observables

Capital roman letters are used for observables, e. g. A . A classical observable A has a fixed value A_τ in every state τ :

$$\tau \rightarrow A_\tau \in \text{Spec } A, \quad (2.5)$$

where the spectrum $\text{Spec } A$ of an observable is the set of all possible measurement values $\{\lambda_A\}$, $\lambda_A \in \mathbb{R}$.

The probability to measure a certain value λ_A is given by:

$$p_{\lambda_A} = \sum_{\tau} \Big|_{A_\tau = \lambda_A} p_\tau , \quad (2.6)$$

where one sums over all states for which $A_\tau = \lambda_A$. A linear combination of two classical observables A and B is again a classical observable:

$$C = f_A A + f_B B \quad \text{with } f_{A,B} \in \mathbb{R}, \quad (2.7)$$

$$C_\tau = f_A A_\tau + f_B B_\tau . \quad (2.8)$$

τ	1	2	3	4
$A^{(2)}$	0	1	0	0
$A^{(3)}$	0	0	1	0

Table 2.1.: A system with $A^{(\sigma)} = \delta_{\sigma\tau}$

And a function of A is an observable, too:

$$\left(f(A)\right)_\tau = f(A_\tau) . \quad (2.9)$$

2.1.4. Expectation value

The expectation value of an observable is defined as

$$\langle A \rangle = \sum_{\tau} p_{\tau} A_{\tau} . \quad (2.10)$$

2.1.5. Dispersion

The dispersion is given by

$$\Delta A^2 = \sum_{\tau} p_{\tau} (A_{\tau} - \langle A \rangle)^2 \quad (2.11)$$

$$= \sum_{\tau} (A_{\tau}^2 - 2 A_{\tau} \langle A \rangle + \langle A \rangle^2) \quad (2.12)$$

$$= \sum_{\tau} p_{\tau} A_{\tau}^2 - \langle A \rangle^2 \quad (2.13)$$

$$= \langle A^2 \rangle - \langle A \rangle^2 . \quad (2.14)$$

2.1.6. Expectation values and probability distribution

The expectation values of a *complete* set of observables determine the probability distribution $\{p_{\tau}\}$ uniquely. Consider Ω observables $A^{(\sigma)}$ with $A_{\tau}^{(\sigma)} = \delta_{\sigma\tau}$, as illustrated in Table 2.1. Then, it is

$$\langle A^{(\sigma)} \rangle = \sum_{\tau} p_{\tau} A_{\tau}^{(\sigma)} = p_{\sigma} . \quad (2.15)$$

The observables $A^{(\sigma)}$ are called *projectors*. They can be considered as a basis for observables. All classical observables can be expressed as linear combinations of projector observables. An association of $A^{(\sigma)}$ with possible events (i. e. that state σ is realized) leads to the Kolmogorov axioms for probabilities.

2.2. Simple examples

2.2.1. Two-state system

Typical two-state systems would be a coin, a spin, occupied / unoccupied states or computer bits. States, possibilities and observables in these are usually denoted by:

$$\tau: \uparrow, \downarrow \text{ or } \langle 1 \rangle, \langle 2 \rangle \text{ or } 1, 0$$

$$p_\tau: p_\uparrow, p_\downarrow \text{ or } p_1, p_2, \\ p_1 + p_2 = 1$$

A : number up – number down, spin, particle number

As an example, let us define the observable A , for which holds that $A_1 = 1$ for state \uparrow and $A_2 = -1$ for state \downarrow . Then, we can express the spin in z -direction in terms of A :

$$S_z = \frac{\hbar}{2} A. \quad (2.16)$$

The number of particles is given by $N = \frac{1}{2}(A + 1)$, which corresponds to $N_1 = 1$ and $N_2 = 0$. The expectation values of A and N are

$$\langle A \rangle = p_1 A_1 + p_2 A_2 \quad (2.17)$$

$$= p_1 - (1 - p_1) \quad (2.18)$$

$$= 2p_1 - 1, \quad (2.19)$$

$$\langle N \rangle = \langle \frac{1}{2}(A + 1) \rangle \quad (2.20)$$

$$= p_1 N_1 + p_2 N_2 \quad (2.21)$$

$$= p_1. \quad (2.22)$$

Their dispersions are easily calculated. For A we first need

$$\langle A^2 \rangle = p_1 A_1^2 + p_2 A_2^2 \quad (2.23)$$

$$= p_1 + (1 - p_1) \quad (2.24)$$

$$= 1, \quad (2.25)$$

where we used $A_1^2 = A_2^2 = 1$. Then,

$$\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2 \quad (2.26)$$

$$= 1 - (2p_1 - 1)^2 \quad (2.27)$$

$$= 4p_1(1 - p_1), \quad (2.28)$$

and therefore

$$\Delta A^2 = 0 \quad \text{for } p_1 = 0 \quad \text{or } p_1 = 1. \quad (2.29)$$

And for N accordingly:

$$\langle N^2 \rangle = 1 \cdot p_1 + 0 \cdot p_2 = p_1, \quad (2.30)$$

$$\Delta N^2 = p_1(1 - p_1). \quad (2.31)$$

$\uparrow\uparrow$	$\uparrow\downarrow$	$\downarrow\uparrow$	$\downarrow\downarrow$
11	10	01	00
$ 1\rangle$	$ 2\rangle$	$ 3\rangle$	$ 4\rangle$

Table 2.2.: States τ in a four-state system

	$ 1\rangle$	$ 2\rangle$	$ 3\rangle$	$ 4\rangle$	
A_τ	2	0	0	-2	total spin
B_τ	1	1	-1	-1	spin 1
C_τ	1	-1	1	-1	spin 2

Table 2.3.: Observables in a four-state system

2.2.2. Four-state system

An example of a four-state system are two particles, each with a spin in two possible directions. Different possibilities to denote these states are listed in Table 2.2. We can now introduce a variable called the total spin A . Other possible observables are the spin of the first particle B and the spin of the second particle C . Values of these observables are given in Table 2.3.

For the particular case of equal probabilities for all states, $p_1 = p_2 = p_3 = p_4 = \frac{1}{4}$, we can compute the expectation values:

$$\langle A \rangle = \langle B \rangle = \langle C \rangle = 0, \quad (2.32)$$

and the dispersions:

$$\langle B^2 \rangle = \langle C^2 \rangle = 1, \quad \langle A^2 \rangle = 2, \quad (2.33)$$

$$\Delta B^2 = \Delta C^2 = 1, \quad \Delta A^2 = 2. \quad (2.34)$$

2.3. Reduced system or subsystem

For most parts of this lecture the basic concepts of section 2.1, classical statistics, are sufficient. So, we want to describe quantum systems, but often use classical statistics. This needs justification! The following parts of chapter 2 motivate why this is possible for many purposes.

The basic observation is that classical statistics and quantum statistics have much in common! Many “quantum features” appear when one considers subsystems of classical statistical systems. When only part of the information in $\{p_\tau\}$ is needed or available, we can “integrate out” the redundant information, which results in a reduced probability distribution.

Examples

1. Consider the previous example of a four-state system again. Suppose we only have information about spin 1. Then, we can combine the probabilities for states with the

same B_τ and arbitrary C_τ . As a consequence, the system turns into an effective two-state system.

Here, $\bar{\tau} = |\bar{1}\rangle$ groups together these “microscopic states” $\tau = |1\rangle$, $\tau = |2\rangle$, in which B_τ has the same values $B_\tau = 1$. Similarly, $\bar{\tau} = |\bar{2}\rangle$ groups together $\tau = |3\rangle$ and $\tau = |4\rangle$ with $B_\tau = -1$. The states of this reduced ensemble are illustrated in Table 2.4. The probabilities of the subsystem states are

$$\bar{p}_1 = p_1 + p_2, \quad (2.35)$$

$$\bar{p}_2 = p_3 + p_4. \quad (2.36)$$

Using $\bar{B}_1 = 1$ and $\bar{B}_2 = -1$, which are the classical observables in the subsystem, the expectation value of B can be computed from the subsystem:

$$\langle B \rangle = \bar{p}_1 \bar{B}_1 + \bar{p}_2 \bar{B}_2 \quad (2.37)$$

$$= \bar{p}_1 - \bar{p}_2 \quad (2.38)$$

$$= p_1 + p_2 - p_3 - p_4. \quad (2.39)$$

As it should be, this is the same as computed from the underlying microscopic system:

$$\langle B \rangle = \sum_{\tau=1}^4 B_\tau p_\tau. \quad (2.40)$$

In this case, B is a “system observable”, which means that the information in the system is enough to compute the distribution of the measurement values. In contrast, A and C are “environment observables”.

2. Consider the case when A is the system observable of the subsystem. Thus, the system becomes an effective three-state system with

$$\bar{A}_1 = 2, \quad \bar{A}_2 = 0, \quad \bar{A}_3 = -2. \quad (2.41)$$

The probabilities for these states are

$$\bar{p}_1 = p_1, \quad \bar{p}_2 = p_2 + p_3, \quad \bar{p}_3 = p_4. \quad (2.42)$$

State $|\bar{2}\rangle$ of the subsystem groups together the states $|2\rangle$ and $|3\rangle$ of the microsystem that both have $A_\tau = 0$. The states $|\bar{1}\rangle$ and $|\bar{3}\rangle$ of the subsystem are identical to the microstates $|1\rangle$ and $|4\rangle$ with $A_\tau = 2$ and $A_\tau = -2$. We will continue examining this example in section 2.4.

\uparrow	\downarrow
$ \bar{1}\rangle$	$ \bar{2}\rangle$

Table 2.4.: States of the reduced ensemble $\bar{\tau}$ in example 1

2.4. Probabilistic observables

In every state τ a *classical observable* has a fixed value A_τ . Is this the most general case? No, more general is the concept of *probabilistic observables*. In every state τ a probabilistic observable takes its possible measurement values λ_B with a certain probability $w_\tau(\lambda_B)$.

Examples

1. Water on earth:

Divide all probes of H_2O -molecules into three states (using some criterion):

$\tau = 1$: solid (ice)

$\tau = 2$: liquid (water)

$\tau = 3$: gas (vapor)

So, p_τ are the probabilities to find water in the form of solid, liquid or gas. One would like to determine the average density of water on earth, and the dispersion. Therefore, one needs the probability distribution of density in each of the three states $w_1(n)$, $w_2(n)$, $w_3(n)$. (the distributions depend on the state, but in each state the density has no fixed value). Then, the average density is given by

$$\langle n \rangle = p_1 \int dn n w_1(n) + p_2 \int dn n w_2(n) + p_3 \int dn n w_3(n) \quad (2.43)$$

$$= p_1 \bar{n}_1 + p_2 \bar{n}_2 + p_3 \bar{n}_3, \quad (2.44)$$

where

$$\bar{n}_\tau = \int dn n w_\tau(n) \quad (2.45)$$

is the mean value in state τ . We use normalized probability distributions: $\int dn w_\tau(n) = 1$.

The Motivation of these concepts is, to characterize “microstates” by density *and* “state” of water. So, we have microstates (n, τ) with probability distribution $P(n, \tau)$, obeying

$$\sum_\tau \int dn P(n, \tau) = 1. \quad (2.46)$$

In this setting, n is a classical observable with fixed value n in each state (n, τ) . It is

$$\langle n \rangle = \int dn P(n, \tau) n, \quad (2.47)$$

$$\langle n^2 \rangle = \int dn P(n, \tau) n^2. \quad (2.48)$$

Define probabilities for states τ of reduced system:

$$p_\tau = \int dn P(n, \tau), \quad (2.49)$$

and probabilities to find n in a state τ :

$$w_\tau(n) = \frac{P(n, \tau)}{p_\tau}. \quad (2.50)$$

Their normalization is given by

$$\int dn w_\tau(n) = 1. \quad (2.51)$$

Then,

$$P(n, \tau) = p_\tau w_\tau(n), \quad (2.52)$$

and we can rewrite (2.47) and (2.48) as

$$\langle n \rangle = \sum_\tau \int dn w_\tau(n) n = \sum_\tau p_\tau \bar{n}_\tau, \quad (2.53)$$

$$\langle n^2 \rangle = \sum_\tau p_\tau \int dn w_\tau(n) n^2 = \sum_\tau p_\tau (\bar{n}^2)_\tau. \quad (2.54)$$

One finds $(\bar{n}^2)_\tau \neq (\bar{n}_\tau)^2$ due to fluctuations of n in state τ . Therefore,

$$\langle n^2 \rangle \neq \sum_\tau (\bar{n}_\tau)^2. \quad (2.55)$$

For $\langle n \rangle$ the specification of \bar{n}_τ for every state is sufficient (at this point no distinction from classical observable), but $\langle n^2 \rangle$ needs *additional information* about distribution of values n in each state τ .

n is a probabilistic observable on the level of states τ .

2. Subsystem of states with total spin:

Probabilistic observables typically occur for subsystems. For example we can identify the two-spin system with the “microstates” (n, τ) of water. Then, the three states τ characterized by total spin A in subsection 2.2.2:

$$\bar{A}_1 = 2, \quad \bar{A}_2 = 0, \quad \bar{A}_3 = -2, \quad (2.56)$$

are analogous to the states ice, water, vapor in the first example.

For this subsystem the first spin B is a probabilistic observable. As $\bar{B}_1 = 1$ and $\bar{B}_3 = -1$, B behaves as a classical observable in the states $\bar{\tau} = 1, 3$. But what about B in state $\bar{\tau} = 2$? In that case, there is no fixed value for B . Instead, we have the probabilities

$$w_2(1) = \frac{p_2}{p_2 + p_3} \quad \text{to find } B = 1, \quad (2.57)$$

$$w_2(-1) = \frac{p_3}{p_2 + p_3} \quad \text{to find } B = -1. \quad (2.58)$$

The mean value of B in state $\bar{\tau} = 2$ is

$$\bar{B}_2 = w_2(1) \cdot 1 + w_2(-1) \cdot (-1) = \frac{p_2 - p_3}{p_2 + p_3}. \quad (2.59)$$

Let us check this:

$$\langle B \rangle = \bar{p}_1 \bar{B}_1 + \bar{p}_2 \bar{B}_2 + \bar{p}_3 \bar{B}_3 \quad (2.60)$$

$$= p_1 + \underbrace{(p_2 + p_3)}_{\bar{p}_2} \cdot \underbrace{\frac{p_2 - p_3}{p_2 + p_3}}_{\bar{B}_2} - p_4 \quad (2.61)$$

$$= p_1 + p_2 - p_3 - p_4 \quad \text{q. e. d.} \quad (2.62)$$

However,

$$(\bar{B}^2)_2 = 1, \quad (2.63)$$

which differs from $(\bar{B}_2)^2$! Thus,

$$\langle B^2 \rangle = 1 \neq \sum_{\bar{\tau}=1}^3 \bar{p}_{\bar{\tau}} \bar{B}_{\bar{\tau}}^2. \quad (2.64)$$

This is a clear sign of a probabilistic observable. And probabilistic observables need additional information about the probability distribution of measurement values in a given state, which we have sketched in Figure 2.1. So, we conclude:

$$B \text{ environment observable} \xrightarrow{\text{additional information}} B \text{ system observable.} \quad (2.65)$$

We have listed all the observables that we just have computed in Table 2.5.

3. Quantum mechanics:

Consider two-state quantum mechanics:

$$\tau = 1 : \quad \tilde{S}_z = 1, \quad (2.66)$$

$$\tau = 2 : \quad \tilde{S}_z = -1, \quad (2.67)$$

where $S_k = \frac{\hbar}{2} \tilde{S}_k$.

Figure 2.1.: Probability distribution of the probabilistic variable B in given state τ

	$ \bar{1}\rangle$	$ \bar{2}\rangle$	$ \bar{3}\rangle$
\bar{p}	p_1	$p_2 + p_3$	p_4
A	2	0	-2
N	2	1	0
\bar{B}	1	$\frac{p_2 - p_3}{p_2 + p_3}$	-1
\bar{B}^2	1	1	1

Table 2.5.: Observables in the reduced system in example 2

The additional observable \tilde{S}_x has no fixed value in the states τ , only probabilities to find $\tilde{S}_x = \pm 1$:

$$w_1(1) = a, \quad w_1(-1) = 1 - a, \quad (2.68)$$

$$w_2(1) = b, \quad w_2(-1) = 1 - b, \quad (2.69)$$

$$(\bar{S}_x)_1 = 2a - 1, \quad (\bar{S}_x)_2 = 2b - 1. \quad (2.70)$$

Its expectation value is given by

$$\langle \tilde{S}_x \rangle = p_1 \cdot (2a - 1) + p_2 \cdot (2b - 1) \quad (2.71)$$

$$= 2p_1a + 2p_2b - 1. \quad (2.72)$$

So,

$$-1 \leq \langle \tilde{S}_x \rangle \leq 1, \quad (2.73)$$

where $\langle \tilde{S}_x \rangle = 1$ for $a = b = 1$ and $\langle \tilde{S}_x \rangle = -1$ for $a = b = 0$.

2.5. Quantum statistics and density matrix

2.5.1. Expectation value

In quantum statistics the expectation value is computed as

$$\langle A \rangle = \text{tr}(\rho \hat{A}), \quad (2.74)$$

where \hat{A} is the operator associated with A and ρ is the density matrix.

Let us look at the particular case of a pure state. For pure states the density matrix can be expressed as

$$\rho_{\sigma\tau} = \psi_{\sigma}\psi_{\tau}^*, \quad (2.75)$$

where ψ is a complex vector field. Then, we find

$$\langle\psi|\hat{A}|\psi\rangle = \psi_{\tau}^* A_{\tau\sigma} \psi_{\sigma} \quad (2.76)$$

$$= \psi_{\sigma}\psi_{\tau}^* A_{\tau\sigma} \quad (2.77)$$

$$= \rho_{\sigma\tau} A_{\tau\sigma} \quad (2.78)$$

$$= \text{tr}(\rho\hat{A}). \quad (2.79)$$

2.5.2. Dispersion

Accordingly, the dispersion is given by

$$\langle A^2 \rangle = \text{tr}(\rho\hat{A}^2), \quad (2.80)$$

$$\Delta A^2 = \langle A^2 \rangle - \langle A \rangle^2. \quad (2.81)$$

2.5.3. Diagonal operators

When A is diagonal:

$$\hat{A} = \begin{pmatrix} A_1 & 0 & \cdots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_N \end{pmatrix}, \quad (2.82)$$

$$\hat{A}_{\tau\sigma} = A_n \delta_{\tau\sigma}, \quad (2.83)$$

then the expectation value is just

$$\langle A \rangle = \text{tr}(\rho\hat{A}) \quad (2.84)$$

$$= \sum_{\tau,\sigma} \rho_{\sigma\tau} \hat{A}_{\tau\sigma} \quad (2.85)$$

$$= \sum_{\tau} \rho_{\tau\tau} A_{\tau}. \quad (2.86)$$

So, one only needs the diagonal elements of the density matrix. They can be associated with probabilities $p_{\tau} = \rho_{\tau\tau}$, where $\rho_{\tau\tau} \geq 0$. This way, we recover classical statistics:

$$\langle f(A) \rangle = \sum_{\tau} p_{\tau} f(A_{\tau}). \quad (2.87)$$

Example

- A two-state system:

Let $\hat{A} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}$ and $\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{12}^* & \rho_{22} \end{pmatrix}$, then:

$$\rho \hat{A} = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{12}^* & \rho_{22} \end{pmatrix} \cdot \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix} = \begin{pmatrix} A_1 \rho_{11} & A_2 \rho_{12} \\ A_1 \rho_{12}^* & A_2 \rho_{22} \end{pmatrix}, \quad (2.88)$$

$$\text{tr}(\rho \hat{A}) = A_1 \rho_{11} + A_2 \rho_{22}, \quad (2.89)$$

$$\langle A \rangle = \sum_n \rho_{nn} A_n = \sum_\tau p_\tau A_\tau. \quad (2.90)$$

2.5.4. Properties of the density matrix

The density matrix ρ is a $N \times N$ matrix with the following properties:

1. Hermiticity: $\rho^\dagger = \rho$, so all eigenvalues λ_n are real
2. Positivity: $\forall n : \lambda_n \geq 0$
3. Normalization: $\text{tr} \rho = 1$, so $\sum_n \lambda_n = 1$ and $\forall n : 0 \leq \lambda_n \leq 1$

2.5.5. Diagonal elements of ρ define the probability distribution

From the properties in subsection 2.5.4 it follows that $\rho_{\tau\tau} \geq 0$ and $\sum_\tau \rho_{\tau\tau} = 1$. So, $\{\rho_{\tau\tau}\}$ has all the properties of a probability distribution.

Example

- Quantum mechanical two-state system:

In this case, the density matrix ρ is a complex 2×2 matrix of the form

$$\rho = \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{12}^* & 1 - \rho_{11} \end{pmatrix}, \quad (2.91)$$

with $\rho_{11} \in \mathbb{R}$, $0 \leq \rho_{11} \leq 1$. Consider the observable

$$\hat{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \quad (2.92)$$

We can find its expectation value:

$$\langle A \rangle = \frac{\hbar}{2} \text{tr} \left[\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{12}^* & 1 - \rho_{11} \end{pmatrix} \right] \quad (2.93)$$

$$= \frac{\hbar}{2} \text{tr} \begin{pmatrix} \rho_{11} & \rho_{12} \\ -\rho_{12}^* & -1 + \rho_{11} \end{pmatrix} \quad (2.94)$$

$$= \frac{\hbar}{2} (2\rho_{11} - 1). \quad (2.95)$$

From the classical statistics we have $\langle A \rangle = p_1 A_1 + p_2 A_2$. If we set $p_1 = \rho_{11}$, $p_2 = 1 - \rho_{11}$, $A_1 = \frac{\hbar}{2}$, $A_2 = -\frac{\hbar}{2}$, then we will get the same result (2.95).

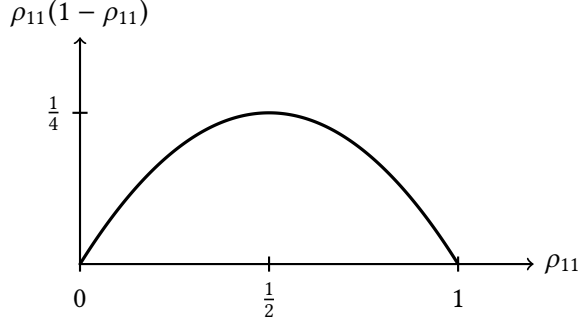


Figure 2.2.: Positivity condition for two state density matrix

2.5.6. Positivity constraints for the elements of ρ

To determine the eigenvalues λ_n of the density matrix (2.91), we have to solve

$$0 = (\rho_{11} - \lambda)(1 - \rho_{11} - \lambda) - f \quad (2.96)$$

$$= \lambda^2 - \lambda + \rho_{11}(1 - \rho_{11}) - f, \quad (2.97)$$

where

$$f = |\rho_{12}|^2 \geq 0. \quad (2.98)$$

We find

$$\lambda_{1,2} = \frac{1}{2} \left(1 \pm \sqrt{1 - 4[\rho_{11}(1 - \rho_{11}) - f]} \right). \quad (2.99)$$

Using the positivity conditions for the eigenvalues, $\lambda_{1,2} \geq 0$, we arrive at

$$\rho_{11}(1 - \rho_{11}) - f \geq 0, \quad (2.100)$$

and therefore,

$$\rho_{11}(1 - \rho_{11}) \geq 0. \quad (2.101)$$

The full positivity condition for the two-state matrix is

$$\boxed{f = |\rho_{12}|^2 \leq \rho_{11}(1 - \rho_{11})}. \quad (2.102)$$

In Figure 2.2 we have plotted $\rho_{11}(1 - \rho_{11})$. From the positivity condition it follows that $|\rho_{12}|^2$ only can take values on or below this curve.

2.5.7. Non-diagonal operators

One may ask why we need the information in ρ_{12} , as for the expectation value it holds $\langle A \rangle = \sum_{\tau} \rho_{\tau\tau} \hat{A}_{\tau}$, provided that \hat{A} is diagonal. The answer is that ρ_{12} carries additional information for the probabilistic observables.

	$ 1\rangle$	$ 2\rangle$
$\langle \tilde{S}_z \rangle$	1	-1
$\langle \tilde{S}_x \rangle$	$\rho_{12} + \rho_{12}^*$	$\rho_{12} + \rho_{12}^*$
$\langle \tilde{S}_x^2 \rangle, \langle \tilde{S}_z^2 \rangle$	1	1

Table 2.6.: Mean values of spin observables in the states $|1\rangle$ and $|2\rangle$ **Example**

- Spin in arbitrary direction $\tilde{S}_i = \frac{2}{\hbar} S_i$:

The spin components \tilde{S}_z and \tilde{S}_x are different types of observables:

\tilde{S}_z : Classical observable

$$(\tilde{S}_z)_1 = 1, \quad (\tilde{S}_z)_2 = -1, \quad (2.103)$$

$$(\tilde{S}_z^2)_1 = 1, \quad (\tilde{S}_z^2)_2 = 1. \quad (2.104)$$

\tilde{S}_x : Probabilistic observable

There is a probability distribution for finding the values $+1$ or -1 in the states $|1\rangle$ and $|2\rangle$:

$$\langle \tilde{S}_x \rangle_1 = \langle \tilde{S}_x \rangle_2 = \rho_{12} + \rho_{12}^*. \quad (2.105)$$

But in both states the square has a fixed value:

$$(\tilde{S}_x^2)_1 = (\tilde{S}_x^2)_2 = 1. \quad (2.106)$$

For a better overview, we have listed these observables in Table 2.6.

The probabilities to find $\tilde{S}_x = \pm 1$ in state $|1\rangle$ are given by

$$w_1(1) = \frac{1}{2}(1 + \rho_{12} + \rho_{12}^*), \quad (2.107)$$

$$w_1(-1) = \frac{1}{2}(1 - \rho_{12} - \rho_{12}^*), \quad (2.108)$$

such that

$$w_1(1) + w_1(-1) = 1. \quad (2.109)$$

With these probabilities we arrive at

$$\langle \tilde{S}_x \rangle_1 = w_1(1) - w_1(-1) = \rho_{12} + \rho_{12}^*. \quad (2.110)$$

Both, $w_1(1)$ and $w_1(-1)$, have to be positive, which can be inferred from the condition

$$|\rho_{12} + \rho_{12}^*| \leq 1, \quad (2.111)$$

or equivalently

$$|\operatorname{Re} \rho_{12}| \leq \frac{1}{2}. \quad (2.112)$$

This follows from (2.102):

$$|\operatorname{Re} \rho_{12}| = \sqrt{|\rho_{12}|^2 - |\operatorname{Im} \rho_{12}|^2} \quad (2.113)$$

$$\leq \sqrt{\rho_{11}(1 - \rho_{11}) - |\operatorname{Im} \rho_{12}|^2} \quad (2.114)$$

$$\leq \sqrt{\rho_{11}(1 - \rho_{11})} \quad (2.115)$$

$$\leq \frac{1}{2}. \quad (2.116)$$

By similar calculations we find the same results for $w_2(1)$ and $w_2(-1)$.

We shall check if classical statistics (with probabilistic observables) and quantum statistics produce discrepancies:

$$\text{Classical: } \langle \tilde{S}_x \rangle = p_1 \langle \tilde{S}_x \rangle_1 + p_2 \langle \tilde{S}_x \rangle_2 \quad (2.117)$$

$$= \rho_{11}(\rho_{12} + \rho_{12}^*) + (1 - \rho_{11})(\rho_{12} + \rho_{12}^*) \quad (2.118)$$

$$= \rho_{12} + \rho_{12}^* \quad (2.119)$$

$$= 2 \operatorname{Re} \rho_{12} \quad (2.120)$$

$$\text{Quantum: } \langle \tilde{S}_x \rangle = \operatorname{tr} \left[\begin{pmatrix} \rho_{11} & \rho_{12} \\ \rho_{12}^* & \rho_{22} \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \right] \quad (2.121)$$

$$= \operatorname{tr} \begin{pmatrix} \rho_{12} & \rho_{11} \\ \rho_{22} & \rho_{12}^* \end{pmatrix} \quad (2.122)$$

$$= \rho_{12} + \rho_{12}^* \quad (2.123)$$

$$= 2 \operatorname{Re} \rho_{12}. \quad (2.124)$$

One can see that (2.120) and (2.124) give exactly the same result.

Actually, this could be generalized for any hermitian quantum operator

$$\hat{A} = \hat{A}^\dagger = \begin{pmatrix} A_{11} & A_{12} \\ A_{12}^* & A_{22} \end{pmatrix}. \quad (2.125)$$

The probabilistic observable has mean values in the states $|1\rangle$ and $|2\rangle$, given by $\langle A \rangle_1$ and $\langle A \rangle_2$:

$$\langle A \rangle_1 = A_{11} + \rho_{12}^* A_{12} + \rho_{12} A_{12}^* \quad (2.126)$$

$$\langle A \rangle_2 = A_{22} + \rho_{12}^* A_{12} + \rho_{12} A_{12}^*. \quad (2.127)$$

It turns out that quantum mechanics can be interpreted as classical statistics, but with additional information for probabilistic observables. As long as we are dealing with diagonal observables, there is no difference between both approaches. In this case, non-diagonal elements of ρ are of no importance. Actually, the last statement is only true for a given point in time or for a static ρ . Non-diagonal elements of ρ are important for the time evolution of the probability distribution:

$$\{p_\tau(t)\} \cong \{\rho_{\tau\tau}(t)\}, \quad (2.128)$$

where the sum over repeated indices is implied.

2.5.8. Pure quantum states

Consider a special form of the density matrix $\rho_{\sigma\tau} = \psi_\sigma \cdot \psi_\tau^*$, where ψ is a complex vector with N components and $\sum_\tau \psi_\tau^* \psi_\tau = 1$. In other words, ψ_τ is a wave function. Then,

$$\langle A \rangle = \text{tr}(\rho \hat{A}) \quad (2.129)$$

$$= \rho_{\sigma\tau} \hat{A}_{\tau\sigma} \quad (2.130)$$

$$= \psi_\tau^* \hat{A}_{\tau\sigma} \psi_\sigma \quad (2.131)$$

$$= \langle \psi | \hat{A} | \psi \rangle, \quad (2.132)$$

where, again, the sum over repeated indices is implied, and

$$p_\tau = \rho_{\tau\tau} = |\psi_\tau|^2 \geq 0. \quad (2.133)$$

As already mentioned, only $\rho_{\tau\tau}$ is needed for diagonal observables. So, for diagonal observables there is no difference between pure and mixed states.

We can state the condition for pure states as

$$\rho^2 = \rho, \quad (2.134)$$

because

$$\rho_{\sigma\tau} \rho_{\tau\alpha} = \psi_\sigma \psi_\tau^* \psi_\tau \psi_\alpha^* \quad (2.135)$$

$$= \psi_\sigma \psi_\alpha^* \quad (2.136)$$

$$= \rho_{\sigma\alpha}. \quad (2.137)$$

2.5.9. Change of the basis by unitary transformation

A unitary transformation is of the kind $\rho \rightarrow \rho' = U \rho U^\dagger$, where $U^\dagger U = 1$. ρ' has still all the properties of the density matrix: $(\rho')^\dagger = \rho'$, $\text{tr} \rho = 1$, positivity. Operators also can be transferred unitarily: $\hat{A} \rightarrow \hat{A}' = U \hat{A} U^\dagger$. Nevertheless, the expectation values are invariant under this transformation:

$$\langle A \rangle = \text{tr}(\rho' \hat{A}') \quad (2.138)$$

$$= \text{tr}(U \rho U^\dagger U \hat{A} U^\dagger) \quad (2.139)$$

$$= \text{tr}(U \rho \hat{A} U^\dagger) \quad (2.140)$$

$$= \text{tr}(U^\dagger U \rho \hat{A}) \quad (2.141)$$

$$= \text{tr}(\rho \hat{A}). \quad (2.142)$$

Let us list several important properties:

- Even if \hat{A} is diagonal, $U \hat{A} U^\dagger$ can be non-diagonal in the general case.
- If ρ can be diagonalized, then $\rho' = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_N)$

l_z	1	2	-1
pure states	$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$
density matrix	$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$
probability	p_1	p_2	p_3

Table 2.7.: z-component of the angular momentum

- Pure states are invariant:

$$\rho^2 = \rho \quad \rightarrow \quad \rho'^2 = \rho'. \quad (2.143)$$

One can see this via

$$\rho'^2 = U \rho U^\dagger U \rho U^\dagger \quad (2.144)$$

$$= U \rho^2 U^\dagger \quad (2.145)$$

$$= U \rho U^\dagger \quad (2.146)$$

$$= \rho'. \quad (2.147)$$

From this it follows that $\lambda_n^2 = \lambda_n$, leaving only the possible values $\lambda_n = 0, 1$. This means that $\rho'_{\sigma\tau} = \delta_{\sigma\beta} \delta_{\tau\beta}$ and therefore,

$$\psi'_\sigma = \delta_{\sigma\beta} e^{i\alpha}, \quad (2.148)$$

$$\psi'^*_\tau = \delta_{\tau\beta} e^{-i\alpha}, \quad (2.149)$$

which is a pure state.

2.5.10. Pure and mixed states

Consider an observable $l = 1$, which could be associated with the angular momentum. Then, l_z could be -1, 0 or 1. The system is described in Table 2.7.

Of course, $p_1 + p_2 + p_3 = 1$ must be satisfied. The density matrix of the system is

$$\rho = \begin{pmatrix} p_1 & 0 & 0 \\ 0 & p_2 & 0 \\ 0 & 0 & p_3 \end{pmatrix}. \quad (2.150)$$

This is a mixed state as long as there are $p_i \neq 0, 1$. One could check and see that in this case

$$\rho^2 = \begin{pmatrix} p_1^2 & 0 & 0 \\ 0 & p_2^2 & 0 \\ 0 & 0 & p_3^2 \end{pmatrix} \neq 0. \quad (2.151)$$

Example

- Let us examine a system characterized by the following density matrix:

$$\rho = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{3} & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}. \quad (2.152)$$

Here, all the states $l_z = -1, 0, 1$ have equal probability, so $\langle l_z \rangle = 0$. If we compare this with a pure state:

$$\rho = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (2.153)$$

leading to the density matrix

$$\rho = \frac{1}{3} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}. \quad (2.154)$$

we can see that both, (2.152) and (2.154), have the same diagonal elements. Consequently, $\langle l_z \rangle = 0$, again. They differ, though, in the mean values of the probabilistic observables, because of the different off-diagonal elements.

2.6. Microstates

Back to the quantum-mechanical point of view: provided that $|\psi_\tau\rangle$ is a complete orthonormal basis, then ψ_τ are microstates.

Examples

1. Hydrogen atom:

Here, $\tau = (\hat{n}, l, m, s)$ is a multiple index. We assume that the microstates are countable, although it is commonly otherwise. Often, they are also limited, e. g. only microstates with $E < \frac{E_0}{5}$. The energy of a state with quantum number \hat{n} is given by

$$E = -\frac{E_0}{\hat{n}^2}. \quad (2.155)$$

Let us consider the following states:

$$\psi_{1,2} : \quad \hat{n} = 1, l = 0, m = 0, s = \pm 1; E = -E_0, \quad (2.156)$$

$$\psi_{3,\dots,10} : \quad \hat{n} = 2, l = 0, 1, m, s; E = -\frac{E_0}{4}, \quad (2.157)$$

where $N = 2 + 8 = 10$. The energy operator is given by $\hat{E} = -E_0 \text{diag} \left(1, 1, \frac{1}{4}, \dots, \frac{1}{4} \right)$. Let the distribution be uniform: $\rho = \text{diag} \left(\frac{1}{10}, \dots, \frac{1}{10} \right)$. Then, the expectation value of the energy is

$$\langle E \rangle = \frac{1}{10} \left(2 + \frac{8}{4} \right) (-E_0) = -0.4 E_0. \quad (2.158)$$

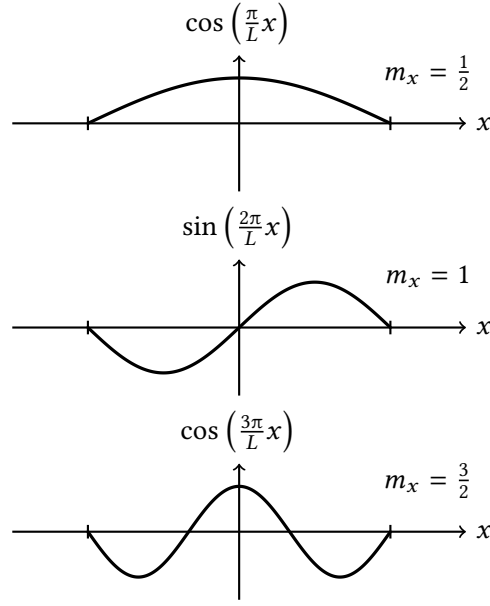


Figure 2.3.: Wave functions of a particle in a box

2. A particle in a box:

The particle is described by the wave equation

$$-\frac{\hbar^2 \Delta}{2M} \psi = E \psi. \quad (2.159)$$

Consider a cubic box with an edge length L . Then, the solution is found to be

$$\psi = e^{\frac{ip_x x}{\hbar}} e^{\frac{ip_y y}{\hbar}} e^{\frac{ip_z z}{\hbar}} \psi_0 \quad (2.160)$$

$$= e^{\frac{i\mathbf{p} \cdot \mathbf{r}}{\hbar}} \psi_0, \quad (2.161)$$

where

$$E = \frac{1}{2M} \mathbf{p}^2. \quad (2.162)$$

The periodical boundary conditions give $\psi(-\frac{L}{2}, y, z) = \psi(\frac{L}{2}, y, z)$ (and the same for y and z). Then $p_{x,y,z}$ have discrete values:

$$\mathbf{p} = \hbar \frac{2\pi}{L} \mathbf{m}, \quad (2.163)$$

with

$$\mathbf{m} = (m_x, m_y, m_z), \quad m_{x,y,z} \in \mathbb{Z}. \quad (2.164)$$

Three examples of these periodical wave functions are illustrated in Figure 2.3.

2.7. Partition function

What is the number of states for a given Energy E of a single particle ($N = 1$) in an one-dimensional ($d = 1$) cubic box with edge length L ? We will denote the number of states in $[E, E + \delta E]$ by $\Omega(E, V)$. For our purposes, assume that $E \gg \frac{\hbar^2 \pi^2}{2ML^2}$. It is

$$\frac{p^2}{2M} = \frac{\pi^2 \hbar^2}{2M} \frac{n^2}{L^2}, \quad (2.165)$$

so,

$$E < \frac{\pi^2 \hbar^2}{2M} \frac{n^2}{L^2} < E + \delta E. \quad (2.166)$$

If $\delta E \ll E$, we can approximate

$$\sqrt{E + \delta E} \approx \sqrt{E} \left(1 + \frac{1}{2} \frac{\delta E}{E} \right) \quad (2.167)$$

$$= \sqrt{E} + \frac{1}{2} \frac{\delta E}{\sqrt{E}}, \quad (2.168)$$

leading to

$$\frac{\sqrt{2ME}}{\pi \hbar} L < n < \frac{\sqrt{2ME}}{\pi \hbar} L + \frac{1}{2\pi \hbar} \left(\frac{2M}{E} \right)^{1/2} L \delta E. \quad (2.169)$$

Therefore, the number of states $\Omega(E, V)$ is given by

$$\boxed{\Omega(E, V) = \frac{V}{2\pi \hbar} \left(\frac{2M}{E} \right)^{1/2} \delta E,} \quad (2.170)$$

which means that $\Omega \sim V!$

2.8. Continuous distribution and continuous variables

We will examine the continuous description of “closely situated” discrete States. We need to define the probability density $p(x)$ with:

$$\boxed{\begin{aligned} \int dx p(x) &= 1 \\ \int dx p(x) A(x) &= \langle A \rangle. \end{aligned}} \quad (2.171)$$

We will use the following definitions:

$I(x)$: interval of states τ which belong to the interval $\left[x - \frac{dx}{2}, x + \frac{dx}{2} \right]$,

$\bar{p}(x)$: mean probability in $I(x)$,

$\Omega(x)$: number of states belonging to $I(x)$,

$A(x)$: $\langle A \rangle_x$, i. e, mean value of A in interval $I(x)$ (probabilistic variable).

We can express the probability of the reduced system with states x as

$$p(x) dx = \sum_{\tau \in I(x)} p_\tau = \bar{p}(x) \Omega(x), \quad (2.172)$$

and $A(x)$ is given by

$$A(x) = \langle A \rangle_x = \frac{\sum_{\tau \in I(x)} p_\tau A_\tau}{\sum_{\tau \in I(x)} p_\tau}. \quad (2.173)$$

Let us check if the continuous case is equivalent to the discrete one:

$$\int dx p(x) A(x) = \sum_x dx p(x) A(x) \quad (2.174)$$

$$= \sum_x \underbrace{\left(\sum_{\tau \in I(x)} p_\tau \right)}_{dx p(x)} \frac{\sum_{\tau \in I(x)} p_\tau A_\tau}{\sum_{\tau \in I(x)} p_\tau} \quad (2.175)$$

$$= \sum_x \sum_{\tau \in I(x)} p_\tau A_\tau \quad (2.176)$$

$$= \sum_\tau p_\tau A_\tau \quad \text{q. e. d.} \quad (2.177)$$

Note that $\langle A^2 \rangle = \int dx p(x) A^2(x) = \sum_\tau p_\tau A_\tau$ holds only, if the dispersion of A vanishes in the continuum limit $dx \rightarrow 0$.

We shall see that $p(x)$ depends on the choice of the variables x . For that purpose, consider the explicit coordinate transformation

$$x = f(x'), \quad (2.178)$$

$$dx = \frac{df}{dx'} dx'. \quad (2.179)$$

The expectation value of A in the new coordinates becomes:

$$\langle A \rangle = \int dx p(x) A(x) = \int dx' \underbrace{\frac{df}{dx'} p[f(x')]}_{p'(x')} \underbrace{A[f(x')]}_{A'(x')}. \quad (2.180)$$

This can be generalized for many variables:

$$\langle A \rangle = \int dx_1 \cdots dx_n p(x_1, \dots, x_n) A(x_1, \dots, x_n) \quad (2.181)$$

where $p(x)$ is transformed by the Jacobi determinant:

$$p'(x') = \left| \frac{df}{dx'} \right| \cdot p[f(x')], \quad (2.182)$$

$$x_i = f_i(x'_j), \quad (2.183)$$

$$\left| \frac{df}{dx'} \right| = \det \begin{pmatrix} \frac{df_1}{dx'_1} & \cdots & \frac{df_n}{dx'_1} \\ \vdots & \ddots & \vdots \\ \frac{df_1}{dx'_n} & \cdots & \frac{df_n}{dx'_n} \end{pmatrix}. \quad (2.184)$$

2.9. Probability density of energy

The probability density of energy can be inferred from

$$p(E) dE = \bar{p}(E) \frac{\Omega(E)}{\delta E} dE, \quad (2.185)$$

via

$$\frac{\Omega(E)}{\delta E} \xrightarrow{\text{change of symbols}} \frac{\delta \Omega(E)}{\delta E} \xrightarrow{\text{continuum limit}} \frac{\partial \Omega(E)}{\partial E}, \quad (2.186)$$

leading us to

$$\boxed{p(E) dE = \bar{p}(E) \frac{\partial \Omega}{\partial E} dE.} \quad (2.187)$$

An example would be one particle in a linear box ($d = 1, N = 1$).

2.9.1. Equal probability for all values τ

We know that

$$\delta \Omega(E) \equiv \Omega(E) = \frac{V}{2\pi\hbar} \left(\frac{2M}{E} \right)^{1/2} \delta E. \quad (2.188)$$

So, we can conclude that

$$\frac{\delta \Omega}{\delta E} \rightarrow \frac{V}{2\pi\hbar} \left(\frac{2M}{E} \right)^{1/2}. \quad (2.189)$$

Then,

$$p_j(E) = \begin{cases} \frac{V}{2\pi\hbar} \left(\frac{2M}{E} \right)^{1/2} Z^{-1} & \text{for } E < E_{\max}, \\ 0 & \text{else.} \end{cases} \quad (2.190)$$

Here, Z is a constant factor, which could be determined by the normalization condition $\int dE p(E) = 1$:

$$Z = \frac{V}{2\pi\hbar} \int_0^{E_{\max}} dE \left(\frac{2M}{E} \right)^{1/2}. \quad (2.191)$$

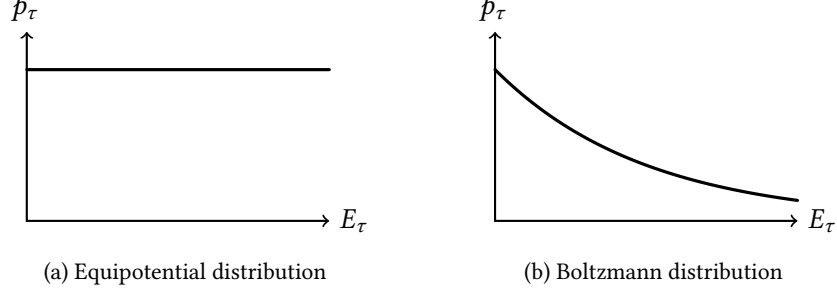


Figure 2.4.: Comparison of equipotential and Boltzmann distribution

2.9.2. Boltzmann distribution for microstates τ

The Boltzmann distribution is given by the formula:

$$p_\tau = Z^{-1} e^{-\frac{E_\tau}{k_B T}} = Z^{-1} e^{-\frac{E_\tau}{T}}. \quad (2.192)$$

We have set $k_B = 1$, which means that we measure the temperature in energy units. E_τ is the energy of state τ . In Figure 2.4 the Boltzmann distribution is compared to an equipotential one.

We find

$$p_B(E) = \frac{V}{2\pi\hbar} \left(\frac{2M}{E} \right)^{1/2} e^{-\frac{E}{T}} Z^{-1}, \quad (2.193)$$

$$\langle E \rangle = \int_0^{E_{\max}} dE p(E) E, \quad (2.194)$$

$$\langle E^2 \rangle = \int_0^{E_{\max}} dE p(E) E^2. \quad (2.195)$$

It should be marked that $p(E)$ is the macroscopic probability distribution and p_τ is the microscopic one.

2.9.3. Number of one particle states in a momentum interval

Number of one particle states in a momentum interval $\Delta p_x, \Delta p_y, \Delta p_z$

First, the one-dimensional case:

$$\Omega(p) = \frac{L}{2\pi\hbar} \Delta p_x. \quad (2.196)$$

Periodical boundary conditions would give

$$p_x < \frac{2\pi\hbar}{L} n < p_x + \Delta p_x. \quad (2.197)$$

For an arbitrary dimension d we have $\Omega(\mathbf{p}) = \frac{V}{(2\pi\hbar)^d} \Delta p_1 \cdots \Delta p_d$, which in the most usual case of $d = 3$ is

$$\boxed{\frac{\partial^3 \Omega}{\partial p_x \partial p_y \partial p_z} = \frac{V}{(2\pi\hbar)^3} \Delta p_1 \Delta p_2 \Delta p_3.} \quad (2.198)$$

Number of one particle states with energy E for arbitrary d

Using the usual energy-momentum relation, we find

$$\frac{\partial \Omega}{\partial E} = \int d^d p \frac{\partial^d \Omega}{\partial p_1 \cdots \partial p_d} \delta \left(\frac{\mathbf{p}^2}{2M} - E \right), \quad (2.199)$$

$$\boxed{\frac{\partial \Omega}{\partial E} = \frac{V}{(2\pi\hbar)^d} \int d^d p \delta \left(\frac{\mathbf{p}^2}{2M} - E \right).} \quad (2.200)$$

For $d = 1$ it is

$$\int_{-\infty}^{\infty} dp \delta \left(\frac{p^2}{2M} - E \right) = \sqrt{2M} \int_0^{\infty} d\varepsilon \varepsilon^{-1/2} \delta(\varepsilon - E) \quad (2.201)$$

$$= \left(\frac{2M}{E} \right)^{1/2}, \quad (2.202)$$

with $\varepsilon = \frac{p^2}{2M}$. We have used that $|p| = \sqrt{2ME}$ and

$$\int_{-\infty}^{\infty} dp = 2 \int_0^{\infty} d|p| \quad (2.203)$$

$$= 2 \int_0^{\infty} d\varepsilon \sqrt{2M} \frac{1}{2} \varepsilon^{-1/2}. \quad (2.204)$$

So, we arrive at

$$\frac{\partial \Omega}{\partial E} = \frac{L}{2\pi\hbar} \left(\frac{2M}{E} \right)^{1/2}. \quad (2.205)$$

For $d = 3$ we have

$$\int d^3 p = 4\pi \int_0^{\infty} d|p| |p|^2, \quad (2.206)$$

with $\varepsilon = \frac{|\mathbf{p}|^2}{2M}$. So, we find

$$\int d^d p \delta \left(\frac{\mathbf{p}^2}{2M} - E \right) = 4\pi \int_0^{\infty} d\varepsilon \delta(\varepsilon - E) \underbrace{\sqrt{\frac{M}{2}} \varepsilon^{-1/2}}_{\frac{\partial |\mathbf{p}|}{\partial \varepsilon}} \underbrace{2M\varepsilon}_{|\mathbf{p}|^2} \quad (2.207)$$

$$= 4\pi \sqrt{2} M^{3/2} E^{1/2}, \quad (2.208)$$

and the number of states per energy is given by

$$\boxed{\frac{\partial \Omega}{\partial E} = \frac{V}{(2\pi\hbar)^3} 4\pi \sqrt{2} M^{3/2} E^{1/2}.} \quad (2.209)$$

Boltzmann probability distribution for one particle

This Boltzmann probability distribution for *one* particle is given by

$$p(E) = \frac{\partial \Omega}{\partial E} e^{-\frac{E}{T}} Z^{-1}, \quad (2.210)$$

where Z is the normalization factor. To find it, we use

$$\int dE p(E) = 1. \quad (2.211)$$

So, we get

$$Z = \int_0^\infty dE \frac{\partial \Omega}{\partial E} e^{-\frac{E}{T}}, \quad (2.212)$$

where $\frac{\partial \Omega}{\partial E} = F E^{1/2}$. Therefore,

$$Z = F \int dE E^{1/2} e^{-E/T} = F \tilde{Z} \quad (2.213)$$

and

$$p(E) = \tilde{Z}^{-1} E^{1/2} e^{-\frac{E}{T}} \quad (2.214)$$

So, it is possible to find an expression for this probability, where \hbar does not longer appear.

2.10. Correlation functions

If there are two observables A and B , what is the probability that a measurement of A and B yields specific values λ_A and λ_B ? Imagine that B is measured after A , then the probability becomes $p(\lambda_B, \lambda_A)$.

2.10.1. Conditional probability $p(\lambda_B|\lambda_A)$

The probability to find λ_B , if λ_A is measured for A , is called conditional probability and is written as $p(\lambda_B|\lambda_A)$. It is

$$p(\lambda_B, \lambda_A) = p(\lambda_B|\lambda_A) p(\lambda_A). \quad (2.215)$$

2.10.2. Product observable $B \circ A$

The product observable is defined as

$$\langle B \circ A \rangle = \sum_{(\lambda_B, \lambda_A)} p(\lambda_B, \lambda_A) \lambda_B \lambda_A \quad (2.216)$$

$$= \sum_{\lambda_B} \sum_{\lambda_A} p(\lambda_B|\lambda_A) p(\lambda_A) \lambda_B \lambda_A. \quad (2.217)$$

Its spectrum is given by

$$\text{Spec}(B \circ A) = \{\lambda_B \lambda_A\}. \quad (2.218)$$

$\langle B \circ A \rangle$ depends on conditional probabilities, not only on $p(\lambda_B)$ and $p(\lambda_A)$.

2.10.3. Correlation function

The (connected) correlation function is defined as

$$\langle B \circ A \rangle_c = \langle B \circ A \rangle - \langle B \rangle \langle A \rangle . \quad (2.219)$$

If $\langle B \circ A \rangle_c \neq 0$, then B and A are correlated.

Example

- A falling marble:

Imagine a marble that is falling out of pipe on a prism, as illustrated in Figure 2.5a. Moreover, it can roll down on any of the sides of the prism with equal probability. However, there are two detectors A and B on the left side. The corresponding measurement values are $\lambda_A = 1$ and $\lambda_B = 1$ if the marble goes through the detectors and 0 otherwise.

The probability to measure $\lambda_A = 1$ is $\frac{1}{2}$. It is easy to see that $p(\lambda_B|\lambda_A) = 1$. Therefore, $p(\lambda_B, \lambda_A) = \frac{1}{2}$. With this we calculate

$$\langle B \circ A \rangle = \frac{1}{2}, \quad (2.220)$$

$$\langle A \rangle = \sum_{\lambda_A} p_A \lambda_A = \frac{1}{2}, \quad (2.221)$$

$$\langle B \rangle = \sum_{\lambda_B} \underbrace{p_B}_{\frac{1}{2}} \lambda_B = \frac{1}{2}, \quad (2.222)$$

$$\langle B \rangle \langle A \rangle = \frac{1}{4}, \quad (2.223)$$

$$\langle B \circ A \rangle_c = \frac{1}{2} - \frac{1}{4} = \frac{1}{4} \quad (2.224)$$

Is $\langle B \circ A \rangle = \langle A \circ B \rangle$ true? Although it is true in this case, in general it is not. A simple example is shown in Figure 2.5b.

2.10.4. Independent observables

If the measurement of A does not influence B , then the conditional probability is independent of λ_A :

$$\boxed{p(\lambda_B|\lambda_A) = p(\lambda_B)} . \quad (2.225)$$

So, *only* for independent observables it is

$$\boxed{p(\lambda_B, \lambda_A) = p(\lambda_B) p(\lambda_A)} . \quad (2.226)$$

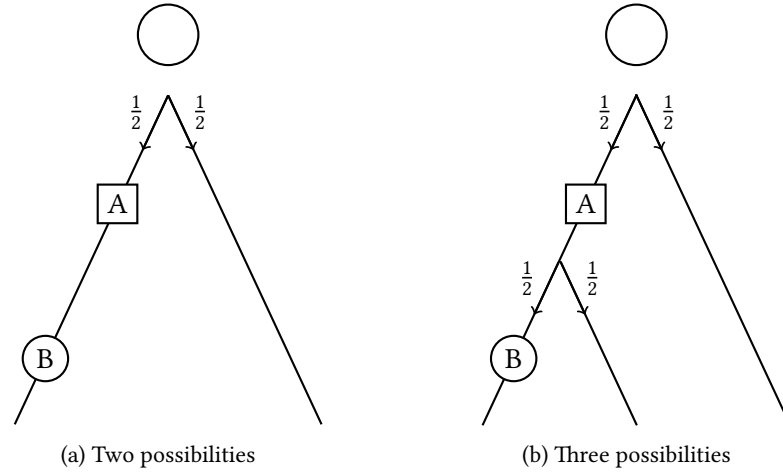


Figure 2.5.: A marble falling down on a prism

But we must keep in mind that *this is not true in general!* In this case, the correlation function can be written as

$$\langle B \circ A \rangle = \sum_{\lambda_B} \sum_{\lambda_A} p(\lambda_B) p(\lambda_A) \lambda_B \lambda_A \quad (2.227)$$

$$= \left(\sum_{\lambda_B} p(\lambda_B) \lambda_B \right) \left(\sum_{\lambda_A} p(\lambda_A) \lambda_A \right) \quad (2.228)$$

$$= \langle B \rangle \langle A \rangle . \quad (2.229)$$

Then,

$$\langle B \circ A \rangle_c = 0 . \quad (2.230)$$

It is important to note that the reverse statement is not true!

2.10.5. Classical correlation function

It is

$$(B \circ A)_\tau = B_\tau A_\tau . \quad (2.231)$$

The classical correlation function is

$$\langle BA \rangle_c = \underbrace{\sum_{\tau} p_{\tau} B_{\tau} A_{\tau}}_{\langle BA \rangle} - \underbrace{\left(\sum_{\tau} p_{\tau} B_{\tau} \right)}_{\langle B \rangle} \underbrace{\left(\sum_{\tau} p_{\tau} A_{\tau} \right)}_{\langle A \rangle} . \quad (2.232)$$

In general, $\langle BA \rangle_c \neq 0$. Setting $B = A$, we find

$$\langle AA \rangle_c = \Delta A^2 . \quad (2.233)$$

There is an easy recipe for calculating conditional probabilities $p(\lambda_B | \lambda_A)$:

1. We eliminate all τ for which $A_\tau \neq \lambda_A$.
2. Now we have a new system with states τ' and

$$\bar{p}_{\tau'} = \frac{p_{\tau'}}{\sum_{\tau'} p_{\tau'}} . \quad (2.234)$$

It is $A_{\tau'} = \lambda_A$ and $B_{\tau'} = B_\tau$.

3. The result is:

$$p(\lambda_B, \lambda_A) = \sum_{\substack{\tau \\ A_\tau = \lambda_A \\ B_\tau = \lambda_B}} p_\tau . \quad (2.235)$$

One should note that after the measurement of A the relative probabilities $\frac{p_{\tau'}}{p_{\rho'}}$ are the same as $\frac{p_\tau}{p_\rho}$ before the measurement of A :

$$\frac{p_{\tau'}}{p_{\rho'}} = \frac{p_\tau}{p_\rho} . \quad (2.236)$$

In this lecture we will be mostly concerned with classical products of two observables, i. e. $BA = AB$, and with classical correlation functions, which are special correlation functions. However, if the relative probabilities change after the measurement of A , then we have different conditional probabilities: $B \circ A \neq BA$. Note that every consistent choice of conditional probabilities $p(\lambda_B | \lambda_A)$ leads to the definition of a product of two observables $B \circ A$. Many different products and correlation functions are possible. In general, the correct choice depends on the way how measurements are performed. This choice is important for the statistical description of systems with few states like atoms (with bounded energy), where the measurement of a Fermi observable has a strong influence on the possible outcome of the measurement of a second observable.

2.10.6. Quantum correlations

We now have

$$\langle B \circ A \rangle = \text{tr} \left(\rho \frac{\hat{A}\hat{B} + \hat{B}\hat{A}}{2} \right) , \quad (2.237)$$

with \hat{A}, \hat{B} being diagonal: $\hat{A} = \text{diag}(A_\tau)$. Therefore,

$$\langle B \circ A \rangle = \sum_{\tau} \rho_{\tau\tau} A_\tau B_\tau . \quad (2.238)$$

The quantum product coincides with the classical one, in this case. However, if \hat{A} and \hat{B} do not commute, this is no longer correct, because in that case \hat{A} and \hat{B} cannot simultaneously be diagonal.

3. Statistical description of many-particles systems

\mathcal{N} is the particle number. Typically $\mathcal{N} \approx N_{\text{Avogadro}} \approx 6 \cdot 10^{23}$.

3.1. Chain with \mathcal{N} lattice sites

3.1.1. Probability distribution

We consider \mathcal{N} fermions being distributed over the chain. If site j is occupied by a fermion, we denote this by $s_j = +1 \cong \uparrow$. If it is not, we denote it by $s_j = -1 \cong \downarrow$. The possible states of such a system is given by $\tau = \{s_j\}$ where $s_j = \pm 1$ for every j . So $\Omega = 2^{\mathcal{N}}$. A possible configuration is shown in Figure 3.1. Later, in subsection 3.2.1, we will interpret this as a chain of spins. In the following, for notational convenience, we will denote any quantity f , which depends on the complete microscopical state $\{s_j\}$, as $f\{s_j\} \equiv f(\{s_j\})$.

Let us define the number of fermions N . For every collection $\{s_j\}$, the corresponding value of the fermion number is

$$N\{s_j\} = \sum_j \frac{1}{2}(s_j + 1). \quad (3.1)$$

Lattice sites are independent and we say that the probability of a site being occupied is q . So, $1 - q$ is the probability of it being unoccupied. Then

$$p\{s_j\} = \prod_{j=1}^{\mathcal{N}} \left[\frac{1}{2} + \left(q - \frac{1}{2} \right) s_j \right] \quad (3.2)$$

$$= q^N (1 - q)^{\mathcal{N} - N}. \quad (3.3)$$

3.1.2. Macroscopic probability

We want to compute the macroscopic probability to find N fermions (particles), that is $p(N)$:

$$p(N) = q^N (1 - q)^{\mathcal{N} - N} \Omega(N) \quad (3.4)$$

where

$$\Omega(N) = \frac{\mathcal{N}!}{N! (\mathcal{N} - N)!} \quad (3.5)$$

is the number of microscopical states with N particles.

3. Statistical description of many-particles systems

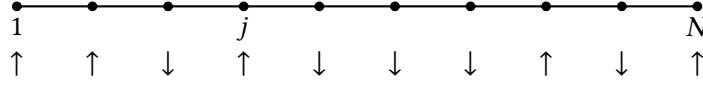


Figure 3.1.: Fermions on a chain

It is given by the number of possible ways to distribute N indistinguishable particles on \mathcal{N} different sites:

Particle 1	\mathcal{N} sites	\rightarrow	$\Omega(1) = \mathcal{N}$
Particle 2	$(\mathcal{N} - 1)$ sites	\rightarrow	$\Omega(2) = \frac{\mathcal{N}(\mathcal{N} - 1)}{2}$
Particle 3	$(\mathcal{N} - 2)$ sites	\rightarrow	$\Omega(3) = \frac{\mathcal{N}(\mathcal{N} - 1)(\mathcal{N} - 2)}{2 \cdot 3}$
			\vdots

All in all, we find the probability to be given by the binomial distribution

$$p(N) = q^N (1 - q)^{\mathcal{N}-N} \frac{\mathcal{N}!}{N! (\mathcal{N} - N)!}. \quad (3.6)$$

Using the general relation

$$(q + r)^{\mathcal{N}} = \sum_{N=0}^{\mathcal{N}} \frac{\mathcal{N}!}{N! (\mathcal{N} - N)!} q^N r^{\mathcal{N}-N}, \quad (3.7)$$

we can prove that $p(N)$ has the correct normalization

$$\sum_{N=0}^{\mathcal{N}} p(N) = 1 \quad (3.8)$$

by setting $r = 1 - q$.

We conclude that there are two different levels at which one can characterize this system. One can consider the $2^{\mathcal{N}}$ microstates $\tau_{\text{micro}} = \{s_j\}$ or the $\mathcal{N} + 1$ macrostates $\tau_{\text{macro}} = N$.

3.1.3. Expectation value

The expectation value is given by

$$\langle N \rangle = \sum_{N=0}^N p(N) N \quad (3.9)$$

$$= \sum_{N=0}^N \frac{N!}{N! (N-N)!} N q^N (1-q)^{N-N} \quad (3.10)$$

$$= \sum_{N=0}^N \frac{N!}{N! (N-N)!} q \frac{\partial}{\partial q} q^N r^{N-N} \Big|_{r=1-q} \quad (3.11)$$

$$= q \frac{\partial}{\partial q} \sum_{N=0}^N \frac{N!}{N! (N-N)!} q^N r^{N-N} \Big|_{r=1-q} \quad (3.12)$$

$$= q \frac{\partial}{\partial q} (q+r)^N \Big|_{r=1-q} \quad (3.13)$$

$$= N(q+r)^{N-1} q \Big|_{r=1-q} \quad (3.14)$$

$$= Nq. \quad (3.15)$$

To arrive at this result, we used the following trick: We expressed N via a q -derivative acting on q^N . Thereto we considered r to be independent of q . Only in the end, after using (3.7) and performing the q -derivative, we substituted $r = 1 - q$ again.

3.1.4. Dispersion

The dispersion can be computed using the same trick as above:

$$\langle N^2 \rangle = \sum_{N=0}^N p(N) N^2 \quad (3.16)$$

$$= \left(q \frac{\partial}{\partial q} \right)^2 (q+r)^N \Big|_{r=1-q} \quad (3.17)$$

$$= q \frac{\partial}{\partial q} \left(Nq (q+r)^{N-1} \right) \Big|_{r=1-q} \quad (3.18)$$

$$= qN \left[(q+r)^{N-1} + (N-1) q (q+r)^{N-2} \right] \Big|_{r=1-q} \quad (3.19)$$

$$= qN + q^2 N(N-1) \quad (3.20)$$

$$= q^2 N^2 + N(q - q^2). \quad (3.21)$$

From the expectation value and the dispersion we can calculate the variance

$$\Delta N^2 = \langle N^2 \rangle - \langle N \rangle^2 \quad (3.22)$$

$$= \mathcal{N}q(1-q). \quad (3.23)$$

and the relative standard deviation

$$\frac{\Delta N}{\langle N \rangle} = \frac{\sqrt{\mathcal{N}q(1-q)}}{\mathcal{N}q} \quad (3.24)$$

$$= \frac{1}{\sqrt{\mathcal{N}}} \sqrt{\frac{1-q}{q}} \quad (3.25)$$

$$= \frac{1}{\sqrt{\langle N \rangle}} \sqrt{1-q}. \quad (3.26)$$

Notice that the relative standard deviation is proportional to $\frac{1}{\sqrt{N}}$! So, the distribution gets sharper with increasing N . For example:

$$N = 10^{20} \quad \rightarrow \quad \frac{\Delta N}{\langle N \rangle} \sim 10^{-10} ! \quad (3.27)$$

This is one of the main reasons why statistical physics has such a predictive power. The statistical uncertainty is so small that one virtually can calculate deterministically by using mean values.

3.2. Uncorrelated probability distribution

Here, uncorrelated means that the probabilities q do not depend on other lattice sites. It does not mean that correlation functions between different sites have to vanish.

3.2.1. \mathcal{N} spins on a chain

We consider a chain of \mathcal{N} spins. Therefore the total spin S , which is a macroscopic observable, is given by

$$S = \sum_j s_j. \quad (3.28)$$

In quantum mechanics the spin is $\frac{\hbar}{2}S$. The number of \uparrow -spins (compare to (3.1)) is

$$N = \sum_{j=1}^{\mathcal{N}} \frac{1}{2}(s_j + 1) \quad (3.29)$$

$$= \frac{1}{2}S + \frac{1}{2}\mathcal{N} \quad (3.30)$$

and therefore

$$S = 2N - \mathcal{N} \quad (3.31)$$

$$= \mathcal{N}(2q - 1). \quad (3.32)$$

Spin expectation value, dispersion, variance and relative standard deviation are then found to be

$$\langle S \rangle = 2 \langle N \rangle - \mathcal{N} \quad (3.33)$$

$$= \mathcal{N}(2q - 1), \quad (3.34)$$

$$\langle S^2 \rangle = 4 \langle N^2 \rangle - 4 \mathcal{N} \langle N \rangle + \mathcal{N}^2, \quad (3.35)$$

$$\Delta S^2 = 4 \left(\langle N^2 \rangle - \langle N \rangle^2 \right) \quad (3.36)$$

$$= 4 \Delta N^2 \quad (3.37)$$

$$= 4 \mathcal{N} q (1 - q), \quad (3.38)$$

$$\frac{\Delta S}{\langle S \rangle} \sim \frac{1}{\sqrt{\mathcal{N}}}, \quad (3.39)$$

where the last equation only holds if $\langle S \rangle \neq 0$, i. e. $q \neq \frac{1}{2}$.

Now, using equation (3.6), the macroscopic probability is given by

$$p(S) = p \left(N = \frac{N+S}{2} \right) \quad (3.40)$$

$$= q^{\frac{N+S}{2}} (1-q)^{\frac{N-S}{2}} \frac{\mathcal{N}!}{\left(\frac{N+S}{2} \right)! \left(\frac{N-S}{2} \right)!} \quad (3.41)$$

and the microscopic probability by

$$p\{s_j\} = \prod_{j=1}^{\mathcal{N}} \left[\frac{1}{2} + \left(q - \frac{1}{2} \right) s_j \right] \quad (3.42)$$

$$= \prod_{j=1}^{\mathcal{N}} P(s_j). \quad (3.43)$$

The latter is a product of independent probabilities for each spin s_j . That is because we assumed “uncorrelated spins”, i. e. the spin s_j does not know anything about its neighbours. In the following this property is used to compute correlation functions.

3.2.2. Correlation functions

Let us compute the correlation function for $l \neq k$ in a microstate $\tau = \{s_j\}$:

$$\langle s_l s_k \rangle = \sum_{\{s_j\}} s_l s_k p\{s_j\} \quad (3.44)$$

Using (3.43) and the fact that the sum over all micro-states can be written as

$$\sum_{\{s_j\}} = \prod_j \sum_{s_j=\pm 1} \quad (3.45)$$

$$= \sum_{s_1=\pm 1} \sum_{s_2=\pm 1} \cdots \sum_{s_N=\pm 1}, \quad (3.46)$$

we find

$$\langle s_l s_k \rangle = \left(\prod_j \sum_{s_j=\pm 1} \right) \cdot \left(\prod_j P(s_j) \right) s_l s_k \quad (3.47)$$

$$= \prod_j \left(\sum_{s_j=\pm 1} P(s_j) \right) s_l s_k \quad (3.48)$$

$$= \prod_{j \neq l, k} \sum_{s_j=\pm 1} P(s_j) \left(\sum_{s_l=\pm 1} \sum_{s_k=\pm 1} P(s_l) P(s_k) s_l s_k \right). \quad (3.49)$$

Moreover, we can use the normalisation of $p\{s_j\}$

$$\prod_j \sum_{s_j=\pm 1} P(s_j) = 1 \quad (3.50)$$

to compute

$$\prod_{j \neq l, k} \sum_{s_j=\pm 1} P(s_j) = \frac{1}{\sum_{s_l=\pm 1} \sum_{s_k=\pm 1} P(s_l) P(s_k)}. \quad (3.51)$$

With this, we finally arrive at

$$\langle s_l s_k \rangle = \frac{\sum_{s_l=\pm 1} \sum_{s_k=\pm 1} P(s_l) P(s_k) s_l s_k}{\sum_{s_l=\pm 1} \sum_{s_k=\pm 1} P(s_l) P(s_k)}. \quad (3.52)$$

Notice, that this is the same expression as for a system consisting of only two spins! We conclude that in the case of uncorrelated spins the spins s_j , $j \neq l, k$, do not matter for $\langle s_l s_k \rangle$.

By a similar calculation we also get an expression for the expectation value:

$$\langle s_k \rangle = \frac{\sum_{s_k=\pm 1} P(s_k) s_k}{\sum_{s_k=\pm 1} P(s_k)}. \quad (3.53)$$

So, obviously

$$\langle s_l s_k \rangle = \langle s_l \rangle \langle s_k \rangle, \quad (3.54)$$

which leads to the observation that the connected correlation function, defined in (2.219) as

$$\langle s_l s_k \rangle_c = \langle s_l s_k \rangle - \langle s_l \rangle \langle s_k \rangle, \quad (3.55)$$

vanishes for the system we are looking at:

$$\langle s_l s_k \rangle_c = 0. \quad (3.56)$$

This resembles the fact that we are dealing with uncorrelated spins here.

In contrast to that, an example for a system with correlated spins is the Ising model. Its microscopic probability is given by

$$p\{s_j\} = e^{-\kappa \sum_j (s_j s_{j+1})}. \quad (3.57)$$

In this case there are next neighbour interactions, so the spins know about each other. We will come back to this later.

3.2.3. Independent measurement

Consider a sequence of *independent* measurements of *one* spin. We denote the individual measurements by $j = 1, \dots, \mathcal{N}$. Then such a sequence is denoted as

$$\{s_j\} = \{s_1, s_2, \dots, s_{\mathcal{N}}\}. \quad (3.58)$$

If the probability for one spin is $P(s_k)$, the probability for a given sequence is

$$p\{s_j\} = \prod_j P(s_j), \quad (3.59)$$

just like throwing dice \mathcal{N} times.

The average of s for a sequence of \mathcal{N} measurements is

$$\langle s \rangle = \frac{\langle S \rangle}{\mathcal{N}} \quad (3.60)$$

$$= 2q - 1. \quad (3.61)$$

Of course, there always is some statistical error for a finite sequence, i. e. the value of the “total spin” S fluctuates between different sequences. The dispersion is computed as

$$\Delta S^2 = \langle S^2 \rangle - \langle S \rangle^2, \quad (3.62)$$

$$\text{where} \quad \langle S^2 \rangle = \sum_{\{s_j\}} p\{s_j\} S^2, \quad (3.63)$$

$$\text{and} \quad S^2 = \left(\sum_m s_m \right)^2. \quad (3.64)$$

From (3.38) we know that

$$\Delta S = 2 \sqrt{\mathcal{N}} \sqrt{q(1-q)}, \quad (3.65)$$

which can be related to Δs via

$$\Delta s = \frac{\Delta S}{\mathcal{N}} \quad (3.66)$$

$$= \frac{2 \sqrt{q(1-q)}}{\sqrt{\mathcal{N}}} \quad (3.67)$$

$$\sim \frac{1}{\sqrt{\mathcal{N}}}. \quad (3.68)$$

Therefore, large sequences give accurate results.

But beware of the fact that the statistical measurement error Δs has nothing to do with the fluctuations Δs_j of the actual spin:

$$\Delta s_j^2 = \langle s_j^2 \rangle - \langle s_j \rangle^2 \quad (3.69)$$

$$= 1 - (2q - 1)^2 \quad (3.70)$$

$$= 4q(1 - q). \quad (3.71)$$

So,

$$\Delta s_j = 2 \sqrt{q(1 - q)} \neq \frac{2 \sqrt{q(1 - q)}}{\sqrt{N}} = \Delta s. \quad (3.72)$$

Δs is an effect of measuring, whereas Δs_j is a quantum mechanical effect as every single measurement either yields $s_j = +1$ or $s_j = -1$ (if $q = \frac{1}{2}$, $\Delta s_j = 1$). Or in other words, s is a macroscopic, whereas s_j is a microscopic variable.

Summary

1. Different physical situations are described by the same probability distribution $p\{s_j\}$:
 - a) Fermions on a lattice of N sites,
 - b) N independent spins,
 - c) Random walk (exercise),
 - d) Sequence of measurements.
2. One major theme in statistical physics is to compute macroscopic probability distributions from microscopic probability distributions.

3.3. Statistical Ensembles

A statistical ensemble is constituted by specifying the states τ and the probability distribution p_τ or the density matrix ρ_τ , respectively.

In such an ensemble different operators / variables can be defined.

We can imagine an ensemble as a collection of infinitely many similar, uncorrelated experimental set-ups in which we can measure the values of the state variables. From this their probability distributions can be inferred ("sequences of measurements with $N \rightarrow \infty$ ").

For example, consider an ensemble of 1-particle systems with spin, i. e. an ensemble of 2-state systems ($\tau = \{\uparrow, \downarrow\}$). This is a sequence of measurements on a single spin and should not be confused with systems with many degrees of freedom (many spins). These would be described by ensembles with many lattice sites or particles.

Of course, real sequences of measurements can only be finite.

3.4. Gaussian distribution

Systems with many uncorrelated degrees of freedom are approximated by a Gaussian distribution. This approximation gets exact in the limit of infinite number of particles $N \rightarrow \infty$.

The continuous description is given by

$$p(x) = c e^{-\frac{(x-\bar{x})^2}{2\sigma^2}}, \quad (3.73)$$

where $\bar{x} = \langle x \rangle$ is the mean value of the macroscopic variable x , $\sigma = \Delta x = \sqrt{\Delta x^2}$ is the variance and $c = c(\sigma)$ is a normalization constant which is determined by demanding

$$1 = \int_{-\infty}^{\infty} dx p(x) \quad (3.74)$$

$$= c(\sigma) \sqrt{2\pi\sigma^2}. \quad (3.75)$$

We find

$$c(\sigma) = \frac{1}{\sqrt{2\pi}} \frac{1}{\sigma}. \quad (3.76)$$

The generalization to several variables x_i is given by:

$$p = c e^{-\frac{1}{2} A_{ij} (x_i - \bar{x}_i)(x_j - \bar{x}_j)}, \quad (3.77)$$

where A_{ij} are the components of a positive definite symmetric Matrix A , i. e. a matrix with only positive eigenvalues. As an example we can think about fermions on a 1-dimensional chain.

We will show that it is the correct large N limit of the binomial distribution

$$p(N) = \frac{N!}{N! (\mathcal{N} - N)!} q^N r^{\mathcal{N}-N}, \quad r = 1 - q. \quad (3.78)$$

Using Stirling's formula

$$\lim_{N \rightarrow \infty} \ln(N!) = \left(N + \frac{1}{2}\right) \ln N - N + \frac{1}{2} \ln(2\pi) + O(N^{-1}), \quad (3.79)$$

we get

$$\alpha(N) = \ln p(N) \quad (3.80)$$

$$\begin{aligned} &= \left(\mathcal{N} + \frac{1}{2}\right) \ln \mathcal{N} - \mathcal{N} + \frac{1}{2} \ln(2\pi) \\ &\quad - \left(N + \frac{1}{2}\right) \ln N + N - \frac{1}{2} \ln(2\pi) \\ &\quad - \left(\mathcal{N} - N + \frac{1}{2}\right) \ln(\mathcal{N} - N) + \mathcal{N} - N - \frac{1}{2} \ln(2\pi) \\ &\quad + N \ln q + (\mathcal{N} - N) \ln(1 - q) \end{aligned} \quad (3.81)$$

$$\begin{aligned} &= \left(\mathcal{N} + \frac{1}{2}\right) \ln \mathcal{N} - \left(N + \frac{1}{2}\right) \ln N - \left(\mathcal{N} - N + \frac{1}{2}\right) \ln(\mathcal{N} - N) \\ &\quad + N \ln q + (\mathcal{N} - N) \ln(1 - q) - \frac{1}{2} \ln(2\pi). \end{aligned} \quad (3.82)$$

It can be shown that

$$\alpha(N) \approx -\frac{(N - \bar{N}^2)}{2 \Delta N^2}, \quad (3.83)$$

where

$$\bar{N} = \mathcal{N}q, \quad (3.84)$$

$$\Delta N^2 = \mathcal{N}q(1 - q), \quad (3.85)$$

for N close to the maximum \bar{N} .

The fact that the maximum is given by $\bar{N} = \mathcal{N}q$, can be proven by computing the first derivative of α :

$$0 = \left. \frac{\partial \alpha}{\partial N} \right|_{N=\bar{N}} \quad (3.86)$$

$$= -\ln \bar{N} - \frac{\bar{N} + \frac{1}{2}}{\bar{N}} + \ln(\mathcal{N} - \bar{N}) + \frac{\mathcal{N} - \bar{N} + \frac{1}{2}}{\mathcal{N} - \bar{N}} + \ln q - \ln(1 - q) \quad (3.87)$$

$$= \ln \frac{\mathcal{N} - \bar{N}}{\bar{N}} - \ln \frac{1 - q}{q} + O(\bar{N}^{-1}). \quad (3.88)$$

So, up to \bar{N}^{-1} corrections we find

$$\frac{\mathcal{N} - \bar{N}}{\bar{N}} = \frac{1 - q}{q}, \quad (3.89)$$

and therefore

$$\bar{N} = \mathcal{N}q. \quad (3.90)$$

Moreover, from the second derivative we can infer the variance:

$$-\frac{1}{\Delta N^2} = \left. \frac{\partial^2 \alpha}{\partial N^2} \right|_{N=\bar{N}} \quad (3.91)$$

$$= \frac{1}{\mathcal{N} - \bar{N}} - \frac{1}{\bar{N}} - \frac{\mathcal{N}}{2} \left(-\frac{1}{\bar{N}^2(\mathcal{N} - \bar{N})} + \frac{1}{(\mathcal{N} - \bar{N})^2 \bar{N}} \right) \quad (3.92)$$

$$= -\frac{\mathcal{N}}{(\mathcal{N} - \bar{N}) \bar{N}} + O(\bar{N}^{-2}). \quad (3.93)$$

So, using $\bar{N} = \mathcal{N}q$, up to \bar{N}^{-2} corrections it is

$$-\frac{1}{\Delta N^2} = -\frac{\mathcal{N}}{(\mathcal{N} - \mathcal{N}q) \mathcal{N}q}, \quad (3.94)$$

and therefore

$$\Delta N = \mathcal{N}q(1 - q). \quad (3.95)$$

Furthermore, the third derivative yields

$$\left. \frac{\partial^3 \alpha}{\partial N^3} \right|_{N=\bar{N}} = \frac{\mathcal{N}^2}{(\mathcal{N} - \bar{N})^2 \bar{N}^2} \quad (3.96)$$

All in all, this leads to the central limit theorem of statistics which states that any uncorrelated ensemble can be described by a Gaussian probability distribution in the limit $N \rightarrow \infty$.

3.5. Thermodynamic limit

The thermodynamic limit is given by $V \rightarrow \infty$ or $N \rightarrow \infty$, respectively. In this limit the Gaussian distribution becomes extremely narrow. For example

$$\bar{N} = 10^{20} \quad \rightarrow \quad \Delta N^2 = 2N = 2 \cdot 10^{20}, \quad (3.97)$$

when $q = 1/2$. If we then consider a small deviation

$$N = \bar{N} (1 + 10^{-5}), \quad (3.98)$$

the relative probability is given by

$$\frac{p(N)}{p(\bar{N})} = e^{-\frac{(N-\bar{N})^2}{2 \Delta N^2}} \quad (3.99)$$

$$= e^{-\frac{10^{-10} \cdot 10^{40}}{4 \cdot 10^{20}}} \quad (3.100)$$

$$= e^{-\frac{1}{4} 10^{10}} \quad (3.101)$$

$$\sim 10^{-10^9}. \quad (3.102)$$

And for a deviation

$$N = \bar{N} (1 + 10^{-10}) \quad (3.103)$$

it is given by

$$\frac{p(N)}{p(\bar{N})} = e^{-\frac{1}{4}} = O(1). \quad (3.104)$$

Therefore, only values

$$\bar{N} - 5 \Delta N \lesssim N \lesssim \bar{N} + 5 \Delta N \quad (3.105)$$

play a role. And in this regime the deviations of the actual probability distribution from a Gaussian distribution are extremely small!

4. Equilibrium Ensembles

Equilibrium states are described by statistical ensembles with a static probability distribution.

If the probability distribution $\{p_\tau\}$ is time independent, then all expectation values of the observables are time independent, too. One example would be an isolated system after sufficiently long time or, respectively, just parts of the isolated system. Another one would be an open system being immersed in a heat bath.

Basic postulate for equilibrium states

An isolated system approaches an equilibrium state. After sufficiently long time an isolated system can be described by an equilibrium ensemble.

This conjecture was made by Boltzmann, but was never proven. Actually, it most probably is not true, as there are two problems with it. First, the necessary time scale could be extremely large. For example think of systems like water, honey or glass. And second, the microscopic equations are invariant under time reversal, but the time direction always points towards equilibrium.

But even though the basic postulate may not be strictly true, and despite the existence of important practical systems which do not equilibrate, the basic postulate is useful for many applications and it often provides very good approximations.

4.1. Microcanonical ensemble

In the following we will be examining an isolated system in a volume V with *fixed* energy E and *fixed* number of particles N .

4.1.1. Fundamental statistical postulate

Microstates

These are the states τ of a statistical ensemble, given by all possible quantum states with specified V, E, N . Note that in quantum mechanics one has to choose a complete basis of eigenstates to E, N in a given volume V . Then one can identify

$$\psi_\tau \cong \tau. \quad (4.1)$$

Or equivalently: An isolated system in equilibrium can be found with equal probability in each of its allowed states.

Microcanonical ensemble

The number of microstates with given E and N is denoted by $\Omega(E, N)$. Then the probability of the system being in one these is given by

$$p_\tau = \frac{1}{\Omega(E, N)} . \quad (4.2)$$

This raises several arguments in favour of the basic postulate for equilibrium states. To begin with, no state τ is preferred to another state τ' . In addition to that, this so-called equipartition of states (i. e. equal p_τ) is constant in time. But one has to be careful. The existence of an asymptotic state for $t \rightarrow \infty$ is only valid for large enough N . Moreover there is an obstruction due to conserved correlations, i. e. a system whose initial correlations differ from the ones of the microcanonical ensemble cannot reach the latter asymptotically.

Note that for a microcanonical ensemble all expectation values and classical correlations for classical observables are uniquely fixed.

4.1.2. Microcanonical partition function

The microcanonical partition function is denoted by Z_{mic} , where

$$Z_{\text{mic}} = \Omega(E) , \quad (4.3)$$

$$p_\tau = Z_{\text{mic}}^{-1} . \quad (4.4)$$

$Z_{\text{mic}}(E, N, V)$ depends on the thermodynamical variables E, N, V . Recall that $\Omega(E) \equiv \delta\Omega(E)$ is the number of states with energy in between E and $E + \delta E$.

4.1.3. Entropy S

For a microcanonical ensemble the entropy is given by:

$$\boxed{S = k_B \ln \Omega} , \quad (4.5)$$

where k_B is the Boltzmann constant. Setting it to 1 simply means to measure the temperature in energy units, e. g. eV. Then we have

$$S(E, N, V) = \ln \Omega(E, N, V) . \quad (4.6)$$

This equation is valid for isolated systems in equilibrium. The entropy S is an extensive thermodynamic potential, i. e. it is additive when uniting subsystems. One can conclude that, in the limit of large N , the entropy is proportional to N :

$$S \sim N \quad \text{for} \quad N \rightarrow \infty . \quad (4.7)$$

Examples

1. Consider N uncorrelated particles where each of them can be in F different states. Furthermore, say the energy is given by $E = cN$, independent of the particles states. Then there are F^N possible microstates.

$$\Omega = F^N \quad (4.8)$$

$$\ln \Omega = N \ln F \quad (4.9)$$

$$S = k_B N \ln F \quad (4.10)$$

2. Consider N uncorrelated particles on \mathcal{N} lattice sites with $E = cN$ and $\mathcal{N} = bV$.

$$\Omega = \frac{\mathcal{N}!}{N! (\mathcal{N} - N)!} \quad (4.11)$$

$$S = \ln \mathcal{N}! - \ln N! - \ln (\mathcal{N} - N)! \quad (4.12)$$

Applying Stirling's formula on $\ln N!$ for large N :

$$\ln N! \approx N \ln N - N \quad (4.13)$$

one finds

$$S = [\mathcal{N} \ln \mathcal{N} - \mathcal{N}] - [N \ln N - N] - [(\mathcal{N} - N) \ln (\mathcal{N} - N) - (\mathcal{N} - N)] \quad (4.14)$$

$$= N \ln \frac{\mathcal{N} - N}{N} + \mathcal{N} \ln \frac{\mathcal{N}}{\mathcal{N} - N} \quad (4.15)$$

$$= N \ln \left(\frac{\mathcal{N}}{N} - 1 \right) - \mathcal{N} \ln \left(1 - \frac{N}{\mathcal{N}} \right) \quad (4.16)$$

$$= N \left[\ln \left(\frac{bV}{N} - 1 \right) - \frac{bV}{N} \ln \left(1 - \frac{N}{bV} \right) \right]. \quad (4.17)$$

Now it is obvious that S is extensive, since one can write it as N times a function of the particle density $n = \frac{N}{V}$:

$$S = N \cdot f(n). \quad (4.18)$$

For small n we can approximate

$$S = N \left[\ln \left(\frac{b}{n} - 1 \right) + 1 \right]. \quad (4.19)$$

4.1.4. Partition function for ideal classical gas

To calculate the partition function for an ideal classical gas, we have to count the number of quantum states of N particles in a box with volume V (e. g. with periodic boundary conditions) for a fixed total energy E . We will approach this step by step.

1) Quantum states for one particle ($N = 1$)

$$\tau \cong \mathbf{p} = (p_x, p_y, p_z) \quad (4.20)$$

Assigning an occupation number $n(\mathbf{p})$ to each of these states, we find

$$1 = \sum_{\mathbf{p}} n(\mathbf{p}), \quad (4.21)$$

$$E = \sum_{\mathbf{p}} n(\mathbf{p}) E(\mathbf{p}), \quad (4.22)$$

$$\text{with } E(\mathbf{p}) = \frac{\mathbf{p}^2}{2M}. \quad (4.23)$$

The partition sum can then be expressed as

$$\Omega = \sum_{\mathbf{p}} \left| \frac{\mathbf{p}^2}{2M} = E \right| = \int d^3p \frac{\partial^3 \Omega}{\partial p^3} \delta \left(\frac{\mathbf{p}^2}{2M} - E \right) \delta E \quad (4.24)$$

$$= \frac{V}{(2\pi\hbar)^3} \int d^3p \delta \left(\frac{\mathbf{p}^2}{2M} - E \right) \delta E \quad (4.25)$$

$$= \frac{V}{(2\pi\hbar)^3} 4\pi \sqrt{2} M^{3/2} E^{1/2} \delta E. \quad (4.26)$$

2) Quantum states for two particles ($N = 2$)

Using the same definitions as in the one particle case, we now get

$$2 = \sum_{\mathbf{p}} n(\mathbf{p}), \quad (4.27)$$

$$E = \sum_{\mathbf{p}} n(\mathbf{p}) E(\mathbf{p}). \quad (4.28)$$

We should discriminate between fermions, for which $n(\mathbf{p}) \in \{0, 1\}$, and bosons, for which $n(\mathbf{p})$ can take arbitrary values. In any case, we can split up the partition sum into two parts:

$$\Omega = \Omega_1 + \Omega_2. \quad (4.29)$$

Here, Ω_1 is the number of states with $n(\mathbf{p}_1) = 1, n(\mathbf{p}_2) = 1$ for some $\mathbf{p}_1 \neq \mathbf{p}_2$, and Ω_2 is the number of states with $n(\mathbf{p}) = 2$ for some \mathbf{p} . These two contributions are given by

$$\Omega_1 = \frac{1}{2} \sum_{\mathbf{p}_1} \sum_{\mathbf{p}_2 \neq \mathbf{p}_1} \left| \frac{\mathbf{p}_1^2 + \mathbf{p}_2^2}{2M} = E \right|, \quad (4.30)$$

$$\Omega_2 = \begin{cases} 0 & \text{for fermions,} \\ \Omega_2^{(b)} = \sum_{\mathbf{p}} \left| \frac{\mathbf{p}^2}{M} = E \right| & \text{for bosons.} \end{cases} \quad (4.31)$$

The factor $\frac{1}{2}$ in (4.30) is due to the fact that we are dealing with indistinguishable quantum mechanical particles. So, swapping both particles does not count as a different state.

The full partition sum then takes the form

$$\Omega = \frac{1}{2} \sum_{\mathbf{p}_1} \sum_{\mathbf{p}_2} \left|_{E=\frac{\mathbf{p}_1^2 + \mathbf{p}_2^2}{2M}} \pm \frac{1}{2} \Omega_2^{(b)} \right. \quad \left(\begin{array}{l} \text{bosons} \\ \text{fermions} \end{array} \right) . \quad (4.32)$$

Note that in (4.32) the \mathbf{p}_2 sum runs over all possible momentum values, whereas in (4.30) it excludes $\mathbf{p}_2 = \mathbf{p}_1$.

If we neglect the $\pm \frac{1}{2} \Omega_2^{(b)}$ term in (4.32) we arrive at the approximate description of a classical gas. Then there is no distinction between fermions and bosons any more. Performing the two momentum sums analogously to the one particle case leads to

$$\Omega = \frac{1}{2} \int d^3 p_1 d^3 p_2 \left(\frac{V}{(2\pi\hbar)^3} \right)^2 \delta \left(\frac{\mathbf{p}_1^2 + \mathbf{p}_2^2}{2M} - E \right) \delta E . \quad (4.33)$$

3) Quantum states for N particles

The extension to the general N particle case is straight-forward:

$$N = \sum_{\mathbf{p}} n(\mathbf{p}) , \quad (4.34)$$

$$E = \sum_{\mathbf{p}} n(\mathbf{p}) E(\mathbf{p}) , \quad (4.35)$$

$$\Omega = \frac{1}{N!} \int d^3 p_1 \cdots d^3 p_n \left(\frac{V}{(2\pi\hbar)^3} \right)^N \delta \left(\sum_{j=1}^N \frac{\mathbf{p}_j^2}{2M} - E \right) \delta E . \quad (4.36)$$

As in the two particle case, equation (4.36) neglects the particular behaviour when two or more momenta commute. Again, this is the classical approximation where no distinction between fermions and bosons is made.

4) Partition function

We define

$$\omega = \frac{\Omega}{\delta E} = \frac{\partial \Omega}{\partial E} . \quad (4.37)$$

To find ω , we have to evaluate a $3N$ -dimensional integral. For this purpose, we will perform a change of variables:

$$p_{j,k} = \sqrt{2ME} x_{j,k} , \quad \frac{\mathbf{p}_j^2}{2M} = \mathbf{x}_j^2 E . \quad (4.38)$$

where the \mathbf{x}_j are chosen such that

$$\sum_j \mathbf{x}_j^2 E = E . \quad (4.39)$$

Then

$$\int_{-\infty}^{\infty} d^3 p_1 \cdots d^3 p_N = (2ME)^{3N/2} \int_{-\infty}^{\infty} d^3 x_1 \cdots d^3 x_N, \quad (4.40)$$

$$\sum_j \frac{\mathbf{p}_j^2}{2M} = \sum_j \mathbf{x}_j^2 E. \quad (4.41)$$

For our purposes, we also introduce the $3N$ -component vector $\mathbf{y} = (x_{1,1}, x_{1,2}, x_{1,3}, x_{2,1}, \dots)$. Now, finally, we can calculate ω :

$$\omega = \frac{1}{N!} \left(\frac{V}{(2\pi\hbar)^3} \right)^N (2ME)^{3N/2} \int_{-\infty}^{\infty} dy_1 \cdots dy_{3N} \delta \left(\sum_i y_i^2 E - E \right) \quad (4.42)$$

$$= \frac{1}{N!} V^N \left(\frac{ME}{2\pi^2\hbar^2} \right)^{3N/2} \frac{1}{E} \cdot F. \quad (4.43)$$

The factor $\frac{1}{E}$ comes from applying the relation $\delta(E \cdot X) = \frac{1}{E} \delta(X)$ and

$$F = \int_{-\infty}^{\infty} dy_1 \cdots dy_{3N} \delta \left(1 - \sum_{i=1}^{3N} y_i^2 \right) \quad (4.44)$$

is the volume of the $(3N - 1)$ -dimensional unit sphere \mathcal{S}^{3N} , which is equal to the surface of a $3N$ -dimensional unit ball. It can be calculated from the general formula

$$F = \frac{2}{\Gamma\left(\frac{3N}{2}\right)} \pi^{3N/2} \quad (4.45)$$

$$= \frac{2}{\left(\frac{3N}{2} - 1\right)!} \pi^{3N/2}. \quad (4.46)$$

As a consistency check we compute (4.46) in the case $N = 1$:

$$F = \frac{2}{\frac{1}{2}!} \pi^{3/2} \quad (4.47)$$

$$= \frac{4}{\sqrt{\pi}} \pi^{3/2} \quad (4.48)$$

$$= 4\pi \quad \text{q. e. d.} \quad (4.49)$$

Here we have used $\frac{1}{2}! = \Gamma\left(\frac{3}{2}\right) = \frac{1}{2} \sqrt{\pi}$ in the second line. Using the definition (4.37), the partition sum is given by

$$\Omega = \frac{2}{N! \left(\frac{3N}{2} - 1\right)!} \left(\frac{M}{2\pi\hbar^2} \right)^{3N/2} E^{3N/2} V^N \frac{\delta E}{E}. \quad (4.50)$$

Finally, using $\frac{\delta E}{E} = \frac{1}{3N}$, we arrive at

$$Z_{\text{mic}} = \Omega = \frac{1}{N! \left(\frac{3N}{2}\right)!} \left(\frac{M}{2\pi\hbar^2}\right)^{3N/2} E^{3N/2} V^N. \quad (4.51)$$

N is typically of the order $6 \cdot 10^{26}$. Therefore, Z_{mic} increases extremely rapidly with E and V !

4.1.5. Entropy of an ideal classical gas

With the help of our favorite approximated Stirling's formula (4.13) we can find the entropy of an ideal classical gas S :

$$S = \frac{3N}{2} \ln \frac{M}{2\pi\hbar^2} + \frac{3N}{2} \ln E + N \ln V - \ln N! - \ln \left(\frac{3N}{2}\right)! \quad (4.52)$$

$$= k_B 3N \left\{ \frac{1}{2} \ln \frac{M}{2\pi\hbar^2} + \frac{1}{3} \ln V + \frac{1}{2} \ln E - \frac{1}{3} \ln N + \frac{1}{3} - \frac{1}{2} \ln \frac{3N}{2} + \frac{1}{2} \right\}. \quad (4.53)$$

So,

$$S = k_B 3N \left\{ \frac{1}{3} \ln \frac{V}{N} + \frac{1}{2} \ln \frac{2}{3} \frac{E}{N} + \frac{1}{2} \ln \frac{M}{2\pi\hbar^2} + \frac{5}{6} \right\}. \quad (4.54)$$

$3N = f$ is the number of degrees of freedom. In a compact form we can write:

$$S = c k_B f \quad (4.55)$$

and therefore

$$\Omega(E) \sim E^{cf}, \quad (4.56)$$

where c depends on the system. So, the partition sum increases rapidly with increasing energy. This form of the entropy and the partition sum applies too many other systems, too.

It is important to note, that the factor $\frac{1}{N!}$ in $\Omega(E)$ was necessary for the entropy to be an extensive quantity. This factor is absent in the classical description! It was first introduced by Max Planck in 1900 and was heavily discussed back then.

4.1.6. Thermodynamic limit

The thermodynamic limit is described by

$$V \rightarrow \infty, \quad N \rightarrow \infty, \quad E \rightarrow \infty. \quad (4.57)$$

In this limit the useful quantities to describe the system are:

$$\frac{N}{V} = n \quad (\text{particle density}), \quad (4.58)$$

$$\frac{E}{V} = \epsilon \quad (\text{energy density}), \quad (4.59)$$

Extensive	Intensive
V	p (pressure)
N	n
E	ϵ
S	$s = \frac{S}{V}$ (entropy density)

Table 4.1.: Examples for extensive and intensive quantities

as these stay constant in the limiting process. So, mathematically more rigorously we define the limiting process via an auxiliary parameter a :

$$V \rightarrow aV, \quad N \rightarrow aN, \quad E \rightarrow aE \quad (4.60)$$

and consider the limit

$$a \rightarrow \infty. \quad (4.61)$$

We make the distinction between extensive quantities (A), where

$$\lim_{V \rightarrow \infty} \frac{\langle A \rangle}{V} = \text{finite} \neq 0, \quad (4.62)$$

and the intensive quantities (B), where

$$\lim_{V \rightarrow \infty} \langle A \rangle = \text{finite}. \quad (4.63)$$

Some extensive and intensive quantities are listed in table Table 4.1.

Again note, that the entropy is an extensive quantity:

$$S = 3k_B n V c(n, \epsilon), \quad (4.64)$$

where

$$c(n, \epsilon) = \frac{1}{2} \ln \left(\frac{2}{3} \frac{\epsilon}{n} \right) - \frac{1}{3} \ln n + \frac{1}{2} \ln \frac{M}{2\pi\hbar^2} + \frac{5}{6}. \quad (4.65)$$

On the other hand, ratios of extensive quantities, such as n , ϵ and s , are intensive quantities.

By setting $k_B = 1$ and $\hbar = 1$ we find the entropy density to be given by

$$s = \frac{S}{V} \quad (4.66)$$

$$= 3n \left\{ \ln n^{-1/3} + \ln \left(\frac{\epsilon}{n} \right)^{1/2} + \ln M^{1/2} + \text{const.} \right\}. \quad (4.67)$$

So

$$s = 3n \left(\ln \sqrt{\frac{M\epsilon}{n^{5/3}}} + \text{const.} \right). \quad (4.68)$$

In general we have to be careful with units, as we can only take logarithms of dimensionless numbers. But a quick dimensional analysis confirms this here:

$$M \sim M^1, \quad (4.69)$$

$$\epsilon \sim M^4, \quad (4.70)$$

$$n \sim M^3. \quad (4.71)$$

4.1.7. The limit $E \rightarrow E_0$

We now denote by E_0 the quantum mechanical energy of the ground state and choose δE to be an energy difference smaller than the difference between E_0 and the lowest excited state E_1 :

$$\delta E < E_1 - E_0 . \quad (4.72)$$

So, if the ground state is non-degenerate, then

$$\lim_{E \rightarrow E_0} \Omega(E) = 1 \quad (4.73)$$

and as a consequence

$$\boxed{\lim_{E \rightarrow E_0} S(E) = 0 .} \quad (4.74)$$

On the other hand, if we have a q -fold degeneracy, then

$$\lim_{E \rightarrow E_0} S = k_B \ln q \ll k_B f . \quad (4.75)$$

So, often the assumption is made that S vanishes in this limit, as $\ln q$ is tiny compared to the usually huge f . Then

$$\frac{S}{V} \rightarrow 0 \quad (4.76)$$

in the thermodynamic limit.

4.1.8. Derivatives of the entropy

Temperature

First recall that up to now, the partition sum and the entropy have been functions of the state variables E , N , and V :

$$Z_{\text{mic}} = \Omega(E, N, V) , \quad (4.77)$$

$$S(E, N, V) = k_B \ln \Omega(E, N, V) . \quad (4.78)$$

Now we take the derivative with respect to E at fixed particle number and volume:

$$\boxed{\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{N, V} .} \quad (4.79)$$

This defines the temperature $T = T(E, N, V)$. It can be used as a new thermodynamic state variable. So, we can make a variable substitution in order to express the entropy not as a function of E but of T : $S = S(T, N, V)$. Moreover, we see that the temperature is an intensive quantity, as S and E are extensive and so their ratio must be intensive.

Pressure

Now we take the derivative with respect to V at fixed particle number and energy:

$$\boxed{\frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E,N}} \quad (4.80)$$

This defines the pressure p . And it can immediately be concluded that the pressure is an intensive quantity.

Let us now compute this for our ideal gas of particles:

$$\frac{\partial S}{\partial V} = k_B \frac{\partial}{\partial V} \left(N \ln \frac{V}{N} + f(E, N) \right) \quad (4.81)$$

$$= k_B \frac{N}{V} \quad (4.82)$$

$$= \frac{p}{T} \quad (4.83)$$

From this we can infer the equation of state for an ideal classical gas:

$$\boxed{pV = k_B N T} \quad (4.84)$$

In thermodynamics T and p are determined by the change of the number of available states when changing E and V .

Chemical potential

Now we take the derivative with respect to N at fixed volume and energy:

$$\mu = -T \left. \frac{\partial S}{\partial N} \right|_{E,V} \quad (4.85)$$

This defines the chemical potential μ .

We see that we can compute all these variables, if we just have the entropy at hand. So, basically the entropy already contains the whole thermodynamics! From here on, one can develop the complete formal machinery of thermodynamics, resting upon the thermodynamic potentials, such as S , E , F and G , and their derivatives (Legendre transformations).

Note that up to now T is only a definition, but it will be motivated in the following.

We conclude this section by giving a little summary about entropy in the case of a microcanonical ensemble description of a system in equilibrium:

- $S = k_B \ln \Omega$
- S is extensive, i. e. $S \sim V$, $S \sim f \sim N$
- $\frac{S}{V} \rightarrow 0$ for $E \rightarrow E_0$ (3rd law of thermodynamics)
- $\frac{1}{T} = \left. \frac{\partial S(E, N, V)}{\partial E} \right|_{N,V}$ (important part of the 2nd law of thermodynamics)

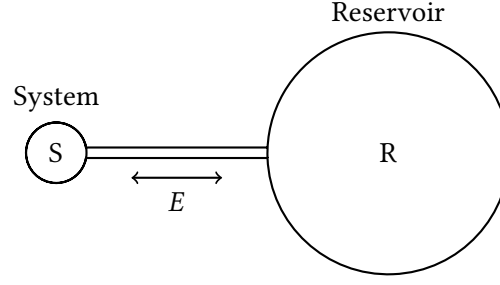


Figure 4.1.: System connected to a reservoir with energy exchange

4.2. Canonical ensemble

4.2.1. System in a heat bath

A canonical ensemble describes a system in a heat bath, intuitively speaking. We consider a closed system with a volume V . So, energy can be exchanged with the environment, but the particle number N remains constant. Additionally, there is a heat bath characterized by the temperature T . In the following we will derive all properties of this compound system.

The system S is connected to the reservoir (heat bath) R with which it can exchange energy, but no particles. This is illustrated in Figure 4.1. But it is important that the total system is isolated. Moreover we demand that

$$V_R \gg V_S, \quad (4.86)$$

$$N_R \gg N_S \gg 1, \quad (4.87)$$

i. e. the reservoir R is much larger than the system S concerning volume as well as particle number.

The total system can be described by a microcanonical ensemble with total energy

$$E_G = E_S + E_R = \text{const.} \quad (4.88)$$

However, E_R and E_S are not fixed individually. More precisely, the total system is described by a Hamiltonian

$$H = H_S + H_R + H_{\text{int}}, \quad (4.89)$$

where H_{int} describes the interaction between S and R . We assume this coupling to be weak:

$$|\langle H_{\text{int}} \rangle| \ll \langle H_S \rangle, \langle H_R \rangle. \quad (4.90)$$

Now we denote by $|\psi_n\rangle$ the quantum states of S and by $|\phi_\nu\rangle$ the quantum states of R . They form a complete basis, i. e. a quantum state (microstate) of the full system can always be expressed as

$$|\chi_\tau\rangle = |\psi_n\rangle |\phi_\nu\rangle, \quad (4.91)$$

where $\tau = (n, \nu)$. Now the energies are given by

$$E_{Sn} = \langle \psi_n | H_S | \psi_n \rangle , \quad (4.92)$$

$$E_{R\nu} = \langle \varphi_\nu | H_R | \varphi_\nu \rangle , \quad (4.93)$$

$$E_{Gn\nu} = E_{Sn} + E_{R\nu} = E_G . \quad (4.94)$$

And from the statistical postulate it follows that

$$p_{n\nu} = \Omega_G^{-1}(E_G) . \quad (4.95)$$

We now consider operators which only depend on n , i. e. they are independent of the properties of the reservoir:

$$\langle A \rangle = \sum_n \sum_\nu p_{n\nu} A_n \quad (4.96)$$

$$= \sum_n p_n A_n , \quad (4.97)$$

where

$$p_n = \sum_\nu p_{n\nu} . \quad (4.98)$$

In other words, we look at the reduced system S. This restriction leads to the concept of a canonical ensemble. Our task is now to compute the reduced probability distribution p_n . These considerations will lead us to the famous Boltzmann factor $e^{-\frac{E_n}{k_B T}}$.

4.2.2. Canonical partition function and Boltzmann factor

From (4.94), (4.95) and (4.98) we conclude that

$$p_n = \sum_\nu \Omega_G^{-1}(E_G) \delta(E_{Sn} + E_{R\nu} - E_G) \delta E_G . \quad (4.99)$$

This means that the sum only runs over such ν which fulfil the constraint (4.94). So, the sum is just the number of states in R. Therefore p_n will depend on E_{Sn} and the properties of the reservoir (T_R) and we conclude that

$$p_n(E_S) = \frac{\Omega_R(E_R)}{\Omega_G(E_G)} \Big|_{E_R=E_G-E_S} . \quad (4.100)$$

As we demanded $E_S \ll E_R$, E_R is close to E_G and we can make a Taylor expansion of $\ln \Omega_R(E_R)$ around E_G :

$$\ln \Omega_R(E_R) \approx \ln \Omega_R(E_G) + \underbrace{(E_R - E_G)}_{-E_S} \underbrace{\frac{\partial}{\partial E_G} \ln \Omega_R(E_G)}_{\beta} . \quad (4.101)$$

For $\frac{E_R}{E_S} \rightarrow \infty$ this gets exact ($\ln \Omega_R \sim f_R \ln E_R$).

We have defined

$$\beta = \left. \frac{\partial}{\partial E} \ln \Omega_R(E) \right|_{E=E_R}, \quad (4.102)$$

where the evaluation at E_R instead of E_G is fine, as this difference is a higher order effect in the Taylor expansion. Note that $\beta \geq 0$ for monotonously increasing $\Omega_R(E)$.

So, β just depends on the density of states of the reservoir:

$$\beta = \left. \frac{1}{k_B} \frac{\partial S_R}{\partial E} \right|_{E=E_R} \quad (4.103)$$

and we infer that

$$\beta = \frac{1}{k_B T_R}, \quad (4.104)$$

with T_R being the temperature of the reservoir. Ω_R is now given by

$$\Omega_R(E_R) = \Omega_R(E_G) e^{\beta(E_R - E_G)}, \quad (4.105)$$

and inserting this into (4.100) leads to

$$p_n(E_s) = Z_{\text{can}}^{-1} e^{-\beta E_s}. \quad (4.106)$$

Here,

$$Z_{\text{can}} = \frac{\Omega_G(E_G)}{\Omega_R(E_G)} \quad (4.107)$$

is the canonical partition function. It only depends on properties of the reservoir. It can be computed from the normalisation of p_n :

$$\sum_n p_n = 1. \quad (4.108)$$

So,

$$Z_{\text{can}} = \sum_n e^{-\beta E_n}. \quad (4.109)$$

Or, summing over energy states instead:

$$Z_{\text{can}} = \sum_E \Omega(E) e^{-\beta E} \quad (4.110)$$

$$= \int dE \omega(E) e^{-\beta E}. \quad (4.111)$$

So all in all,

$$Z_{\text{can}} = Z_{\text{can}}(\beta, N, V), \quad (4.112)$$

where N and V are the particle number and the volume of S. So, Z_{can} only depends on R via β .

4.2.3. Relation between temperature and energy

Let us next compute the mean energy of our statistical ensemble. First, we recall the probability to be given by

$$p_n = Z_{\text{can}}^{-1} e^{-\beta E_n}, \quad (4.113)$$

where

$$Z_{\text{can}} = \sum_n e^{-\beta E_n} \quad (4.114)$$

and, using the convention $k_B = 1$,

$$\beta = \frac{1}{T} = \frac{1}{T_R}. \quad (4.115)$$

Then we find

$$\bar{E} = \langle E \rangle = \sum_n p_n E_n \quad (4.116)$$

$$= \frac{\sum_n e^{-\beta E_n} E_n}{\sum_n e^{-\beta E_n}} \quad (4.117)$$

$$= -\frac{\partial}{\partial \beta} \ln \sum_n e^{-\beta E_n}. \quad (4.118)$$

So, the mean energy is given by

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln Z_{\text{can}}. \quad (4.119)$$

4.2.4. Energy distribution

The probability density to find an energy E is

$$w(E) = \omega(E) e^{-\beta E} \quad (4.120)$$

$$\sim E^{c_s f_s} e^{-\beta E}, \quad (4.121)$$

where f_s , the number of degrees of freedom, is usually a huge number. Then, this distribution describes an extremely sharp peak, as sketched in Figure 4.2.

Say, that $w(E)$ is given by a Gaussian distribution:

$$w(E) = \text{const.} \cdot e^{-\frac{(E - \bar{E})^2}{2 \Delta E^2}}. \quad (4.122)$$

Then we can determine \bar{E} from the maximum of w or, respectively $\ln w$:

$$\ln w = -\frac{(E - \bar{E})^2}{2 \Delta E^2} + \text{const.} \quad (4.123)$$

$$= c_s f_s \ln E - \beta E, \quad (4.124)$$

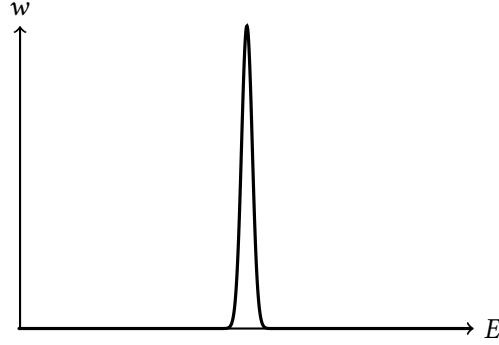


Figure 4.2.: Energy distribution

where we inserted (4.121) in the second line. We get

$$0 = \left. \frac{\partial \omega}{\partial E} \right|_{\bar{E}} = \frac{c_s f_s}{\bar{E}} - \beta. \quad (4.125)$$

This means that

$$\bar{E} = \frac{c_s f_s}{\beta} = c_s f_s T, \quad (4.126)$$

and therefore

$$T = \frac{\bar{E}}{c_s f_s}, \quad (4.127)$$

so the temperature can be interpreted as something like the energy per degree of freedom.

The energy dispersion can be inferred from the second derivative:

$$\frac{1}{\Delta E^2} = - \left. \frac{\partial^2 \omega}{\partial E^2} \right|_{\bar{E}} = \frac{c_s f_s}{\bar{E}^2}. \quad (4.128)$$

So,

$$\Delta E = \frac{\bar{E}}{\sqrt{c_s f_s}}, \quad (4.129)$$

$$\frac{\Delta E}{\bar{E}} = \frac{1}{\sqrt{c_s f_s}}. \quad (4.130)$$

In the thermodynamic limit the relative energy fluctuations are negligible. So, we can assume $E = \bar{E}$ and associate statistical quantities with actual numbers, this way. This means that we can actually speak of the “energy in a canonical ensemble”. Then, microcanonical and canonical ensemble are equivalent!

Microcanonical	Canonical
E_{mic}	β
$Z_{\text{mic}}(E) = \Omega(E)$	$Z_{\text{can}}(\beta)$

Table 4.2.: Related quantities in microcanonical and canonical ensemble

4.2.5. Equivalence of microcanonical and canonical ensemble

Different quantities, characterizing the microcanonical or canonical ensemble, can be related, as shown in Table 4.2. Note that the energy E is fixed in this case.

They are related via

$$\bar{E} = -\frac{\partial}{\partial \beta} \ln Z_{\text{can}}(\beta), \quad (4.131)$$

$$\beta_{\text{mic}} = \frac{\partial}{\partial E} \ln Z_{\text{mic}}(E). \quad (4.132)$$

The question is, if we identify $E_{\text{mic}} = E_{\text{can}}$, can we also conclude $\beta_{\text{mic}} = \beta_{\text{can}}$, i. e. $T_s = T_R$?

4.2.6. Temperature

Microcanonical ensemble ($T = T_s$)

$$\beta_{\text{mic}} = \left. \frac{\partial \ln \Omega}{\partial E} \right|_{N,V} = \frac{1}{k_B T} \quad (4.133)$$

Here, T is an absolute temperature. So, its units are Kelvin or eV, if we set $k_B = 1$. For example $300 \text{ K} \approx \frac{1}{40} \text{ eV}$. Recall that T is an intensive quantity and that it can be interpreted as an energy per degree of freedom:

$$T = \beta^{-1} = \frac{E_s}{c_s f_s}. \quad (4.134)$$

In the case of an ideal gas we can compute T from the entropy

$$S = \frac{3N}{2} \ln \frac{2E}{3N} + f(V, N) \quad (4.135)$$

via

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{N,V} = \frac{3N}{2E}. \quad (4.136)$$

This means that

$$\boxed{\frac{E}{3N} = \frac{E}{f} = \frac{1}{2} T}, \quad (4.137)$$

and therefore the average (kinetic) energy per particle is

$$\boxed{\frac{E}{N} = \frac{3}{2} T}, \quad (4.138)$$

as each particle has 3 degrees of freedom.

Canonical ensemble ($T = T_R$)

T is the characteristic quantity for the canonical ensemble, i. e. it is a property of the total system. So to speak, we can “forget” the concept of the reservoir. T can be “measured” by measuring the mean energy $\langle E \rangle$ in the system.

As a next step, we will show that $T_R = T_S$.

4.2.7. Temperature and energy distribution in subsystems

Consider a system that is divided into two arbitrary sized subsystems s_1 and s_2 with energies E_1 and E_2 , respectively. Using the notation $E = E_1$ and $E_G = E_1 + E_2$, the probability to find the system in a state with a certain energy E is

$$w(E) = \Omega_1(E) \Omega_2(E_G - E) \Omega_G^{-1}(E_G). \quad (4.139)$$

As said before, $w(E)$ is a sharply peaked distribution, dominated by $E_{\max} = \bar{E}$. It is

$$0 = \left. \frac{\partial}{\partial E} \ln w \right|_{E=\bar{E}} \quad (4.140)$$

$$= \left. \frac{\partial}{\partial E} \ln \Omega_1(E) \right|_{E=\bar{E}} + \left. \frac{\partial}{\partial E} \ln \Omega_2(E_G - E) \right|_{E=\bar{E}} \quad (4.141)$$

$$= \beta_1 - \beta_2 \quad (4.142)$$

$$= \frac{1}{T_1} - \frac{1}{T_2}, \quad (4.143)$$

where we used the definition of β in the third equation. The minus sign in front of β_2 arises from the derivative of the argument $E_G - E$ of Ω_2 with respect to E . So, we find

$$\boxed{T_1 = T_2.} \quad (4.144)$$

We conclude, that we can define the temperature of any ensemble with small fluctuations of E around \bar{E} via

$$\boxed{\frac{1}{T} = \left. \frac{\partial}{\partial E} \ln \Omega(E) \right|_{E=\bar{E}}.} \quad (4.145)$$

But note, that in general the temperature is determined by the energy dependency of the partition sum.

4.2.8. Thermal equilibrium

0th law of thermodynamics

If two systems are in thermal equilibrium with a third system, they are also in mutual thermal equilibrium. There exists a variable characterising the equilibrium state, the temperature, that has the same value for the two systems.

Of course this is also valid for several systems in thermal equilibrium ($T_1 = T_2 = \dots = T_n$).

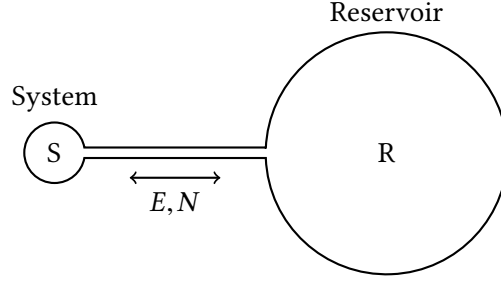


Figure 4.3.: System connected to a reservoir with energy and particle exchange

4.3. Grand canonical ensemble

The three different types of statistical ensembles are, next to the volume V , characterized by different quantities:

Microcanonical:	N, E
Canonical:	N, T
Grand canonical:	μ, T

The grand canonical ensemble describes an open system S in equilibrium that is able to exchange energy and particles with a reservoir R , as illustrated in Figure 4.3. For example one can consider a part of a gas volume.

We will treat it in the same manner as we treated the canonical ensemble. The probability for a certain state is now given by

$$p_n(E_n, N_n) = Z_{\text{gc}}^{-1} e^{-(\beta E_n + \alpha N_n)}, \quad (4.146)$$

where

$$\beta = \frac{\partial}{\partial E} \ln \Omega_R(E_G, N_G), \quad (4.147)$$

$$\alpha = \frac{\partial}{\partial N} \ln \Omega_R(E_G, N_G). \quad (4.148)$$

We define the chemical potential μ as

$$\mu = -\frac{\alpha}{\beta}. \quad (4.149)$$

Then, the probability takes the form

$$p_n(E_n, N_n) = Z_{\text{gc}}^{-1} e^{-\beta(E_n - \mu N_n)}, \quad (4.150)$$

the grand canonical partition function is given by

$$Z_{\text{gc}} = \sum_E \sum_N \Omega(E, N, V) e^{-\beta(E - \mu N)} \quad (4.151)$$

$$= \int dE \int dN \omega(E, N, V) e^{-\beta(E - \mu N)}, \quad (4.152)$$

and the partition sum is

$$\Omega(E, N, V) = \omega(E, N, V) \delta E \delta N. \quad (4.153)$$

The averaged quantities can again be extracted via derivatives of the logarithm of the partition function:

$$\bar{E} - \mu \bar{N} = -\frac{\partial}{\partial \beta} \ln Z_{\text{gc}}(\beta, \mu, V), \quad (4.154)$$

$$\bar{N} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z_{\text{gc}}(\beta, \mu, V). \quad (4.155)$$

4.4. Relations for entropy

4.4.1. Entropy in canonical ensemble

We can approximate the probability density $\omega(E)$ in the partition function

$$Z_{\text{can}} = \int dE \omega(E) e^{-\beta E} \quad (4.156)$$

by means of a saddle point expansion:

$$Z_{\text{can}} = \omega(\bar{E}) e^{-\beta \bar{E}} \underbrace{\int dE e^{-\frac{a}{2}(E-\bar{E})^2}}_{\sqrt{\frac{2\pi}{a}}} + \dots \quad (4.157)$$

where

$$a = \frac{cf}{\bar{E}^2}. \quad (4.158)$$

Then the partition function takes the form

$$Z_{\text{can}} = \Omega(\bar{E}) e^{-\beta \bar{E}} \underbrace{\frac{\bar{E}}{\delta E} \sqrt{\frac{2\pi}{cf}}}_{O(1)} \quad (4.159)$$

and its logarithm is, using $E = \bar{E}$,

$$\ln Z_{\text{can}} = \underbrace{\ln \Omega(E)}_{\sim f} - \underbrace{\beta E}_{\sim f} + \dots \quad (4.160)$$

Then, altogether, we find the entropy to be given by

$$S = \ln Z_{\text{can}} + \beta E = \left(1 - \beta \frac{\partial}{\partial \beta}\right) \ln Z_{\text{can}}. \quad (4.161)$$

Note that, as said before, the difference between the fixed energy E for the microcanonical and $E = \langle E \rangle = \bar{E}$ for the canonical ensemble is negligible. It vanishes in the thermodynamic limit.

4.4.2. Entropy in grand canonical ensemble

By the same reasoning as before we get

$$\ln Z_{\text{gc}} = \ln \Omega(E) - \beta(E - \mu N) \quad (4.162)$$

and therefore the entropy is

$$S = \ln Z_{\text{gc}} + \beta(E - \mu N) = \left(1 - \beta \frac{\partial}{\partial \beta}\right) \ln Z_{\text{gc}}. \quad (4.163)$$

4.4.3. General expression for entropy

In general, the entropy can be expressed in the following form:

$$S = - \sum_{\tau} p_{\tau} \ln p_{\tau}, \quad (4.164)$$

where we had $\tau \cong n$ in the previous cases. The validity of this expression can be checked easily:

In the microcanonical case it is

$$p_{\tau} = \frac{1}{\Omega}, \quad (4.165)$$

so we have

$$S = - \sum_{\tau} \frac{1}{\Omega} \ln \frac{1}{\Omega} \quad (4.166)$$

$$= \ln \Omega \left(\frac{1}{\Omega} \underbrace{\sum_{\tau} 1}_{\Omega} \right) \quad (4.167)$$

$$= \ln \Omega \quad \text{q. e. d.} \quad (4.168)$$

And in the canonical case it is

$$p_{\tau} = \frac{1}{Z_{\text{can}}} e^{-\beta E_{\tau}}, \quad (4.169)$$

$$\ln p_{\tau} = -\ln Z_{\text{can}} - \beta E_{\tau}, \quad (4.170)$$

so we have

$$S = - \sum_{\tau} p_{\tau} \ln p_{\tau} \quad (4.171)$$

$$= \sum_{\tau} p_{\tau} (\ln Z_{\text{can}} + \beta E_{\tau}) \quad (4.172)$$

$$= \ln Z_{\text{can}} + \beta E \quad \text{q. e. d.} \quad (4.173)$$

Similarly, it can be checked in the grand canonical case.

4.4.4. Entropy and density matrix

The entropy is related to the density matrix via

$$\boxed{S = \text{tr } \rho \ln \rho .} \quad (4.174)$$

This can also be considered as a general definition of the entropy, as it coincides with $S = \ln \Omega(E)$ for microcanonical, canonical and grand canonical ensembles.

Most importantly, this expression is valid in any basis. Therefore it also holds when diagonalizing ρ , which proves (4.164) more generally:

$$S = - \sum_{\tau} \rho_{\tau\tau} \ln \rho_{\tau\tau} \quad (4.175)$$

$$= - \sum_{\tau} p_{\tau} \ln p_{\tau} \quad \text{q. e. d.} \quad (4.176)$$

5. Thermodynamics

5.1. Thermodynamic potentials

5.1.1. State variables, equations of state

Classical thermodynamics describes equilibrium states and transitions from one equilibrium state to another. Historically it was developed before statistical physics, but it can be derived as an effective macroscopic theory from the microscopic statistical theory. Without any doubt it has an immense importance regarding technical aspects. It is mostly axiomatic, based on the laws of thermodynamics and their consequences. Classical thermodynamics only uses a few macroscopic variables (both extensive and intensive) such as E, T, S, N, μ, V, p . We distinguish state variables or independent thermodynamic variables, which characterise the ensemble, and dependent thermodynamic variables. For the different types of ensembles (systems) this is illustrated in Table 5.1

The thermodynamic limit provides a functional relationship between the state variables. Usually, no average values are used, because $\langle E \rangle \equiv E$, etc.

There are several practical question, which we will try to answer in this chapter:

- How much energy should be deposited in a system, so that its temperature rises by a certain amount, provided the volume stays constant?
- How much will the pressure increase during this process?
- How is “heat” transformed into mechanical energy? What is the efficiency of machines?
- When is a process reversible and when not?

The dependence of the dependent variables on the state variables is called *equation of state*.

Examples

1. Equation of state for an ideal gas:

$$pV = RTn_M = k_B TN \quad (5.1)$$

n_M : number of moles

R : gas constant ($R = 8.31 \frac{\text{J}}{\text{mol K}}$)

This follows from

$$\left. \frac{\partial S}{\partial V} \right|_{E, N} = \frac{p}{T} . \quad (5.2)$$

	State variables	Dependent variables
microcanonical ensemble (isolated system)	$\ln Z_{\text{mic}}(E, N, V)$	S, T, μ, p
canonical ensemble (closed system)	$\ln Z_{\text{can}}(T, N, V)$	S, E, μ, p
grand canonical ensemble (open system)	$\ln Z_{\text{gc}}(T, \mu, V)$	S, E, N, p

Table 5.1.: Different kinds of ensembles (systems) and their corresponding state and dependent variables

It can be concluded that

$$p = Tn . \quad (5.3)$$

2. Van der Waals equation of state for a real gas:

$$(p + an^2) = \frac{nT}{1 - bn} \quad (5.4)$$

a, b : constants specific to the gas

The task of the statistical physics is to explain the equations of state and to derive their functional form. On the other hand, classical thermodynamics examines the consequences of their existence. There are seven variables and only 4 equations, which automatically makes three of the variables independent (state variables), but there also are additional thermodynamic conditions. As already mentioned, variables split into two types: extensive (E, N, V, S) and intensive (T, μ, p). However, the choice of the state variables cannot be arbitrary – at least one of them has to be extensive.

Example

- Grand canonical ensemble with state variables T, μ, V :

$$S = Vf_S(T, \mu) \quad (5.5)$$

$$E = Vf_E(T, \mu) \quad (5.6)$$

$$N = Vf_N(T, \mu) \quad (5.7)$$

$$p = f_p(T, \mu) \quad (5.8)$$

Therefore, (p, T, μ) cannot be used as three independent variables at the same time!

If the particle number is not conserved (e. g. photon), there are only two state variables, as $\mu = 0$. Then,

$$p = f_p(T) . \quad (5.9)$$

So, there is a relation between pressure and temperature. Defining $\epsilon = \frac{E}{V}$, we find

$$p = \tilde{f}_p(\epsilon) , \quad (5.10)$$

and in the special case of photons it is

$$p = \frac{1}{3}\epsilon . \quad (5.11)$$

5.1.2. Thermodynamic potentials and partition functions

$$S = S(E, N, V) = k_B \ln Z_{\text{mic}}(E, N, V) \quad (\text{entropy}) \quad (5.12)$$

$$F = F(T, N, V) = -k_B T \ln Z_{\text{can}}(T, N, V) \quad (\text{free energy}) \quad (5.13)$$

$$J = J(T, \mu, V) = -k_B T \ln Z_{\text{gc}}(T, \mu, V) \quad (\text{Gibbs potential}) \quad (5.14)$$

$$\left(\frac{\partial S}{\partial E} \right)_{N, V} = \frac{1}{T} \quad (5.15)$$

$$\left(\frac{\partial S}{\partial N} \right)_{E, V} = -\frac{\mu}{T} \quad (5.16)$$

$$\left(\frac{\partial S}{\partial V} \right)_{E, N} = \frac{p}{T} \quad (5.17)$$

At this point, these are just definitions. Their physical meaning will be explained later. Note that, traditionally, in thermodynamics one denotes the fixed quantities explicitly. These 3 relations together with $S(E, N, V)$ lead to 4 relations for the 7 variables in total.

We now use

$$S = \ln Z_{\text{can}} + \beta E \quad (5.18)$$

and $\beta = \frac{1}{T}$ to find

$$-\frac{F}{T} = \ln Z_{\text{can}} = S - \frac{E}{T}, \quad (5.19)$$

and therefore

$$\boxed{F = E - TS.} \quad (5.20)$$

On the other hand, using

$$S = \ln Z_{\text{gc}} + \beta(E - \mu N), \quad (5.21)$$

we have

$$-\frac{J}{T} = \ln Z_{\text{gc}} = S - \frac{E - \mu N}{T}, \quad (5.22)$$

and therefore

$$J = E - \mu N - TS \quad (5.23)$$

or

$$\boxed{J = F - \mu N.} \quad (5.24)$$

5.1.3. Differential relations

Recall the following definitions:

$$\frac{1}{T} = \left. \frac{\partial S}{\partial E} \right|_{N,V}, \quad (5.25)$$

$$-\frac{\mu}{T} = \left. \frac{\partial S}{\partial N} \right|_{E,V}, \quad (5.26)$$

$$\frac{p}{T} = \left. \frac{\partial S}{\partial V} \right|_{E,N}. \quad (5.27)$$

With these, one arrives at the following expressions for the differentials of S :

$$\boxed{dS = \frac{1}{T}dE - \frac{\mu}{T}dN + \frac{p}{T}dV,} \quad (5.28)$$

and of E :

$$\boxed{dE = \underbrace{TdS}_{\substack{\text{heat} \\ \delta Q}} + \underbrace{\mu dN}_{\substack{\text{added particles} \\ \delta E_N}} - \underbrace{pdV}_{\substack{\text{work} \\ \delta A}}.} \quad (5.29)$$

1st law of thermodynamics: “conservation of energy” (Helmholtz, Mayer)

A thermodynamic system can be characterized by its energy. The energy can be changed by “adding” heat, work or matter. For an isolated system the energy is conserved.

It is very important to note that heat, work and added particles are *not* properties of an equilibrium system! Statements like “Heat is added / lost” make no sense! But these quantities may describe differences between two equilibrium systems.

Heat

For example, the heat is defined as

$$\delta Q = E_2 - E_1 \quad \text{for fixed } V, N. \quad (5.30)$$

Therefore it accounts for a change in temperature:

$$\delta Q = E(T_2, V, N) - E(T_1, V, N). \quad (5.31)$$

Again, note that this is no property of a system. Neither is it a thermodynamic variable. But it can be used as a thermodynamic definition of the entropy:

$$\delta Q = TdS. \quad (5.32)$$

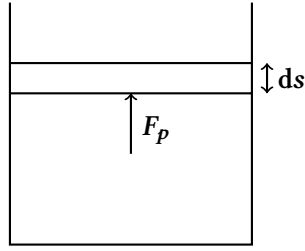


Figure 5.1.: Mechanical work

Mechanical work

A definition of mechanical work can be given via a slow change of volume at fixed S, N (compare Figure 5.1):

$$\delta A = -F_p \cdot ds \quad (5.33)$$

$$= -|F_p| ds \quad (5.34)$$

$$= -p a ds \quad (5.35)$$

$$= -p dV, \quad (5.36)$$

where p is the pressure and a the area.

We can identify different signs of δA with different physical situations:

$$\delta A < 0 : \quad \text{system "does" work,} \quad (5.37)$$

$$\delta A > 0 : \quad \text{systems absorbs mechanical energy.} \quad (5.38)$$

In general, δA describes the change in energy by work at fixed S, N . For example, it also describes the change of an external magnetic field instead of the change of volume.

We note that the statistical definition of pressure coincides with the mechanical definition “force per area”, but the statistical one is more general!

5.1.4. Legendre transform

The change from S to F at fixed V, N is one example of a Legendre transform. In this case, the conjugate variables are E and $\beta = \frac{1}{T}$. To be formally more precise, its actually a transform between the following quantities:

$$S(E) \leftrightarrow \beta F(\beta) = \tilde{F}(\beta). \quad (5.39)$$

Let us look at this in more detail. The new variable is β :

$$\frac{\partial S}{\partial E} = \beta. \quad (5.40)$$

Then, the Legendre transform of S is given as

$$\tilde{F} = \beta E - S, \quad (5.41)$$

and therefore

$$F = E - TS . \quad (5.42)$$

The derivative of \tilde{F} with respect to β is

$$\frac{\partial \tilde{F}}{\partial \beta} = E + \beta \frac{\partial E}{\partial \beta} - \frac{\partial S}{\partial \beta} \quad (5.43)$$

$$= E + \beta \frac{\partial E}{\partial \beta} - \underbrace{\frac{\partial S}{\partial E}}_{\beta} \frac{\partial E}{\partial \beta} \quad (5.44)$$

$$= E , \quad (5.45)$$

the original variable, as it should be.

On the other hand, we find

$$\left. \frac{\partial F}{\partial T} \right|_{N,V} = -S , \quad (5.46)$$

using (5.20). And therefore

$$\boxed{dF = -SdT + \mu dN - pdV .} \quad (5.47)$$

Similarly, it is

$$\boxed{dJ = -SdT - Nd\mu - pdV ,} \quad (5.48)$$

which can be concluded, using

$$\frac{\partial}{\partial \mu} \ln Z_{\text{gc}} = \beta N , \quad (5.49)$$

$$-\beta J = \ln Z_{\text{gc}} , \quad (5.50)$$

so,

$$\frac{\partial J}{\partial \mu} = -N , \quad (5.51)$$

as well as

$$S = \left(1 - \beta \frac{\partial}{\partial \beta} \right) \ln Z_{\text{gc}} \quad (5.52)$$

$$= \left(1 - \beta \frac{\partial}{\partial \beta} \right) (-\beta J) \quad (5.53)$$

$$= -\beta J + \beta J + \beta^2 \frac{\partial J}{\partial \beta} \quad (5.54)$$

$$= -\frac{\partial J}{\partial T} . \quad (5.55)$$

Now we have everything at hand to construct thermodynamics:

1. As the basis: $S = \ln \Omega$, the central probabilistic quantity
(equivalent to $S = -\sum_{\tau} p_{\tau} \ln p_{\tau} = -\text{tr}(\rho \ln \rho)$)
2. Definitions of T, μ, p as derivatives of S :

$$dS = \frac{1}{T}dE - \frac{\mu}{T}dN + \frac{p}{T}dV \quad (5.56)$$

3. Computation of the partition function for the given system $S(E, N, V)$, or equivalently $F(T, N, V)$ or $J(T, \mu, V)$
(aim of statistical physics, “learn counting”)

Note that no other than these 3 axioms are needed for thermodynamics! And all fundamental laws of thermodynamics can be derived from these! We already have shown the latter for the 0th and the 1st law.

5.1.5. Properties of entropy

1) Additivity

Consider two systems with energies E_1 and E_2 (neglect again H_{int}), then

$$\Omega_G = \Omega_1 \cdot \Omega_2, \quad (5.57)$$

$$\ln \Omega_G = \ln \Omega_1 + \ln \Omega_2, \quad (5.58)$$

$$S_G = S_1 + S_2. \quad (5.59)$$

2) Maximum principle for entropy

We permit energy exchange between these two systems, with the constraint

$$E_1 + E_2 = E_G. \quad (5.60)$$

We define $E_1 = E$ and want to find the condition for the maximum of the probability distribution $W(E)$, which describes the equilibrium state:

$$w(E) = \frac{\Omega_1(E_1) \Omega_2(E_2)}{\Omega_G(E_G)} \Big|_{E_1+E_2=E_G}. \quad (5.61)$$

From the considerations in 1) we can conclude that a maximum of $w(E)$ corresponds to a maximum of $\ln \Omega_1 + \ln \Omega_2$, which corresponds to a maximum of $S_1 + S_2$ and therefore to a maximum of S_G . This means that in equilibrium the system chooses E_1 such that the total entropy is maximal, with the constraint $E_1 + E_2 = E_G$.

In equilibrium entropy is maximal, subject to constraints.

Of course, the individual energies E_1, E_2 can be computed from the maximum of $S_1 + S_2$ with the condition $E_1 + E_2 = E_G$.

Another property of the entropy can be inferred from this. Let us start with two isolated systems in equilibrium, with arbitrary energies $E_1^{(0)}, E_2^{(0)}$ and $E_G = E_1^{(0)} + E_2^{(0)}$. They have entropies $S_1^{(0)}, S_2^{(0)}$. Then we bring them into thermal contact (i. e. we allow energy exchange). This leads to an energy flow $E_1^{(0)} \rightarrow E_1, E_2^{(0)} \rightarrow E_2$ until $S_1 + S_2$ is maximal for a given E_G . This means that

$$\boxed{S_1 + S_2 \geq S_1^{(0)} + S_2^{(0)}} \quad (5.62)$$

Entropy can only increase!

This is a key ingredient of the 2nd law of thermodynamics!

5.1.6. Zero of absolute temperature

Approaching its ground state, the partition sum of a system will approach a value of the order of 1:

$$\Omega \xrightarrow{E \rightarrow E_0} \Omega(E_0) = \Omega_0 = \mathcal{O}(1), \quad (5.63)$$

and therefore the entropy will approach 0:

$$S \xrightarrow{E \rightarrow E_0} 0 \quad (5.64)$$

(or $S \rightarrow S_0, \frac{S_0}{V} \rightarrow 0$).

Using

$$\left. \frac{\partial \ln \Omega}{\partial E} \right|_{E_0} = \frac{1}{\Omega_0} \frac{\Delta \Omega}{\Delta E}, \quad (5.65)$$

and therefore, by the definition of the absolute temperature,

$$T = \left(\frac{\partial \ln \Omega}{\partial E} \right)^{-1} \approx \Omega_0 \frac{\Delta E}{\Delta \Omega}, \quad (5.66)$$

we find

$$\frac{\Delta \Omega}{\Delta E} \xrightarrow{E \rightarrow E_0} \infty, \quad (5.67)$$

as, choosing $E_0 \approx 0$,

$$\Delta E \approx E \sim f, \quad (5.68)$$

$$\Delta \Omega \approx e^f \gg f. \quad (5.69)$$

This means that in the thermodynamic limit the ground state temperature approaches 0:

$$T \sim f e^{-f} \xrightarrow{f \rightarrow \infty} 0, \quad (5.70)$$

or, in other words,

$$T(E = E_0) = 0. \quad (5.71)$$

The properties (5.63) and (5.67) of $\Omega(E)$ are illustrated in Figure 5.2.

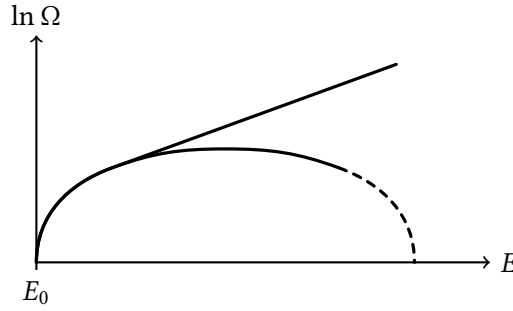
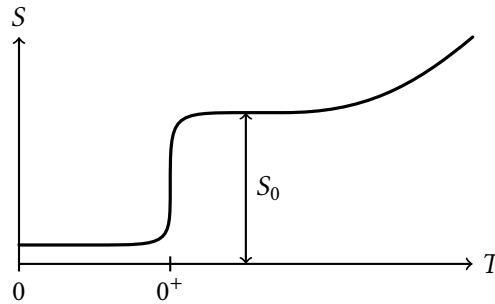


Figure 5.2.: Behaviour of the partition sum at zero temperature

Figure 5.3.: Entropy for $T \rightarrow 0^+$

5.1.7. 3rd law of thermodynamics

3rd law of thermodynamics (empirical by Planck (Nernst))

For $T \rightarrow 0^+$ the entropy obeys $S \rightarrow S_0$, where S_0 is independent of the thermodynamic variables of the system.

So, the entropy for $T \rightarrow 0^+$ will behave as sketched in Figure 5.3. An example for this would be the hyper-fine splitting in atoms.

5.2. Reversible and irreversible processes

5.2.1. Subsystems and constraints

We consider an *isolated* system in equilibrium with fixed energy E (or interval $[E, E + \delta E]$, respectively). Furthermore we recall that $\Omega = Z_{\text{mic}}$ is the number of (micro-)states with energy E , which the system can occupy, and that the entropy is defined as $S = k_B \ln \Omega$.

In general, the system is determined by some constraints. These correspond to particular parameters:

$$\Omega = \Omega(E, y_1, \dots, y_k) \quad (5.72)$$

(or intervals $[y_1, y_1 + \delta y_1], \dots, [y_k, y_k + \delta y_k]$).

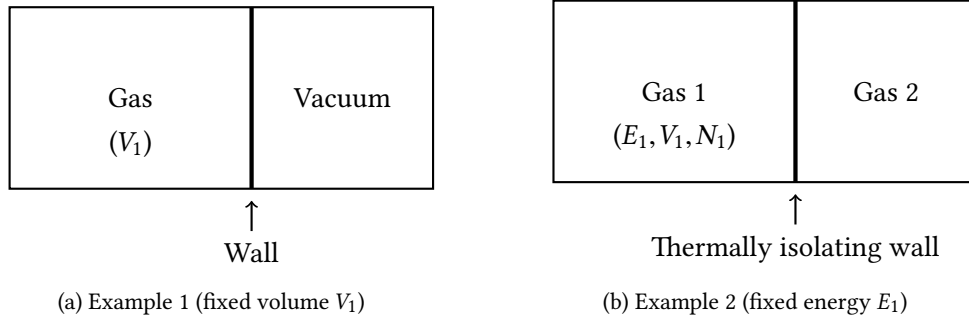


Figure 5.4.: Examples of constraints

Examples

1. Fixed volume V_1 (Figure 5.4a):
 V_1 : volume occupied by the gas
2. Fixed energy of a subsystem E_1 (Figure 5.4b):
 E_1 : energy of gas 1
 V_1 : volume of gas 1
 N_1 : particle number of gas 1

This example can be considered as two isolated subsystems:

$$\Omega_G(E, E_1, V, V_1, N, N_1) = \Omega(E_1, V_1, N_1) \cdot \Omega(E - E_1, V - V_1, N - N_1). \quad (5.73)$$

5.2.2. Removing the constraint

After sufficient time the system will reach a new equilibrium state:

$$\Omega_i \rightarrow \Omega_f. \quad (5.74)$$

Thereby, one always has

$$\boxed{\Omega_f \geq \Omega_i.} \quad (5.75)$$

All the states compatible with the constraint are still possible. But additionally, new states with fixed energy E become possible, which previously have been excluded by the constraint.

In general, removing the constraint in an isolated system leads to a final state whose entropy is larger than or equal to the entropy of the initial state:

$$S_f \geq S_i. \quad (5.76)$$

(The final and initial states are equilibrium states. It is not important what happens in between.)

Examples

1. Removing the wall:

$$\Omega_i \sim V_1^N \quad (5.77)$$

$$\Omega_f \sim V^N \quad (5.78)$$

$$\Omega_f \gg \Omega_i \quad \text{for } V > V_1 \quad (5.79)$$

2. Wall becomes thermally conductive (subsystem is a canonical ensemble):

$$\Omega_i = \Omega_{i1} \Omega_{i2} \quad (5.80)$$

$$\Omega_{i1} \approx \frac{1}{N_1! \left(\frac{3N_1}{2}\right)!} \left(\frac{m}{2\pi\hbar^2}\right)^{3N_1/2} V_1^{3N_1/2} E_1^{3N_1/2} \frac{\delta E_1}{E_1} \quad (\text{diluted gas at high energy}) \quad (5.81)$$

$$\Omega_f \approx \Omega_i \Big|_{E_i \rightarrow \bar{E}_i} \quad (\Omega_i \sim w(E_1), \Omega_f \sim w(\bar{E}_1)) \quad (5.82)$$

$$\Omega_f \gg \Omega_i \quad \text{for } E_1 \neq E_{1\max} = \bar{E}_1 \quad (\text{recall Figure 4.2}) \quad (5.83)$$

$$\Omega_f \approx \Omega_i \quad \text{for } T_1 = T_2 \quad (5.84)$$

5.2.3. Increase of entropy by removing the constraints

The probability that the system without constraints, after sufficiently long time, still stays in its initial state is

$$w_i = \frac{\Omega_i}{\Omega_f}. \quad (5.85)$$

That is because $\frac{1}{\Omega_f}$ is the probability that the system is in *one* particular macrostate (of the system without constraints) and Ω_i is the number of microstates satisfying the initial conditions.

Of course,

$$w_i \ll 1 \quad \text{if } \Omega_i \ll \Omega_f. \quad (5.86)$$

Examples

1. Probability that the gas still occupies a volume V_1 after removing the wall for $V = 2V_1$:

$$w_i = \left(\frac{1}{2}\right)^N \quad (5.87)$$

(factor $\frac{1}{2}$ for each molecule)

2. Probability that the energy of gas 1 remains E_1 after the wall becomes thermally conductive:

$$w_i = w(E_1) \delta E \quad (5.88)$$

For the interval $[E_{\max} - \frac{1}{2}\delta E, E_{\max} + \frac{1}{2}\delta E]$ with $\frac{\delta E}{E} \geq \frac{1}{\sqrt{N}}$ it is

$$w(E_{\max}) \delta E = w(\bar{E}) \delta E \approx 1. \quad (5.89)$$

For $E_1 \neq \bar{E}$ we find $w(E_1)$ to be exponentially suppressed.

Let us look at the effect of removing the constraint on a general variable y (e. g. $y = E_1$, the energy in the subvolume V_1). Beforehand, it has a fixed initial value y_i . Afterwards, it satisfies a probability distribution characterized by

$$w(y)\delta y \sim \Omega(y) , \quad (5.90)$$

where, in general, $\bar{y} \neq y_i$, with \bar{y} being the position of the maximum of $w(y)$, $\frac{\partial w}{\partial y} \bar{y} = 0$. (The generalization to multiple variables y_1, \dots, y_k is straightforward.)

Now we consider the limit

$$N \rightarrow \infty . \quad (5.91)$$

This will lead to a sharp value $y_f = \bar{y} = y_{\max}$ (again, generally $y_f \neq y_i$). y_f is determined by the entropy becoming maximal ($\Omega(y_{\max})$):

$$\boxed{S(y_f) \text{ maximal.}} \quad (5.92)$$

Without constraints the parameters will take such values that the entropy becomes maximal (at given boundary conditions like V, N, E of the system).

Example

- y is the energy in a subvolume:

Beforehand, $S(E, N, V; y)$ is calculated with given constraints. Afterwards, we have to satisfy $\left. \frac{\partial S}{\partial y} \right|_{E, N, V} = 0$ to find y_{\max} .

The entropy of a system grows when approaching the equilibrium state.

Note that this is a direct consequence of our two postulates! (It is not necessarily valid if one of the basic postulates is not obeyed!)

Initial states with the probability distribution not being a uniform distribution over all possible microstates (equilibrium distribution) are very unlikely for large N :

$$w_i = \frac{\Omega_i}{\Omega_f} \ll 1 , \quad (5.93)$$

$$S_i < S_f . \quad (5.94)$$

So, from the equilibrium state point of view approaching the equilibrium state corresponds to the transition from an unlikely to a likely state (content of the first postulate).

5.2.4. 2nd law of thermodynamics

The entropy of an isolated system cannot decrease.

Historical formulations:

Clausius:

Heat cannot flow from lower to higher temperature.

Kelvin:

It is impossible to produce work continuously by lowering the temperature of a system, without other changes in the environment.

In other words, the latter formulation states the impossibility of perpetual motion machines.

Short summary of the laws of thermodynamics

0th law: $T, T_1 = T_2$

1st law: E , conservation of E

2nd law: S , S cannot decrease

3rd law: $S_{\min} = 0$

5.2.5. Reversible processes for isolated systems

A process (i) \rightarrow (f) in an isolated system is reversible, if the process (f) \rightarrow (i) is possible, too. Otherwise, the process is irreversible.

In an isolated system an increase of entropy indicates an irreversible process. On the other hand, a reversible process is, concerning the isolated total system, adiabatic (the entropy of the total system stays constant).

5.2.6. Reversible processes for open systems

It must be possible, after performing some process, to perform its reversed process, so that the system gets back into its initial state, without changing the environment. The Carnot cycle is an example of such a reversible process (if it is quasi-static). The temperature difference between the reservoirs is lowered and work is performed. Its reversed process describes a fridge: work is performed in order to enlarge the temperature difference between the reservoirs. The $p - V$ diagram of the Carnot cycle is shown in Figure 5.5.

Irreversible processes, on the other hand, are dissipative processes. Typically, in such a process the energy is being distributed over several degrees of freedom (conversion of “mechanical energy” into heat). Examples are mechanical friction, eddy currents, viscose friction and balancing the temperature of two heat reservoirs.

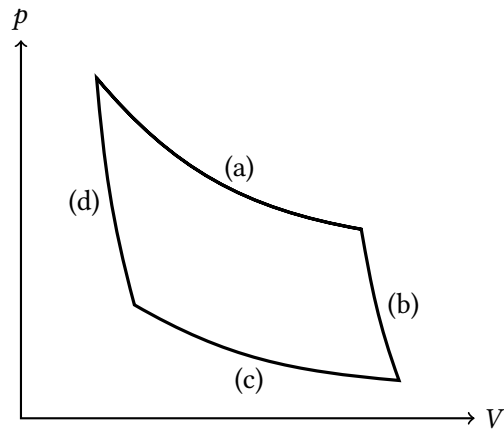


Figure 5.5.: Carnot cycle

5.2.7. Thermodynamics and non-equilibrium states

The course of an irreversible process cannot be described by thermodynamic potentials. But its final and initial states can be described, if they are equilibrium states.

Part II.

Statistical Systems

6. Ideal quantum gas

In an ideal gas the interactions between its particles can be neglected. This can be used as an approximation for several kinds of physical systems like:

- molecules
- photons (radiation)
- electrons
- quasi-particles in solids (phonons, magnetons ...)

It provides a good approximation especially at small densities, but also for many other problems, e. g. electrons in solids.

6.1. Occupation numbers and functional integral

We consider some one-particle states ν . In the case of momentum states, we have $\nu = \mathbf{p} = (p_1, p_2, p_3)$. To each of these states we can assign an energy $E_\nu = E(\mathbf{p})$ and an occupation number $n_\nu = n(\mathbf{p})$. A given sequence of occupation numbers $\{n_\nu\} = \{n(\mathbf{p})\}$ corresponds to a given function $n(\mathbf{p})$. Depending on the type of particles we consider, the occupation numbers can attain different value ranges:

$$\text{fermions:} \quad n_\nu = 0, 1 \quad (6.1)$$

$$\text{bosons:} \quad n_\nu = 0, 1, 2, \dots, \infty \quad (6.2)$$

A microstate τ of the full system is given by the sequence of occupation numbers:

$$\tau \cong \{n_\nu\}. \quad (6.3)$$

This can be used to express the sum over all microstates via the one-particle states:

$$\sum_{\tau} \cong \sum_{\{n_\nu\}} \quad (6.4)$$

$$= \sum_{n_1} \sum_{n_2} \cdots \sum_{n_k} \cdots \quad (6.5)$$

$$= \prod_{\nu} \sum_{n_\nu}. \quad (6.6)$$

This way, every possible sequence appears exactly once! For the different types of particles this yields:

$$\text{fermions:} \quad \sum_{\{n_\nu\}} = \prod_\nu \left(\sum_{n_\nu=0}^1 \right) \quad (6.7)$$

$$\text{bosons:} \quad \sum_{\{n_\nu\}} = \prod_\nu \left(\sum_{n_\nu=0}^{\infty} \right) \quad (6.8)$$

As said before, we can relate a sequence of occupation numbers $\{n_\nu\}$ to a function $n(\mathbf{p})$. This means that we also can identify the sum over all possible sequences with a sum over all possible functions $n(\mathbf{p})$:

$$\sum_{\{n_\nu\}} \cong \int \mathcal{D} n(\mathbf{p}). \quad (6.9)$$

This is a functional integral!

6.2. Partition function for the grand canonical ensemble

Using the results from the previous section, we can express the grand canonical partition function as

$$Z_{\text{gc}} = \sum_{\tau} e^{-\beta (E_\tau - \mu N_\tau)} \quad (6.10)$$

$$= \sum_{\{n_\nu\}} e^{-\beta (\sum_\nu E_\nu n_\nu - \mu \sum_\nu n_\nu)} \quad (6.11)$$

$$= \sum_{\{n_\nu\}} e^{-\beta \sum_\nu n_\nu (E_\nu - \mu)} \quad (6.12)$$

$$= \sum_{\{n_\nu\}} \prod_\nu e^{-\beta n_\nu (E_\nu - \mu)} \quad (6.13)$$

$$= \left(\prod_{\nu'} \sum_{n_{\nu'}} \right) \left(\prod_\nu e^{-\beta n_\nu (E_\nu - \mu)} \right) \quad (6.14)$$

$$= \prod_\nu \left(\sum_{n_\nu} e^{-\beta n_\nu (E_\nu - \mu)} \right). \quad (6.15)$$

The last step may seem confusing, but by explicitly writing it out, it is easy to see that every summand in (6.14) appears exactly once and with the same weight in (6.15), too.

The logarithm of the partition function is given by

$$\ln Z_{\text{gc}} = \sum_\nu \ln \left(\sum_{n_\nu} e^{-\beta n_\nu (E_\nu - \mu)} \right). \quad (6.16)$$

For fermions this yields

$$\boxed{\ln Z_{\text{gc}}^{(\text{F})} = \sum_\nu \ln \left(1 + e^{-\beta (E_\nu - \mu)} \right)}, \quad (6.17)$$

whereas for bosons we find ($j = n_\nu$)

$$\ln Z_{\text{gc}}^{(\text{B})} = \sum_\nu \ln \left[\sum_{j=0}^{\infty} \left(e^{-\beta(E_\nu - \mu)} \right)^j \right]. \quad (6.18)$$

Applying the formula for the geometric series:

$$1 + x + x^2 + x^3 + \dots = \frac{1}{1 - x}, \quad (6.19)$$

we arrive at

$$\boxed{\ln Z_{\text{gc}}^{(\text{B})} = \sum_\nu \ln \frac{1}{1 - e^{-\beta(E_\nu - \mu)}}.} \quad (6.20)$$

Now we only have to carry out sums over *one-particle states*! This is an enormous simplification as it is much easier to perform \sum_ν rather than $\sum_\tau = \sum_{\{n_\nu\}}$! But note that this just works for free particles without any interactions.

6.3. Mean occupation numbers

6.3.1. Bosons

In general it is

$$\bar{N} = \sum_\tau N_\tau p_\tau \quad (6.21)$$

$$= \sum_\tau \frac{N_\tau e^{-\beta(E_\tau - \mu N_\tau)}}{Z_{\text{gc}}} \quad (6.22)$$

$$= \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z_{\text{gc}}. \quad (6.23)$$

So, using (6.20), for bosons we find

$$\bar{N} = \sum_\nu \left[-\frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \left(1 - e^{-\beta(E_\nu - \mu)} \right) \right] \quad (6.24)$$

$$= \sum_\nu \frac{1}{\beta} \frac{\frac{\partial}{\partial \mu} e^{-\beta(E_\nu - \mu)}}{1 - e^{-\beta(E_\nu - \mu)}} \quad (6.25)$$

$$= \sum_\nu \underbrace{\frac{e^{-\beta(E_\nu - \mu)}}{1 - e^{-\beta(E_\nu - \mu)}}}_{\bar{n}_\nu^{(\text{B})}}, \quad (6.26)$$

and therefore

$$\boxed{\bar{N} = \sum_\nu \bar{n}_\nu^{(\text{B})},} \quad (6.27)$$

6. Ideal quantum gas

with $\bar{n}_v^{(B)}$ being the mean occupation number of the state v :

$$\bar{n}_v^{(B)} = \frac{1}{e^{\beta(E_v - \mu)} - 1} . \quad (\text{Bose-Einstein statistics}) \quad (6.28)$$

If we also allow for state dependent μ_v , this generalizes to

$$Z(\beta; \mu_1, \dots, \mu_v) = \sum_{n_1} \dots \sum_{n_v} e^{-\beta \sum_v (E_v n_v - \mu_v n_v)} \quad (6.29)$$

and

$$\bar{n}_v = \frac{1}{\beta} \frac{\partial}{\partial \mu_v} \ln Z \Big|_{\mu_v = \mu} . \quad (6.30)$$

For bosons this yields

$$\ln Z(\beta; \mu_v) = \sum_v \ln \frac{1}{1 - e^{-\beta(E_v - \mu_v)}} \quad (6.31)$$

and $\bar{n}_v^{(B)}$ as before.

6.3.2. Fermions

As for bosons, the mean occupation number of fermions is given by the sum of the mean one-particle occupation numbers:

$$\bar{N} = \sum_v \bar{n}_v^{(F)} , \quad (6.32)$$

where in this case

$$\bar{n}_v^{(F)} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln \left(1 + e^{-\beta(E_v - \mu)} \right) \quad (6.33)$$

$$= \frac{e^{-\beta(E_v - \mu)}}{1 + e^{-\beta(E_v - \mu)}} . \quad (6.34)$$

And therefore

$$\bar{n}_v^{(F)} = \frac{1}{e^{\beta(E_v - \mu)} + 1} . \quad (\text{Fermi-Dirac-statistics}) \quad (6.35)$$

From this it is easy to see that $\bar{n}_v^{(F)}$ only can take values in the following range:

$$0 \leq \bar{n}_v^{(F)} \leq 1 . \quad (6.36)$$

Moreover, we can look at some special cases:

$$\bar{n}_v^{(F)} \rightarrow 0 \quad \text{for} \quad E_v - \mu \gg T , \quad (6.37)$$

$$\bar{n}_v^{(F)} \rightarrow 1 \quad \text{for} \quad E_v - \mu < 0 \quad \text{and} \quad |E_v - \mu| \gg T . \quad (6.38)$$

6.3.3. Mean energy

In general, the mean energy is computed as

$$\bar{E} = \left(-\frac{\partial}{\partial \beta} + \frac{\mu}{\beta} \frac{\partial}{\partial \mu} \right) \ln Z_{\text{gc}} \quad (6.39)$$

$$= -\frac{\partial}{\partial \beta} \ln Z_{\text{gc}} + \mu \bar{N} . \quad (6.40)$$

For bosons it is

$$\bar{E} = \sum_v \frac{(E_v - \mu) e^{-\beta(E_v - \mu)}}{1 - e^{-\beta(E_v - \mu)}} + \mu \bar{N} \quad (6.41)$$

$$= \sum_v (E_v - \mu) \bar{n}_v^{(\text{B})} + \mu \sum_v \bar{n}_v^{(\text{B})} \quad (6.42)$$

$$= \sum_v E_v \bar{n}_v^{(\text{B})} . \quad (6.43)$$

Performing the same calculation for fermions, we find that both are given by

$$\boxed{\bar{E} = \sum_v E_v \bar{n}_v} . \quad (6.44)$$

Similar formulas can be found for other macroscopic quantities, too! Note that \bar{n}_v itself already is a macroscopic quantity!

6.3.4. Chemical potential

One question we have not discussed yet is: Which value should we choose for μ ?

1) Conserved number of particles

Examples for conserved numbers of particles are:

N_e : electric charge and lepton number

N_i : number of molecules of a certain type, if chemical transitions are negligible

N_B : baryon number (in the early universe)

In general,

$$\bar{N} = \frac{1}{\beta} \frac{\partial}{\partial \mu} \ln Z_{\text{gc}}(\beta, \mu) = \bar{N}(\beta, \mu) , \quad (6.45)$$

so we conclude

$$\mu = \mu(\bar{N}, T) . \quad (6.46)$$

In the thermodynamic limit the grand canonical and the canonical ensemble are equivalent, which leads to

$$\mu = \mu(N, T) . \quad (6.47)$$

Note that, if there are several conserved quantum numbers N_i , there also are several independent μ_i .

2) No conserved number of particles

Examples for cases where the number of particles is not conserved are photons and protons. In these cases the canonical ensemble does not depend on N , i. e.

$$\mu \equiv 0. \quad (6.48)$$

But there is no restriction of the summands in $\sum_{\{n_\nu\}}$, anyway, as N is no conserved quantum number. Here, the mean occupation numbers characterize the canonical ensemble.

3) Approximatively conserved quantum numbers

If one considers time scales, which are much shorter than the decay time of the particles, then one can consider the corresponding particle number as approximatively conserved and assume

$$\mu \neq 0. \quad (6.49)$$

6.3.5. Boltzmann statistics

In the regime of large energies (compared to T):

$$E_\nu - \mu \gg T \quad (6.50)$$

the mean occupation number of a state ν is approximately

$$\boxed{\bar{n}_\nu \approx e^{-\beta(E_\nu - \mu)}}, \quad (\text{Maxwell-Boltzmann statistics for "classical particles"}) \quad (6.51)$$

which holds for bosons and fermions, equally. This recovery of the “classical” description in the large energy limit is an example of the correspondence principle between QM and classical mechanics.

We also can write this as

$$\bar{n}_\nu = C e^{-\frac{E_\nu - \mu}{T}} \quad (6.52)$$

$$= \tilde{C} e^{-\frac{E_\nu}{T}} \quad (6.53)$$

to emphasize the famous Boltzmann factor $e^{-\frac{E_\nu}{T}}$.

Note that E_ν is the energy of the one-particle state ν – it is not to be mistaken for the energy E_τ of a microstate τ like, for example, in $p_\tau = Z_{\text{can}}^{-1} e^{-\beta E_\tau}$.

As a comparison recall the computation of $\Omega(E, N)$ earlier in this lecture, with the approximation that Ω is dominated by microstates where each ν is occupied at most once, which corresponds to

$$\bar{n}_\nu \ll 1. \quad (6.54)$$

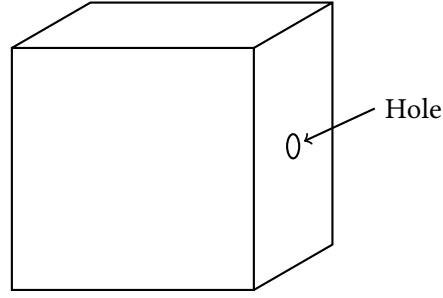


Figure 6.1.: Hole in a box

6.4. Photon gas

Think of a box with a small hole in it, like in Figure 6.1. The electromagnetic radiation within the box will be in equilibrium, with the temperature T given by the temperature of the walls. The hole is needed for observation. If there is no radiation from the outside entering through the hole, this system is an example of a “black body”.

The concept of a black body was crucial in the development of modern physics, as classical field theory yields wrong results (large deviations from the measured frequency distribution of the radiation). Historically, two of the most important milestones in understanding this as a photon gas in thermal equilibrium have been the quantisation of electromagnetic energy by Planck in 1900 (\hbar) and Einstein's idea of quantising the electromagnetic radiation itself as photons in 1905 (explanation of the photoelectric effect).

From a modern statistical physics point of view, a photon gas is a gas of non-interacting particles whose particle number is not conserved ($\mu = 0$). So it is described by a canonical ensemble. Note that in the case of free photons $n(\mathbf{p}) = n_\nu$ is conserved for every \mathbf{p} (ν). Such an ensemble with $\mu_\nu \neq 0$ can be very different from black body radiation ($\mu_\nu = 0$), e. g. radio waves.

6.4.1. Black body radiation

Consider a gas of photons in thermal equilibrium with a temperature T . As $\mu = 0$, we immediately can conclude

$$\bar{n}_\nu = \frac{1}{e^{\beta E_\nu} - 1}. \quad (6.55)$$

So, to proceed, one-photon states ν and their energies E_ν are necessary.

To get them, we start with the Maxwell equation (wave equation for the electric field):

$$\frac{1}{c^2} \frac{\partial^2 \mathbf{E}}{\partial t^2} = \nabla^2 \mathbf{E}. \quad (6.56)$$

Assuming

$$\mathbf{E} = \mathbf{E}(\mathbf{r}) e^{-i\omega t}, \quad (6.57)$$

it takes the form

$$\nabla^2 \mathbf{E}(\mathbf{r}) + \frac{\omega^2}{c^2} \mathbf{E} = 0, \quad (6.58)$$

which is solved by

$$E(r) = A e^{i\mathbf{k}r} . \quad (6.59)$$

The exact form of boundary conditions is not important, so we choose periodic ones:

$$\mathbf{k} = \frac{2\pi}{L} \mathbf{n} \quad \text{with} \quad n_{x,y,z} \in \mathbb{Z} . \quad (6.60)$$

From this we can conclude the dispersion law

$$\mathbf{k}^2 = \frac{\omega^2}{c^2}, \quad \omega = c |\mathbf{k}| . \quad (6.61)$$

Now, using the fact that the diversion of the electric field vanishes (no sources):

$$\nabla E(r) = 0 , \quad (6.62)$$

we can directly infer the transversality of electromagnetic waves:

$$\mathbf{k} \cdot \mathbf{A} = 0 . \quad (6.63)$$

This means that there are 2 degrees of freedom for every \mathbf{k} (massless particle with spin 1 and helicity $H = \pm 1$).

Quantisation

For an electromagnetic wave, energy and momentum are given by

$$E = \hbar\omega , \quad (6.64)$$

$$\mathbf{p} = \hbar\mathbf{k} . \quad (6.65)$$

Therefore, the one-photon state energy is, using (6.60),

$$E_\nu = \hbar c |\mathbf{k}| = \frac{2\pi\hbar c}{L} , \quad (6.66)$$

where the states ν are given by

$$\nu = (n_x, n_y, n_z, H) \quad \text{with} \quad n_{x,y,z} \in \mathbb{Z}, H = \pm 1 . \quad (6.67)$$

We also can split up the states as

$$\nu = (\nu_T, \nu_I) , \quad (6.68)$$

into

$$\nu_T \equiv (n_x, n_y, n_z) \quad \text{translational degrees of freedom,} \quad (6.69)$$

$$\nu_I \equiv H \quad \text{inner degrees of freedom.} \quad (6.70)$$

Note that the translational degrees of freedom are the same for all kinds of particles, e. g. atoms, molecules.

Density of the translational states

For large volumes all translational states lie very densely, as $E_\nu \sim \frac{1}{L}$. Therefore, we can replace the sums over the states by integrals:

$$\sum_{\nu_\tau} \rightarrow V \int \frac{d^3 p}{(2\pi\hbar)^3}, \quad (6.71)$$

$$\sum_\nu \rightarrow 2V \int \frac{d^3 k}{(2\pi\hbar)^3}, \quad (6.72)$$

where we defined the volume $V = L^3$. To arrive at (6.72), we used $\mathbf{p} = \hbar\mathbf{k}$ and the fact that E_ν is independent of H . So, the factor of 2 arises due to the two possible values of the helicity $H = \pm 1$.

6.4.2. Spectrum of black body radiation

The mean number of photons per volume (photon density \tilde{n}) in the momentum interval $[\mathbf{k}, \mathbf{k} + d\mathbf{k}]$ is, using (6.55), (6.72) and $E = \hbar\omega$,

$$\tilde{n}(\mathbf{k}) d^3 k = \frac{2}{e^{\beta\hbar\omega} - 1} \frac{d^3 k}{(2\pi)^3}. \quad (6.73)$$

With the dispersion law $\omega = c|\mathbf{k}|$ we can deduce

$$d^3 k = 4\pi|\mathbf{k}|^2 d|\mathbf{k}| \quad (6.74)$$

$$= \frac{4\pi}{c^3} \omega^2 d\omega, \quad (6.75)$$

so the mean number of photons per volume in the frequency interval $[\omega, \omega + d\omega]$ is

$$\tilde{n}(\omega) d\omega = \frac{8\pi}{(2\pi c)^3} \frac{\omega^2}{e^{\beta\hbar\omega} - 1} d\omega. \quad (6.76)$$

Note that there is no divergence for $\omega \rightarrow 0$!

Using $E = \hbar\omega$, we can easily infer the mean photon energy per volume in the frequency interval $[\omega, \omega + d\omega]$ as

$$\bar{u}(\omega, T) d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{e^{\beta\hbar\omega} - 1} d\omega. \quad (6.77)$$

This is the frequency spectrum of black body radiation. It is plotted in Figure 6.2. Defining $\eta = \beta\hbar\omega = \frac{\hbar\omega}{k_B T}$, we recover the original form

$$\bar{u}(\eta, T) d\eta = \frac{(k_B T)^4}{\pi^2 (\hbar c)^3} \frac{\eta^3}{e^\eta - 1} d\eta, \quad (6.78)$$

postulated by Planck in 1900. Its maximum is roughly at $\hbar\omega \approx 3k_B T$ or $\eta \approx 3$, respectively.

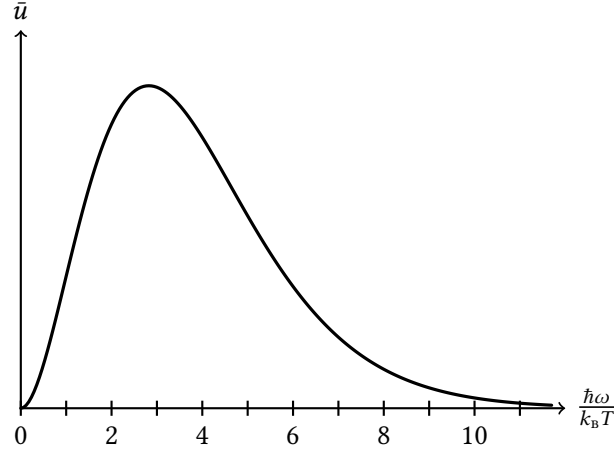


Figure 6.2.: Frequency spectrum of black body radiation

Of course, we also can deduce the wavelength spectrum, using $\omega = \frac{2\pi c}{\lambda}$:

$$\bar{u}(\lambda, T) d\lambda = 16\pi^2 \hbar c \frac{1}{e^{\frac{2\pi\hbar c}{k_B T \lambda}} - 1} \frac{d\lambda}{\lambda^5}. \quad (6.79)$$

Let us look at an example: the spectrum of sunlight. Its maximum is at $\lambda_{\max} \approx 500$ nm which is just what one would expect from a black body radiation spectrum at the sun's surface temperature $T \approx 5800$ K. This maximum lies in the regime of green visible light which, as an interesting side note, is also the maximum of the sensitivity of the human eye!

Furthermore, we can consider the low frequency limit $\hbar\omega \ll k_B T$, where the frequency spectrum can be approximated by

$$\bar{u}(\omega, T) d\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\beta \hbar \omega} d\omega \quad (6.80)$$

$$= \frac{k_B T}{\pi^2 c^3} \omega^2 d\omega, \quad (6.81)$$

which recovers classical electrodynamics (Einstein 1905). In this limit \hbar does not appear, any more.

The extrapolation to large frequencies $\hbar\omega \gg k_B T$, on the other hand, leads to a divergence:

$$\int_0^\infty d\omega \bar{u}(\omega, T) \rightarrow \infty. \quad (6.82)$$

This demonstrates the incompleteness of the theory! So, classical electrodynamics has to be modified. A solution is given by quantum statistics of photons, which accounts for the particle-wave duality (for large occupation number we are in the regime of classical fields, whereas for small occupation numbers we are in the regime of classical particles).

6.4.3. Energy density

The energy density is given by

$$\rho = \frac{E}{V} = \int_0^\infty d\omega \bar{u}(\omega, T) \quad (6.83)$$

$$= \int_0^\infty d\eta \bar{u}(\eta, T) \quad (6.84)$$

$$= \frac{(k_B T)^4}{\pi^2 c^3 \hbar^3} \underbrace{\int_0^\infty d\eta \frac{\eta^3}{e^\eta - 1}}_{\frac{\pi^4}{15}}. \quad (6.85)$$

So, we conclude that

$$\boxed{\rho = \frac{\pi^2}{15} \frac{(k_B T)^4}{(c\hbar)^3}}. \quad (\text{Stefan-Boltzmann law}) \quad (6.86)$$

This law contains several fundamental constants:

c : unification of space and time: special relativity

\hbar : quantum physics

k_B : associates T to energy, microphysical understanding of thermodynamics

Natural units

A frequently used system of units are the natural units. They are defined via

$$\hbar = c = k_B = 1. \quad (6.87)$$

This has some consequences:

$k_B = 1$: Temperature is measured in units of energy, e. g.

$$1 \text{ K} = 8.617 \cdot 10^{-5} \text{ eV}. \quad (6.88)$$

$c = 1$: Mass is measured in units of energy, e. g.

$$m_e = 511 \text{ keV}, \quad (6.89)$$

$$1 \text{ eV} = 1.783 \cdot 10^{-33} \text{ g}, \quad (6.90)$$

and distances are measured in units of time, e. g. light seconds.

$\hbar = 1$: Time and distances are measured in inverse units of energy, e. g.

$$1 \text{ fm} = 10^{-15} \text{ m} = \frac{1}{197 \text{ MeV}}, \quad (6.91)$$

$$1 \text{ m} = 5.068 \cdot 10^6 (\text{eV})^{-1}. \quad (6.92)$$

In Table 6.1 some physical quantities and constants and their respective mass (energy) dimensions in natural units are listed.

Quantity	Mass dimension
E	M
T	M
L, t	M^{-1}
V	M^{-3}
N	M^0
μ	M
p	M^4
$Z_{\text{mic, can, gc}}$	M^0
F	M
S	M^0
e	M^0
Q	M^0
α	M^0

Table 6.1.: Mass dimension of some physical quantities

6.4.4. Canonical partition function

The logarithm of the canonical partition function for a photon gas is computed as

$$\ln Z_{\text{can}}(\beta, V) = \sum_{\nu} \ln \frac{1}{1 - e^{-\beta E_{\nu}}} . \quad (6.93)$$

Together with (6.72) and $E_{\nu} = \frac{\eta}{\beta}$ this yields

$$\ln Z_{\text{can}} = \frac{V}{\pi^2} \left(\frac{k_{\text{B}} T}{\hbar c} \right)^3 \underbrace{\int_0^{\infty} d\eta \eta^2 \ln \frac{1}{1 - e^{-\eta}}}_{\frac{\pi^4}{45}} \quad (6.94)$$

$$= \frac{\pi^2}{45} V \beta^{-3} , \quad (6.95)$$

where we used natural units in the last step again. Let us check, if this gives the correct result for the energy density:

$$E = -\frac{\partial}{\partial \beta} \ln Z_{\text{can}} \quad (6.96)$$

$$= \frac{\pi^2}{15} V \beta^{-4} \quad (6.97)$$

and therefore

$$\rho = \frac{E}{V} = \frac{\pi^2}{15} T^4 \quad (6.98)$$

which agrees with (6.86), using natural units.

Remarks

1. Not only E itself is a macroscopic observable, but the energy distribution as a function of the frequency is one, too.
2. From Z_{can} other quantities, e. g. the pressure of the photon gas, easily can be attained.
 \rightarrow thermodynamic potentials \rightarrow classical thermodynamics

6.4.5. Free energy and equation of state

Using (6.95), the free energy is found to take the form

$$F = -k_B T \ln Z_{\text{can}} \quad (6.99)$$

$$= -\frac{\tau}{3} VT^4, \quad (6.100)$$

where we defined

$$\tau = \frac{k_B^4 \pi^2}{15c^3 \hbar^3} \quad (6.101)$$

or, using natural units,

$$\tau = \frac{\pi^2}{15}, \quad (6.102)$$

respectively. Then, the entropy, energy and pressure are given by

$$S = -\left(\frac{\partial F}{\partial T}\right)_V = \frac{4}{3}\tau VT^3, \quad (6.103)$$

$$E = F + TS = \tau VT^4, \quad (6.104)$$

$$p = -\left(\frac{\partial F}{\partial T}\right)_V = \frac{\tau}{3} T^4. \quad (6.105)$$

Combining the expressions for E and p and using $\rho = \frac{E}{V}$, we arrive at the equation of state for a photon gas:

$$\boxed{p = \frac{1}{3}\rho}, \quad (6.106)$$

where, as a reminder, the energy density ρ is given by

$$\boxed{\rho = \frac{\pi^2}{15} T^4}. \quad (6.107)$$

6.5. Degenerate Fermi gas**6.5.1. Quasiparticles**

At low temperature and high pressure quantum statistics become important! A typical example of this are electrons in metals. They can be considered as quasi free and are, due to their spin of $s = \frac{1}{2}$, fermions. Recall that the mean fermionic occupation number is given by

$$\tilde{n}_v^{(\text{F})} = \frac{1}{e^{\beta(E_v - \mu)} + 1}. \quad (6.108)$$

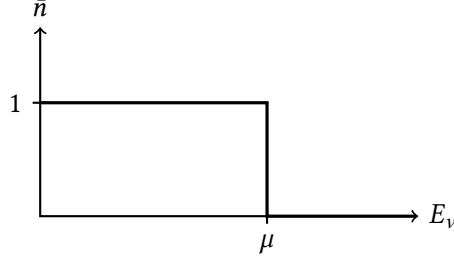


Figure 6.3.: Mean occupation number in zero temperature limit

6.5.2. The limit $T \rightarrow 0$

We consider the low temperature regime:

$$k_B T \ll |E_v - \mu|. \quad (6.109)$$

Thereby, the zero temperature limit of the mean occupation number depends on the relative size of E_v and μ :

$$\bar{n} \xrightarrow{k_B T \rightarrow 0} 0 \quad \text{for } E_v > \mu, \quad (6.110)$$

$$\bar{n} \xrightarrow{k_B T \rightarrow 0} 1 \quad \text{for } E_v < \mu. \quad (6.111)$$

This describes a step function as plotted in Figure 6.3.

We define the Fermi energy ε_F via

$$\varepsilon_F = \lim_{T \rightarrow 0} \mu(N, T), \quad (6.112)$$

and the Fermi momentum p_F or k_F , respectively, via the usual energy momentum relation:

$$\varepsilon_F = \frac{p_F^2}{2M} = \frac{\hbar^2 k_F^2}{2M}, \quad (6.113)$$

with M being the effective mass of the quasiparticle.

6.5.3. Computation of the Fermi energy

We consider electrons or atoms with 2 internal degrees of freedom ($g_i = 2$). Then, the number of particles is computed as

$$N = 2 \frac{V}{(2\pi)^3} \int_{k^2 < k_F^2} d^3 k \quad (6.114)$$

$$= \frac{V}{\pi^2} \int_0^{k_F} d|\mathbf{k}| |\mathbf{k}|^2 \quad (6.115)$$

$$= \frac{V}{3\pi^2} k_F^3. \quad (6.116)$$

So, using $n = \frac{N}{V}$, the Fermi momentum and energy are given by

$$k_F = (3\pi^2 n)^{1/3}, \quad (6.117)$$

$$\varepsilon_F = \frac{1}{2M} (3\pi^2 n)^{2/3}. \quad (6.118)$$

This way, we also know $\mu(n)$ at $T = 0$.

For free non-relativistic particles one also can define the Fermi temperature:

$$k_B T_F = \varepsilon_F. \quad (6.119)$$

For example the Fermi temperature of electrons in copper is approximately 80 000 K. So, we see that in this case (and also in general) $k_B T \ll \varepsilon_F$ is a very good approximation.

6.5.4. Thermodynamics

for our purposes the energy $E(T, V, N)$ is the most suitable thermodynamic potential. To calculate it, we perform an expansion in $\frac{T}{T_F}$. In the lowest order approximation the energy is given by the ground state energy:

$$E = E_0. \quad (6.120)$$

6.5.5. Ground state energy

The ground state energy is computed as

$$E_0 = \lim_{T \rightarrow 0} E \quad (6.121)$$

$$= 2 \frac{V}{(2\pi)^3} \int_{k^2 < k_F^2} d^3k \frac{k^2}{2M} \quad (6.122)$$

$$= \frac{V}{\pi^2} \frac{1}{2M} \int_0^{k_F} d|k| |k|^4 \quad (6.123)$$

$$= \frac{V}{10\pi^2 M} k_F^5 \quad (6.124)$$

$$= \frac{3}{10} \frac{k_F^2 N}{M}, \quad (6.125)$$

where we used (6.116) in the last step. Also using (6.113), leads us to

$$E_0 = \frac{3}{5} N \varepsilon_F. \quad (6.126)$$

So, in the low temperature regime we have

$$\frac{E}{N} = \frac{3}{5} \mu + \dots \quad (6.127)$$

One might wonder if $\frac{E}{N}$ is constant? The answer is no! It depends on the density due to

$$\mu = \varepsilon_F = \frac{\hbar^2}{2M} (3\pi^2 n)^{2/3}. \quad (6.128)$$

An additional particle always has the energy ε_F ! (not the mean energy of the already present particles)

So far, we have found the zero temperature equation of state for our Fermi gas:

$$E(T = 0, N, V) = \frac{3}{10M} (3\pi^2 n)^{2/3} N. \quad (6.129)$$

6.5.6. One-particle density of states

We would like to find a more general type of dispersion relation

$$E_\nu \equiv \varepsilon(\mathbf{k}) \equiv \hbar\omega(\mathbf{k}), \quad (6.130)$$

which also considers the interaction with the lattice. Note that ε denotes the one-particle energy here, not the energy density $\frac{E}{V}$. It is given by

$$\varepsilon = \frac{\hbar^2}{2M} |\mathbf{k}|^2 \quad (6.131)$$

for a single atomic non-relativistic gas and by

$$\varepsilon = \hbar c |\mathbf{k}| \quad (6.132)$$

for a relativistic Fermi gas. In principle, this does not change anything regarding the computation of k_F , now we just allow for a general function $\varepsilon_F(k_F)$. But nevertheless,

$$N = g_I \frac{V}{6\pi^2} k_F^3 \quad (6.133)$$

still holds, as it is a general result.

So, the energy is now computed as

$$E = g_I \frac{V}{(2\pi)^3} \int_{k^2 < k_F^2} d^3k \varepsilon(\mathbf{k}) \bar{n}(\mathbf{k}) \quad (6.134)$$

$$= g_I \frac{V}{(2\pi)^3} \int_{-\infty}^{\infty} d\varepsilon \int d^3k \delta[\varepsilon - \varepsilon(\mathbf{k})] \bar{n}(\varepsilon) \varepsilon, \quad (6.135)$$

and therefore

$$E = N \int_{-\infty}^{\infty} d\varepsilon \Omega_1(\varepsilon) \bar{n}(\varepsilon) \varepsilon, \quad (6.136)$$

where we have defined

$$\Omega_1(\varepsilon) = g_I \frac{V}{N} \int \frac{d^3k}{(2\pi)^3} \delta[\varepsilon - \varepsilon(\mathbf{k})], \quad (6.137)$$

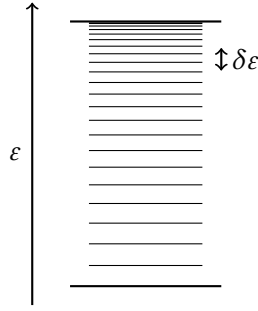


Figure 6.4.: Energy states

the number of states per particle for given ε or rather $\delta\varepsilon$, as sketched in Figure 6.4. (It corresponds to the quantity $\omega(E)$ for statistics of one particle in the first part of this lecture, as it has mass dimension M^{-1} .) It is an important microscopic property of a specific Fermi gas, as it influences the conductivity of metals, for example.

6.5.7. $\Omega_1(\varepsilon)$ for free non-relativistic fermions

Inserting the non-relativistic dispersion relation (6.131) into the expression (6.137) for Ω_1 leads to

$$\Omega_1(\varepsilon) = \frac{g_i V}{2\pi^2 N} \int_0^\infty d|\mathbf{k}| |\mathbf{k}|^2 \delta\left(\varepsilon - \underbrace{\frac{\hbar^2 |\mathbf{k}|^2}{2M}}_{\varepsilon(\mathbf{k})}\right). \quad (6.138)$$

The result is

$$\Omega_1(\varepsilon) = \frac{3}{2} \varepsilon_F^{-3/2} \varepsilon^{1/2} \Theta(\varepsilon), \quad (6.139)$$

which is only non-vanishing if $\varepsilon > 0$.

Let us check if this reproduces the correct ground state energy:

$$\frac{E_0}{N} = \int_0^{\varepsilon_F} d\varepsilon \varepsilon \Omega_1(\varepsilon) \quad (6.140)$$

$$= \frac{3}{2} \varepsilon_F^{-3/2} \cdot \frac{2}{5} \varepsilon_F^{5/2} \quad (6.141)$$

$$= \frac{3}{5} \varepsilon_F \quad \text{q. e. d.} \quad (6.142)$$

Can we determine $\Omega_1(\varepsilon)$ from thermodynamics, too? Often it is difficult to calculate it from first principles! But while the whole function $\Omega_1(\varepsilon)$ may be difficult to compute, one can infer some important properties. For example, one would like to know how thermodynamic variables, as specific heat, depend on $\Omega_1(\varepsilon)$, i. e. how they depend on the microphysics.

One uses thermodynamic observables for the measurement of microphysical properties.

The method, used for this, is to compute how observables depend on $\Omega_1(\varepsilon)$ by leaving $\Omega_1(\varepsilon)$ a free function.

6.5.8. $E(T, N)$ for arbitrary temperature

We already know that

$$E(T, \mu) = N \int_{-\infty}^{\infty} d\varepsilon \Omega_1(\varepsilon) \frac{\varepsilon}{e^{\beta(\varepsilon-\mu)} + 1}, \quad (6.143)$$

which we obtain by inserting (6.108) into (6.136). It holds for arbitrary T . So, in principle, this yields a general equation of state for $\frac{E}{N}$, but only if we know the form of $\mu(T)$. To compute it, we start with the general expression for the number of particles:

$$N = \sum_{\nu} \bar{n}_{\nu} \quad (6.144)$$

$$= \int_{-\infty}^{\infty} d\varepsilon \bar{n}(\varepsilon) g_1 V \int \frac{d^3 k}{(2\pi)^3} \delta[\varepsilon - \varepsilon(\mathbf{k})] \quad (6.145)$$

$$= N \int_{-\infty}^{\infty} d\varepsilon \Omega_1(\varepsilon) \bar{n}(\varepsilon). \quad (6.146)$$

From this we can read off the normalisation condition

$$\int_{-\infty}^{\infty} d\varepsilon \Omega_1(\varepsilon) \frac{1}{e^{\beta(\varepsilon-\mu)} + 1} = 1, \quad (6.147)$$

which, for known $\Omega_1(\varepsilon)$, determines $\mu(T)$.

As a cross check, we look at the $T = 0$ case again. In this case it is

$$\bar{n}(\varepsilon) = \Theta(\varepsilon_F - \varepsilon), \quad (6.148)$$

leading to

$$E_0 = N \int_0^{\varepsilon_F} d\varepsilon \Omega_1(\varepsilon) \varepsilon \quad (6.149)$$

which agrees with (6.140).

6.5.9. Smoothing of Fermi surface

If we look at temperatures $T > 0$, the Fermi surface is not a sharp step function, any more. But as long as we are still considering low temperatures, it just deviates slightly from this, as illustrated in Figure 6.5, i. e. it becomes a smooth step function, as only electrons close to the Fermi edge play a role at these small temperatures. This can be seen by looking at the deviation of the energy at $T \neq 0$ from the ground state energy ($T = 0$) for fixed N, V :

$$E - E_0 = N \int_{-\infty}^{\infty} d\varepsilon \varepsilon \Omega_1(\varepsilon) \left[\frac{1}{e^{\beta(\varepsilon-\mu)} + 1} - \Theta(\varepsilon_F - \varepsilon) \right]. \quad (6.150)$$

For $T \ll T_F$ the main contribution to this deviation comes from ε values close to ε_F . Note that $\varepsilon_F = \varepsilon_F(n)$, which follows from the calculation at $T = 0$, whereas $\mu = \mu(T)$, as $\mu \neq \varepsilon_F$ for $T \neq 0$.

A remark towards nomenclature: If electrons have an energy larger than ε_F , we call them (excited) particles. If they have an energy less than ε_F , we can call them holes.

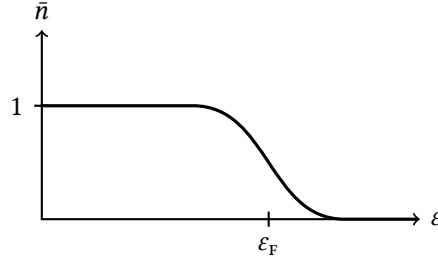


Figure 6.5.: Smooth Fermi surface

6.5.10. Computation of $\mu(T)$ from the normalisation condition

We know that

$$\int_{-\infty}^{\infty} d\epsilon \Omega_1(\epsilon) \left[\frac{1}{e^{\beta(\epsilon-\mu)} + 1} - \Theta(\epsilon_F - \epsilon) \right] = 0. \quad (6.151)$$

To calculate it, we will perform the so-called Sommerfeld approximation: We perform a Taylor expansion of $\Omega_1(\epsilon)$ around ϵ_F :

$$\Omega_1(\epsilon) = \Omega_1(\epsilon_F) + \Omega_1'(\epsilon_F) (\epsilon - \epsilon_F) + \dots \quad (6.152)$$

and define

$$\mu = \epsilon_F + \Delta\mu, \quad (6.153)$$

as well as as the new variable

$$x = \beta(\epsilon - \epsilon_F), \quad (6.154)$$

such that

$$\beta(\epsilon - \mu) = \beta(\epsilon - \epsilon_F) - \beta\Delta\mu \quad (6.155)$$

$$= x - \beta\Delta\mu. \quad (6.156)$$

This leads to

$$0 = \int_{-\infty}^{\infty} dx \left[\Omega_1(\epsilon_F) + \frac{\Omega_1'(\epsilon_F) x}{\beta} \right] \left[\frac{1}{e^x e^{-\beta\Delta\mu} + 1} - \Theta(-x) \right] + \dots \quad (6.157)$$

Expanding it for small $\beta\Delta\mu$:

$$\frac{1}{e^x e^{-\beta\Delta\mu} + 1} \approx \frac{1}{e^x (1 - \beta\Delta\mu) + 1} \quad (6.158)$$

$$\approx \frac{1}{(e^x + 1) \left(1 - \beta\Delta\mu \frac{e^x}{e^x + 1} \right)} \quad (6.159)$$

$$\approx \frac{1}{(e^x + 1)} + \beta\Delta\mu \frac{e^x}{(e^x + 1)^2}, \quad (6.160)$$

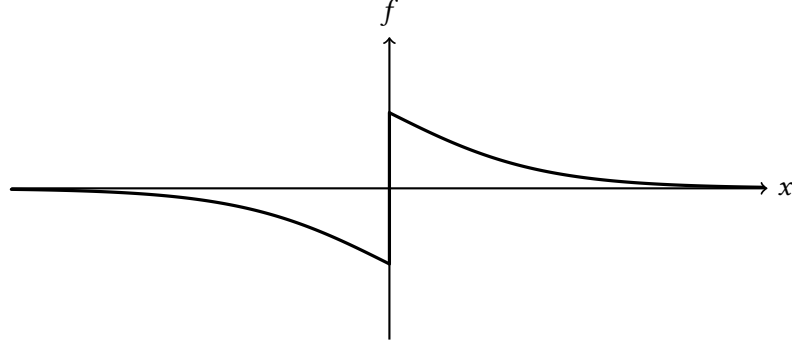


Figure 6.6.: The difference in between the occupation numbers at $T \neq 0$ and $T = 0$

yields

$$0 = \int_{-\infty}^{\infty} dx \left[\Omega_1(\epsilon_F) + \frac{\Omega'_1(\epsilon_F) x}{\beta} \right] \left[f(x) + \beta \Delta\mu \frac{e^x}{(e^x + 1)^2} \right], \quad (6.161)$$

where we defined

$$f(x) = \frac{1}{e^x + 1} - \Theta(-x), \quad (6.162)$$

Now, (6.161) is just a linear expression in $\Delta\mu$, but the x -integrations still have to be performed.

6.5.11. Occupation numbers for particles and holes

Note that the function $f(x)$, which characterizes the difference in between the occupation numbers at $T \neq 0$ and $T = 0$, is odd:

$$f(-x) = \frac{1}{e^{-x} + 1} - \Theta(x) \quad (6.163)$$

$$= \frac{e^x}{e^x + 1} + \Theta(-x) - 1 \quad (6.164)$$

$$= \frac{e^x - (1 + e^x)}{e^x + 1} + \Theta(-x) \quad (6.165)$$

$$= -\frac{1}{e^x + 1} + \Theta(-x) \quad (6.166)$$

$$= -f(x). \quad (6.167)$$

This means that at $T \neq 0$ the occupation number for the particles is enhanced in the same extent as it is lowered for the holes, when compared to the occupation number at $T = 0$. It is easy to this by looking at the plot of $f(x)$ in Figure 6.6.

We also can immediately conclude that

$$\int_{-\infty}^{\infty} dx f(x) = 0. \quad (6.168)$$

Moreover, some straightforward calculations yield

$$\int_{-\infty}^{\infty} dx \, x f(x) = \frac{\pi^2}{6}, \quad (6.169)$$

$$\int_{-\infty}^{\infty} dx \, \frac{e^x}{(e^x + 1)^2} = 1, \quad (6.170)$$

$$\int_{-\infty}^{\infty} dx \, x \frac{e^x}{(e^x + 1)^2} = 0, \quad (6.171)$$

where the last one follows directly from $\frac{e^x}{(e^x+1)^2}$ being even.

6.5.12. Solve for $\mu(T)$

Inserting (6.168) to (6.171) into (6.161) yields

$$\beta \Delta\mu \Omega_1(\varepsilon_F) = -\frac{\Omega'_1(\varepsilon_F)}{\beta} \frac{\pi^2}{6}. \quad (6.172)$$

By reinserting $\mu = \varepsilon_F + \Delta\mu$, we arrive at an expression for μ :

$$\mu = \varepsilon_F - \frac{\pi^2}{6} \frac{\Omega'_1(\varepsilon_F)}{\Omega_1(\varepsilon_F)} (k_B T)^2. \quad (6.173)$$

Let us compute this explicitly for the example of free non-relativistic particles:

$$\Omega_1 = \frac{3}{2} \varepsilon_F^{-3/2} \varepsilon^{1/2} \quad \rightarrow \quad \Omega_1(\varepsilon_F) = \frac{3}{2} \varepsilon_F^{-1}, \quad (6.174)$$

$$\Omega'_1 = \frac{3}{4} \varepsilon_F^{-3/2} \varepsilon^{-1/2} \quad \rightarrow \quad \Omega'_1(\varepsilon_F) = \frac{3}{4} \varepsilon_F^{-2}. \quad (6.175)$$

Therefore it is

$$\frac{\Omega'_1(\varepsilon_F)}{\Omega_1(\varepsilon_F)} = \frac{1}{2} \varepsilon_F^{-1}, \quad (6.176)$$

and we end up with

$$\mu = \varepsilon_F \left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\varepsilon_F} \right)^2 + \dots \right). \quad (6.177)$$

A similar calculation can be performed for the energy:

$$\frac{E - E_0}{N} = \int_{-\infty}^{\infty} d\varepsilon \, \Omega_1(\varepsilon) \left[\frac{1}{e^{\beta(\varepsilon - \varepsilon_F - \Delta\mu)} + 1} - \Theta(\varepsilon_F - \varepsilon) \right] \quad (6.178)$$

$$\approx \Delta\mu \varepsilon_F \Omega_1(\varepsilon_F) + \frac{\pi^2}{6} (k_B T)^2 [\Omega_1(\varepsilon_F) + \varepsilon_F \Omega'_1(\varepsilon_F)]. \quad (6.179)$$

Using $\Delta\mu = \mu - \varepsilon_F$ and inserting (6.173) leads to

$$E = E_0 + \frac{\pi^2}{6} (k_B T)^2 \Omega_1(\varepsilon_F) N. \quad (6.180)$$

	T_F [K]	T [K]
^3He gas (single atomic)	0.006	300
^3He liquid	12	3
electrons in Na	37 000	300
white dwarfs (electrons)	10^{11}	10^4
neutron stars	$4 \cdot 10^{12}$	–

Table 6.2.: Fermi temperature T_F vs. characteristic temperature T of different materials

By defining the Fermi temperature in a more general context via

$$k_B T_F = \frac{3}{2\Omega_1(\varepsilon_F)}, \quad (6.181)$$

which also reproduces the result $k_B T_F = \varepsilon_F$ for free non-relativistic particles, we can bring this into the form

$$E = E_0 + \frac{\pi^2 (k_B T)^2}{4 k_B T_F} N. \quad (6.182)$$

This way, one also can consider T_F as a measure for $\Omega_1(\varepsilon_F)$.

With an expression for the energy at hand we now can calculate the specific heat

$$c_V = \left. \frac{\partial E}{\partial T} \right|_{V,N}. \quad (6.183)$$

It is given by

$$c_V = \frac{\pi^2}{2} k_B N \frac{T}{T_F}, \quad (6.184)$$

which is linear in T ! Indeed, this is in good agreement with experimental results for metals at low temperatures! (the lattice contribution is $c_V \sim T^3$) Moreover, the measurement of T_F allows us to determine the density of states at the Fermi edge $\Omega_1(\varepsilon_F)$. This as an example of getting microphysical information out of the macroscopic thermodynamic behaviour of a system! In Table 6.2 we have listed the Fermi temperature in comparison to the characteristic temperature of some materials.

6.6. Non-relativistic bosons

6.6.1. Chemical potential

We are considering single atomic bosonic gases without any internal degrees of freedom excited ($g_i = 1$), e. g. hydrogen. So, the mean occupation number is

$$\bar{n}_\nu = \frac{1}{e^{\beta(E_\nu - \mu)} - 1}. \quad (6.185)$$

We demand $\bar{n}_v \geq 0$. It immediately follows that $\mu < E_0$, if E_0 is the lowest energy. Without loss of generality we set $E_0 = 0$, i. e. $\mu < 0$ (otherwise rescale E_v and μ by some additive shift, as only their difference $E_v - \mu$ matters, which is shift invariant), for example $E_v = \frac{p^2}{2M}$.

6.6.2. Number of bosons in the ground state

The mean occupation of the ground state is given by

$$\bar{n}_0 = \frac{1}{e^{-\beta\mu} - 1}. \quad (6.186)$$

For small $|\beta\mu|$ this becomes

$$\bar{n}_0 = \frac{1}{1 - \beta\mu - 1} \quad (6.187)$$

$$= -\frac{1}{\beta\mu}, \quad (6.188)$$

which can become very large for $\mu \rightarrow 0^-$. Macroscopically, \bar{n}_0 corresponds to Bose-Einstein condensation, i. e. $\frac{\bar{n}_0}{N}$ denotes the fraction of condensed bosons (for superfluids $\frac{\bar{n}_0}{N}$ is the superfluid density). In the thermodynamic limit $V \rightarrow \infty$ we distinguish the following cases:

$$\frac{\bar{n}_0}{N} \rightarrow 0 \quad \text{no condensate,} \quad (6.189)$$

$$\frac{\bar{n}_0}{N} \rightarrow n_s > 0 \quad \text{Bose-Einstein condensate,} \quad (6.190)$$

6.6.3. Qualitative behaviour of $\bar{n}(E_v)$ for $T \rightarrow 0$

For smaller T at fixed N there are fewer particles at high E_v due to stronger Boltzmann suppression. On the other hand, of course, this has to be compensated by more particles at small E_v . We have plotted this in Figure 6.7. It follows that the low temperature limit of μ is

$$\mu(T, N) \xrightarrow{T \rightarrow 0} 0. \quad (6.191)$$

6.6.4. Gibbs potential

The Gibbs potential was defined as

$$J = -T \ln Z_{\text{gc}} = -pV. \quad (6.192)$$

In this case, it is given by

$$J = -T \sum_v \ln \frac{1}{1 - e^{-\beta(E_v - \mu)}} \quad (6.193)$$

$$= T g_t V \int \frac{d^3k}{(2\pi)^3} \ln \left[1 - e^{-\frac{1}{T} \left(\frac{\hbar^2 k^2}{2M} - \mu \right)} \right]. \quad (6.194)$$

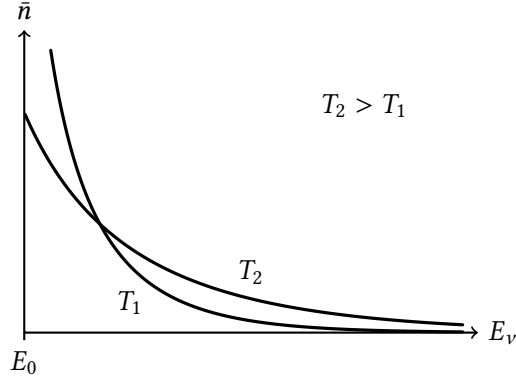


Figure 6.7.: Occupation number for different temperatures at fixed N

6.7. Classical ideal gas

6.7.1. Classical approximation

The classical approximation is given by

$$E_v - \mu \gg T, \quad (6.195)$$

or, inserting E_v ,

$$\frac{\hbar^2 \mathbf{k}^2}{2M} - \mu \gg T. \quad (6.196)$$

Regarding the occupation number, this leads to

$$\bar{n}_v = \frac{1}{e^{\beta(E_v - \mu)} \mp 1} \rightarrow e^{-\beta(E_v - \mu)} \quad (6.197)$$

for bosons and fermions, equally. So, the classical approximation is only valid for $\bar{n}_v \ll 1$.

6.7.2. Gibbs potential

Using this approximation, the Gibbs potential is given by

$$J = -T g_1 V \int \frac{d^3 k}{(2\pi)^3} e^{-\frac{1}{T} \left(\frac{\hbar^2 \mathbf{k}^2}{2M} - \mu \right)} \quad (6.198)$$

$$= \frac{-T g_1 V}{2\pi^2} \int_0^\infty d|\mathbf{k}| |\mathbf{k}|^2 e^{-\frac{1}{T} \left(\frac{\hbar^2 |\mathbf{k}|^2}{2M} - \mu \right)}, \quad (6.199)$$

which, of course, is again the same for bosons and fermions. So,

$$J = -T g_1 \frac{V}{2\pi^2} e^{\frac{\mu}{T}} I_3, \quad (6.200)$$

where

$$I_3 = \int_0^\infty dx x^2 e^{-bx^2} \quad (6.201)$$

$$= \frac{\partial}{\partial b} \int_0^\infty dx e^{-bx^2} \quad (6.202)$$

$$= \frac{\partial}{\partial b} \left(\frac{1}{2} \sqrt{\frac{\pi}{b}} \right) \quad (6.203)$$

$$= \frac{1}{4} \pi^{1/2} b^{-3/2}, \quad (6.204)$$

with

$$b = \frac{\hbar^2}{2Mk_B T}. \quad (6.205)$$

6.7.3. Thermal de Broglie wavelength

We define the thermal de Broglie wavelength as

$$\lambda(T) = \left(\frac{2\pi\hbar^2}{Mk_B T} \right)^{1/2}, \quad (6.206)$$

or, using natural units, as

$$\lambda(T) = \left(\frac{2\pi}{MT} \right)^{1/2}. \quad (6.207)$$

Due to the relation between temperature and kinetic energy:

$$\frac{\mathbf{p}^2}{2M} \sim T \quad (6.208)$$

we also can associate it to the momentum via

$$MT \sim \mathbf{p}^2. \quad (6.209)$$

Recalling the definition (6.205) of b , we find

$$\lambda^2 = 4\pi b, \quad (6.210)$$

and therefore

$$b^{-3/2} = \lambda^{-3} (4\pi)^{3/2}. \quad (6.211)$$

This way, we can express the factor I_3 in the Gibbs potential (6.200) in terms of the thermal de Broglie wavelength:

$$I_3 = 2\pi^2 \lambda^{-3}. \quad (6.212)$$

6.7.4. Thermodynamics

The Gibbs potential is now given by

$$\boxed{J = -T g_1 V \lambda^{-3} e^{\frac{\mu}{T}},} \quad (6.213)$$

as well as

$$J = -pV. \quad (6.214)$$

So, the number of particles is computed as

$$N = -\left(\frac{\partial J}{\partial \mu}\right)_{T,V} \quad (6.215)$$

$$= -\frac{1}{T} J \quad (6.216)$$

$$= \frac{pV}{T}, \quad (6.217)$$

leading to the equation of state:

$$pv = NT. \quad (6.218)$$

6.7.5. Validity of the approximation for a diluted gas (classical gas)

We want to check the validity of the classical approximation

$$\frac{p^2}{2M} - \mu \gg T. \quad (6.219)$$

For $\mu < 0$ it is sufficient to check if

$$-\frac{\mu}{T} \gg 1. \quad (6.220)$$

We know that

$$N = \frac{pV}{T} \quad (6.221)$$

$$= -\frac{J}{T} \quad (6.222)$$

$$= g_1 V \lambda^{-3} e^{\frac{\mu}{T}}, \quad (6.223)$$

leading to

$$\boxed{e^{\frac{\mu}{T}} = \frac{N\lambda^3}{g_1 V},} \quad (6.224)$$

and therefore

$$-\frac{\mu}{T} = \ln g_1 + \ln \frac{V}{N\lambda^3}. \quad (6.225)$$

From this we can read off that the approximation is valid if

$$\frac{V}{N} \gg \lambda^3(T). \quad (6.226)$$

Therefore, recalling that $\lambda \sim T^{-1/2}$, the classical approximation is valid for a diluted gas at sufficiently high temperature:

$$\boxed{n \ll \lambda^{-3}(T)}, \quad (6.227)$$

i. e. with far fewer than 1 particle per volume $\lambda^3(T)$.

6.8. Bose-Einstein condensation

We want to examine what happens, if we keep n fixed, while lowering T .

6.8.1. $n(\mu)$

The density is given by

$$n = \frac{N}{V} = \int \frac{d^3k}{(2\pi)^3} \frac{1}{e^{\beta \left(\frac{k^2}{2M} - \mu \right)} - 1}. \quad (6.228)$$

Using

$$\int \frac{d^3k}{(2\pi)^3} = \frac{1}{2\pi^2} \int d|\mathbf{k}| |\mathbf{k}|^2 \quad (6.229)$$

$$= \frac{1}{4\pi^2} \int d|\mathbf{k}|^2 \sqrt{|\mathbf{k}|^2} \quad (6.230)$$

and defining

$$z = \frac{|\mathbf{k}|^2}{2MT} \quad (6.231)$$

$$= \frac{\lambda(T)^2 |\mathbf{k}|^2}{4\pi}, \quad (6.232)$$

this becomes

$$n = \frac{1}{4\pi^2} \left(\frac{4\pi}{\lambda^2} \right)^{3/2} \int_0^\infty dz \sqrt{z} \frac{1}{e^{z-\beta\mu} - 1}, \quad (6.233)$$

and therefore

$$\boxed{n\lambda^3 = \frac{2}{\sqrt{\pi}} \int dz \sqrt{z} \frac{1}{e^{z-\beta\mu} - 1}}. \quad (6.234)$$

As μ is restricted to $\mu \leq 0$, we are interested in the limit $\mu \rightarrow 0^-$, i. e. $-\beta\mu \rightarrow 0^+$. One can read off that $n\lambda^3$ increases in this limit.

6.8.2. Critical temperature

Let us compute $n\lambda^3$ for $\mu = 0$ (its maximal value):

$$(n\lambda^3)_c = \frac{2}{\sqrt{\pi}} \underbrace{\int dz \sqrt{z} \frac{1}{e^z - 1}}_{\Gamma\left(\frac{3}{2}\right) \zeta\left(\frac{3}{2}\right)}. \quad (6.235)$$

Using $\Gamma\left(\frac{3}{2}\right) = \frac{1}{2}\sqrt{\pi}$, we arrive at

$$(n\lambda^3)_c = \zeta\left(\frac{3}{2}\right). \quad (6.236)$$

So, the critical wavelength is given by

$$\lambda_c = \zeta\left(\frac{3}{2}\right)^{1/3} n_c^{-1/3}, \quad (6.237)$$

which, using

$$\lambda_c^2 = \frac{2\pi}{MT_c}, \quad (6.238)$$

leads us to an expression for the critical temperature:

$$T_c = \frac{2\pi}{M} \zeta\left(\frac{3}{2}\right)^{-2/3} n_c^{2/3}. \quad (6.239)$$

Using $\zeta\left(\frac{3}{2}\right) \approx 2.61$, gives us the approximate result

$$T_c \approx 3.31 \frac{1}{M} n_c^{2/3}. \quad (6.240)$$

So, if we lower T until it reaches T_c for a given n , μ reaches 0. But what happens for $T < T_c$? Is there no solution?

6.8.3. Ground state of bosonic gas

At $T = 0$ the gas is in its ground state and has minimal energy. Therefore, *all* bosons have to be in the state with the lowest energy E_0 . In a box with periodic boundary conditions E_0 is given by

$$E_0 = \frac{\hbar^2 \mathbf{k}_0^2}{2M}, \quad (6.241)$$

with $\hbar \mathbf{k}_0$ being the lowest possible discrete momentum. And in a cavity it is given by

$$E_0 = \hbar \omega_0, \quad (6.242)$$

with ω_0 being the lowest resonance frequency.

But if we now demand

$$\bar{n}_{\nu=0}(T=0) = N, \quad (6.243)$$

we have a problem, as this diverges in the thermodynamic limit. The macroscopic number of particles in the ground state is

$$\frac{\bar{n}_{\nu=0}}{V} = \frac{N}{V} = n_0. \quad (6.244)$$

So, the atoms in the ground state are not properly counted in the transition from the discrete sum \sum_{ν} to the integral $\int \frac{d^3k}{(2\pi)^3}$. The total number number of atoms is given by

$$n = n_0(T) + \tilde{n}(T) \quad (6.245)$$

where $\tilde{n}(T)$ is the number of non-condensed thermal atoms. If $n_0 \neq 0$ we speak of Bose-Einstein condensation (BEC).

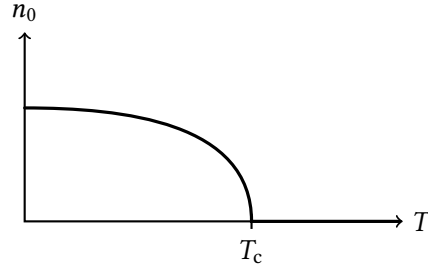


Figure 6.8.: Number of condensed atoms

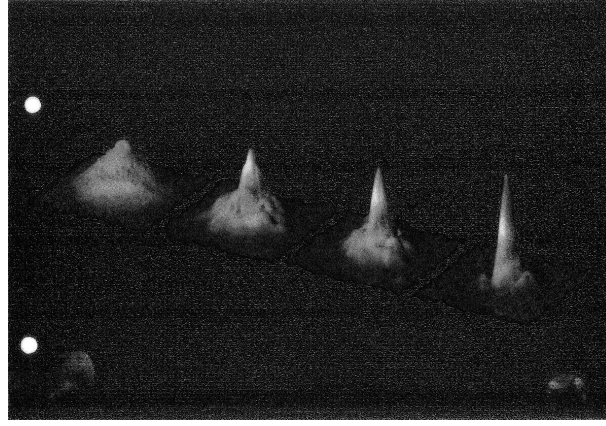


Figure 6.9.: Bose-Einstein condensation

6.8.4. Bose-Einstein condensation

At $T < T_c$ it is $\mu = 0$ and the number of non-condensed thermal atoms is given by

$$\tilde{n}(T) = \zeta\left(\frac{3}{2}\right) \lambda(T)^{-3} < n, \quad (6.246)$$

and therefore the remaining atoms have to form a Bose-Einstein condensate:

$$n_0(T) = n - \tilde{n}(T) > 0. \quad (6.247)$$

The qualitative behaviour of this function is plotted in Figure 6.8. At the critical temperature it is

$$n_0(T_c) = 0 \quad (6.248)$$

and it occurs a second order phase transition! For $n_0 \neq 0$ we also observe superfluidity! Experimental images of a BEC are shown in Figure 6.9.

We can express the number of condensed atoms in terms of the critical temperature. To do so, we use

$$\frac{\tilde{n}(T)}{\tilde{n}(T_c)} = \frac{\lambda^3(T_c)}{\lambda^3(T)} = \left(\frac{T}{T_c}\right)^{3/2}, \quad (6.249)$$

6. *Ideal quantum gas*

as $\lambda(T) \sim T^{-1/2}$. So, as $\tilde{n}(T_c) = n$, it follows

$$\tilde{n}(T) = \left(\frac{T}{T_c} \right)^{3/2} n, \quad (6.250)$$

and therefore

$$n_0(T) = \left[1 - \left(\frac{T}{T_c} \right)^{3/2} \right] n. \quad (6.251)$$

This has consequences for the total energy, too, as only particles with $E_v \neq E_0 = 0$ contribute if $T < T_c$:

$$E = \frac{g_1 M^{3/2} T^{5/2} V}{\sqrt{2} \pi^2 \hbar^3} \underbrace{\int_0^\infty dz \frac{z^{3/2}}{e^z - 1}}_{\Gamma\left(\frac{5}{2}\right) \zeta\left(\frac{5}{2}\right)}. \quad (6.252)$$

So, using

$$\Gamma\left(\frac{5}{2}\right) \zeta\left(\frac{5}{2}\right) = \frac{3}{4} \sqrt{\pi} \zeta\left(\frac{5}{2}\right) \approx 1.341, \quad (6.253)$$

we arrive at

$$\frac{E}{NT} \approx 0.77 \left(\frac{T}{T_c} \right)^{3/2}. \quad (6.254)$$

For comparison, for a classical gas we had $\frac{E}{NT} = \frac{3}{2}$.

7. Gas of interacting particles

7.1. Interactions and complexity

So far we only considered free particles. Adding interactions between these particles is a major change that will lead to a considerable increase in complexity! For example we will encounter new concepts like phase transitions and long range correlations.

Interactions are dealt with differently, depending on the considered temperature regime. At low temperatures only few states are excited. Collective modes of these states can be described as interactions between quasi-particles. Examples are acoustic waves and phonons in solids or magnons and spin waves in magnetic systems. At high temperatures, on the other hand, one introduces interactions as perturbations of the free theory.

We will find that many systems show some kind of microscopical order at low temperatures, whereas they are widely disordered at high temperatures. We are going to examine transitions between these two phases.

The aim of statistical physics is to describe all of these phenomena. Starting from known microphysics (atomic physics), it serves as a bridge to macrophysics. The key elements to achieve this, are statistical fluctuations and correlations.

7.2. Real gases

Contrary to ideal gases, one has to consider the interactions between atoms or molecules, respectively, to describe real gases. A typical form of the potential $V(R)$ between these particles is the Lenard-Jones potential:

$$V(R) = V_0 \left[\left(\frac{R_0}{R} \right)^{12} - 2 \left(\frac{R_0}{R} \right)^6 \right], \quad (7.1)$$

with R being the distance between two particles. This potential is used to describe Van-der-Waals forces at long distances, as well as the effect of Pauli exclusion at short distances. We have plotted it in Figure 7.1a. A simpler approximation would be

$$V(R) = \begin{cases} \infty & \text{for } R < R_0, \\ -V_0 \left(\frac{R_0}{R} \right)^6 & \text{for } R > R_0. \end{cases} \quad (7.2)$$

This describes hard spheres with an attractive interaction. It is sketched in Figure 7.1b.

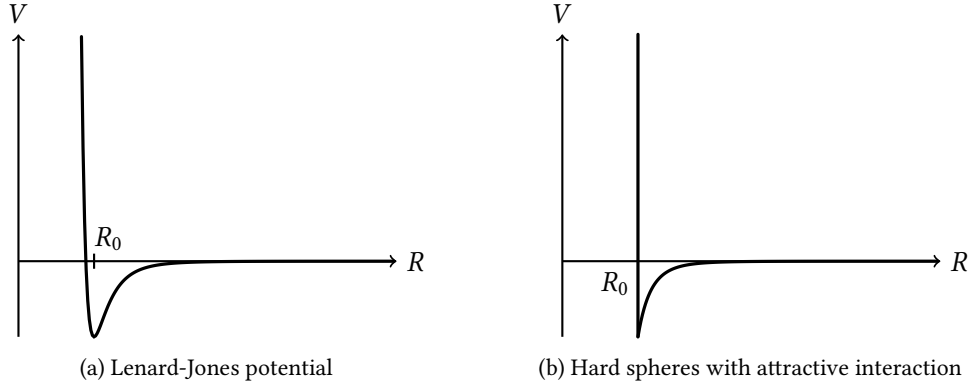


Figure 7.1.: Typical interaction potentials

Virial expansion

An important technique, used to describe real gases, is the virial expansion. It is an expansion of the equation of states in terms of the density $n = \frac{N}{V}$:

$$\frac{pV}{Nk_B T} = 1 + B_2(T) n + B_3(T) n^2 + \dots, \quad (7.3)$$

where the B_i are the so-called virial coefficients. Using the grand canonical partition function, we find

$$\frac{1}{V} \ln Z_{gc} = \frac{p}{k_B T} \quad (7.4)$$

$$= f(n, T) \quad (7.5)$$

$$= n (1 + B_2(T) n + B_3(T) n^2 + \dots). \quad (7.6)$$

7.3. Mean field approximation

Another important approach is the mean field approximation. In this approximation one considers a molecule to “see” an effective potential, generated by all the other molecules. This way, the complicated interacting system is reduced to an one-particle description with an effective Hamiltonian. Note that this is a self-consistent description and that it does not depend on a small density as expansion parameter! Moreover, we will work in classical statistics.

7.3.1. Classical statistics in phase space

The general Hamiltonian of a system of interacting particles is given by

$$H = \sum_i \frac{\mathbf{p}_i^2}{2M} + \sum_{ij} V(\mathbf{r}_i - \mathbf{r}_j), \quad (7.7)$$

where \sum_{ij} means the sum over pairs of particles. The problem here is that in quantum mechanics \mathbf{p}_i and \mathbf{r}_i do not commute. So, it is not easy to diagonalize H ! Therefore, we will consider the classical limit for fixed N . In this limit, the sum over microstates becomes

$$\sum_{\tau} \rightarrow \frac{1}{N!} \prod_{i=1}^N \int d^3 r_i \frac{d^3 p_i}{(2\pi\hbar)^3}. \quad (7.8)$$

The factor $\frac{1}{N!}$ takes care of the fact that we are dealing with identical particles. It can be shown that this is the correct limit by using the Wigner representation of the quantum density matrix.

If we had free particles, then the spatial integration would just lead to a volume factor for each particle:

$$\sum_{\tau} = \frac{1}{N!} V^N \prod_{i=1}^N \int \frac{d^3 p_i}{(2\pi\hbar)^3}. \quad (7.9)$$

This is an exact result!

7.3.2. Average potential \cong mean field

With the interaction potential in the Hamiltonian (7.7) one can not perform the spatial integration that easy, any more. At this point, the mean field approximation comes into play by replacing

$$\sum_{ij} V(\mathbf{r}_i - \mathbf{r}_j) \rightarrow \sum_i U_{\text{eff}}(\mathbf{r}_i). \quad (7.10)$$

Here, U_{eff} is an average potential “seen” by one particle, generated by the interaction with all other particles.

With this replacement the canonical partition function factorizes:

$$Z_{\text{can}} = \frac{1}{N!} \prod_i \int \frac{d^3 p_i d^3 r_i}{(2\pi\hbar)^3} e^{-\beta \left[\sum_j \frac{\mathbf{p}_j^2}{2M} + \sum_{jk} V(\mathbf{r}_j - \mathbf{r}_k) \right]} \quad (7.11)$$

$$= \frac{1}{N!} \prod_i \int \frac{d^3 p_i d^3 r_i}{(2\pi\hbar)^3} e^{-\beta \left[\frac{\mathbf{p}_i^2}{2M} + U_{\text{eff}}(\mathbf{r}_i) \right]}. \quad (7.12)$$

7.3.3. Expansion of canonical partition function

The canonical partition function can now be expressed as

$$Z_{\text{can}} \approx \frac{1}{N!} \left[\int \frac{d^3 p d^3 r}{(2\pi\hbar)^3} e^{-\beta \left[\frac{\mathbf{p}^2}{2M} + U_{\text{eff}}(\mathbf{r}) \right]} \right]^N \quad (7.13)$$

$$= Z_{\text{ideal}} V^{-N} \left[\int d^3 r e^{-\beta U_{\text{eff}}(\mathbf{r})} \right]^N \quad (7.14)$$

$$= Z_{\text{ideal}} \left(Z_{\text{ww}}^{(1)} \right)^N, \quad (7.15)$$

with Z_{ideal} being the canonical partition function for an ideal gas. Using this, the free energy becomes

$$F = -k_B T \ln Z_{\text{can}} \quad (7.16)$$

$$= -k_B T \ln Z_{\text{ideal}} - N k_B T \ln Z_{\text{ww}}^{(1)} \quad (7.17)$$

$$= F_{\text{ideal}} - N k_B T \ln Z_{\text{ww}}^{(1)}, \quad (7.18)$$

where F_{ideal} is the free energy of an ideal gas:

$$F_{\text{ideal}} = -N k_B T \left(1 - \ln \left[\frac{N}{V} \left(\frac{2\pi\hbar^2}{m k_B T} \right)^{3/2} \right] \right), \quad (7.19)$$

or, splitting up the logarithm,

$$F_{\text{ideal}} = -N k_B T + N k_B T \ln \frac{N}{V} - \frac{3}{2} N k_B T \ln \frac{m k_B T}{2\pi\hbar^2}. \quad (7.20)$$

7.3.4. Computation of $Z_{\text{ww}}^{(1)}$

To compute this deviation $Z_{\text{ww}}^{(1)}$ from the ideal gas partition function, we have to consider two effects:

1. excluded volume,
2. attractive interaction.

It is

$$Z_{\text{ww}}^{(1)} = \frac{1}{V} \int d^3r e^{-\beta U_{\text{eff}}(r)}. \quad (7.21)$$

This expression will be modified due to the two mentioned effects:

1. V_{ex} is the region, where $U_{\text{eff}} \rightarrow \infty$.
2. In the remaining space we exchange U_{eff} by a mean potential \bar{W} :

$$e^{-\beta \bar{W}} = \frac{1}{V - V_{\text{ex}}} \int_{V - V_{\text{ex}}} d^3r e^{-\beta U_{\text{eff}}(r)}. \quad (7.22)$$

Together, this leads to

$$Z_{\text{ww}}^{(1)} = \frac{V - V_{\text{ex}}}{V} e^{-\beta \bar{W}}. \quad (7.23)$$

7.3.5. Excluded volume

We have to determine V_{ex} , the mean excluded volume for an atom. To begin with, the total excluded volume is $N V_{\text{ex}}$, which can be expressed as

$$N V_{\text{ex}} = \underbrace{\frac{1}{2} N (N - 1)}_{\substack{\text{number of pairs} \\ \approx \frac{1}{2} N}} \underbrace{\frac{4\pi}{3} a^3}_{\substack{\text{excl. volume} \\ \text{per pair}}}, \quad (7.24)$$

with a being the minimal distance between two particles, i. e. $a = R_0$ in Figure 7.1b. So, we find

$$V_{\text{ex}} \approx \frac{1}{2} N \frac{4\pi}{3} a^3 \quad (7.25)$$

$$= \frac{2\pi}{3} a^3 N \quad (7.26)$$

$$= b_0 N, \quad (7.27)$$

where we defined the virial coefficient

$$b_0 = \frac{2\pi}{3} a^3. \quad (7.28)$$

7.3.6. Mean potential

Moreover, we have to compute the \bar{W} , the mean potential “seen” by a given atom. We will do this in a similar fashion. The potential energy of all atoms is given by

$$N \bar{W} = \frac{1}{2} N (N - 1) \bar{u}, \quad (7.29)$$

with \bar{u} being the mean potential energy per pair. It is

$$\bar{u} = 4\pi \int_a^\infty dR R^2 W(R) (V - V_{\text{ex}})^{-1}, \quad (7.30)$$

and therefore

$$\bar{W} = \frac{1}{2} N \bar{u} \quad (7.31)$$

$$= \frac{2\pi N}{V - V_{\text{ex}}} \int_a^\infty dR R^2 W(R). \quad (7.32)$$

Introducing another virial coefficient b_1 , we can write this as

$$\boxed{\bar{W} = -\frac{N}{V} b_1.} \quad (7.33)$$

Note that $\bar{W} \sim n!$

Altogether, we arrive at

$$Z_{\text{ww}}^{(1)} = \frac{V - b_0 N}{V} e^{-\frac{\bar{W}}{k_B T}}, \quad (7.34)$$

$$\ln Z_{\text{ww}}^{(1)} = \ln \frac{V - b_0 N}{V} - \frac{\bar{W}}{k_B T} \quad (7.35)$$

$$= \ln \frac{V - b_0 N}{V} + \frac{N}{V k_B T} b_1. \quad (7.36)$$

7.3.7. Free energy

Using the result (7.36), the free energy becomes

$$F = F_{\text{ideal}} - N k_B T \ln Z_{\text{ww}}^{(1)} \quad (7.37)$$

$$= -N k_B T + N k_B T \ln \frac{N}{V} - \frac{3}{2} N k_B T \ln \frac{m k_B T}{2\pi \hbar^2} - N k_B T \ln \frac{V - b_0 N}{V} - N k_B T \frac{N b_1}{k_B T V}. \quad (7.38)$$

So, we arrive at

$$F = -N k_B T + N k_B T \ln \frac{N}{V - b_0 N} - \frac{3}{2} N k_B T \ln \frac{m k_B T}{2\pi \hbar^2} - b_1 \frac{N^2}{V}. \quad (7.39)$$

7.4. Equation of state

Using the general relation

$$p = - \left. \frac{\partial F}{\partial N} \right|_{N, T}, \quad (7.40)$$

we arrive at the equation of state:

$$p = \frac{N k_B T}{V - b_0 N} - b_1 \left(\frac{N}{V} \right)^2. \quad (7.41)$$

Note that here only the structural form is important. The computation of b_0 and b_1 can be improved by a self-consistent computation of $U_{\text{eff}}(\mathbf{r})$.

Often, it is more useful to express these equations just in terms of the density $n = \frac{N}{V}$ and T . So, instead of F we look at $\frac{F}{V}$:

$$\frac{F}{V} = -n k_B T + n k_B T \ln \frac{n}{1 - b_0 n} - \frac{3}{2} n k_B T \ln \frac{m k_B T}{2\pi \hbar^2} - b_1 n^2, \quad (7.42)$$

and the equation of state becomes

$$p = \frac{n k_B T}{1 - b_0 n} - b_1 n^2. \quad (7.43)$$

This is the van der Waals equation. One finds it to describe the phase transition from gas to liquid qualitatively correct!

7.5. Specific heat of real gases

The specific heat was defined as

$$c_V = T \left. \frac{\partial S}{\partial T} \right|_{T,N} \quad (7.44)$$

$$= -T \left. \frac{\partial^2 F}{\partial T^2} \right|_{V,N} \quad (7.45)$$

$$= V \left. \frac{\partial}{\partial \ln T} \left(\frac{S}{V} \right) \right|_n, \quad (7.46)$$

with $\frac{S}{V}$ the entropy density:

$$\frac{S}{V} = - \left. \frac{\partial}{\partial T} \left(\frac{F}{V} \right) \right|_n \quad (7.47)$$

$$= k_B n - k_B n \ln \frac{n}{1 - b_0 n} + \frac{3}{2} k_B n \ln \frac{m k_B T}{2\pi \hbar^2} + \frac{3}{2} k_B n. \quad (7.48)$$

So,

$$\boxed{\frac{S}{V} = k_B n \left(\frac{5}{2} - \ln \frac{n}{1 - b_0 n} + \frac{3}{2} \ln \frac{m k_B T}{2\pi \hbar^2} \right)}, \quad (7.49)$$

which is independent of b_1 .

This leads to

$$\frac{c_V}{V} = \left. \frac{\partial}{\partial \ln T} \left(\frac{S}{V} \right) \right|_n = \frac{3}{2} k_B n, \quad (7.50)$$

and we finally arrive at

$$\boxed{c_V = \frac{3}{2} k_B N}. \quad (7.51)$$

So, regarding the specific heat there is no difference between a van der Waals gas and an ideal gas!

7.6. Equipartition theorem

At high temperatures one finds the specific heat to be proportional to N with a system dependent prefactor, for example:

$$\text{Monatomic (real) gas:} \quad c_V = \frac{3}{2} k_B N, \quad (7.52)$$

$$\text{Phonons:} \quad c_V = 3 k_B N. \quad (7.53)$$

In general, it is

$$\boxed{c_V = \frac{1}{2} k_B N f_1}, \quad (7.54)$$

with f_1 being the number of excited degrees of freedom per particle. The total number of excited degrees of freedom is $f = N f_1$. So, for example one finds:

$$\text{Monatomic gas:} \quad f_1 = 3 \quad (\text{translations}), \quad (7.55)$$

$$\text{Phonons:} \quad f_1 = 6 \quad (\text{position and velocity}), \quad (7.56)$$

$$\text{Diatomic molecule:} \quad f_1 = 3 + 2 + (2) \quad (\text{translations, rotations and vibrations}). \quad (7.57)$$

Regarding the diatomic molecule: note that one (longitudinal) vibration corresponds to $f_1 = 2$. Alternatively, one can split up its total $f_1 = 7$ into 6 velocity components and the distance between the atoms.

The relation (7.54) is always true, if the quadratic approximation for the energy holds:

$$E = E_0 + \sum_{\rho} b_{\rho} q_{\rho}^2. \quad (7.58)$$

Here, the q_{ρ} are the excited degrees of freedom of the system and the b_{ρ} are suitable prefactors. ρ labels the excited degrees of freedom and therefore $\sum_{\rho} = N f_1$. We denote the non-excited degrees of freedom by \tilde{q} . So, for a general degree of freedom q_r one finds $q_r \in \{q_{\rho}, \tilde{q}\}$.

The average energy of such a system is computed as

$$\langle E \rangle = \frac{\int \mathcal{D}q_r (E_0 + \sum_{\rho'} b_{\rho'} q_{\rho'}^2) e^{-\beta E_0 - \beta \sum_{\rho} b_{\rho} q_{\rho}^2}}{\int \mathcal{D}q_r e^{-\beta E_0 - \beta \sum_{\rho} b_{\rho} q_{\rho}^2}}, \quad (7.59)$$

where

$$\int \mathcal{D}q_r = \int dq_1 \cdots dq_r \cdots dq_N \quad (7.60)$$

$$= \int \mathcal{D}q_{\rho} \mathcal{D}\tilde{q}. \quad (7.61)$$

It is easy to see that the $\mathcal{D}\tilde{q}$ integration and the factor $e^{-\beta E_0}$ in numerator and denominator cancel. Moreover, we can pull the E_0 out of the remaining $\mathcal{D}q_{\rho}$ integration. We end up with

$$\langle E \rangle = \frac{\prod_{\rho} \left(\int dq_{\rho} e^{-\beta b_{\rho} q_{\rho}^2} \right) \sum_{\rho'} b_{\rho'} q_{\rho'}^2}{\prod_{\rho} \left(\int dq_{\rho} e^{-\beta b_{\rho} q_{\rho}^2} \right)}. \quad (7.62)$$

For each summand only one integration is non-trivial:

$$\langle E \rangle - E_0 = \sum_{\rho'} \frac{\int dq_{\rho'} b_{\rho'} q_{\rho'}^2 e^{-\beta b_{\rho'} q_{\rho'}^2}}{\int dq_{\rho'} e^{-\beta b_{\rho'} q_{\rho'}^2}} \quad (7.63)$$

$$= \sum_{\rho'} \left(-\frac{\partial}{\partial \beta} \right) \ln \int dq_{\rho'} e^{-\beta b_{\rho'} q_{\rho'}^2} \quad (7.64)$$

$$= \sum_{\rho'} \left(-\frac{\partial}{\partial \beta} \right) \underbrace{\ln \sqrt{\frac{\pi}{\beta b_{\rho'}}}}_{-\frac{1}{2} \ln \beta + \text{const.}} \quad (7.65)$$

$$= \sum_{\rho'} \frac{1}{2\beta} . \quad (7.66)$$

And therefore,

$$\langle E \rangle = E_0 + \frac{1}{2} k_B T \cdot N f_1 . \quad (7.67)$$

Note that we implicitly assumed a classical treatment, i. e. we neglected the discreteness of energy levels. This is only valid if the difference between the ground state energy and the energy levels in the ρ -channel is small compared to $k_B T$:

$$k_B T \gg E_{1,\rho} - E_0 . \quad (7.68)$$

8. Magnetic systems

8.1. Magnetization

8.1.1. Degrees of freedom

We consider a lattice with points \mathbf{x} . In this case it is a cubic lattice:

$$\mathbf{x} = \mathbf{n}a, \quad (8.1)$$

$$\mathbf{n} = (n_1, n_2, n_3) \quad \text{with} \quad n_k \in \mathbb{Z}, \quad (8.2)$$

where a is the lattice spacing. There are classical spins $\chi_a(\mathbf{x})$ at each lattice site \mathbf{x} which satisfy

$$\sum_{a=1}^3 \chi_a(\mathbf{x}) \chi_a(\mathbf{x}) = 1. \quad (8.3)$$

The microstates of this system are given by the possible spin configurations $\tau \cong \chi_a(\mathbf{x})$, the so-called “lattice fields”. Moreover, we have a Hamiltonian

$$H = H[\chi_a(\mathbf{x})], \quad (8.4)$$

where we identify $H \equiv E_\tau$. Of course, this can be generalized to N -component fields:

$$\chi_a(\mathbf{x}) \quad \text{with} \quad a = 1, \dots, N, \quad (8.5)$$

$$\sum_{a=1}^N \chi_a(\mathbf{x}) \chi_a(\mathbf{x}) = 1. \quad (8.6)$$

One example of such a system is the Ising model, which is characterized by $N = 1$ and $\chi(\mathbf{x}) = \pm 1$ (discrete number of microstates). So, its microstates are given by

$$\tau : \{\chi(\mathbf{x})\} \cong \{s_j\} \quad \text{with} \quad s_j = \pm 1 \quad (8.7)$$

and the corresponding probabilities take the form

$$p_\tau = \frac{e^{-\beta E_\tau}}{Z} = \frac{e^{-\beta H}}{Z}, \quad (8.8)$$

as we are describing it as a canonical ensemble. Note that we are not dealing with uncorrelated probabilities as in chapter 3, any more!

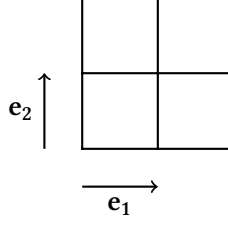


Figure 8.1.: Unit lattice vectors

8.1.2. Hamiltonian

We are considering the following Hamiltonian:

$$H = - \underbrace{\sum_x \sum_\mu \sum_a \chi_a(\mathbf{x}) \chi_a(\mathbf{x} + \mathbf{e}_\mu)}_{H_{\text{int}}} - c_B \sum_x \sum_a \chi_a(\mathbf{x}) B_a . \quad (8.9)$$

Here, c_B is the magnetic moment, for which we assume

$$c_B > 0 . \quad (8.10)$$

\mathbf{e}_μ is the unit lattice vector in μ -direction, as shown in Figure 8.1. Of course, the sum over μ runs from 1 to the dimension d of the lattice ($d = 3$ for a cubic lattice):

$$\sum_\mu = \sum_{\mu=1}^d . \quad (8.11)$$

So, the properties of H_{int} are that it only considers next-neighbour interactions and that it tries to align the spins. On the other hand, B is a magnetic field which tries to align the spins in the same direction as B .

The ground state of the system is characterized by the limit $\beta \rightarrow \infty$. In this limit the spin expectation values will be proportional to the magnetic field:

$$\langle \chi_a(\mathbf{x}) \rangle \sim B_a . \quad (8.12)$$

Note that the coupling multiplying H_{int} can be absorbed into $\beta \rightarrow \beta'$.

8.1.3. Magnetization

The magnetization is defined as

$$M_a = \sum_x \langle \chi_a(\mathbf{x}) \rangle . \quad (8.13)$$

Since B_a is an intensive quantity, M_a is an extensive one. Now, the question is: How does the magnetization depend on the magnetic field?

$$\boxed{M(B) = ?} \quad (8.14)$$

8.1.4. Thermodynamic potential

As always, the canonical partition function is given by

$$Z = \sum_{\tau} e^{-\beta H} = \sum_{\tau} e^{-\beta E_{\tau}}. \quad (8.15)$$

From (8.9) we can read off that it includes a source term, which is linear in the field $\chi(\mathbf{x})$. Componentwise, the corresponding source field is given by

$$j_a = \beta c_B B_a. \quad (8.16)$$

So, instead of considering the partition function as B -dependent, we also can consider it as j -dependent:

$$Z(T, B) = Z(T, j), \quad (8.17)$$

which is a mere change of notation. But we will stick to the notation using j , as it is the same notation as in QFT. Note that the volume V and the particle number N do not play a role here – they are no thermodynamic variables!

Again, we can perform the transition from a sum over states to a “functional integral”:

$$\sum_{\tau} \equiv \prod_x \left[\left(\prod_a \int_{-\infty}^{\infty} d\chi_a(\mathbf{x}) \right) \delta \left(\sum_a \chi_a(\mathbf{x}) \chi_a(\mathbf{x}) - 1 \right) \right]. \quad (8.18)$$

On the other hand, for the discrete Ising model ($N = 1$) we just have

$$\sum_{\tau} = \prod_x \sum_{\chi(x)=\pm 1}. \quad (8.19)$$

We define

$$W(T, j) = \ln Z(T, j), \quad (8.20)$$

which is related to the free energy in the following way:

$$F(T, B) = -TW(T, j). \quad (8.21)$$

This way, the basic relation for the magnetization takes the form

$$\boxed{\frac{\partial W}{\partial j_a} = M_a} \quad (8.22)$$

which can be seen by

$$\frac{\partial W}{\partial j_a} = \frac{1}{Z} \frac{\partial Z}{\partial j_a} \quad (8.23)$$

$$= \frac{1}{Z} \frac{\partial}{\partial j_a} \sum_{\tau} e^{-\beta H_{\text{int}} + \sum_x \chi_a(\mathbf{x}) j_a(\mathbf{x})} \quad (8.24)$$

$$= \left\langle \sum_x \chi_a(\mathbf{x}) \right\rangle \quad (8.25)$$

$$= M_a. \quad (8.26)$$

Therefore, one also can compute the magnetization from the free energy:

$$M \sim \frac{\partial F}{\partial B}, \quad (8.27)$$

i. e. we can identify some analogies to our description of gases:

$$B \cong \mu, \quad (8.28)$$

$$M \cong N, \quad (8.29)$$

$$F \cong J \quad (\text{Gibbs potential}). \quad (8.30)$$

So, in the following we will use

$$W(T, j) \quad \text{as thermodynamic potential}, \quad (8.31)$$

$$T, j \quad \text{as thermodynamic variables}. \quad (8.32)$$

8.1.5. Effective potential

We denote the number of lattice sites by N . Then $M_a \sim N$. So, we introduce the magnetization per lattice site ϕ_a via

$$\phi_a = \frac{M_a}{N}. \quad (8.33)$$

Moreover, we define the effective potential $U(\phi_a, T)$:

$$U = \frac{1}{N} \left[-W(j, T) + M_a j_a \right], \quad (8.34)$$

where the sum over repeated indices is implied. Obviously,

$$U = -\frac{W(j, T)}{N} + \phi_a j_a, \quad (8.35)$$

which is a Legendre transformation.

Note that the sources j_a depend on the ϕ_b . The exact dependence can be inferred by inverting the relation

$$\phi_a(j_b) = \frac{1}{N} \frac{\partial W}{\partial j_b}, \quad (8.36)$$

leading to

$$\boxed{\frac{\partial U}{\partial \phi_a} = j_a}. \quad (8.37)$$

This is the basic relation between magnetization and magnetic field, which can be proven by direct computation:

$$\frac{\partial U}{\partial \phi_a} = j_a + \frac{\partial j_b}{\partial \phi_a} \left(\underbrace{\phi_b - \frac{\partial}{\partial j_b} \frac{W(j, T)}{N}}_{\phi_b} \right). \quad (8.38)$$

In the following we want to compute $U(\phi)$!

8.1.6. Minimal principle

We define

$$\bar{U}(\phi, j) = U(\phi) - j\phi. \quad (8.39)$$

The magnetization is determined by the *minimum* of $\bar{U}(\phi, j)$ with respect to ϕ at fixed j :

$$\left. \frac{\partial \bar{U}}{\partial \phi} \right|_j = 0. \quad (8.40)$$

We denote this minimum by ϕ_0 . Then it is

$$M_a = \mathcal{N} \phi_{0,a}. \quad (8.41)$$

We can relate \bar{U} and F via

$$\mathcal{N} \bar{U} = U - jM \quad (8.42)$$

$$= -W \quad (8.43)$$

$$= \frac{F}{T}, \quad (8.44)$$

where $F(T, j; M)$ is the free energy with the “constraint” that M takes the given value (by inverting $M(j)$, one gets the free energy $F(T, j)$ in standard form). This is similar to the entropy of a system with constraint y ($S(y)$ maximal $\rightarrow F(y)$ minimal).

What is the probability to find a given M ? To answer that, we need the probability distribution $w(M)$:

$$w(M) = \sum_{\tau} \left| \sum_{\mathbf{x}} \chi_a(\mathbf{x}) = M_a \right| e^{-\beta H[\chi_a(\mathbf{x})]} \quad (8.45)$$

$$\approx \Omega(M) e^{-\beta E(M)} \quad (8.46)$$

$$= e^{S(M) - \frac{1}{T} E(M)} \quad (8.47)$$

$$= e^{-\frac{F(M)}{T}}, \quad (8.48)$$

where we used $F = E - TS$ in the last step. From this we conclude that the maximum of $w(M)$ corresponds to the minimum of $F(M)$ and therefore to the minimum of $\bar{U}(\phi)$. For $j = 0$ this then corresponds to a minimum of U :

$$\frac{\partial U}{\partial \phi_a} = 0. \quad (8.49)$$

8.2. Spontaneous symmetry breaking

8.2.1. Symmetries

Obviously, H is invariant under lattice translations. Moreover, H_{int} is invariant under lattice rotations and reflections. But these symmetries are (at least partially) broken by $B \neq 0$ ($j \neq 0$).

$O(N)$ -symmetry

$\chi_a(\mathbf{x})$ and $j_a(\mathbf{x})$ are N -component vectors. So, H_{int} is

- invariant under N -dimensional rotations, i. e. “ $SO(N)$ -invariant”,
- invariant under $\chi_a(\mathbf{x}) \rightarrow -\chi_a(\mathbf{x})$ (internal reflection).

From a group theoretical point of view the symmetry group $SO(N)$ together with reflections yields the symmetry group $O(N)$, which describes all linear transformations that leave $\sum_a (\chi_a \chi_a)$ invariant. Therefore, H_{int} is $O(N)$ -invariant.

But the source term breaks this $O(N)$ -symmetry for fixed $j \neq 0$. So, only *simultaneous* rotations of $\chi_a(\mathbf{x})$ and $j_a(\mathbf{x})$ leave H invariant. From this we can conclude that $W(j, T)$ is invariant under $O(N)$ rotations of j_a and that, as ϕ_a also is a vector under rotations, $U(\phi_a)$ is invariant under $O(N)$ rotations of ϕ_a .

This symmetry of $U(N)$ leads to a simplification of the computations. As already said, U is invariant with respect to $O(N)$ -symmetry and the explicit breaking by a magnetic field appears only in the source term. Therefore, the solution of

$$\frac{\partial U}{\partial \phi_a} = j_a \quad (8.50)$$

is not $O(N)$ -symmetric for a given $j_a \neq 0$. This means, that the magnetic field singles out a particular direction.

8.2.2. Vanishing magnetic field

What happens for $B = 0$? Or more precisely, what happens in the limit $j_a \rightarrow 0$? There are two possibilities. *Either* there is a non-vanishing magnetization $\phi_a \neq 0$ in an arbitrary direction, as no direction is preferred (“ground-state degeneracy”). If this happens, we speak of “spontaneous symmetry breaking” (SSB). *Or* there is no magnetization ($\phi_a = 0$). In this case the $O(N)$ -symmetry is preserved.

Example

- In the Ising model, $\phi = 1$ and $\phi = -1$ are equivalent. So the magnetisation has to choose one of the two possible values. In the ground state, the minimum of H , all spins are aligned, so

$$\text{either } \chi(\mathbf{x}) = 1 \quad \text{for all } \mathbf{x}, \quad \phi = 1, \quad (8.51)$$

$$\text{or } \chi(\mathbf{x}) = -1 \quad \text{for all } \mathbf{x}, \quad \phi = -1. \quad (8.52)$$

8.2.3. Domains

A domain is a continuous region of spins pointing in the same direction. In Figure 8.2 we sketched three neighbouring domains. At the boundary between two domains there is a positive

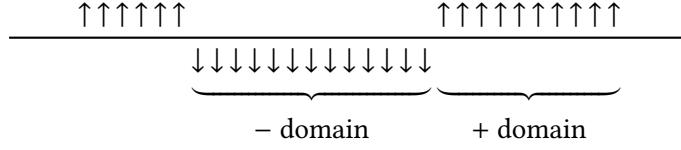


Figure 8.2.: Spin domains

interaction energy. Therefore, the probability p_τ of a “domain-microstate” is smaller than the probability p_τ of the ground state. But on the other hand, there are more domain-states than ground states (only two in the Ising model).

The magnetization is given by

$$M = \sum_{\tau} p_{\tau} M_{\tau} . \quad (8.53)$$

But what wins? p_τ or the number of contributions? We expect, that for low T the system will have near ground state order with $\phi \neq 0$ (SSB), and that for high T the system will be completely disordered with $\phi = 0$ (SYM). There will be a phase transition between these two cases at the critical temperature T_c .

Note that in practice domains are visible as “Weiss domains” in ferromagnets.

8.2.4. General form of U

As we have already pointed out, the effective potential $U(\phi_a)$ must be invariant under rotations. Let us define

$$\rho = \frac{1}{2} \phi_a \phi_a . \quad (8.54)$$

This ρ is the simplest expression, which is invariant under rotations ($O(N)$ includes all transformations that leave $\phi_a \phi_a$ invariant). Therefore, U has to be a function of ρ . This simplifies the problem, as now U only depends on the one variable ρ and not explicitly on N variables ϕ_a .

8.3. Landau theory and phase transition

8.3.1. Quartic expansion of U

We expand U up to order ϕ^4 (ρ^2):

$$U = U_0 + \mu^2 \rho + \frac{1}{2} \lambda \rho^2 , \quad (8.55)$$

where U_0 is irrelevant for the magnetization, as it drops out when performing the ϕ derivative:

$$\frac{\partial U}{\partial \phi_a} = j_a . \quad (8.56)$$

We assume $\lambda > 0$, as otherwise U has no absolute minimum.

Derivatives with respect to ϕ_a can be expressed in terms of derivatives with respect to ρ :

$$\frac{\partial U}{\partial \phi_a} = \frac{\partial U}{\partial \rho} \frac{\partial \rho}{\partial \phi_a} = \frac{\partial U}{\partial \rho} \phi_a . \quad (8.57)$$

8.3.2. Phase transition for $j = 0$

We look for the minimum of U (as $U = \bar{U}$ for $j = 0$):

$$0 = \frac{\partial U}{\partial \phi_a} = \mu^2 \phi_a + \lambda \rho \phi_a \quad (8.58)$$

$$= (\mu^2 + \frac{1}{2} \lambda \phi_b \phi_b) \phi_a . \quad (8.59)$$

There are two cases:

$\mu^2 > 0$: Only solution is $\phi_a = 0$ (no magnetisation, SYM, Figure 8.3).

$\mu^2 < 0$: Minimum of U for $\rho = \rho_0$ (SSB, Figure 8.4), where

$$\rho_0 = -\frac{\mu^2}{\lambda}, \quad \phi_0 \phi_0 = -\frac{2\mu^2}{\lambda} . \quad (8.60)$$

This is the so-called “mexican hat potential”.

In between, at $\mu^2 = 0$ there is a phase transition (Figure 8.5) At this point there is no quadratic term and the potential is just given by

$$U = \frac{\lambda}{8} (\phi_a \phi_a)^2 . \quad (8.61)$$

Because $|\phi_0|(\mu^2)$ is a *continuous* function, as one can see in Figure 8.6, we speak of a second order phase transition.

Critical temperature

$$\mu^2 = \mu^2(T) , \quad (8.62)$$

$$\lambda = \lambda(T) \quad (8.63)$$

are continuous functions of the temperature. Thereby, $\mu^2(T)$ is monotonically increasing and we know that

$$\mu^2(T = 0) < 0 , \quad (8.64)$$

$$\mu^2(T \text{ large}) > 0 . \quad (8.65)$$

The critical temperature is defined via

$$\mu^2(T = T_c) = 0 , \quad (8.66)$$

i. e. it is the temperature where the phase transition occurs. The resulting phase diagram is sketched in Figure 8.7.

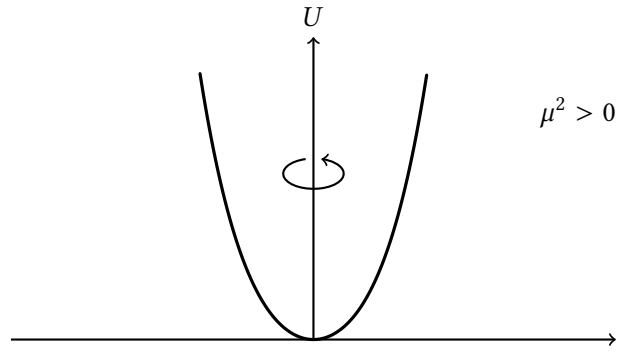


Figure 8.3.: Effective potential in the symmetric phase (SYM)

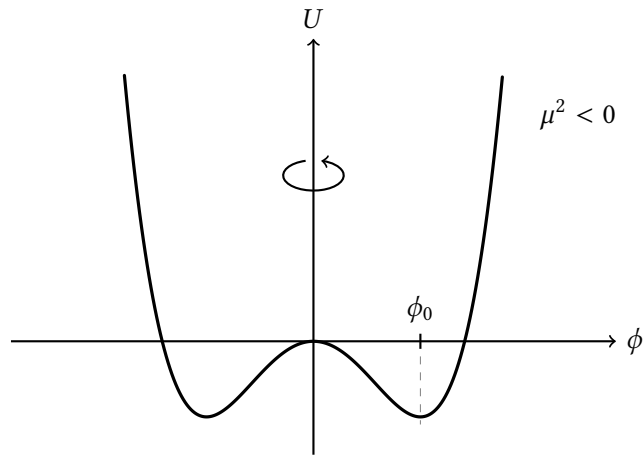


Figure 8.4.: Effective potential in the spontaneously broken phase (SSB)

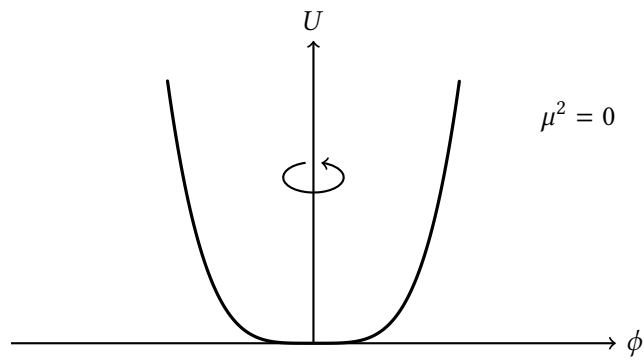


Figure 8.5.: Effective potential at the phase transition

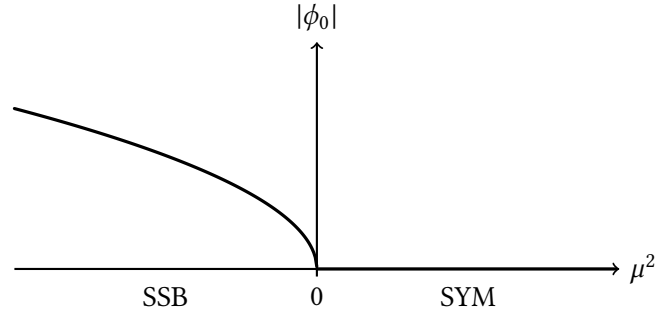
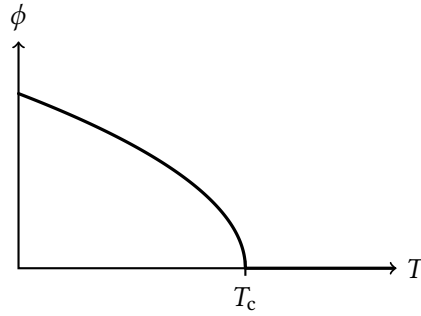
Figure 8.6.: $|\phi_0|$ as a function of μ^2 

Figure 8.7.: Phase diagram of Landau theory

8.3.3. Discontinuity in magnetization for non-zero B

Now, we are considering the general case $j \neq 0$. So, up to quartic order in ϕ , \bar{U} is given by

$$\bar{U} = \frac{1}{2}\mu^2\phi^2 + \frac{1}{8}\phi^4 - j\phi, \quad (8.67)$$

which we have plotted in Figure 8.8 for different cases of j . So, we are looking for the solutions of

$$0 = \frac{\partial \bar{U}}{\partial \phi} = \mu^2\phi + \frac{1}{2}\phi^3 - j. \quad (8.68)$$

This is a cubic equation, i. e. there are either one or three extrema.

The position ϕ of the absolute minimum is called the order parameter. In this case, when varying B (i. e. j), there is a jump in this order parameter, a discontinuity, as you can see from Figure 8.9. What is the origin of this? The physical phenomenon behind this is hysteresis.

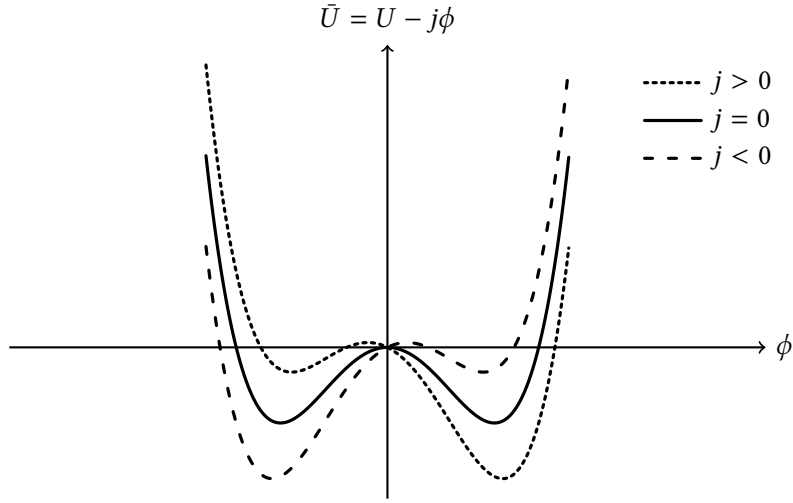
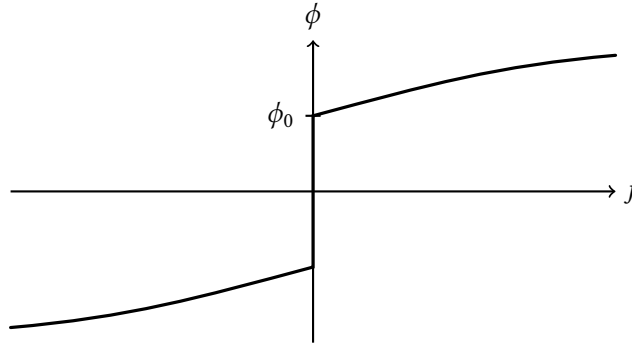
Figure 8.8.: Modified effective potential \bar{U} for different cases of j 

Figure 8.9.: Jump of the order parameter

8.4. Magnetic susceptibility

The magnetic susceptibility is the isothermal change of magnetization with magnetic field.

$$N = 1$$

It is given by

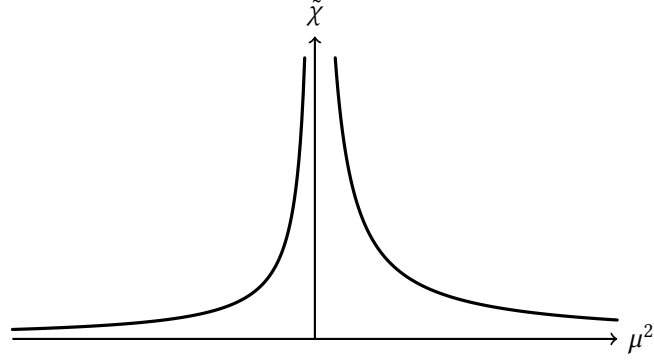
$$\chi_T = \left. \frac{\partial \phi}{\partial B} \right|_T, \quad (8.69)$$

or equivalently,

$$\tilde{\chi} = \frac{\partial \phi}{\partial j}. \quad (8.70)$$

By inverting it:

$$\tilde{\chi}^{-1} = \frac{\partial j}{\partial \phi} \quad (8.71)$$


 Figure 8.10.: Magnetic susceptibility for $j = 0$

and using $j = \frac{\partial U}{\partial \phi}$, we find ($j = 0$)

$$\tilde{\chi}^{-1} = \left. \frac{\partial^2 U}{\partial \phi^2} \right|_{\phi=\phi_0}. \quad (8.72)$$

In the two different phases this is computed as

$$\text{SYM-phase } (\mu^2 > 0): \quad \tilde{\chi}^{-1} = \mu^2 \quad \rightarrow \quad \tilde{\chi} = \frac{1}{\mu^2}, \quad (8.73)$$

$$\text{SSB-phase } (\mu^2 < 0): \quad \tilde{\chi}^{-1} = \mu^2 + \left. \frac{3}{2} \lambda \phi^2 \right|_{\phi=\phi_0} \quad (8.74)$$

$$= \mu^2 + 3\lambda\rho_0. \quad (8.75)$$

For $j = 0$ the latter can be evaluated, using $\rho_0 = -\frac{\mu^2}{\lambda}$ (for $\mu^2 < 0$):

$$\tilde{\chi}^{-1} = 2\lambda\rho_0 = -2\mu^2 \quad \rightarrow \quad \tilde{\chi} = \frac{1}{2\lambda\rho_0} = -\frac{1}{2\mu^2}. \quad (8.76)$$

So, by combining (8.73) and (8.76), we can plot the magnetic susceptibility for $j = 0$ over the complete μ^2 range in Figure 8.10. We see that it *diverges* at the phase transition ($\mu^2 \rightarrow 0$, $T \rightarrow T_c$). For $j \neq 0$ we would find a finite susceptibility at T_c .

$N > 1$

In this case, we have a “susceptibility tensor”:

$$\tilde{\chi}_{ab} = \frac{\partial \phi_a}{\partial j_b}. \quad (8.77)$$

Defining the “mass matrix”

$$m_{ab}^2 = \frac{\partial^2 U}{\partial \phi_a \partial \phi_b}, \quad (8.78)$$

we find the matrix identity

$$\tilde{\chi} = (m^2)^{-1}. \quad (8.79)$$

This is easily proven by

$$\tilde{\chi}_{ab} m_{bc}^2 = \frac{\partial \phi_a}{\partial j_b} \frac{\partial}{\partial \phi_c} \left(\frac{\partial U}{\partial \phi_b} \right) \quad (8.80)$$

$$= \frac{\partial \phi_a}{\partial j_b} \frac{\partial j_b}{\partial \phi_c} \quad (8.81)$$

$$= \frac{\partial \phi_a}{\partial \phi_c} \quad (8.82)$$

$$= \delta_{ac}. \quad (8.83)$$

For $j = 0$ the mass matrix is given by

$$m_{ab}^2 = \left. \frac{\partial^2 U}{\partial \phi_a \partial \phi_b} \right|_{\phi_0}, \quad (8.84)$$

with

$$U = \mu^2 \rho + \frac{1}{2} \rho^2. \quad (8.85)$$

In the SYM-phase this leads to

$$m_{ab}^2 = \mu^2 \delta_{ab}, \quad (8.86)$$

which has N equal eigenvalues μ^2 . So, the susceptibility tensor then has N equal eigenvalues μ^{-2} . In the SSB-phase on the other hand, it is

$$m_{ab}^2 = (\mu^2 + \lambda \rho_0) \delta_{ab} + \lambda \phi_{0,a} \phi_{0,b}. \quad (8.87)$$

Take

$$\phi_{0,1} = \phi_0, \quad \phi_{0,a \neq 1} = 0 \quad \rightarrow \quad \rho_0 = \frac{1}{2} \phi_0^2, \quad (8.88)$$

which leads to

$$m_{11}^2 = 2\lambda \rho_0, \quad m_{ab}^1 = 0 \quad \text{for } a \neq 1, b \neq 1, \quad (8.89)$$

and therefore

$$m^2 = \begin{pmatrix} 2\lambda \rho_0 & 0 & \cdots & 0 \\ 0 & 0 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}. \quad (8.90)$$

So, there are

$$N - 1 \quad \text{“massless modes” (Goldstone modes) and} \quad (8.91)$$

$$1 \quad \text{“massive mode” (radial mode).} \quad (8.92)$$

This means that in the SSB- phase the susceptibility tensor has $N - 1$ diverging eigenvalues for $j \rightarrow 0, \rho_0 \rightarrow \text{const.} \neq 0$ (N diverging eigenvalues for $j \rightarrow 0, \rho_0 \rightarrow 0$).

In the special case of $N = 2$ this corresponds to a complex field

$$\phi = \frac{1}{\sqrt{2}}(\phi_1 + i\phi_2), \quad (8.93)$$

with

$$\rho = \phi^* \phi = \phi = \frac{1}{2}(\phi_1^2 + \phi_2^2). \quad (8.94)$$

In this case the system experiences superfluidity in the SSB-phase, which is related to the Goldstone mode.

The case of $N = 4$ corresponds to two complex fields. Its relativistic version is the Higgs mechanism (Goldstone bosons become longitudinal components of massive vector bosons in presence of local gauge symmetry).

9. Phase transitions

9.1. Order of phase transition

We speak of phase transitions when we observe non-analytic behaviour of thermodynamic potentials or their derivatives, while varying thermodynamic variables.

9.1.1. First order phase transition

A first order phase transition is characterized by a discontinuous jump of the expectation value of a thermodynamic potential or its derivative. In chapter 9 we already encountered an example of this: the discontinuity of the magnetization if the magnetic field changes direction, as sketched in Figure 8.9 (recall that $M \sim \phi$ and $B \sim j$). In this example, the thermodynamic potential \bar{U} , that needs to be minimized, has two separated local extrema. The absolute minimum changes from one to the other at $B = 0$, as illustrated in Figure 8.8. At $B = 0$ both minima have equal \bar{U} .

Of course, this is only the special case of a magnetic system. In general, first order phase transitions may occur when varying some other quantity of a system, like temperature or pressure. For example, this is the case for the liquid-gas phase transition. There, the absolute minimum of the corresponding thermodynamic potential changes from liquid to gas phase at $T = T_c$, as we will work out later in this chapter. It is illustrated in Figure 9.2.

9.1.2. Second order phase transition

Contrary to a first order phase transition, a second order phase transition is characterized by a continuous behaviour of the expectation value. But susceptibilities or other derivatives diverge. Again, we already encountered an example of this: the magnetization for $j = 0$ ($B = 0$). In that case, the effective potential was given by

$$U = \mu^2(T)\rho + \frac{1}{2}\lambda\rho^2. \quad (9.1)$$

Thereby, $\mu^2(T)$ increases monotonically and $\mu^2(T_c) = 0$. For temperatures T near the critical temperature T_c it behaves like

$$\mu^2 = a(T - T_c) + \dots, \quad (9.2)$$

with some constant a . From this the behaviours of ρ_0 and ϕ_0 follow directly:

$$\rho_0(T) = -\frac{\mu^2(T)}{\lambda} = \frac{a(T_c - T)}{\lambda}, \quad (9.3)$$

$$\phi_0(T) = \sqrt{2\rho_0(T)} = c\sqrt{T_c - T}, \quad (9.4)$$

where $c = \sqrt{\frac{2a}{\lambda}}$. We already had sketched this behaviour in Figure 8.7.

In general, one finds ϕ_0 to behave like a power law near the critical temperature:

$$\phi_0 \sim (T_c - T)^\beta, \quad (9.5)$$

with β being a so-called critical exponent. In this case $\beta = \frac{1}{2}$ (mean field value).

For our example the derivative of ϕ_0 with respect to T is given by

$$\frac{\partial \phi_0}{\partial T} \sim \frac{1}{\sqrt{T_c - T}}. \quad (9.6)$$

Indeed, this diverges for $T \rightarrow T_c$.

Assuming $j \neq 0$ again, the effective potential at T_c is computed as

$$U = \frac{1}{2} \lambda \rho^2, \quad (9.7)$$

and therefore

$$\bar{U} = \frac{1}{8} \lambda (\phi_a \phi_a)^2 - j_a \phi_a. \quad (9.8)$$

What is the form of $\phi_0(j)$ in this case? To compute it, we take $j_a = j \delta_{a1}$, leading to $\phi_{0,a} = \phi_0 \delta_{a1}$. Then, we find

$$j = \frac{\partial U}{\partial \phi_1} = \frac{\lambda}{2} \phi_0^3, \quad (9.9)$$

and therefore

$$\phi_0 \sim j^{1/3}, \quad M \sim B^{1/3}. \quad (9.10)$$

In general, one finds M to behave like

$$M \sim B^{1/\delta}, \quad (9.11)$$

with δ being another critical exponent. In this case $\delta = 3$ (mean field value).

The derivative of M with respect to B is given by

$$\frac{\partial M}{\partial B} \sim B^{-2/3}, \quad (9.12)$$

which diverges for $B \rightarrow 0$.

The origin of all of this is that the minimum of the thermodynamic potential U splits up into two minima at $T = T_c$, as shown in Figures 8.3 to 8.5 (recall that $\mu^2 < 0$, $\mu^2 = 0$ and $\mu^2 > 0$ correspond to $T < T_c$, $T = T_c$ and $T > T_c$, respectively). This is typical for spontaneous symmetry breaking and can occur in other situations, too.

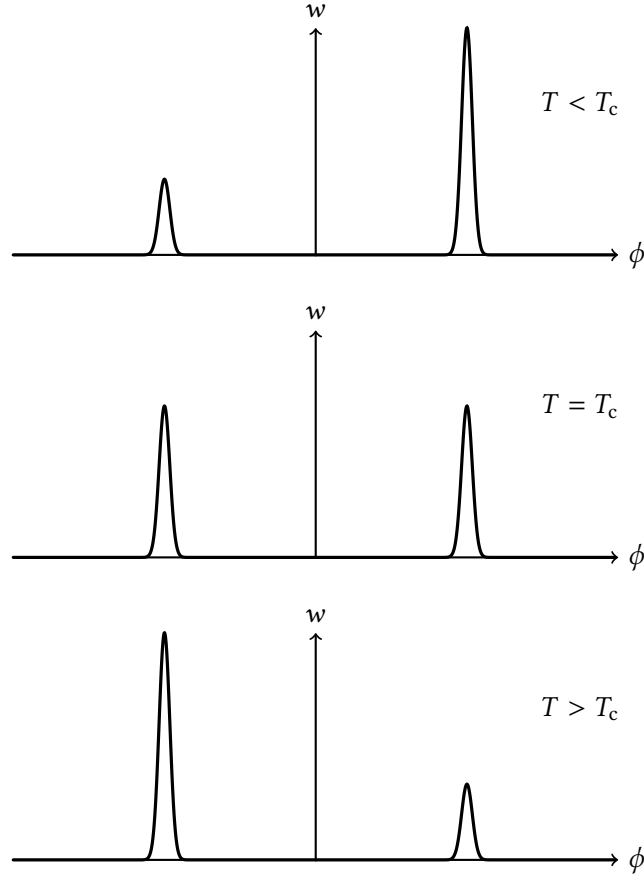


Figure 9.1.: Probability distribution for different temperatures

9.1.3. Probabilistic interpretation

The probability to find some value ϕ is given by

$$w(\phi) \sim e^{-c\bar{U}(\phi)}, \quad (9.13)$$

where $c \rightarrow \infty$ in the thermodynamic limit, as $c \sim V$ or $c \sim \mathcal{N}$ for lattice sites. The two non-degenerate local minima of \bar{U} correspond to two local maxima of the probability distribution, as sketched in Figure 9.1. Thereby, the relative probability of these two maxima depends on the temperature.

“A different state becomes most probable at a first order phase transition.”

There is a striking phenomenon: The microscopic properties of water molecules are the same for $T < T_c$ and $T > T_c$, but a very small change in the average kinetic energy (temperature) produces qualitatively different properties of water and vapor!

Magnetic systems	Real gas
ϕ	v
j	$-p$
U	\tilde{f}
\bar{U}	\tilde{g}

Table 9.1.: Magnetic systems vs. real gas

9.2. Liquid-gas phase transition

9.2.1. Effective potential

The Van der Waals equation of state for real gases is given by

$$p = \frac{k_B N T}{V - b_0 N} - b_1 \left(\frac{N}{V} \right)^2, \quad (9.14)$$

with b_0, b_1 being the so-called virial coefficients. The free energy F is related to the pressure via

$$\frac{\partial F}{\partial V} = -p, \quad (9.15)$$

and has the explicit form

$$F = k_B T N \left[\ln \frac{N}{V - b_0 N} - 1 - \frac{3}{2} \ln \frac{m k_B T}{2 \pi \hbar^2} \right] - b_1 \frac{N^2}{V}. \quad (9.16)$$

In the following we will use the intensive variable $v = \frac{V}{N} = \frac{1}{n}$ (volume per particle). Correspondingly, we define

$$\tilde{f} = \frac{F}{N}, \quad (9.17)$$

such that

$$\frac{\partial \tilde{f}}{\partial v} = -p, \quad (9.18)$$

and

$$\tilde{g}(v, T, p) = \tilde{f}(v, T) + p v. \quad (9.19)$$

\tilde{g} plays the role of \bar{U} for magnetic systems, i. e. the maximal probability $w(v)$ corresponds to the minimum of $\tilde{g}(v)$:

$$\left. \frac{\partial \tilde{g}}{\partial v} \right|_{T, p} = 0. \quad (9.20)$$

In Table 9.1 we have listed the corresponding quantities of magnetic systems and real gases. But note that there is one important difference: there is no $O(N)$ -symmetry for the real gas.

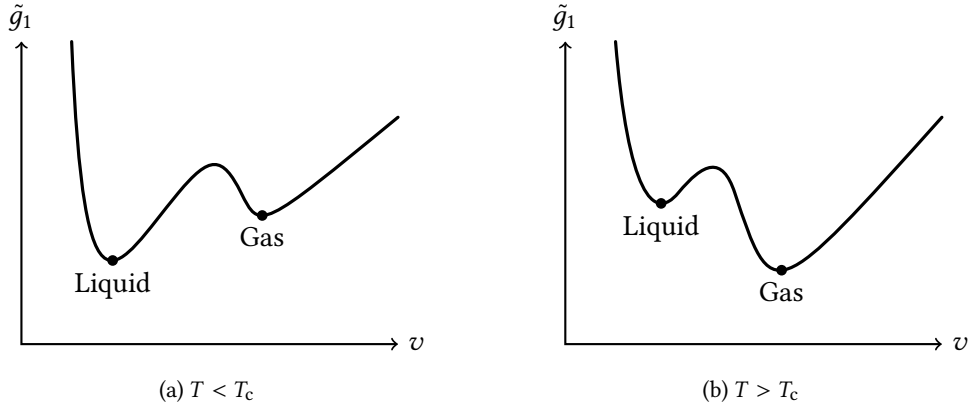
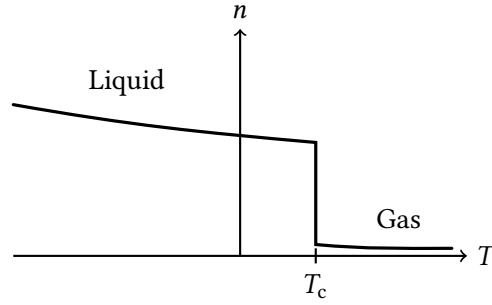
Figure 9.2.: Effective potential \tilde{g}_1 above and below the critical temperature

Figure 9.3.: Jump in density during liquid-gas transition

9.2.2. First order transition

We look for local minima of $\tilde{g}(v)$ for fixed T, p . Therefore, we split up $\tilde{g}(v)$ into a v -dependent and a non- v -dependent part:

$$\tilde{g}(v) = \tilde{g}_0(T) + \tilde{g}_1(v, T, p), \quad (9.21)$$

with $(k_B = \hbar = 1)$

$$\tilde{g}_1 = -T \ln(v - b_0) - \frac{b_1}{v} + pv, \quad (9.22)$$

leading to

$$0 = \frac{\partial \tilde{g}_1}{\partial v} = -\frac{T}{v - b_0} + \frac{b_1}{v^2} + p. \quad (9.23)$$

We find that the equation of state has three solutions for given T, p (2 minima, 1 maximum), as sketched in Figure 9.2. For $T = T_c$ both local minima have equal \tilde{g}_1 . Moreover, there is a jump in density at $T = T_c$, as shown in Figure 9.3.

9.2.3. Compressibility

The isothermal compressibility κ_T is defined as

$$\frac{1}{\kappa_T} = -V \left. \frac{\partial p}{\partial V} \right|_{T,N}. \quad (9.24)$$

Note that $\frac{\partial p}{\partial V} < 0$, as larger pressure corresponds to smaller volume. Therefore $\kappa_T > 0$ (if $\kappa_T < 0$, this would be an unstable situation).

So, in terms of intensive variables,

$$\frac{1}{\kappa_T} = -v \frac{\partial p}{\partial v}, \quad (9.25)$$

and therefore

$$-\left. \frac{\partial p}{\partial v} \right|_T = \frac{\partial}{\partial v} \left(-\frac{T}{v - b_0} + \frac{b_1}{v^2} \right) \Big|_T. \quad (9.26)$$

By comparing this to

$$\left. \frac{\partial^2 \tilde{g}}{\partial v^2} \right|_{T,p} = \frac{\partial}{\partial v} \left(-\frac{T}{v - b_0} + \frac{b_1}{v^2} \right) \Big|_T, \quad (9.27)$$

we find

$$\frac{1}{\kappa_T} = v \left. \frac{\partial^2 \tilde{g}}{\partial v^2} \right|_{v_0}. \quad (9.28)$$

As $\kappa_T > 0$, this corresponds to a minimum of \tilde{g} . Moreover, we conclude that $\kappa_T v$ corresponds to the susceptibility in magnetic systems.

9.2.4. Chemical potential

The chemical potential is given by

$$\mu = \tilde{g}(v_0), \quad (9.29)$$

as one can see as follows:

$$\tilde{g} = \tilde{f} + p v_0 \quad (9.30)$$

$$= \frac{1}{N} (F + pV) \quad (9.31)$$

$$= \frac{1}{N} (F - J_{\text{Gibbs}}) \quad (9.32)$$

$$= \frac{1}{N} (\mu N) \quad (9.33)$$

$$= \mu. \quad (9.34)$$

Recall that

$$J_{\text{Gibbs}} = F - \mu N = -pV \quad (9.35)$$

and note that $N\tilde{g}$ is the free enthalpy

$$G(N, T, p) = F + pV. \quad (9.36)$$

We are considering chemical reactions where N atoms of a given species can be bound either in molecule A or molecule B. The corresponding variable is the concentration, e. g.

$$c_B = \frac{\text{\#molecules B}}{\text{\#molecules A} + \text{\#molecules B}}. \quad (9.37)$$

It replaces v . Then, the probability to find some value c_B is proportional to $e^{-\frac{N\tilde{g}(c_B)}{T}}$.

9.3. Phase diagram

9.3.1. Vapor pressure curve

The vapor pressure curve describes the coexistence of fluid and gas at the critical value of (p, T) . There are two conditions:

1. $\frac{\partial \tilde{g}}{\partial v} = 0$, from which we get the equation of state and the solution for the two minima v_1, v_2 as a function of p ,
2. $\tilde{g}(v_1) = \tilde{g}(v_2)$, which is equivalent to $\mu_l = \mu_g$ (liquid-gas coexistence condition) and from which we get the vapor pressure curve $p_c(T)$.

So, there is only *one* free parameter (for fixed N, V), e. g. T or p :

$$\text{given } p \rightarrow T_c, \quad (9.38)$$

$$\text{given } T \rightarrow p_c. \quad (9.39)$$

9.3.2. $p - T$ diagram

The free enthalpies of the two phases are

$$G_l = N_l \mu_l(p, T), \quad G_g = N_g \mu_g(p, T). \quad (9.40)$$

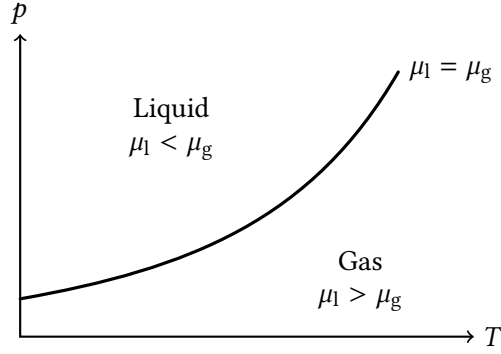
We distinguish the following cases:

$$\mu_l < \mu_g : \quad \text{all particles in liquid phase } (N_l = N, N_g = 0), \quad (9.41)$$

$$\mu_l > \mu_g : \quad \text{all particles in gas phase } (N_l = 0, N_g = N), \quad (9.42)$$

$$\mu_l = \mu_g : \quad \text{transition line } (N_l, N_g \neq 0, N_l + N_g = N). \quad (9.43)$$

The transition line is described by $p = p_c(T)$, which we have sketched in Figure 9.4. So, coexistence in equilibrium is possible. Note that this is a universal curve, as it is independent of N_l, N_g .


 Figure 9.4.: $p - T$ transition line between liquid and gas phase

9.3.3. Clausius-Clapeyron relation

We can set up a differential equation for the transition line by combining the coexistence conditions of two infinitesimally distant points on the transition line (changing the index notation from l,g to 1,2):

$$\mu_1(T, p) = \mu_2(T, p), \quad (9.44)$$

$$\mu_1(T + dT, p + dp) = \mu_2(T + dT, p + dp), \quad (9.45)$$

to arrive at

$$d\mu_1 = d\mu_2, \quad (9.46)$$

with

$$d\mu_i = \underbrace{\frac{\partial \mu_i}{\partial T}}_{-\frac{S}{N}} dT + \underbrace{\frac{\partial \mu_i}{\partial p}}_{\frac{V}{N}} dp. \quad (9.47)$$

The latter is the so-called Duhem-Gibbs relation:

$$d\mu = \frac{V}{N} dp - \underbrace{\frac{S}{N}}_{\tilde{s}} dT. \quad (9.48)$$

When expressing it in terms of intensive quantities and combining it with (9.46) we arrive at

$$v_1 dp - \tilde{s}_1 dT = v_2 dp - \tilde{s}_2 dT, \quad (9.49)$$

which we can rewrite as

$$\underbrace{(\tilde{s}_2 - \tilde{s}_1)}_{\Delta \tilde{s}} dT = \underbrace{(v_2 - v_1)}_{\Delta v} dp, \quad (9.50)$$

to find the Clausius-Clapeyron equation:

$$\frac{dp_c}{dT} = \frac{\Delta \tilde{s}}{\Delta v}. \quad (9.51)$$

It tells us that the slope of $p_c(T)$ is given by the ratio of entropy difference and volume difference.

9.3.4. Latent heat

There is a jump in entropy at the phase transition. So, heat has to be absorbed or released (the phase transition can also be adiabatic if the change in T is slow). It is called “latent heat” and is given by

$$L_{12} = T\Delta S = T(S_2 - S_1) . \quad (9.52)$$

Combining it with the Clausius-Clapeyron equation gives us

$$L_{12} = \frac{dT}{dp_c} T \Delta V . \quad (9.53)$$

The latent heat is necessary to separate the water molecules during the transition from the liquid to the gas phase or to break up the crystal structure during the transition from the ice (solid) to the liquid phase, respectively.

9.3.5. Superheating, supercooling

Let us introduce the following time scales:

t_A : typical time scale of establishing the equilibrium in the liquid (relaxation time),

t_B : typical time scale of establishing the equilibrium in the gas,

t_T : time scale of the liquid-gas transition (forming of droplets / bubbles).

For $T < T_c$ the probability to be in the gas phase is much smaller than the probability to be in the liquid phase:

$$w_B \ll w_A , \quad (9.54)$$

as sketched in Figure 9.5. But for the states in between these two phases it is even smaller:

$$w(v) \ll w_B \quad \text{for } v_A < v < v_B . \quad (9.55)$$

Therefore, small fluctuations near B, e. g. at x , tend to return to B (stability of B regarding small fluctuations). An equivalent explanation would be to say that B is a local minimum of G , even if $T < T_c$, as illustrated in Figure 9.6.

The typical time scale of a fluctuation returning to B is given by t_B , which is usually much shorter than t_T (overcoming a barrier; similar to tunneling in QM: exponentially suppressed probabilities). Let us now consider an experimental time scale t_E in between these:

$$t_B \ll t_E \ll t_T . \quad (9.56)$$

In this case, the system has not enough time for the transition $B \rightarrow A$, but instead stays in the “local” equilibrium B even beyond the critical point. This phenomenon is called supercooling (one speaks of superheating if during the transition $A \rightarrow B$ the systems stays in A beyond the critical point). Eventually, the local instability will become extremal when B becomes a saddle point and finally a maximum. In practice one can observe the coexistence of water and vapor not only at T_c .

For magnets the same effect appears in the form of hysteresis. Indeed, this is a characteristic effect for *first* order transitions in general.

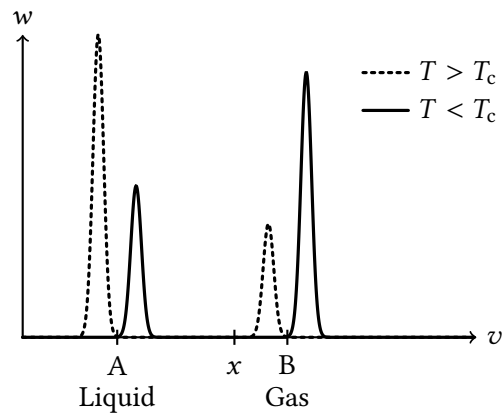


Figure 9.5.: Probability distributions for liquid and gas phase

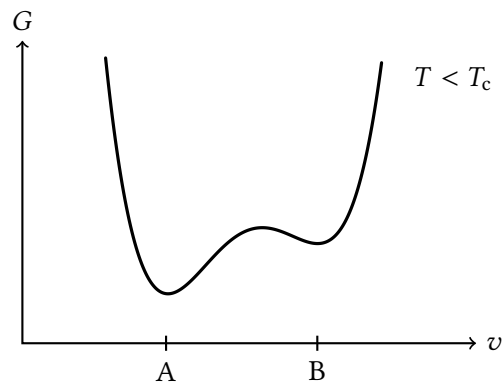


Figure 9.6.: Free enthalpy in liquid phase

10. Critical Phenomena

10.1. Critical exponents

For second order phase transitions the effective potential describes correctly the qualitative behaviour of the system in the regime $T \rightarrow T_c$. But a closer examination of the singularities yields values for the critical exponents, which deviate from the simple “mean field” approximation we used in subsection 9.1.2.

In the following we assume $B = 0$ for magnetic systems or $p = p_*$, the pressure at the endpoint of the critical line, for liquid-gas transitions. These are just the conditions to observe second order phase transitions. Moreover, we parametrize the behaviour for $T \rightarrow T_c$ by power laws, with “critical exponents”.

10.1.1. Specific heat

The power law for the specific heat reads

$$c(T) = c_{\text{reg}} + A |T - T_c|^{-\alpha}, \quad (10.1)$$

with c_{reg} being the simplest approximation expression for the specific heat, α being a critical exponent, and A being some constant. We have plotted (10.1) in Figure 10.1 for different cases of α .

10.1.2. Order parameter as function of temperature

The order parameter of magnetic systems is the magnetization per lattice site $\phi = \frac{M}{N}$. Its temperature dependence was illustrated in Figure 8.7. Close to the critical temperature one has

$$\left| \frac{M}{N} \right| \sim (T_c - T)^\beta \quad (10.2)$$

for $T < T_c$ (compare (9.5)) and $M = 0$ for $T > T_c$. Here, β is another critical exponent.

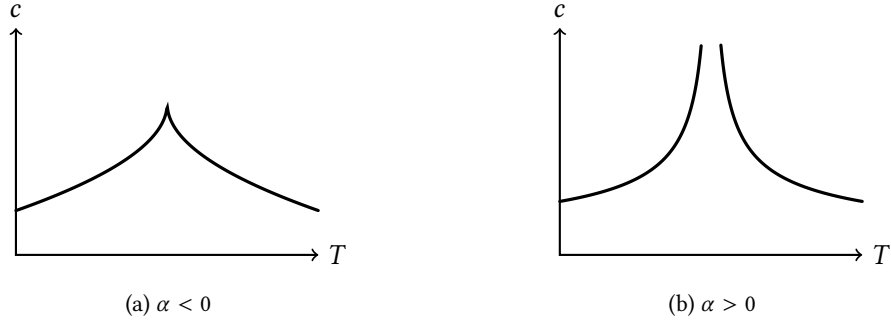
10.1.3. Susceptibility

The susceptibility χ in magnetic systems corresponds to the compressibility κ_T in fluids. At $B = 0$ / $p = p_*$ both behave the same close to the critical temperature:

$$\chi \sim |T - T_c|^{-\gamma}, \quad (10.3)$$

$$\kappa_T \sim |T - T_c|^{-\gamma}. \quad (10.4)$$

This hold for $T > T_c$ as well as for $T < T_c$. They both share the same critical exponent γ .

Figure 10.1.: Specific heat for different values of α

10.1.4. Order parameter at critical temperature as function of the “source”

We can take a look at the “source”-dependence of the order parameters, too. For magnetic systems this source is the external magnetic field B , for fluids it is the pressure p (or rather its deviation from the critical value p_*). Again, both types of systems show the same behaviour:

$$|\Delta n(p)| \sim |p - p_*|^{1/\delta}, \quad (10.5)$$

$$\left| \frac{M}{N}(B) \right| \sim |B|^{1/\delta}, \quad (10.6)$$

with some critical exponent δ . For magnetic systems we already found this relation in (9.11).

10.1.5. Mean field exponents

Let us first calculate / collect the mean field values of the critical exponents. In general, they are extracted from an effective potential of the form

$$U = \frac{1}{2} A (T - T_c) \phi^2 + \frac{\lambda}{8} \phi^4 + U_0, \quad (10.7)$$

as already used in (9.1). We will need its first two derivatives:

$$\frac{\partial U}{\partial \phi} = A (T - T_c) \phi + \frac{\lambda}{2} \phi^3, \quad (10.8)$$

$$\frac{\partial^2 U}{\partial \phi^2} = A (T - T_c) + \frac{3}{2} \phi^2. \quad (10.9)$$

The first derivative also directly gives us the position of the minimum for SSB:

$$\frac{1}{2} \phi_0^2 = \frac{A (T_c - T)}{\lambda}, \quad (10.10)$$

as we already had found in (9.4). For SYM it is $\phi_0 = 0$.

1) α

The specific heat is defined as

$$c = -T \frac{\partial^2 F}{\partial T^2}, \quad (10.11)$$

where

$$F = T N \bar{U}(\phi_0), \quad (10.12)$$

as we had found in (8.44). So, by inserting ϕ_0 into U , we find

$$F \sim T U(\phi_0) = \begin{cases} T U_0 & \text{for SYM,} \\ T U_0 - \frac{1}{2} \frac{A^2 T}{\lambda} (T_c - T)^2 & \text{for SSB.} \end{cases} \quad (10.13)$$

And by applying the definition of c , we arrive at

$$c \sim \begin{cases} c_{\text{reg}} & \text{for SYM,} \\ c_{\text{reg}} - \frac{A^2 T}{\lambda} (3T - 2T_c) & \text{for SSB,} \end{cases} \quad (10.14)$$

with

$$c_{\text{reg}} \sim T^2 \frac{\partial^2 U_0}{\partial T^2} + T \frac{\partial U_0}{\partial T}. \quad (10.15)$$

There appears no $T - T_c$ in the expression for c , so $\alpha = 0$. Or, in other words, c has no pole at $T = T_c$. But there is a jump in specific heat at the phase transition:

$$c_{\text{SYM}} - c_{\text{SSB}} \Big|_{T_c} = \frac{A^2 T_c}{\lambda}. \quad (10.16)$$

2) β

From

$$\phi_0 \sim (T_c - T)^{1/2} \quad (10.17)$$

we can read off $\beta = \frac{1}{2}$.

3) γ

By combining (8.72) and (10.9) we find

$$\chi \sim \left(\frac{\partial^2 U}{\partial \phi^2} \right)^{-1} \Big|_{\phi_{hi_0}} \sim (T - T_c)^{-1}. \quad (10.18)$$

So, it follows $\gamma = 1$.

4) δ

At $T = T_c$ the effective potential is given by

$$U = \frac{\lambda}{8} \phi^4. \quad (10.19)$$

Moreover, we know that

$$\frac{\partial U}{\partial \phi}(\phi_0) = j. \quad (10.20)$$

So, we can immediately conclude that $\phi_0 \sim j^{1/3}$ and therefore $\delta = 3$.

We see that these mean field exponents are independent of the dimension d or the number of components N (in case of $O(N)$ -symmetry). Experimental results show that the actual values of the critical exponents differ from the mean field values and that they depend on d and N . But otherwise they are *universal*! For example, one finds, even experimentally, the same exponents for liquid-gas transitions and uniaxial magnets. Note that both are described by an Ising model with $N = 1$.

So, we conclude that the critical exponents are independent of the precise details of a system (e. g. the exact form of the equation of state for fluids, i. e. b_0, b_1 and so on). Furthermore, they are the same for different types of systems, if these share the same dimension and symmetries (d, N). We say, they are in the same “Ising universality class”. So, for example, liquid-gas transitions, uniaxial magnets and many other systems belong to the same class.

As these universality classes only depend on d and N , there are only few of them. But why? Why are 2 parameters sufficient to fix four critical exponents? And why do they only depend on d and N ? The understanding of this needs the profound concept of “renormalization”. Part of the answer arises from the existence of additional “scaling laws” that relate the critical exponents, e. g.

$$\beta = \frac{\gamma}{\delta - 1}, \quad (10.21)$$

$$2 = \alpha + 2\beta + \gamma. \quad (10.22)$$

The concepts which are necessary to understand where the scaling relations come from and to understand this universal behaviour of systems with equal symmetries, in general, have been developed by Kadanoff, Wilson and Wegner. We will introduce these concepts in the following sections.

10.1.6. Universal critical exponents

In Table 10.1 we have summarized the mean field results of the previous subsection and compare them to more sophisticated theoretical calculations as well as to experimental values. Note that the results of the two-dimensional Ising model are exact. It is the only interacting model that has been solved exactly, so far. This solution was found by Onsager in 1944.

	α	β	γ	δ
Theoretical results:				
Mean field	0	0.5	1	3
Ising model in $d = 2$	0	0.125	≈ 1.75	≈ 15
Renormalization group in $d = 3$				
$n = 1$	0.11	0.3249	1.2371(4)	4.817
$n = 3$	-0.1159	0.3647	1.3866	4.802
Experimental results:				
Liquid-gas / uniaxial magnet ($n = 1$)	0.1	0.34	1.2	4.4
Isotropic magnet ($n = 3$)	-0.15	0.36	1.25	4.3

Table 10.1.: Universal critical exponents

10.2. Correlation functions

10.2.1. Correlation length

The question we would now like to answer is: What does the density at one point \mathbf{r}_1 know about the density at another point \mathbf{r}_2 , as the local density fluctuates even in equilibrium? And from which distance on are $n(\mathbf{r}_1)$ and $n(\mathbf{r}_2)$ uncorrelated, i. e. what is the correlation length of the system? For magnetic systems we would ask the same questions for the magnetisation, but we will stick to the fluid notation in the following.

To answer these questions, let us look at the correlation function (2-point function):

$$G(\mathbf{r}_1, \mathbf{r}_2) = \left\langle \left[n(\mathbf{r}_1) - \bar{n} \right] \left[n(\mathbf{r}_2) - \bar{n} \right] \right\rangle \quad (10.23)$$

$$= \left\langle n(\mathbf{r}_1) n(\mathbf{r}_2) \right\rangle - \bar{n}^2. \quad (10.24)$$

Note that, due to the translational and rotational symmetry of our system, G can only depend on the distance between two points:

$$G(\mathbf{r}_1, \mathbf{r}_2) = G(|\mathbf{r}_1 - \mathbf{r}_2|) = G(|\mathbf{r}|) = G(r), \quad (10.25)$$

with $\mathbf{r} = \mathbf{r}_1 - \mathbf{r}_2$. Often, the correlation function shows a very typical behaviour for large r :

$$G(r) \sim e^{-\frac{r}{\xi}}, \quad (10.26)$$

where ξ is the correlation length.

$G(r)$ is the central quantity to describe the local distribution of fluctuations – regardless of whether one deals with the amorphous structure of polymers or the galaxy distribution in the universe, for example. Moreover, it has to play an important role in phase transitions. As we have already seen, second order phase transitions correspond to a order-disorder transition, e. g.

$$T > T_c : \quad M = 0, \quad (10.27)$$

$$T < T_c : \quad |M| > 0 \quad (10.28)$$

in magnetic systems at $B = 0$. To make this transition possible, the spins have to “know” of each other, i. e. they have to be correlated. But not only that, the correlation length has to be a function of temperature, too, to explain these different behaviours above and below the critical temperature.

10.2.2. Correlation function and generating functional

Let us assume a very general form of a partition function for a magnetic system:

$$Z[j(\mathbf{r})] = \sum_{\tau} e^{-\beta H[\phi(\mathbf{r})] + \sum_{\mathbf{r}} \phi(\mathbf{r}) j(\mathbf{r})}, \quad (10.29)$$

where $j(\mathbf{r})$ are artificial *local* sources. They resemble an arbitrary inhomogeneous magnetic field. Moreover, $H[\phi(\mathbf{r})]$ and $Z[j(\mathbf{r})]$ are functionals, i. e. functions of functions. Note that they are not just functions of some constant ϕ or j .

Now, the so-called generating functional is defined as

$$W = \ln Z. \quad (10.30)$$

Its first derivative with respect to $j(\mathbf{r})$ is the expectation value of ϕ :

$$\frac{\partial W}{\partial j(\mathbf{r})} = \frac{1}{Z} \sum_{\tau} \phi(\mathbf{r}) e^{-\beta H + \sum \phi j} \quad (10.31)$$

$$= \langle \phi(\mathbf{r}) \rangle, \quad (10.32)$$

and its second derivative is exactly the 2-point correlation function:

$$\frac{\partial^2 W}{\partial j(\mathbf{r}) \partial j(\mathbf{r}')} = - \frac{1}{Z^2} \left[\sum_{\tau} \phi(\mathbf{r}) e^{-\beta H + \sum \phi j} \right] \left[\sum_{\tau} \phi(\mathbf{r}') e^{-\beta H + \sum \phi j} \right] \quad (10.33)$$

$$+ \frac{1}{Z} \sum_{\tau} \phi(\mathbf{r}) \phi(\mathbf{r}') e^{-\beta H + \sum \phi j}$$

$$= \langle \phi(\mathbf{r}) \phi(\mathbf{r}') \rangle - \langle \phi(\mathbf{r}) \rangle \langle \phi(\mathbf{r}') \rangle \quad (10.34)$$

$$= \langle [\phi(\mathbf{r}) - \langle \phi(\mathbf{r}) \rangle] [\phi(\mathbf{r}') - \langle \phi(\mathbf{r}') \rangle] \rangle \quad (10.35)$$

$$= G(\mathbf{r}, \mathbf{r}'). \quad (10.36)$$

If we now demand translational symmetry, it is

$$\langle \phi(\mathbf{r}) \rangle = \bar{\phi}, \quad (10.37)$$

$$G(\mathbf{r}, \mathbf{r}') = G(\mathbf{r} - \mathbf{r}'). \quad (10.38)$$

And if we further demand symmetry under spatial rotations, it is again

$$G(\mathbf{r}, \mathbf{r}') = G(|\mathbf{r} - \mathbf{r}'|). \quad (10.39)$$

10.2.3. Correlation function and susceptibility

Now, let us look at what happens, if we replace $j(\mathbf{r})$ by j :

$$j(\mathbf{r}) \rightarrow j, \quad (10.40)$$

with j being independent of \mathbf{r} . Then, the first two derivatives of W with respect to j yield

$$\frac{\partial W}{\partial j} = \sum_{\mathbf{r}} \langle \phi(\mathbf{r}) \rangle \quad (10.41)$$

$$= \mathcal{N} \bar{\phi}, \quad (10.42)$$

$$\frac{\partial^2 W}{\partial j^2} = \sum_{\mathbf{r}, \mathbf{r}'} \langle [\phi(\mathbf{r}) - \bar{\phi}] [\phi(\mathbf{r}') - \bar{\phi}] \rangle \quad (10.43)$$

$$= \sum_{\mathbf{r}, \mathbf{r}'} G(\mathbf{r}, \mathbf{r}') \quad (10.44)$$

$$= \mathcal{N} \sum_{\mathbf{r}} G(\mathbf{r}, 0) \quad (10.45)$$

$$= \mathcal{N} \sum_{\mathbf{r}} G(\mathbf{r}), \quad (10.46)$$

where we used translational and rotational symmetry in the last two steps.

On the other hand, the susceptibility is defined as

$$\chi = \frac{\partial \bar{\phi}}{\partial j} \quad (10.47)$$

$$= \frac{\partial}{\partial j} \frac{1}{\mathcal{N}} \frac{\partial W}{\partial j} \quad (10.48)$$

$$= \frac{1}{\mathcal{N}} \frac{\partial^2 W}{\partial j^2}. \quad (10.49)$$

So, the susceptibility and the correlation function are related via

$$\chi = \sum_{\mathbf{r}} G(\mathbf{r}). \quad (10.50)$$

While we have considered a magnetic system, so far, these calculations can be performed completely analogously for fluids. In that case, one ends up with a relation between the compressibility and the correlation function:

$$\kappa_T = \frac{1}{n^2 T} \int d^3 r G(\mathbf{r}). \quad (10.51)$$

10.2.4. Critical exponent of correlation length and anomalous dimension

As we have already mentioned, at large r the correlation function behaves like an exponential cutoff:

$$\lim_{r \rightarrow \infty} G(r) \sim e^{-\frac{r}{\xi}}. \quad (10.52)$$

So, using the relations between susceptibility / compressibility and the correlation function, we can conclude that χ and κ_T are finite as long as ξ remains finite. But as we have seen in section 8.4, the susceptibility diverges at the critical temperature (e. g. recall Figure 8.10). The same holds for the compressibility. We can immediately infer that the correlation length has to diverge at $T = T_c$, too:

$$\chi, \kappa_T \xrightarrow{T \rightarrow T_c} \infty, \quad (10.53)$$

$$\xi \xrightarrow{T \rightarrow T_c} \infty. \quad (10.54)$$

To describe this behaviour, we introduce another critical exponent ν :

$$\xi \sim |T - T_c|^{-\nu}. \quad (10.55)$$

In the mean field approximation one finds $\nu = \frac{1}{2}$. Moreover, the large r limit of G at $T = T_c$ can, in general, be expressed in the following way:

$$\lim_{r \rightarrow \infty} G(r) \sim r^{-(d-2+\eta)}, \quad (10.56)$$

with the so-called “anomalous dimension” η being yet another critical exponent. Its mean field value is $\eta = 0$.

10.2.5. Critical opalescence

Considering a fluid, density fluctuations will lead to fluctuations in the refractive index. It follows that light with a wavelength close to the correlation length will experience stronger scattering. As we have seen that the correlation length becomes large in the vicinity of a second order phase transition, it will eventually reach the wavelength of visible light, causing the usually transparent fluid to appear cloudy. This phenomenon is known as critical opalescence.

10.3. Universality

10.3.1. Macroscopic laws for second order phase transition

At second order phase transitions the correlation length diverges, leading to long-range phenomena. This is the reason for universality: the long-range behaviour does *not* depend on details of the microscopic interactions, any more, except for very few properties, like the dimension, the symmetries and the number of relevant degrees of freedom. These properties determine the universality class of the system.

We have seen that the mean field theory is not capable of describing this long-range behaviour properly. But nevertheless, exact statements are possible! One finds laws which are only loosely connected to the microphysics and mainly arise from the statistics of many body systems. These effective laws strongly depend on the relevant distance scale. For example, particle and atomic physics are suitable to describe systems at very small scales, whereas these systems are better described by the laws of chemistry and biology at way larger, macroscopic scales.

So, what we are looking for, is a framework to explain this qualitatively different behaviour on different scales – to describe the transition from microscopic laws on small scales to macroscopic laws on large scales. This framework is the so-called *renormalization group*. It describes the distance scale dependence of the effective laws and makes it possible to cast the qualitative description of these central ideas into a mathematically rigorous form.

In the following subsections, we will introduce this concept via the block spin description, developed by Kadanoff.

10.3.2. Ising model in one dimension

Let us first look at the Ising model in one dimension. Using the abbreviation

$$\beta J = K, \quad (10.57)$$

its partition function takes the form

$$Z = \sum_{\{s_i\}} e^{K \sum_{i=1}^N s_i s_{i+1}} \quad (10.58)$$

$$= \sum_{\{s_i\}} \prod_{i=1}^N e^{K s_i s_{i+1}}. \quad (10.59)$$

Now, we introduce the concept of a block lattice. Therefore, we relabel the spins in the following way:

$$\bar{s}_j : \quad j = \frac{i+1}{2} \quad \text{for } i \text{ odd} \quad (\text{spins on the block lattice, “block-spins”}), \quad (10.60)$$

$$s'_j : \quad j = \frac{i}{2} \quad \text{for } i \text{ even} \quad (\text{intermediate spins}). \quad (10.61)$$

It is illustrated in Figure 10.2.

The idea now is to “integrate out” the intermediate spins, to obtain a new Ising model for the block spins \bar{s}_j . We achieve this by always grouping two subsequent factors of the product in (10.59) into one block factor. This way, we arrive at a product of blocks:

$$Z = \sum_{\{s_i\}} \prod_{j=1}^{N/2} \left(e^{K \bar{s}_j s'_j} e^{K \bar{s}_{j+1} s'_j} \right) \quad (10.62)$$

$$= \sum_{\{s_i\}} \prod_j e^{K (\bar{s}_j s'_j + \bar{s}_{j+1} s'_j)} \quad (10.63)$$

$$= \sum_{\{s_i\}} \prod_j e^{K s'_j (\bar{s}_j + \bar{s}_{j+1})}. \quad (10.64)$$

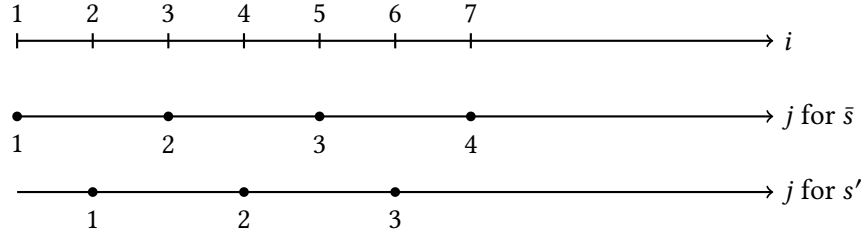


Figure 10.2.: One-dimensional block lattice

Now, we can split the configuration sum, as s'_j only appears in block j :

$$Z = \sum_{\{\bar{s}_j\}} \prod_j \left(\sum_{s'_j=\pm 1} e^{K s'_j (\bar{s}_j + \bar{s}_{j+1})} \right). \quad (10.65)$$

Next, we perform the inner sum (“integrate out the intermediate spins”):

$$Z = \sum_{\{\bar{s}_j\}} \prod_j \left(e^{K(\bar{s}_j + \bar{s}_{j+1})} + e^{-K(\bar{s}_j + \bar{s}_{j+1})} \right). \quad (10.66)$$

This way, we arrive at a new Ising model for the block spins:

$$Z = \sum_{\{\bar{s}_j\}} \prod_j e^{K' \bar{s}_j \bar{s}_{j+1} + C}, \quad (10.67)$$

with a suitably defined K' and some constant C .

To determine K' , we use the condition that for all possible spin configurations the two expressions (10.66) and (10.67) for Z have to be identical:

$$(\bar{s}_j, \bar{s}_{j+1}) : \quad e^{K(\bar{s}_j + \bar{s}_{j+1})} + e^{-K(\bar{s}_j + \bar{s}_{j+1})} = e^{K' \bar{s}_j \bar{s}_{j+1} + C}, \quad (10.68)$$

$$++ : \quad e^{2K} + e^{-2K} = e^{K' + C}, \quad (10.69)$$

$$+- : \quad 2 = e^{-K' + C}, \quad (10.70)$$

$$-+ : \quad 2 = e^{-K' + C}, \quad (10.71)$$

$$-- : \quad e^{-2K} + e^{2K} = e^{K' + C}. \quad (10.72)$$

Inverting (10.70) leads to

$$\frac{1}{2} = e^{K' - C}, \quad (10.73)$$

and multiplying this with (10.69) gives us

$$\frac{1}{2} (e^{2K} + e^{-2K}) = e^{2K'}. \quad (10.74)$$

So, our new coupling K' is given by

$$K' = \frac{1}{2} \ln \left[\frac{1}{2} (e^{2K} + e^{-2K}) \right] = \frac{1}{2} \ln \cosh 2K. \quad (10.75)$$

The additive constant C is irrelevant and can be neglected. With that, (10.67) has exactly the same form as (10.59). So, if the original lattice had lattice spacing a , the new partition function is describing a lattice with lattice spacing $2a$ and a coupling K' . But, of course, both expressions of the partition function are the same – we are still describing the same model! In other words, (10.67) is describing the original model, just on a larger length scale. And on this length scale it behaves like a model on a lattice with doubled lattice spacing and *effective* coupling K' :

Renormalization: $K \rightarrow K'$, Lattice spacing: $a \rightarrow 2a$.	(10.76)
--	---------

Repeating the renormalization step (10.76) leads to a length scale dependent $K(l)$:

$$K \rightarrow K' \rightarrow K'' \rightarrow K''' \rightarrow \dots \quad (10.77)$$

$$K(a) \rightarrow K(2a) \rightarrow K(4a) \rightarrow K(8a) \rightarrow \dots \quad (10.78)$$

This procedure describes the renormalization group. It is a group, whose elements are all possible sequences of renormalization steps. Actually, if you want to be mathematically more rigorous, it is just a semigroup, as it does not contain the inverses of its elements.

Now, let us look at the different cases for K . If $K > 0$, then

$$e^{-2K} < e^{2K} , \quad (10.79)$$

so,

$$e^{2K'} = \frac{1}{2} (e^{2K} + e^{-2K}) < e^{2K} , \quad (10.80)$$

and therefore,

$$0 < K' < K , \quad (10.81)$$

i. e. the coupling decreases as the length scale increases. So, the interaction goes to zero, if one is going to larger and larger length scales:

$$K(l) \xrightarrow{l \rightarrow \infty} 0 . \quad (10.82)$$

This means that it is possible to treat this system perturbatively, as an expansion in K .

In the other case, $K < 0$, we speak of an antiferromagnet. But then (10.75) tells us that $K' > 0$, nevertheless. So, from there on it just behaves the same as the $K > 0$ case.

10.3.3. Ising model in $d > 1$ dimensions

New interactions

If we now consider a lattice in $d > 1$ dimensions, integrating out the microscopic modes (the “intermediate spins”) leads to way more complicated interactions, as we have sketched in Figure 10.3 for $d = 2$. And performing further renormalization steps will lead to even more and more complicated interactions. So, in the following we will have to find a systematic way to do so.

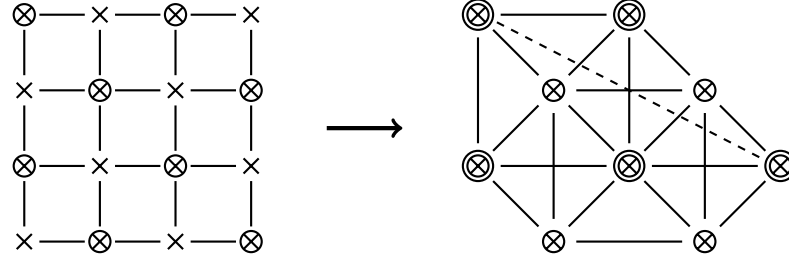


Figure 10.3.: Complicated interactions in two-dimensional block lattice. The symbol \otimes denotes block spins in the first renormalization step. Links denote effective interactions. In addition to next-neighbour interactions one now also finds “diagonal interactions”. A second renormalization step leads to the block spins \otimes . This induces even more complicated interactions, as demonstrated by the dashed link.

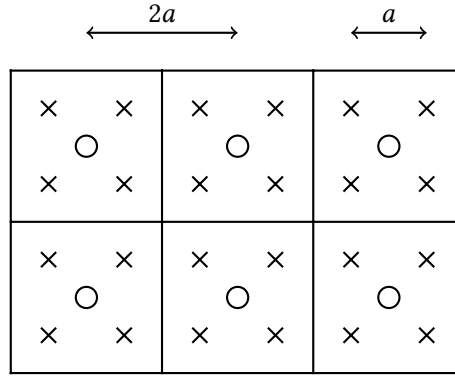


Figure 10.4.: Block spins in two-dimensional lattice

Block spins

Let us generalize the concept of block spins, that we have introduced in the one-dimensional case. A block will now consist out of 2^d spins, as illustrated in Figure 10.4 for $d = 2$. We define the spin of the k -th block as

$$\hat{s}_k = \sum_{i \in \text{block}(k)} \frac{s_i}{\sqrt{2^d}}, \quad (10.83)$$

where we have chosen an arbitrary normalisation. Now, these block spins represent our new degrees of freedom, i. e. we have reduced the number of degrees of freedom.

The partition function is given by

$$Z = \sum_{\{s_i = \pm 1\}} e^{\beta J \sum_{\langle ij \rangle} s_i s_j} \quad (10.84)$$

$$= \sum_{\{s_i = \pm 1\}} e^{-\beta H_0[s_i]}. \quad (10.85)$$

Due to (10.83) we can write

$$\int_{-\infty}^{\infty} d\hat{s}_k \delta\left(\hat{s}_k - 2^{-d/2} \sum_{\text{block}} s_i\right) = 1. \quad (10.86)$$

For each block we can now insert such a 1 into the partition function:

$$Z = \prod_{k \in \text{block lattice}} \int d\hat{s}_k \underbrace{\sum_{\substack{\{s_i = \pm 1\} \\ \text{original lattice}}} \delta\left(\hat{s}_k - 2^{-d/2} \sum_{\text{block}} s_i\right) e^{-\beta H_0[s_i]}}_{e^{-\beta H_1[\hat{s}_k]}} \quad (10.87)$$

$$= \prod_k \int d\hat{s}_k e^{-\beta H_1[\hat{s}_k]}. \quad (10.88)$$

As in the one-dimensional case, we end up with an equivalent partition function, formulated on a block lattice.

Renormalization flow of the coupling

In the same way, as we have done it in the one-dimensional Ising model, we can repeat this blocking procedure, i. e. this renormalization step:

$$H_0 \rightarrow H_1 \rightarrow H_2 \rightarrow H_3 \rightarrow \dots. \quad (10.89)$$

This way, we end up with a so-called renormalization flow $J(L)$ of the coupling, where $L = 2^n a$. Moreover, new couplings are generated by this procedure.

New variables

The sum of the spins in a block k , $\sum_{i \in \text{block}(k)} s_i$, can take values from -2^d to 2^d . For our normalisation (10.83) the spacing between the values \hat{s}_k amounts to $2^{-d/2}$, instead of 1 for the original spins. The range of \hat{s}_k is given by

$$\hat{s}_k \in [-2^{d/2}, 2^{d/2}], \quad (10.90)$$

instead of $s_i \in [-1, 1]$ for the original spins. Thus, both range and spacing of the block variable is different from the original variable.

After many renormalization steps we can replace our block spin by a continuous function ϕ :

$$\hat{s}_k \rightarrow \phi(x), \quad (10.91)$$

where $-\infty < \phi(x) < \infty$.

Potential

The interactions between the spins of one given block spin are described by the potential $U(\hat{s}_k)$. They are called “ultra local” or “on-site” interactions. Due to the symmetry under $\hat{s}_k \rightarrow -\hat{s}_k$ it can only depend on the square of the spin:

$$U(\hat{s}_k) = U(\hat{s}_k^2). \quad (10.92)$$

For the microscopic Ising spins there is only one possible value $s_i^2 = 1$, so there are no on-site interactions; i. e. microscopic spins s_i do not interact with themselves. For block spins, on the other hand, $U(\hat{s}_k^2)$ is not trivial any more, as \hat{s}_k^2 can take values between 0 and 2^d . And in the continuum limit, $U[\phi(x)]$ is a function of a continuous variable ϕ .

Let us denote the potential at step n of the blocking procedure, i. e. at effective lattice spacing $l = 2^n a$, by $U[\phi; l]$. So, the potential changes as a function of length scale. Now, the Hamiltonian at length scale l can be expressed as a potential for the on-site interactions plus additional terms ΔH :

$$\beta H[\phi(x); l] = \sum_x U[\phi(x); l] + \beta \Delta H. \quad (10.93)$$

Here, ΔH vanishes if all $\phi(x)$ are equal, i. e. if we are considering a homogeneous field $\phi(x) = \phi$.

Let us assume

$$\Delta H \geq 0. \quad (10.94)$$

Then, $H[\phi(x); l]$ is minimal for a homogeneous field ϕ_0 , with ϕ_0 being the minimum of $U[\phi(x); l]$. In general, this minimum changes due to fluctuations. So it is length-scale dependent: $\phi_0(l)$.

10.3.4. Renormalization flow

Principle of the renormalization flow

$H[\phi(x); l]$ describes a family of lattice models with different l . But they all correspond to the *same* microscopic model. So, we can conclude that the physical laws take different forms on different length scales l . For example, the form of the on-site interaction potential U might depend on l in the way sketched in Figure 10.5. Its interpretation would be that small domains favour magnetisation (non-zero position of the minimum of U), whereas they get averaged out for large domains.

The general principle behind this is, that for $H[\phi; l]$ the fluctuations with wavelength smaller than l are averaged / integrated out. On the other hand, fluctuations with wavelength larger than l remain to be done.

Momentum space

All of this can be formulated more elegantly in momentum space. The momentum scale, belonging to the length scale l , is

$$k \approx \frac{\pi}{l}. \quad (10.95)$$

So, fluctuations with momenta $q^2 > k^2$ are integrated out.

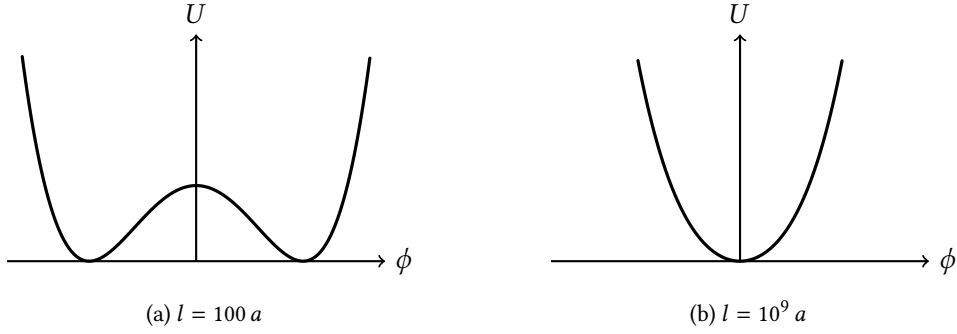


Figure 10.5.: Example of an potential on different length scales

In momentum space, Hamiltonian and potential are denoted by $H[\phi; k]$ and $U[\phi; k]$. The microscopic limit of the Hamiltonian is given by

$$H[\phi; k = \pi/a] = H_0[\phi], \quad (10.96)$$

and its $k \rightarrow 0$ limit can be written as

$$\beta H[\phi; k = 0] = \mathcal{N}U[\phi] + \beta \Delta H. \quad (10.97)$$

In the latter limit, all fluctuations are integrated out. The potential in this limit is called the *effective potential*:

$$U[\phi] = U[\phi; k \rightarrow 0]. \quad (10.98)$$

Flow equation

The central equation of the renormalization group is the flow equation:

$$\frac{\partial U[\phi; k]}{\partial \ln k} = \dots, \quad (10.99)$$

where the r. h. s. still remains to be computed. This equation describes the change of the potential U when varying the momentum scale. To extract the free energy for constant ϕ (magnetisation), it is sufficient to solve this equation in the $k \rightarrow 0$ limit. Equivalently, we can consider this equation in real space:

$$\frac{\partial U[\phi; l]}{\partial l} = \dots, \quad (10.100)$$

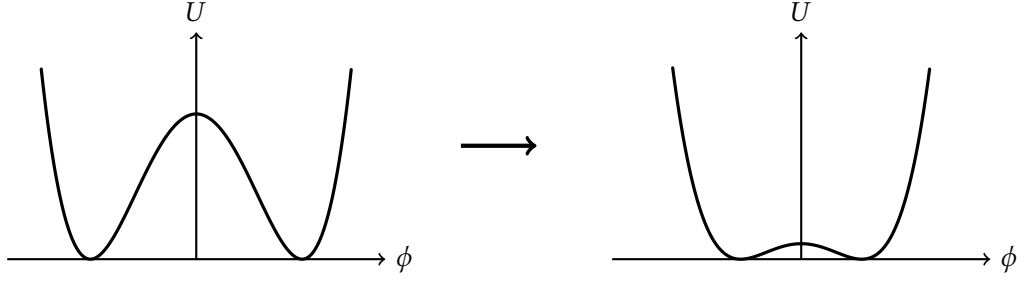
and solve it in the $l \rightarrow \infty$ limit.

As a simplification, we restrict the potential to a quartic polynomial at every scale l :

$$U[\phi; l] = \frac{1}{2} \lambda_l (\rho - \rho_{0,l})^2, \quad (10.101)$$

with

$$\rho = \frac{1}{2} \phi^2. \quad (10.102)$$

Figure 10.6.: Flow of the potential towards larger l

Here, λ_l and $\rho_{0,l}$ depend on l and $\rho_{0,l}$ is the minimum of $U[\phi; l]$.

Now we want to set up flow equations for $\rho_{0,l}$ and λ_l . Assuming $a^2 \ll l^2 \ll (\lambda_l \rho_{0,l})^{-1}$ as an approximation, one finds in three dimensions

$$\frac{\partial}{\partial l} \rho_0 = -\frac{3c_1}{2\pi^2} \frac{1}{l^2}, \quad (10.103)$$

$$\frac{\partial}{\partial l} \lambda = -\frac{9c_2}{4\pi^2} \lambda^2, \quad (10.104)$$

with some constants c_1, c_2 of order $O(1)$. (This result requires a computation not covered by this lecture.) We can directly infer that ρ_0 and λ decrease for increasing l . The qualitative change of U due to that has been sketched in Figure 10.6.

Solution of flow equation

Let us first consider the *running minimum*. The equation we have to solve, is

$$d\rho_0 = -\frac{3c_1}{2\pi^2} \frac{dl}{l^2} \quad (10.105)$$

$$= \frac{3c_1}{2\pi^2} d(l^{-1}). \quad (10.106)$$

The solution to this differential equation is given by

$$\rho_0 = \hat{\rho}_0 + \frac{3c_1}{2\pi^2} \left(\frac{1}{l} - \frac{1}{l_0} \right), \quad (10.107)$$

with initial value $\hat{\rho}_0 = \rho_0(l_0)$, which depends on T and J .

Let us examine its large length scale behaviour:

$$\rho_0(l) \xrightarrow{l \rightarrow \infty} \hat{\rho}_0 - \frac{3c_1}{2\pi^2 l_0} = \hat{\rho}_0 - \rho_{0,c}. \quad (10.108)$$

So, we can distinguish three different cases:

$\hat{\rho}_0 > \rho_{0,c}$: In this case, $\rho_0(l \rightarrow \infty) > 0$. So, it describes the ordered phase with SSB.

$\hat{\rho}_0 < \rho_{0,c}$: In this case, $\rho_0(l)$ reaches 0 before $l \rightarrow \infty$. So, it describes the disordered phase, where the $\phi \rightarrow -\phi$ symmetry is unbroken (SYM).

$\hat{\rho}_0 = \rho_{0,c}$: At this point, the phase transition occurs! The corresponding critical temperature can be deduced from $\rho_0(T_c) = \rho_{0,c}$.

Now, let us consider the *running coupling*. Its differential equation is given by

$$\frac{d\lambda}{\lambda^2} = -d(\lambda^{-1}) = -\frac{9c_2}{4\pi^2} dl. \quad (10.109)$$

Integrating this, yields

$$\frac{1}{\lambda} - \frac{1}{\lambda_0} = \frac{9c_2}{4\pi^2} (l - l_0), \quad (10.110)$$

where $\lambda_0 = \lambda(l_0)$. Solving this for λ gives us

$$\lambda = \left(\frac{1}{\lambda_0} + \frac{9c_2}{4\pi^2} (l - l_0) \right)^{-1} \quad (10.111)$$

$$= \lambda_0 \left(1 + \frac{9c_2\lambda_0}{4\pi^2} (l - l_0) \right)^{-1}. \quad (10.112)$$

In the $l \rightarrow \infty$ limit, this becomes

$$\lambda = \frac{4\pi^2}{9c_2} \frac{1}{l}. \quad (10.113)$$

Note that this is independent of the initial value! This is an example of how the memory of the microscopic physics is lost at macroscopic scales!

10.3.5. Fixed points

Now, we want to look for fixed points of the system, i. e. points, where the flow of λ and ρ_0 vanishes. Actually, we want to look for fixed points of the dimensionless quantities

$$\tilde{\lambda} = \lambda l, \quad (10.114)$$

$$\kappa = \rho_0 l, \quad (10.115)$$

with respect to the dimensionless scale

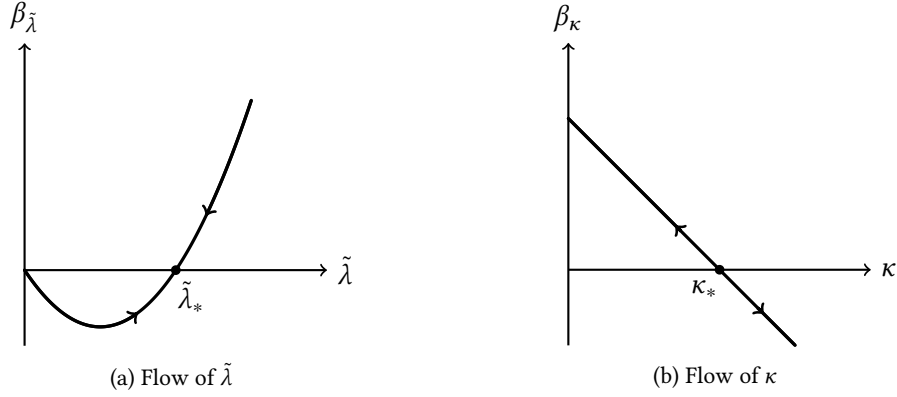
$$t = -\ln \frac{l}{l_0}. \quad (10.116)$$

This is necessary to find truly scale-independent fixed points of the system!

For κ we find

$$\frac{d}{dt}\kappa = -l \frac{d}{dl}(\rho_0 l) \quad (10.117)$$

$$= -\kappa - l^2 \frac{d\rho_0}{dl}. \quad (10.118)$$

Figure 10.7.: Flow of the dimensionless quantities towards larger l

So, using (10.107), we arrive at

$$\beta_\kappa = \partial_t \kappa = -\kappa + \frac{3c_1}{2\pi^2}. \quad (10.119)$$

Similarly, we compute the t -derivative of $\tilde{\lambda}$:

$$\frac{d}{dt} \tilde{\lambda} = -\tilde{\lambda} - l^2 \frac{d\lambda}{dl}, \quad (10.120)$$

and, using (10.112), find

$$\beta_{\tilde{\lambda}} = \partial_t \tilde{\lambda} = -\tilde{\lambda} + \frac{9c_2}{4\pi^2} \tilde{\lambda}^2. \quad (10.121)$$

Here, we adopted the usual QFT notation of β -functions. Note that no mass or length scale appears in these flow equations!

We have sketched the flow of $\tilde{\lambda}$ in Figure 10.7a. Thereby, the arrows indicate the flow towards larger l , i. e. smaller t . We see that, independent of the starting point, $\tilde{\lambda}$ flows to the fixed point $\tilde{\lambda}_*$ for $l \rightarrow \infty$. From

$$\beta_{\tilde{\lambda}}(\tilde{\lambda}_*) = 0 \quad (10.122)$$

it follows

$$\tilde{\lambda}_* = \frac{4\pi}{9c_2}. \quad (10.123)$$

So, we can confirm the large length scale behaviour found in (10.113):

$$\lambda \xrightarrow{l \rightarrow \infty} \frac{\tilde{\lambda}_*}{l} = \frac{4\pi}{9c_2 l}. \quad (10.124)$$

This flow of the effective couplings towards an “infrared” fixed point is related to the observed loss of memory of the microscopic physics. The details of the microphysics determine the starting point of the flow of the coupling. But as we have seen, for larger (macroscopic) length scales the

coupling flows towards this fixed point independent of the specific starting point! This is also the deep origin of the universality of the critical exponents. They are related to the behaviour of the β -functions around the fixed point. This behaviour is determined by the form of the potential U . And this, in turn, is determined by the dimension and the symmetries of our system.

Universality follows from the flow!

Contrary to the coupling, the flow of κ , as sketched in Figure 10.7b, shows the existence of an “ultraviolet” fixed point κ_* . Every other starting point than $\kappa = \kappa_*$ leads to a divergent flow for $l \rightarrow \infty$. So, the critical point is described by $\kappa = \kappa_*$:

$$\kappa = \kappa_* \quad \cong \quad T = T_c . \quad (10.125)$$

That is because the critical point is characterized by $\rho_0 = \frac{\kappa}{l} \rightarrow 0$ for $l \rightarrow \infty$. And this is only possible for $\kappa = \kappa_*$.

So, we see that at the critical point all quantities are on a fixed point. Therefore, our system is scale-invariant at the critical point! This behaviour is characteristic for second order phase transitions. But note that this needs some tuning for κ , in contrast to the infrared fixed point for $\tilde{\lambda}$.