Statistical Physics

ETH Zürich – Herbstsemester 2008

Manfred Sigrist, HPZ G13 (Tel.: 3-2584,
Email: sigrist@itp.phys.ethz.ch)

Literature:

- *Statistical Mechanics*, K. Huang, John Wiley & Sons (1987).
- *Equilibrium Statistical Physics*, M. Plischke and B. Bergersen, Prentice-Hall International (1989).
- *Theorie der Wärme*, R. Becker, Springer (1978).
- *Thermodynamik und Statistik*, A. Sommerfeld, Harri Deutsch (1977).
- *Statistische Physik, Vol. V*, L.D. Landau and E.M Lifschitz, Akademie Verlag Berlin (1975).
- *Statistische Mechanik*, F. Schwab, Springer (2000).
- *Elementary Statistical Physics*, Charles Kittel,John Wiley & Sons (1967).
- Many other books and texts mentioned throughout the lecture

Contents

Chapter 1

Laws of Thermodynamics

Thermodynamics is a phenomenological, empirically derived, macroscopic description of the equilibrium properties of physical systems with enormously many particles or degrees of freedom. These systems are usually considered big enough such that the boundary effects do not play an important role (we call this the "thermodynamic limit"). The macroscopic states of such systems can be characterized by several macroscopically measurable quantities whose mutual relation can be cast into equations and represent the theory of thermodynamics.

Much of the body of thermodynamics has been developed already in the 19^{th} century before the microscopic nature of matter, in particular, the atomic picture has been accepted. Thus, the principles of thermodynamics do not rely on the input of such specific microscopic details. The three laws of thermodynamics constitute the essence of thermodynamics.

1.1 State variables and equations of state

Equilibrium states described by thermodynamics are states which a system reaches after a long waiting (relaxation) time. Macroscopically systems in equilibrium to not show an evolution in time. Therefore time is not a variable generally used in thermodynamics. Most of processes discussed are considered to be *quasistatic*, i.e. the system is during a change practically always in a state of equilibrium (variations occur on time scales much longer than relaxation times to the equilibrium).

1.1.1 State variables

In thermodynamics equilibrium states of a system are uniquely characterized by measurable macroscopic quantities, generally called *state variables* (or *state functions*). We distinguish intensive and extensive state variables.

- intensive state variable: homogeneous of degree 0, i.e. independent of system size
- extensive state variable: homogeneous of degree 1, i.e. proportional to the system size

Examples of often used state variables are:

	intensive		extensive
T	temperature	S	entropy
p	pressure	V	volume
H	magnetic field	M	magnetization
E	electric field	P	dielectric polarization
μ	chemical potential	N	particle number

Intensive and extensive variables form pairs of conjugate variables, e.g. temperature and entropy or pressure and volume, etc.. Their product lead to energies which are extensive quantities (in the above list each line gives a pair of conjugate variables). Intensive state variables can often be used as equilibrium variable. Two systems are equilibrium with each other, if all their equilibrium variable are identical.
The state variables determines uniquely the equilibrium state, independent of the way this state was reached. Starting from a state A the state variable Z of state B is obtained by

$$Z(B) = Z(A) + \int_{\gamma_1} dZ = Z(A) + \int_{\gamma_2} dZ \ , \tag{1.1}$$

where γ_1 and γ_2 represent different paths in space of states (Fig. 1.1).

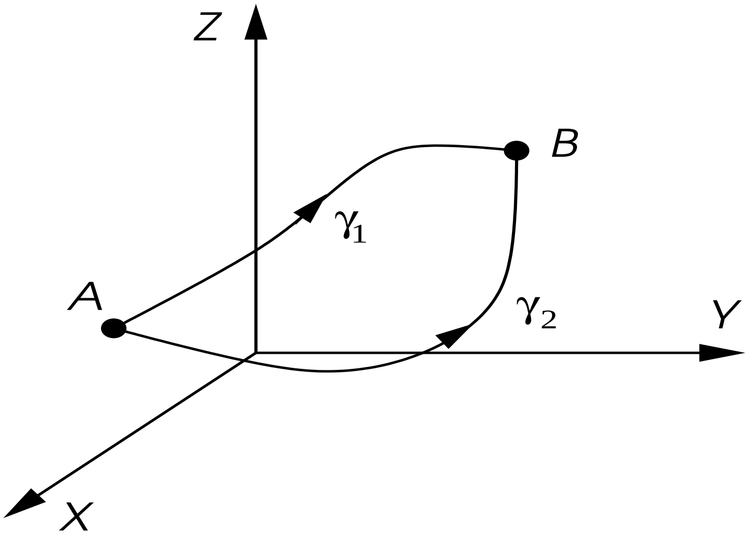


Fig.1.1: Space of states: A and B are equilibrium states connected by paths γ_1 or γ_2. The coordinates X , Y and Z are state variables of the system.

From this follows for the closed path $\gamma = \gamma_1 - \gamma_2$:

$$\oint_{\gamma} dZ = 0 \ . \tag{1.2}$$

This is equivalent to the statement that Z is an *exact differential*, i.e. it is single-valued in the space of states. If we can write Z as a function of two independent state variables X and Y (see Fig.1.1) then we obtain

$$dZ = \left(\frac{\partial Z}{\partial X}\right)_Y dX + \left(\frac{\partial Z}{\partial Y}\right)_X dY \qquad \text{with} \qquad \left[\frac{\partial}{\partial Y}\left(\frac{\partial Z}{\partial X}\right)_Y\right]_X = \left[\frac{\partial}{\partial X}\left(\frac{\partial Z}{\partial Y}\right)_X\right]_Y \ . \tag{1.3}$$

The variable at the side of the brackets indicates the variables to be fixed for the partial derivative. This can be generalized to many, say n, variables X_i $(i = 1, ..., n)$. Then $n(n-1)/2$ conditions are necessary to ensure that Z is an exact differential:

$$dZ = \sum_{i=1}^{n} \left(\frac{\partial Z}{\partial X_i}\right)_{X_{j\neq i}} dX_i \quad \text{with} \quad \left[\frac{\partial}{\partial X_k}\left(\frac{\partial Z}{\partial X_i}\right)_{X_{j\neq i}}\right]_{X_{j\neq k}} = \left[\frac{\partial}{\partial X_i}\left(\frac{\partial Z}{\partial X_k}\right)_{X_{j\neq k}}\right]_{X_{j\neq i}} \ . \tag{1.4}$$

As we will see later not every measurable quantity of interest is an exact differential. Examples of such quantities are the different forms of work and the heat as a form of energy.

1.1.2 Temperature as an equilibrium variable

Temperature is used as an equilibrium variable which is so essential to thermodynamics that the definition of temperature is sometimes called "0^{th} law of thermodynamics". Every macroscopic thermodynamic system has a temperature T which characterizes its equilibrium state. The following statements apply:

- In a system which is in its equilibrium the temperature is everywhere the same. A system with an inhomogeneous temperature distribution, $T(\vec{r})$ is not in its equilibrium (a heat current proportional to $\vec{\nabla} T(\vec{r})$ flows in order to equilibrate the system; in such a system state variables evolve in time).

- For two systems, A and B, which are each independently in thermal equilibrium, we can always find the relations: $T_A < T_B$, $T_A > T_B$ or $T_A = T_B$.

- The order of systems according to their temperature is transitive. We find for systems A, B and C, that $T_A > T_B$ and $T_B > T_C$ leads to $T_A > T_C$.

- The equilibrium state of two systems (A and B) in thermal contact is characterized by $T_A = T_B = T_{A\cup B}$. If before equilibration $T'_A > T'_B$, then the equilibrium temperature has to lie between the two temperatures: $T'_A > T_{A\cup B} > T'_B$.

Analogous statements can be made about other intensive state variables such as pressure and the chemical potential which are equilibrium variables.

1.1.3 Equation of state

The equilibrium state of a thermodynamic system is described completely by a set of independent state variables $(Z_1, Z_2, \ldots, Z_n)$ which span the space of all possible states. Within this space we find a subspace which represents the equilibrium states and can be determined by the *thermodynamic equation of state*

$$f(Z_1, \ldots, Z_n) = 0 \; . \tag{1.5}$$

The equation of state can be determined from experiments or from microscopic models through statistical physics, as we will see later.

Ideal gas: As one of the simplest examples of a thermodynamic system is the ideal gas, microscopically, a collection of independent identical atoms or molecules. In real gases the atoms/molecules interact with each other such that at low enough temperature ($T < T_B$ boiling temperature) the gas turns into a liquid. The ideal gas conditions is satisfied for real gases at temperatures $T \gg T_B$.[1]

The equation of state is given through the Boyle-Mariotte-law ($\sim$ 1660 - 1675) und the Gay-Lussac-law (1802). The relevant space of states consists of the variables pressure p, volume V and temperature T and the equation of state is

$$pV = nRT = Nk_BT \tag{1.6}$$

where n is the number of moles of the gas. One mole contains $N_A = 6.022 \cdot 10^{23}$ particles (atoms or molecules).[2] N is the number of gas particles. The gas constant R and the Boltzmann constant k_B are then related and given by

$$R = 8.314 \frac{J}{KMol} \quad \text{oder} \quad k_B = \frac{R}{N_A} = 1.381 \cdot 10^{-23} \frac{J}{K} \; . \tag{1.7}$$

The temperature T entering the equation of state of the ideal gas defines the absolute temperature in Kelvin scale with $T = 0K$ at $-273.15°C$.

---

[1] Ideal gases at very low temperatures are subject to quantum effects and become ideal quantum gases which deviate from the classical ideal gas drastically.

[2] At a pressure of 1 atm = 760 Torr = 1.01 bar = $1.01 \cdot 10^5$ Pa (Nm^{-2}) the volume of one mole of gas is $V = 22.4 \cdot 10^{-3} m^3$.

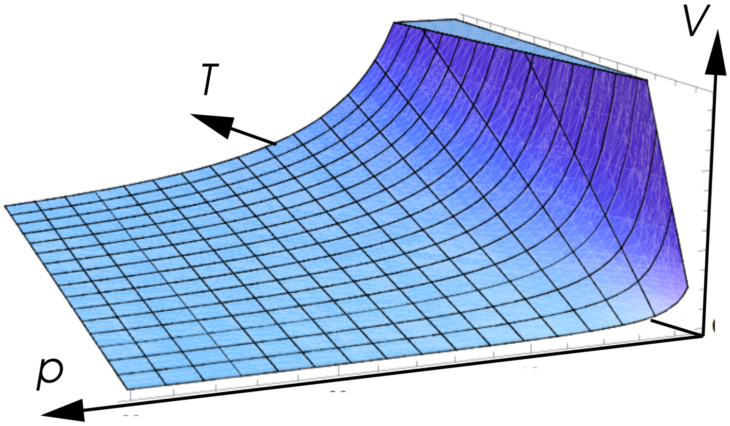


Fig.1.2: Equilibrium states of the ideal gas in the $p - V - T$-space.

1.1.4 Response function

There are various measurable quantities which are based on the reaction of a system to the change of external parameters. We call them response functions. Using the equation of state for the ideal gas we can define the following two response functions:

$$\begin{array}{lll} \text{isobar thermal expansion coefficient} & \alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p & = \frac{1}{T} \\ \text{isothermal compressibility} & \kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T & = \frac{1}{p} \end{array}$$

i.e. the change of volume (a state variable) is measured in response to change in temperature T and pressure p respectively.[3] As we consider a small change of these external parameters only, the change of volume is linear: $\Delta V \propto \Delta T, \Delta p$, i.e. *linear response*.

1.2 First Law of Thermodynamics

The first law of thermodynamics declares heat as a form of energy, like work, potential or electromagnetic energy. The law has been formulated around 1850 by the three scientists Julius Robert Mayer, James Prescott Joule and Hermann von Helmholtz.

1.2.1 Heat and work

In order to change the temperature of a system while fixing state variables such as pressure, volume and so on, we can transfer a certain amount of heat δQ from or to the systems by putting it into contact with a heat reservoir of certain temperature. In this context we can define the specific heat C_p (C_V) for fixed pressure (fixed volume),

$$\delta Q = C_p dT \tag{1.8}$$

Note that C_p and C_V are response functions. Heat Q is not a state variable, indicated in the differential by δQ instead of dQ.

---

[3] *"Isobar":* fixed pressure; *"isothermal":* fixed temperature.

The work W is defined analogous to the classical mechanics, introducing a generalized coordinate q and the generalized conjugate force F,

$$\delta W = F dq \; . \tag{1.9}$$

Analogous to the heat, work is not a state variable. By definition $\delta W < 0$ means, that work is extracted from the system, it "does" work. A form of δW is the mechanical work for a gas,

$$\delta W = -p dV \; , \tag{1.10}$$

or magnetic work

$$\delta W = H dM \; . \tag{1.11}$$

The total amount of work is the sum of all contributions

$$\delta W = \sum_i F_i dq_i \; . \tag{1.12}$$

The first law of thermodynamics may now be formulated in the following way: Each thermodynamic system possesses an internal energy U which consists of the absorbed heat and the work,

$$dU = \delta Q + \delta W = \delta Q + \sum_i F_i dq_i \tag{1.13}$$

In contrast to Q and W, U is a state variable and is conserved in an *isolated* system. Alternative formulation: There is no perpetuum mobile of the first kind, i.e.; there is no machine which produces work without any energy supply.

Ideal gas: The Gay-Lussac-experiment (1807)[4] leads to the conclusion that the internal energy U is only a function of temperature, but not of the volume or pressure, for an ideal gas: $U = U(T)$. Empirically or in derivation from microscopic models we find

$$U(T) = \begin{cases} \dfrac{3}{2} N k_B T + U_0 & \text{single-atomic gas} \\[2ex] \dfrac{5}{2} N k_B T + U_0 & \text{diatomic molecules ,} \end{cases} \tag{1.14}$$

Here U_0 is a reference energy which can be set zero. The difference for different molecules originates from different internal degrees of freedom. Generally, $U = \frac{f}{2} N k_B T$ with f as the number of degrees of freedom. This is the equipartition law.[5] The equation $U = U(T, V, ...)$ is called *caloric equation of state* in contrast to the *thermodynamic equation of state* discussed above.

1.2.2 Thermodynamic transformations

In the context of heat and work transfer which correspond to transformations of a thermodynamic system, we introduce the following terminology to characterize such transformations:

- *quasistatic:* the external parameters are changed slowly such that the system is always approximatively in an equilibrium state.

---

[4] Consider an isolated vessel with two chambers separated by a wall. In chamber 1 we have a gas and chamber 2 is empty. After the wall is removed the gas expands (uncontrolled) in the whole vessel. In this process the temperature of the gas does not change. In this experiment neither heat nor work has been transfered.

[5] A single-atomic molecule has only the three translational degrees of freedom, while a diatomic molecule has additionally two independent rotational degrees of freedom with axis perpendicular to the molecule axis connecting to two atoms.

- *reversible:* a transformation can be completely undone, if the external parameters are changed back to the starting point in the reversed time sequence. A reversible transformation is always quasistatic. But not every quasistatic process is reversible (e.g. plastic deformation).
- *isothermal:* the temperature is kept constant. Isothermal transformations are generally performed by connecting the thermodynamic system to an "infinite" heat reservoir of given temperature T, which provides a source or sink of heat in order to fix the temperature.
- *adiabatic:* no heat is transfered into or out of the system, $\delta Q = 0$. Adiabatic transformations are performed by decoupling the system from any heat reservoir. Such a system is *thermally isolated*.
- *cyclic:* start and endpoint of the transformation are identical, but generally the path in the space of state variables does not involve retracing.

1.2.3 Applications of the first law

The specific heat of a system is determined via the heat transfer δQ necessary for a given temperature change. For a gas this is expressed by

$$\delta Q = dU - \delta W = dU + pdV = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV + pdV \ . \tag{1.15}$$

The specific heat at constant volume ($dV = 0$) is then

$$C_V = \left(\frac{\delta Q}{dT}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V \ . \tag{1.16}$$

On the other hand, the specific heat at constant pressure (isobar) is expressed as

$$C_p = \left(\frac{\delta Q}{dT}\right)_p = \left(\frac{\partial U}{\partial T}\right)_V + \left[\left(\frac{\partial U}{\partial V}\right)_T + p\right]\left(\frac{\partial V}{\partial T}\right)_p \ . \tag{1.17}$$

which leads to

$$C_p - C_V = \left\{\left(\frac{\partial U}{\partial V}\right)_T + p\right\}\left(\frac{\partial V}{\partial T}\right)_p = \left\{\left(\frac{\partial U}{\partial V}\right)_T + p\right\} V\alpha \ . \tag{1.18}$$

Here α is the isobar thermal expansion coefficient. For the ideal gas we know that U does not depend on V. With the equation of state we obtain

$$\left(\frac{\partial U}{\partial V}\right)_T = 0 \qquad \text{und} \qquad \left(\frac{\partial V}{\partial T}\right)_p = \frac{nR}{p} \qquad \Rightarrow \qquad C_p - C_V = nR = Nk_B \ . \tag{1.19}$$

As a consequence it is more efficient to change the temperature of a gas by heat transfer at constant volume, because for constant pressure some of the heat transfer is turned into mechanical work $\delta W = -p\, dV$.

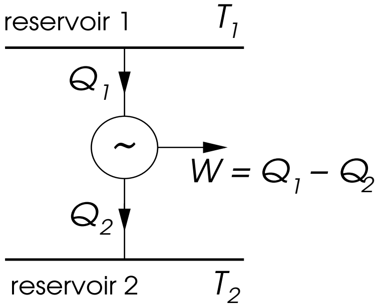
1.3 Second law of thermodynamics

While the energy conservation stated in the first law is quite compatible with the concepts of classical mechanics, the second law of thermodynamics is much less intuitive. It is a statement about the energy transformation. There are two completely compatible formulations of this law:

- *Rudolf Clausius:* There is no cyclic process, whose only effect is to transfer heat from a reservoir of lower temperature to one with higher temperature.
- *William Thomson (Lord Kelvin):* There is no cyclic process, whose effect is to take heat from a reservoir and transform it completely into work. "There is no perpetuum mobile of the second kind."

1.3.1 Carnot engine

A Carnot engine performs cyclic transformations transferring heat between reservoirs of different temperature absorbing or delivering work. This type of engine is a convenient theoretical tool to understand the implications of the second law of thermodynamics. The most simple example is an engine between two reservoirs.



This machine extracts the heat Q_1 from reservoir 1 at temperature T_1 and passes the heat Q_2 on to reservoir 2 at temperature 2, while the difference of the two energies is emitted as work: $W = Q_1 - Q_2$ assuming $T_1 > T_2$. Note, that also the reversed cyclic process is possible whereby the heat flow is reversed and the engine absorbs work. We define the *efficiency* η, which is the ratio between the work output W and the heat input Q_1:

$$\eta = \frac{W}{Q_1} = \frac{Q_1 - Q_2}{Q_1} < 1 \ . \tag{1.20}$$

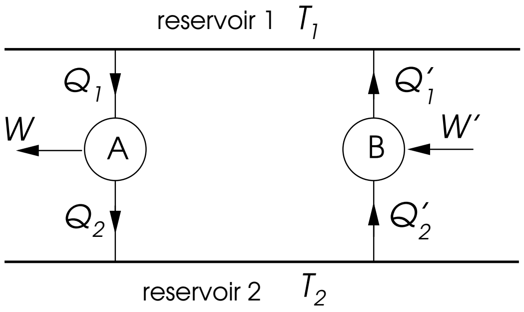
Theorem by Carnot: (a) For reversible cyclic processes of engines between two reservoirs the ratio

$$\frac{Q_1}{Q_2} = \frac{T_1}{T_2} > 0 \tag{1.21}$$

is universal. This can be considered as a definition of the absolute temperature scale and is compatible with the definition of T via the equation of state of the ideal gas.
(b) For an arbitrary engine performing a cyclic process, it follows that

$$\frac{Q_1}{Q_2} \leq \frac{T_1}{T_2} \tag{1.22}$$

Proof: Consider two engines, A and B, between the same two reservoirs



The engine B is a reversible engine (= Carnot engine) while A is an arbitrary cyclic engine. The cycles of both engines are chosen so that $Q_2 = Q'_2$ leaving reservoir 2 unchanged. Thus,

$$\left.\begin{array}{l} W = Q_1 - Q_2 \\[2ex] W' = Q'_1 - Q'_2 \end{array}\right\} \quad \Rightarrow \quad W - W' = Q_1 - Q'_1 \tag{1.23}$$

According to the second law (Kelvin) we get into a conflict, if we assume $W > W'$, as we would extract work from the total engine while only taking heat from reservoir 1. Thus, the statement compatible with the second law is $W \leq W'$ and $Q_1 \leq Q_1'$ such that,

$$\frac{Q_1}{Q_2} \leq \frac{Q_1'}{Q_2'} = \frac{T_1}{T_2} \; . \tag{1.24}$$

The equal sign is obtained when we assume that both engines are reversible, as we can apply the given argument in both directions so that $W \leq W'$ and $W \geq W'$ leads to the conclusion $W = W'$. This proves also the universality. From this we can also conclude that a reversible engine has the maximal efficiency.

1.3.2 Entropy

Heat is not a state variable. Looking at the cycle of a Carnot engine immediately shows this

$$\oint \delta Q \neq 0 \; , \tag{1.25}$$

so δQ is not an exact differential. We can introduce an integrating factor to arrive at a state variable:

$$dS = \frac{\delta Q}{T} \quad \text{with} \quad \oint dS = 0 \tag{1.26}$$

which works for the simple cyclic Carnot process:

$$\oint dS = \frac{Q_1}{T_1} + \frac{Q_2}{T_2} = 0 \tag{1.27}$$

and may be extended straightforwardly to any reversible cyclic process. S is the *entropy.*

Clausius' theorem: For any cyclic transformation the following inequality holds:

$$\oint dS \leq 0 \tag{1.28}$$

and the equality is valid for a reversible transformation.

The proof is simply an extension of Carnot's theorem. The entropy as a state variable satisfies the following relation for a path of a reversible transformation from state A to B:

$$\int_A^B dS = S(B) - S(A) \; . \tag{1.29}$$

In the case of an arbitrary transformation we find

$$\int_A^B \frac{\delta Q}{T} \leq S(B) - S(A) \; . \tag{1.30}$$

Ideal gas: We first look at a reversible expansion of an ideal gas from volume V_1 to V_2 while keeping contact to a heat reservoir of temperature T. Thus the expansion is isothermal, and the internal energy U of the gas does not change as it does only depend on temperature. During the expansion process the ideal gas draws the heat Q which then is transformed into work:

$$\Delta U = 0 \quad \Rightarrow \quad Q = -W = \int_{V_1}^{V_2} p dV = N k_B T \ln\left(\frac{V_2}{V_1}\right) \; . \tag{1.31}$$

where we have used the equation of state. The change of entropy during this change is then easily obtained as

$$(\Delta S_{\text{rev}})_{\text{gas}} = \int_1^2 \frac{\delta Q_{\text{rev}}}{T} = \frac{Q}{T} = N k_B \ln\left(\frac{V_2}{V_1}\right) \; . \tag{1.32}$$

Analogously the entropy of the reservoir changes,

$$(\Delta S_{\text{rev}})_{\text{res}} = -\frac{Q}{T} = -(\Delta S_{\text{rev}})_{\text{gas}} \tag{1.33}$$

such that the entropy of the complete system (ideal gas + reservoir) stays unchanged, $\Delta S_{total} = \Delta S_{res} + \Delta S_{gas} = 0$. In turn the gained work could be stored, e.g. as potential energy.
Turning now to an irreversible transformation between the same initial and final state as before, we consider the free expansion of the gas as in the Gay-Lussac experiment. Thus there is no work gained in the expansion. The entropy change is the same, as S is a state variable. Since the system is isolated the temperature and the internal energy have not changed. No heat was received from a reservoir, $(\Delta S)_{\text{res}} = 0$. Thus the total entropy has increased

$$\Delta S_{total} = \Delta S_{gas} > 0 \tag{1.34}$$

This increase of entropy amounts to a waste of work stored in the initial state. The free expansion does not extract any work from the increase of volume.

1.3.3 Applications of the first and second law

The existence of the entropy S as a state variable is the most important result of the second law of thermodynamics. We consider now some consequences for a gas, such that work is simply $\delta W = -pdV$. For a reversible process we find

$$TdS = \delta Q = dU - \delta W = dU + pdV \ . \tag{1.35}$$

Both dS and dU are exact differentials. We rewrite

$$dS = \frac{1}{T}dU + \frac{p}{T}dV = \left(\frac{\partial S}{\partial U}\right)_V dU + \left(\frac{\partial S}{\partial V}\right)_U dV \ . \tag{1.36}$$

From these we derive the caloric equation of state

$$\left(\frac{\partial S}{\partial U}\right)_V = \frac{1}{T} \quad \Rightarrow \quad T = T(U,V) \quad \Rightarrow \quad U = U(T,V) \ , \tag{1.37}$$

and the thermodynamic equation of state

$$\left(\frac{\partial S}{\partial V}\right)_U = \frac{p}{T} \quad \Rightarrow \quad p = Tf(T,V) \ . \tag{1.38}$$

Taking $S = S(T,V)$ and $U = U(T,V)$ leads us to

$$dS = \left(\frac{\partial S}{\partial T}\right)_V dT + \left(\frac{\partial S}{\partial V}\right)_T dV = \frac{1}{T}dU + \frac{p}{T}dV \ . \tag{1.39}$$

Using

$$dU = \left(\frac{\partial U}{\partial T}\right)_V dT + \left(\frac{\partial U}{\partial V}\right)_T dV \tag{1.40}$$

we obtain

$$dS = \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_V dT + \frac{1}{T}\left\{\left(\frac{\partial U}{\partial V}\right)_T + p\right\} dV \ . \tag{1.41}$$

The comparison of expressions shows that

$$T\left(\frac{\partial S}{\partial T}\right)_V = \left(\frac{\partial U}{\partial T}\right)_V = C_V \ . \tag{1.42}$$

Since dS und dU are exact differentials it follows that

$$\frac{1}{T}\left[\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial T}\right)_V\right]_T = -\frac{1}{T^2}\left[\left(\frac{\partial U}{\partial V}\right)_T + p\right] + \frac{1}{T}\left\{\left[\frac{\partial}{\partial T}\left(\frac{\partial U}{\partial V}\right)_T\right]_V + \left(\frac{\partial p}{\partial T}\right)_V\right\} . \tag{1.43}$$

leading eventually to

$$\left(\frac{\partial U}{\partial V}\right)_T = T\left(\frac{\partial p}{\partial T}\right)_V - p = T^2\left(\frac{\partial}{\partial T}\frac{p}{T}\right)_V . \tag{1.44}$$

Knowing the thermodynamic equation of state, $p = p(T, V, ...)$ allows us to calculate the volume dependence of the internal energy. Interestingly for the ideal gas we find from this,

$$\left(\frac{\partial U}{\partial V}\right)_T = 0 \qquad \Leftrightarrow \qquad p = Tf(V) \tag{1.45}$$

This result was previously derived based on the outcome of the Gay-Lussac experiment. Now it appears as a consequence of the second law of thermodynamics.

1.4 Thermodynamic potentials

Generally we encounter the question of the suitable state variables to describe a given thermodynamic system. We introduce here several *thermodynamic potentials* for the different sets of variables, which have convenient properties.
In order to understand the issue here we consider first the internal energy U:

$$dU = TdS + \sum_i F_i dq_i + \sum_j \mu_j dN_j \tag{1.46}$$

where we use again generalized forces F_i and coordinates q_i in order to describe the work. Additionally we consider the possible change of the amount of matter in the system, i.e. the change of the particle number N_i. The corresponding conjugate variable of N is the *chemical potential* μ_i, the energy to add a particle to the system.
This differential provides immediately the relations

$$\left(\frac{\partial U}{\partial S}\right)_{q_i,N_j} = T \ , \quad \left(\frac{\partial U}{\partial q_i}\right)_{S,N_j,q_{i'\neq i}} = F_i \ , \quad \left(\frac{\partial U}{\partial N_j}\right)_{S,q_i,N_{j'\neq j}} = \mu_j \ . \tag{1.47}$$

These simple forms qualify the set S, q_i and N_j as the natural variables for U. Note that the relations would more complicated for the variables (T, q_i, μ_i) for example.
In these variables it is now also easy to obtain response functions. The specific heat results from

$$\left(\frac{\partial^2 U}{\partial S^2}\right)_{q_i,N_j} = \left(\frac{\partial T}{\partial S}\right)_{q_i,N_j} = \frac{T}{C_{q_i,N_j}} \quad \Rightarrow \quad C_{q_i,N_j} = T\left[\left(\frac{\partial^2 U}{\partial S^2}\right)_{q_i,N_j}\right]^{-1} . \tag{1.48}$$

or for a gas ($\delta W = -pdV$) we find

$$\left(\frac{\partial^2 U}{\partial V^2}\right)_{S,N_j} = -\left(\frac{\partial p}{\partial V}\right)_{S,N_j} = \frac{1}{V\kappa_S} \quad \Rightarrow \quad \kappa_S = \frac{1}{V}\left[\left(\frac{\partial^2 U}{\partial V^2}\right)_{S,N_j}\right]^{-1} , \tag{1.49}$$

providing the adiabatic compressibility ($dS = 0$ no heat transfer).
There are also important differential relations between different variables based on the fact that dU is an exact differential. These are called Maxwell relations. For simple one-atomic gas they are expressed as

$$\left[\frac{\partial}{\partial V}\left(\frac{\partial U}{\partial S}\right)_{V,N}\right]_{S,N} = \left[\frac{\partial}{\partial S}\left(\frac{\partial U}{\partial V}\right)_{S,N}\right]_{V,N} \qquad \Rightarrow \qquad \left(\frac{\partial T}{\partial V}\right)_{S,N} = -\left(\frac{\partial p}{\partial S}\right)_{V,N} . \tag{1.50}$$

Analogously we obtain

$$\left(\frac{\partial T}{\partial N}\right)_{S,V} = \left(\frac{\partial \mu}{\partial S}\right)_{V,N} \quad , \quad \left(\frac{\partial p}{\partial N}\right)_{S,V} = -\left(\frac{\partial \mu}{\partial V}\right)_{S,N} \quad . \tag{1.51}$$

The internal energy as $U(S, q_i, N_i)$ yields simple and convenient relations, and is in this form called a thermodynamical potential.

Note that also the entropy can act as a thermodynamic potential, if we use $U = U(S, V, N)$ and solve for S:

$$S = S(U, V, N) \quad \Rightarrow \quad dS = \frac{1}{T}dU - \sum_i \frac{F_i}{T}dq_i - \sum_i \frac{\mu_i}{T}dN_i \; . \tag{1.52}$$

1.4.1 Legendre transformation to other thermodynamical potentials

By means of the Legendre transformation we can derive other thermodynamical potentials of different variables starting from $U(S, q_i, N_i)$.[6] The thermodynamic system is again a gas so that

$$U = U(S, V, N) \quad \Rightarrow \quad dU = T\,dS - p\,dV + \mu\,dN \tag{1.56}$$

Helmholtz free energy: We replace the entropy S by the temperature T:

$$F = F(T, V, N) = \inf_S \left\{ U - S\left(\frac{\partial U}{\partial S}\right)_{V,N} \right\} = \inf_S \{U - ST\} \tag{1.57}$$

leading to the differential

$$dF = -SdT - pdV + \mu dN \qquad \Rightarrow \quad S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} , \; p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} , \; \mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} . \tag{1.58}$$

The resulting response functions are

$$C_V = -T\left(\frac{\partial^2 F}{\partial T^2}\right)_{V,N} \quad , \quad \kappa_T = \frac{1}{V}\left[\left(\frac{\partial^2 F}{\partial V^2}\right)_{T,N}\right]^{-1} \quad . \tag{1.59}$$

Moreover, following Maxwell relations apply:

$$\left(\frac{\partial p}{\partial T}\right)_{V,N} = \left(\frac{\partial S}{\partial V}\right)_{T,N} \; , \; \left(\frac{\partial S}{\partial N}\right)_{T,V} = -\left(\frac{\partial \mu}{\partial T}\right)_{V,N} \; , \; \left(\frac{\partial p}{\partial N}\right)_{T,V} = -\left(\frac{\partial \mu}{\partial V}\right)_{T,N} \; . \tag{1.60}$$

Enthalpy: We obtain the enthalpy by replacing the variable V by pressure p:

$$H = H(S, p, N) = \inf_V \left\{ U - V\left(\frac{\partial U}{\partial V}\right)_{S,N} \right\} = \inf_V \{U + pV\} \tag{1.61}$$

---

[6] Legendre transformation: We consider a function L of independent variables $x, y, z, ...$ with the exact differential

$$dL = Xdx + Ydy + Zdz + \cdots \; . \tag{1.53}$$

with $X, Y, Z, ...$ being functions of $x, y, z, ...$ We perform a variable transformation

$$\begin{aligned} L \quad &\rightarrow \quad \bar{L} = \inf_x \{L - Xx\} = \inf_x \left\{ L - x\left(\frac{\partial L}{\partial x}\right)_{y,z,...} \right\} \\ x, y, z, \ldots \quad &\rightarrow \quad X, y, z, \ldots \end{aligned} \tag{1.54}$$

from which we obtain

$$d\bar{L} = -xdX + Ydy + Zdz + \cdots \; . \tag{1.55}$$

The variable x is replaced by its conjugate variable X.

with the exact differential

$$dH = TdS + Vdp + \mu dN \qquad \rightarrow \qquad T = \left(\frac{\partial H}{\partial S}\right)_{p,N} , \ V = \left(\frac{\partial H}{\partial p}\right)_{S,N} , \ \mu = \left(\frac{\partial H}{\partial N}\right)_{S,p} . \tag{1.62}$$

An example of a response function we show the adiabatic compressibility

$$\kappa_S = -\frac{1}{V}\left(\frac{\partial^2 H}{\partial p^2}\right)_{S,N} . \tag{1.63}$$

The Maxwell relations read

$$\left(\frac{\partial T}{\partial p}\right)_{S,N} = \left(\frac{\partial V}{\partial S}\right)_{p,N} , \quad \left(\frac{\partial T}{\partial N}\right)_{S,p} = \left(\frac{\partial \mu}{\partial S}\right)_{p,N} , \quad \left(\frac{\partial V}{\partial N}\right)_{S,p} = \left(\frac{\partial \mu}{\partial p}\right)_{S,N} . \tag{1.64}$$

Gibbs free energy (free enthalpy): A further potential is reached by the choice of variables T, p, N:

$$G = G(T,p,N) = \inf_V \left\{ F - V\left(\frac{\partial F}{\partial V}\right)_{T,N} \right\} = \inf_V\{F + pV\} = \inf_{V,S}\{U - TS + pV\} = \mu N \tag{1.65}$$

with

$$dG = -SdT + Vdp + \mu dN \qquad \rightarrow \qquad -S = \left(\frac{\partial G}{\partial T}\right)_{p,N} , V = \left(\frac{\partial G}{\partial p}\right)_{T,N} , \mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} , \tag{1.66}$$

In this case simple response functions are

$$C_p = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_{p,N} , \quad \kappa_T = -\frac{1}{V}\left(\frac{\partial^2 G}{\partial p^2}\right)_{T,N} . \tag{1.67}$$

and the Maxwell relations have the form,

$$\left(\frac{\partial S}{\partial p}\right)_{T,N} = -\left(\frac{\partial V}{\partial T}\right)_{p,N} , \left(\frac{\partial S}{\partial N}\right)_{T,p} = -\left(\frac{\partial \mu}{\partial T}\right)_{p,N} , \left(\frac{\partial V}{\partial N}\right)_{T,p} = \left(\frac{\partial \mu}{\partial p}\right)_{T,N} . \tag{1.68}$$

This concludes the four most important thermodynamic potentials defining different sets of natural state variable which can be used as they are needed. We neglect here the replacement of N by μ.

1.4.2 Homogeneity of the potentials

All thermodynamic potentials are extensive. This is obvious for the internal energy . If a system with a homogeneous phase is scaled by a factor λ, then also U would scale by the same factor:

$$\begin{pmatrix} S \rightarrow \lambda S \\ V \rightarrow \lambda V \\ N \rightarrow \lambda N \end{pmatrix} \qquad \Rightarrow \qquad U \rightarrow \lambda U(S,V,N) = U(\lambda S, \lambda V, \lambda N) \tag{1.69}$$

Note that all variables of U are extensive and scale with the system size. Since the conjugate variable to each extensive variable is intensive, it is easy to see that all other thermodynamic potential are extensive too. For example the free energy

$$F \rightarrow \lambda F \ : \qquad S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \qquad \rightarrow \qquad \lambda S = -\left(\frac{\partial \lambda F}{\partial T}\right)_{V,N} \tag{1.70}$$

Note that all state variables which are used as equilibrium parameters $(T, p, \mu, ..)$ are intensive and do not scale with the system size.

Gibbs-Duhem relation: We examine the scaling properties of the Gibbs free energy for a homogeneous system:

$$G(T,p,N) = \frac{d}{d\lambda}\lambda G(T,p,N)|_{\lambda=1} = \frac{d}{d\lambda}G(T,p,\lambda N)|_{\lambda=1} = N\left(\frac{\partial G}{\partial N}\right)_{T,p} = \mu N \tag{1.71}$$

This defines the chemical Potential μ as the Gibbs free energy per particle. This leads to the important Gibbs-Duhem relation

$$dG = -SdT + Vdp + \mu dN = \mu dN + Nd\mu \qquad \Rightarrow \qquad SdT - Vdp + Nd\mu = 0\ . \tag{1.72}$$

which states that the equilibrium parameters T, p and μ cannot be varied independently.

1.4.3 Conditions of the equilibrium states

Depending on the external conditions of a system one of the thermodynamic potential is most appropriate to describe the equilibrium state. In order to determine the conditions for the equilibrium state of the system, it is also useful to see "how the equilibrium is reached". Here Clausius' inequality turns out to be very useful:

$$TdS \geq \delta Q\ . \tag{1.73}$$

Let us now consider a single-atomic gas under various conditions:

Closed system: This means $dU = 0$, $dV = 0$ and $dN = 0$. These are the variables of the potential 'entropy'. Thus, we find

$$\begin{array}{ll} dS \geq 0 & \text{general} \\ & \\ dS = 0 & \text{in equilibrium} \end{array} \tag{1.74}$$

The equilibrium state has a maximal entropy under given conditions. Consider now two subsystems, 1 and 2, which are connected such that internal energy and particles may be exchanged and the volume of the two may be changed, under the condition

$$\begin{array}{lll} U = U_1 + U_2 = const. & \Rightarrow & dU_1 = -dU_2 \\ V = V_1 + V_2 = const. & \Rightarrow & dV_1 = -dV_2 \\ N = N_1 + N_2 = const. & \Rightarrow & dN_1 = -dN_2 \end{array} \tag{1.75}$$

The entropy is additive,

$$S = S(U,V,N) = S(U_1,V_1,N_1) + S(U_2,V_2,N_2) = S_1 + S_2 \quad \text{with} \quad dS = dS_1 + dS_2 \tag{1.76}$$

We can therefore consider the equilibrium state

$$\begin{aligned} 0 = dS \;&= dS_1 + dS_2 \\ &= \left\{\left(\frac{\partial S_1}{\partial U_1}\right)_{V_1,N_1} - \left(\frac{\partial S_2}{\partial U_2}\right)_{V_2,N_2}\right\} dU_1 + \left\{\left(\frac{\partial S_1}{\partial V_1}\right)_{U_1,N_1} - \left(\frac{\partial S_2}{\partial V_2}\right)_{U_2,N_2}\right\} dV_1 \\ &\quad + \left\{\left(\frac{\partial S_1}{\partial N_1}\right)_{U_1,V_1} - \left(\frac{\partial S_2}{\partial N_2}\right)_{U_2,V_2}\right\} dN_1 \\ &= \left(\frac{1}{T_1} - \frac{1}{T_2}\right) dU_1 + \left(\frac{p_1}{T_1} - \frac{p_2}{T_2}\right) dV_1 + \left(-\frac{\mu_1}{T_1} + \frac{\mu_2}{T_2}\right) dN_1\ . \end{aligned} \tag{1.77}$$

Since the differentials dU_1, dV_1 and dN_1 can be arbitrary, all the coefficients should be zero, leading the equilibrium condition:

$$T = T_1 = T_2 \ , \quad p = p_1 = p_2 \quad \text{and} \quad \mu = \mu_1 = \mu_2 \ . \tag{1.78}$$

These three variables are the free equilibrium parameters of this system. As the system can be divided up arbitrarily, these parameters are constant throughout the system in equilibrium.

Analogous statements can be made for other conditions.
Isentropic, isochor transformation of closed system: $dN = dS = dV = 0$ implies

$$dU \leq TdS = 0 \quad \Rightarrow \quad \begin{cases} dU \leq 0 & \text{general} \\ dU = 0 & \text{in equilibrium} \end{cases} \tag{1.79}$$

with (T, p, μ) as free equilibrium parameters.

Systems with various conditions

fixed variables	process direction	equilibrium condition	free equilibrium parameters
U, V, N	S increasing	S maximal	T, p, μ
T, V, N	F decreasing	F minimal	p, μ
T, p, N	G decreasing	G minimal	μ
S, V, N	U decreasing	U minimal	T, p, μ
S, p, N	H decreasing	H minimal	T, μ

1.5 Third law of thermodynamics

Nernst formulated 1905 based on empirical results for reversible electrical cells the law

$$\lim_{T \to 0} S(T, q_i, N_j) = S_0 \tag{1.80}$$

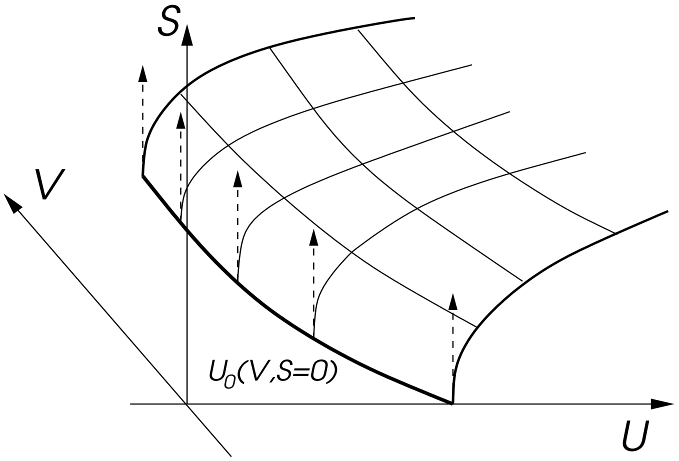
where S_0 does depend neither on q_i nor on N_j, i.e.

$$\lim_{T \to 0} \left(\frac{\partial S}{\partial q_{i'}} \right)_{q_{i \neq i'}, N_j; T} = 0 \qquad \text{and} \qquad \lim_{T \to 0} \left(\frac{\partial S}{\partial N_{j'}} \right)_{q_i, N_{j \neq j'}; T} = 0 \ . \tag{1.81}$$

Planck extended the statement by introducing an absolute scale for the entropy, $S_0 = 0$, the normalization of the entropy. This statement should be true for all thermodynamic systems.[7]
From this we obtain for the entropy as a function of U and V, that the curve for $S = 0$ defines the ground state energy U and the slop in S-direction is infinite ($1/T \to \infty$ for $T \to 0$).

---

[7] Note, that there are exceptions for systems with a residual entropy, as we will see later.



We consider now some consequences for some quantities in the zero-temperature limit for a gas. For the specific heat we find

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_V \quad \Rightarrow \quad S(T,V) = \int_0^T \frac{C_V(T')}{T'}\, dT' \,, \tag{1.82}$$

which yields for $T = 0$ the condition,

$$\lim_{T\to 0} C_V(T) = 0 \,, \tag{1.83}$$

in order to keep C_V/T integrable at $T = 0$, i.e. $C_V(T) \leq aT^\epsilon$ with $\epsilon > 0$ for $T \to 0$. Note that in a solid the lattice vibrations lead to $C_V \propto T^3$ and the electrons of metal to $C_V \propto T$. Analogously we conclude for C_p:

$$S(T,p) = \int_0^T \frac{C_p(T')}{T'}\, dT' \quad \Rightarrow \quad \lim_{T\to 0} C_p(T) = 0 \tag{1.84}$$

Moreover,

$$\lim_{T\to 0} \frac{C_p - C_V}{T} = 0 \,. \tag{1.85}$$

holds, since

$$\frac{C_p - C_V}{T} = \left(\frac{\partial p}{\partial T}\right)_V \left(\frac{\partial V}{\partial T}\right)_p \,. \tag{1.86}$$

When we use the Maxwell relation

$$\left(\frac{\partial p}{\partial T}\right)_V = \left(\frac{\partial S}{\partial V}\right)_T \tag{1.87}$$

we arrive at the result (1.85). Further examples can be found among the response functions such as the thermal expansion coefficient

$$\alpha = \frac{1}{V}\left(\frac{\partial V}{\partial T}\right)_p = -\frac{1}{V}\left(\frac{\partial S}{\partial p}\right)_T \,, \tag{1.88}$$

which obviously leads to $\lim_{T\to 0} \alpha = 0$. Similar results can be found for other response functions.

Chapter 2

Kinetic approach and Boltzmann transport theory

As we mentioned in the last chapter, thermodynamics deals with the behavior and relation of quantities of thermodynamic macroscopic systems which are in equilibrium. Thus time evolution does not enter the discussions in thermodynamics. Considering, however, a macroscopic system as built up from microscopic entities, many degrees of freedom, such as moving atoms in a gas or magnetic moments of a magnet etc, it is very natural to take time evolution into account. In this way time evolution of a macroscopic system can be analyzed providing great insight into some of the basic ideas behind the statistical treatment of macroscopic systems.

In this chapter we will first consider some idealized version of a many-body system to understand how one can view non-equilibrium behavior and how equilibrium is reached and eventually characterized. Entropy will play an essential role to give us the connection to thermodynamics of the equilibrium state. This discussion will also make clear how the enormous magnitude of number of degrees of freedom, N, is essential for the statistical approach and that small systems would not display the laws of thermodynamics in the given way. The latter part of the chapter we will turn to Boltzmann's kinetic gas theory which is one of the best known and useful examples of kinetic theories.

2.1 Time evolution and Master equation

We start by considering a model with N units (atoms, ...) which can be in z different micro-states:

$$\{s_i^\nu\} \qquad \text{with} \quad i = 1, \dots, N; \ \nu = 1, \dots, z \ . \tag{2.1}$$

For simplicity these degrees of freedom are considered to be independent and their z micro-states have the same energy. For the purpose of a simple simulation of the time evolution we introduce discrete time steps t_n with $t_{n+1} - t_n = \Delta t$. During each time step the micro-state of each unit can change from ν to ν' with a probability $p_{\nu\nu'}$ which is connected to the transition rate $\Gamma_{\nu\nu'}$ by

$$p_{\nu\nu'} = \Gamma_{\nu\nu'} \Delta t \ . \tag{2.2}$$

We require that the reverse processes have equal probability due to time reversal symmetry: $p_{\nu\nu'} = p_{\nu'\nu}$ and, thus, $\Gamma_{\nu\nu'} = \Gamma_{\nu'\nu}$.

Among the N units we find at a given time step t_n that N_ν units are in the micro-state ν. Thus, picking at random a unit i we would find it in the micro-state ν with the probability,

$$w_\nu = \frac{N_\nu}{N} \qquad \text{with} \quad \sum_{\nu=1}^{z} w_\nu = 1 \ . \tag{2.3}$$

Let us now discuss the budget of each micro-state ν: (1) the number N_ν is reduced, because some units will have a transition to another micro-state with the rate $\sum_{\nu'} \Gamma_{\nu\nu'} N_\nu$; (2) the number

N_ν increases, since some units in another micro-state undergo a transition to the micro-state ν with the rate $\sum_{\nu'} \Gamma_{\nu'\nu} N_{\nu'}$. Note that these rates are proportional to the number of units in the micro-state which is transformed into another micro-state, because each unit changes independently with the rate $\Gamma_{\nu\nu'}$. The corresponding budget equation is given by

$$N_\nu(t_{n+1}) = N_\nu(t_n) - \Delta t \sum_{\nu' \neq \nu} \Gamma_{\nu\nu'} N_\nu(t_n) + \Delta t \sum_{\nu' \neq \nu} \Gamma_{\nu'\nu} N_{\nu'}(t_n) \; . \tag{2.4}$$

This set of z iterative equations describes the evolution of the N units in a statistical way, whereby we do not keep track of the state of each individual unit, but only of the number of units in each state. Starting from an arbitrary initial configuration $\{N_\nu\}$ of the units at $t_1 = 0$, this iteration moves generally towards a fixed point, $N_\nu(t_{n+1}) = N_\nu(t_n)$ which requires that

$$0 = \sum_{\nu' \neq \nu} \Gamma_{\nu\nu'} N_\nu(t_n) - \sum_{\nu' \neq \nu} \Gamma_{\nu'\nu} N_{\nu'}(t_n) \; . \tag{2.5}$$

There is not any time evolution anymore for $N_\nu(t)$, although the states of the units are naturally changing in time. As this equation is true for all ν we find that independently

$$0 = \Gamma_{\nu\nu'} N_\nu(t_n) - \Gamma_{\nu'\nu} N_{\nu'}(t_n) \; , \tag{2.6}$$

which means that for any pair of micro-states ν and ν' the mutual transitions compensate each other. Equation (2.6) is known as the condition of *detailed balance*. Due to $\Gamma_{\nu\nu'} = \Gamma_{\nu'\nu}$ we find the fixed point condition $N_\nu = N_{\nu'} = N/z$, i.e all micro-states are equally occupied.

We now slightly reformulate the problem taking the limit of $\Delta t \to 0$ and dividing (2.4) by N:

$$\frac{dw_\nu}{dt} = -w_\nu \sum_{\nu' \neq \nu} \Gamma_{\nu\nu'} + \sum_{\nu' \neq \nu} \Gamma_{\nu'\nu} w_{\nu'} \; . \tag{2.7}$$

This is the so-called *master equation* of the system. Also here it is obvious that the detailed balance condition leads to a solution where all probabilities $w_\nu(t)$ are constant:

$$w_\nu \Gamma_{\nu\nu'} = w_{\nu'} \Gamma_{\nu'\nu} \quad \Rightarrow \quad \frac{w_\nu}{w_{\nu'}} = \frac{\Gamma_{\nu'\nu}}{\Gamma_{\nu\nu'}} = 1 \; , w_\nu = \frac{1}{z} \; . \tag{2.8}$$

We can now define a mean value of a property associated with each micro-state ν, which we call α_ν:

$$\langle \alpha \rangle = \sum_\nu \alpha_\nu w_\nu \; . \tag{2.9}$$

The time evolution of this quantity is easily derived,

$$\begin{aligned} \frac{d}{dt}\langle \alpha \rangle &= \sum_\nu \alpha_\nu \frac{dw_\nu}{dt} = \sum_\nu \alpha_\nu \left[-w_\nu \sum_{\nu' \neq \nu} \Gamma_{\nu\nu'} + \sum_{\nu' \neq \nu} \Gamma_{\nu'\nu} w_{\nu'} \right] \\ &= -\frac{1}{2} \sum_{\nu,\nu'} \{\alpha_\nu - \alpha_{\nu'}\}(w_\nu - w_{\nu'}) \Gamma_{\nu\nu'} \; , \end{aligned} \tag{2.10}$$

where the last equality is obtained by symmetrization using the symmetry of $\Gamma_{\nu\nu'}$. Obviously the time evolution stops when detailed balance is reached.

2.1.1 H-function and information

Next we introduce the function $H(t)$ which has the following form

$$H(t) = -\sum_\nu w_\nu \ln w_\nu = -\langle \ln w \rangle \; . \tag{2.11}$$

This function is a measure for the uncertainty of our knowledge about the system. Concretely, if at the time t we pick one atom at random, we may ask how certain we are about finding the atom in a given micro-state. Assume that $w_1 = 1$ and $w_\nu = 0$ for $\nu = 2, \dots, z$. Then we can be sure to find the micro-state 1. In this case $H(t) = 0$. For generic distributions of probabilities $H(t) > 0$. The larger $H(t)$ the less certain the outcome of our picking experiment is, i.e. the less information is available about the system.
The time dependence of $H(t)$ is interesting to see:

$$\frac{dH(t)}{dt} = -\frac{d}{dt}\langle \ln w \rangle = \frac{1}{2}\sum_{\nu,\nu'}(\ln w_\nu - \ln w_{\nu'})(w_\nu - w_{\nu'})\Gamma_{\nu\nu'} \geq 0 \ . \tag{2.12}$$

Note that $(x-y)(\ln x - \ln y) \geq 0$. This implies that $H(t)$ evolves in a specific direction in time, despite the assumption of time reversal symmetry. The condition $dH/dt = 0$ then implies $w_\nu = w_{\nu'}$ (detailed balance). The inequality (2.12) implies that detailed balance corresponds to the macroscopic state of the system with the least certainty and simultaneously maximal H.
Let us consider this from a different view point. For given $w_\nu = N_\nu/N$ in a system of N units, there are

$$\frac{N!}{N_1! N_2! \cdots N_z!} = W(N_\nu) \tag{2.13}$$

realizations among z^N possible configurations. In the large-N limit we may apply Stirlings formula ($\ln n! \approx n \ln n - n + \frac{1}{2}\ln 2\pi n$) which then leads to

$$\begin{aligned} H_W = \ln W(N_\nu) &\approx N \ln N - N - \sum_\nu [N_\nu \ln N_\nu - N_\nu] \\ &= -N \sum_\nu w_\nu \ln w_\nu = N H(t) \ . \end{aligned} \tag{2.14}$$

where we have ignored terms of the order $\ln N$. Maximizing H as a functional of w_ν we introduce the following extended form of H

$$H' = -\sum_\nu w_\nu \ln w_\nu + \lambda \left\{ \sum_\nu w_\nu - 1 \right\} \ , \tag{2.15}$$

where we introduced the constraint of normalization of the probability distribution using a Lagrange multiplier. Thus we get to the equation

$$\frac{dH'}{dw_\nu} = -\ln w_\nu - 1 + \lambda = 0 \qquad \text{and} \qquad \sum_\nu w_\nu = 1 \ , \tag{2.16}$$

which can only be solved by $w_\nu = 1/z$ which is again the condition of detailed balance.
The conclusion of this discussion is that the detailed balance describes the distribution w_ν with the highest probability, i.e. the largest number of different possible realizations yield the same $\{N_1, \dots, N_z\}$. In its time evolution the system moves towards the fixed point with this property and $H(t)$ grows monotonously until it reaches its maximal value.

2.1.2 Simulation of a two-state system

The master equation (2.7) which we have introduced above deals with our system of independent units in a statistical way. It is a rate equation. We would like now to consider the system in a "concrete" realization by investigating a time series of transition events to see how it evolves in time. This will be most illustrative in view of the apparent problem we encounter by the fact that our model is formulated on foundations conserving time reversal symmetry. Nevertheless there seems to be a preferred direction of evolution.

We consider a system of N units with two different micro-states ($z = 2$) and perform a computational simulation. We describe the micro-state of unit i by the quantity

$$s_i = \begin{cases} +1 & \nu = 1 \\ -1 & \nu = 2 \,. \end{cases} \tag{2.17}$$

Thus, the transition corresponds simply to a sign change of s_i, which happens with equal probability in both directions. The algorithm works in the following way. After a time step each unit experiences a transition with a probability p ($0 < p < 1$), whereby the probabilistic behavior is realized by a random number generator. This means generate a uniformly distributed random number $R \in [0, 1]$ and then do the following process for all units:

$$s_i(t_{n+1}) = \begin{cases} s_i(t_n) & \text{if } p < R < 1 \\ -s_i(t_n) & \text{if } 0 \leq R \leq p \,. \end{cases} \tag{2.18}$$

This time series corresponds to a so-called Markov chain which means that the transition at each step is completely independent from previous steps.
After each time step we determine the quantities

$$M(t) = \frac{1}{N} \sum_{i=1}^{N} s_i(t) \qquad \text{and} \qquad H(t) = -\sum_{\nu=1}^{2} w_\nu(t) \ln w_\nu(t) \tag{2.19}$$

where

$$w_1(t) = \frac{1}{2}\{1 + M(t)\} \qquad \text{and} \qquad w_2(t) = \frac{1}{2}\{1 - M(t)\} \,. \tag{2.20}$$

The results of the simulations for both quantities are shown in Fig. 2.1. In the starting configuration all units have the micro-state $s_i = +1$. There is only one such configuration, while the fixed point distribution is characterized by $w_1 = w_2$ with a large number of realizations

$$W = \frac{N!}{\{(N/2)!\}^2} \approx \sqrt{\frac{2}{\pi N}} 2^N \qquad \text{for large } N \,. \tag{2.21}$$

The closer the system is to the fixed point distribution the more realizations are available. We may estimate this by

$$\begin{aligned} W(M) &= \frac{N!}{N_1! N_2!} = \frac{N!}{\{N/2(1+M)\}!\{N/2(1-M)\}!} \\ \Rightarrow \ln W(M) &\approx N \ln 2 - \frac{N}{2}\left\{(1+M)\ln(1+M) + (1-M)\ln(1-M)\right\} - \frac{1}{2}\ln\frac{2}{\pi N} \\ &\approx N \ln 2 - \frac{1}{2}\ln\frac{2}{\pi N} - \frac{M^2 N}{2} \\ \Rightarrow W(M) &\approx 2^N \sqrt{\frac{\pi N}{2}} e^{-M^2 N/2} \,, \end{aligned} \tag{2.22}$$

where we have used the Stirling formula for large N and the expansion for $M \ll 1$ ($\ln(1 + x) = x - \frac{x^2}{2} + \cdots$). The number of realizations drop quickly from the maximal value at $M = 0$ with a width $\Delta M = \sqrt{2/N}$. The larger N the larger the fraction of realizations of the system belonging to the macroscopic state with $M \approx 0$.

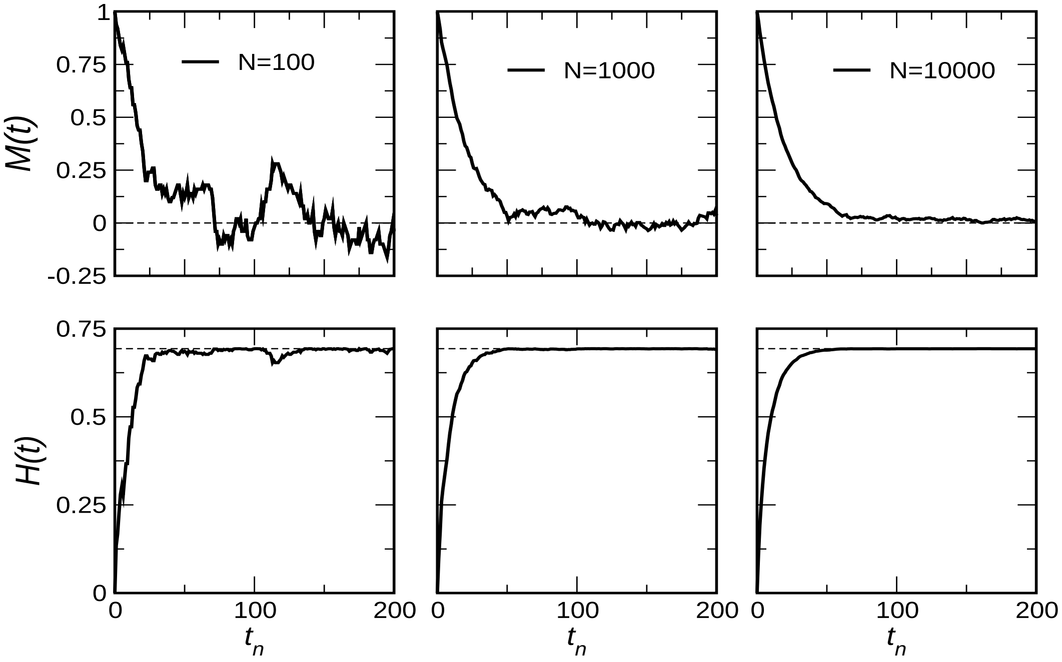


Fig. 2.1: Simulation of a two-state system for different numbers of units N*:* 1^st *column,* $N = 100$*;* 2^{nd} *column,* $N = 1000$*;* 3^rd *column,* $N = 10000$*. Note that for sufficiently large* t *(*$\gg \delta/p$*),* $M(t) \to 0$ *and* $H(t) \to \ln 2$ *.*

This explains the trend seen in the simulations. The "measured" quantities approach the fixed point but show fluctuations. These fluctuations shrink with increasing system size N. States deviating from the fixed point occupy less and less of the available configuration space for increasing N. However, it becomes obvious from the small systems that the relation found from (2.12) is not strictly true when we consider a specific time series of a finite system. The master equations leading to (2.12) are statistical and consider an averaged situation.
The master equations of our two-state system are

$$\frac{dw_1}{dt} = \Gamma(w_2 - w_1) \qquad \text{and} \qquad \frac{dw_2}{dt} = \Gamma(w_1 - w_2)\ , \tag{2.23}$$

where $\Gamma = \Gamma_{12} = \Gamma_{21}$. This leads to the equations

$$\frac{d}{dt}(w_1 + w_2) = 0 \qquad \text{and} \qquad \frac{d}{dt}(w_1 - w_2) = -2\Gamma(w_1 - w_2)\ . \tag{2.24}$$

The first equation states the conservation of the probability $w_1 + w_2 = 1$. The second leads to the exponential approach of $M(t)$ to 0,

$$M(t) = M_0 e^{-2t\Gamma}\ , \tag{2.25}$$

from which we also obtain

$$H(t) \approx \ln 2 - \frac{M_0^2}{2} e^{-t\Gamma}\ . \tag{2.26}$$

We see that the relaxation times for M and H differ by a factor two

$$\tau_M = \frac{1}{2\Gamma} \qquad \text{and} \qquad \tau_H = \frac{1}{\Gamma} \tag{2.27}$$

which can be observed by eye in the results of our simulation (Fig. 2.1.).

2.1.3 Equilibrium of a system

The statistical discussion using the master equation and our simulation of a system with many degrees of freedom shows that the system relaxes towards a fixed point of the probability distribution w_ν which describes its macroscopic state. This probability distribution of the micro-states accounts for some averaged properties, and we may calculate average values of certain quantities. This fixed point is equivalent to the equilibrium of the system. Deviations from this equilibrium yield non-equilibrium states of the system, which then decay towards the equilibrium, if we allow the system to do so. The equilibrium state is characterized in a statistical sense as the state with the maximal number of realizations in terms of configurations of micro-states of the units.

2.2 Analysis of a closed system

We turn now to a system of N units which are not independent anymore. The states have different energy, ϵ_ν. Again we describe the macro-state of the system by means of the probability distribution $w_\nu = N_\nu/N$. The system shall be closed so that no matter and no energy is exchanged with the environment. Thus, the total energy and the number of units is conserved.

$$\langle\epsilon\rangle = \sum_{\nu=1}^{z} w_\nu \epsilon_\nu \qquad \text{and} \qquad 1 = \sum_{\nu=1}^{z} w_\nu \tag{2.28}$$

This defines the internal energy $U = N\langle\epsilon\rangle$ which is unchanged in a closed system and the normalization of the probability distribution.

2.2.1 *H* and the equilibrium thermodynamics

As before we define the function $H(t)$, an increasing function of time, saturating at a maximal value, which we consider as the condition for "equilibrium". The details of the time evolution will be addressed later. Rather we search now for the maximum of H with respect to w_ν under the condition of energy conservation. Thus, we introduce again Lagrange multipliers for these constraints

$$H(w_\nu) = -\sum_\nu w_\nu \ln w_\nu + \lambda\{\sum_\nu w_\nu - 1\} - \frac{1}{\theta}\{\sum_\nu w_\nu\epsilon_\nu - \langle\epsilon\rangle\} \tag{2.29}$$

and obtain the equation

$$0 = \frac{dH}{dw_\nu} = -\ln w_\nu - 1 + \lambda - \frac{\epsilon_\nu}{\theta} \; . \tag{2.30}$$

This equation leads to

$$w_\nu = e^{\lambda - 1 - \epsilon_\nu/\theta} \qquad \text{with} \qquad e^{1-\lambda} = \sum_\nu e^{-\epsilon_\nu/\theta} = Z \tag{2.31}$$

which satisfies the normalization condition for w_ν.
Our aim is it now to give the different quantities a thermodynamic meaning. We begin by multiplying Eq.(2.30) by w_ν and sum then over ν. This leads to

$$0 = -\sum_\nu \left\{ w_\nu \ln w_\nu + w_\nu(1-\lambda) + w_\nu \frac{\epsilon_\nu}{\theta} \right\} = H - 1 + \lambda - \frac{\langle\epsilon\rangle}{\theta} \tag{2.32}$$

which we use to replace $1-\lambda$ in (2.30) and obtain

$$\langle\epsilon\rangle = \theta(\ln w_\nu + H) + \epsilon_\nu \; . \tag{2.33}$$

The differential reads

$$d\langle\epsilon\rangle = (H + \ln w_\nu)d\theta + \theta\left(dH + \frac{dw_\nu}{w_\nu}\right) + d\epsilon_\nu \; . \tag{2.34}$$

After multiplying by w_ν and a ν-summation the first term on the right hand side drops out and we obtain

$$d\langle\epsilon\rangle = \theta dH + \sum_\nu w_\nu d\epsilon_\nu \; . \tag{2.35}$$

Now we view $\langle\epsilon\rangle = u$ as the internal energy per unit and $d\epsilon_\nu = \sum_i \frac{\partial\epsilon_\nu}{\partial q_i} dq_i$ where q_i is a generalized coordinate, such as a volume, a magnetization etc. Thus $\frac{\partial\epsilon_\nu}{\partial q_i} = -F_{iw_\nu}$ is a generalized force, such as pressure, magnetic field etc. Therefore we write (2.35) as

$$du = \theta dH - \sum_{\nu,i} w_\nu F_{i\nu} dq_i = \theta dH - \sum_i \langle F_i\rangle dq_i \quad \Rightarrow \quad dH = \frac{du}{\theta} + \frac{1}{\theta}\sum_i \langle F_i\rangle dq_i \;;. \tag{2.36}$$

This suggests now to make the following identifications:

$$\theta = k_B T \qquad \text{and} \qquad H = \frac{s}{k_B} \; , \tag{2.37}$$

i.e. θ is the temperature and H the entropy density. We may then rewrite (2.33)

$$\ln w_\nu = \frac{\Psi - \epsilon_\nu}{k_B T} \; . \tag{2.38}$$

Here we now take the derivative with respect to T:

$$\frac{\partial w_\nu}{\partial T} = \frac{w_\nu}{k_B T^2}\left\{T\frac{\partial\Psi}{\partial T} - \Psi + \epsilon_\nu\right\} \tag{2.39}$$

Under the sum over ν the left hand side vanishes and using (2.38) we can express the entropy density as

$$s = -\frac{\partial\Psi}{\partial T} \; . \tag{2.40}$$

Analogously we derive

$$\langle F_i\rangle = -\frac{\partial\Psi}{\partial q_i} \; . \tag{2.41}$$

Thus we can identify $\Psi = \Psi(T, q_i)$ as the Helmholtz free energy density $f(T, q_i)$, a thermodynamic potential.[1]

2.2.2 Master equation

Because the micro-states have different energies, the time evolution as we have discussed by the master equation for completely independent units has to be modified in a way that the energy in each time step is conserved. This can be guaranteed only by involving two micro-states of different units to be transformed together conserving the energy. A master equation doing this job has the form

$$\frac{dw_\nu}{dt} = \sum_{\nu_1,\nu_2,\nu_3} \left\{\Gamma_{\nu\nu_1;\nu_2\nu_3} w_{\nu_2} w_{\nu_3} - \Gamma_{\nu_2\nu_3;\nu\nu_1} w_\nu w_{\nu_1}\right\} \; . \tag{2.44}$$

For this purpose we assume that there is no correlation between the states of different units. Here $\Gamma_{\nu\nu_1;\nu_2\nu_3}$ denotes the rate for the transition of the pair of states (ν_2, ν_3) to (ν, ν_1) under the

---

[1] Time dependence of Ψ: From Eq.(2.38) we obtain

$$\Psi = k_B T \sum_\nu w_\nu \ln w_\nu + \langle\epsilon\rangle = -k_B T H + \langle\epsilon\rangle \; . \tag{2.42}$$

The time derivative is then

$$\frac{d\Psi}{dt} = -k_B T \frac{dH}{dt} \leq 0 \; , \tag{2.43}$$

which means that free energy tends towards a minimum, as we have seen in the previous chapter.

condition that $\epsilon_{\nu_2} + \epsilon_{\nu_3} = \epsilon_\nu + \epsilon_{\nu_1}$. Time reversal symmetry and the freedom to exchange two states give rise to the following relations:

$$\Gamma_{\nu\nu_1;\nu_2\nu_3} = \Gamma_{\nu_2\nu_3;\nu\nu_1} = \Gamma_{\nu_1\nu;\nu_2\nu_3} = \Gamma_{\nu\nu_1;\nu_3\nu_2} \; . \tag{2.45}$$

This can be used in order to show that the H-function is only increasing with time,

$$\begin{aligned} \frac{dH}{dt} &= -\sum_\nu \frac{dw_\nu}{dt} \{\ln w_\nu + 1\} \\ &= \frac{1}{4} \sum_{\nu,\nu_1,\nu_2,\nu_3} \Gamma_{\nu,\nu_1;\nu_2\nu_3} (w_\nu w_{\nu_1} - w_{\nu_2} w_{\nu_3}) \{\ln(w_\nu w_{\nu_1}) - \ln(w_{\nu_2} w_{\nu_3})\} \geq 0 \; . \end{aligned} \tag{2.46}$$

The equilibrium is reached when $dH/dt = 0$ and using (2.31) leads to,

$$w_\nu w_{\nu_1} = \frac{e^{-\epsilon_\nu/k_B T} e^{-\epsilon_{\nu_1}/k_B T}}{Z^2} = \frac{e^{-\epsilon_{\nu_2}/k_B T} e^{-\epsilon_{\nu_3}/k_B T}}{Z^2} = w_{\nu_2} w_{\nu_3} \; . \tag{2.47}$$

We now revisit the condition of detailed balance under these new circumstances. On the one hand, we may just set each term in the sum of the right hand side of (2.44) to zero to obtain a detailed balance statement:

$$0 = \Gamma_{\nu\nu_1;\nu_2\nu_3} w_{\nu_2} w_{\nu_3} - \Gamma_{\nu_2\nu_3;\nu\nu_1} w_\nu w_{\nu_1} = \Gamma_{\nu\nu_1;\nu_2\nu_3} \{w_{\nu_2} w_{\nu_3} - w_\nu w_{\nu_1}\} \; , \tag{2.48}$$

which is a consequence of time reversal. On the other hand, we may compress the transition rates in order to reach at a similar form as we had it earlier in (2.6),

$$\Gamma'_{\nu\nu'} w_\nu = \Gamma'_{\nu'\nu} w_{\nu'} \; . \tag{2.49}$$

where we define

$$\Gamma'_{\nu\nu'} = \sum_{\nu_1,\nu_2} \Gamma_{\nu\nu_1;\nu'\nu_2} w_{\nu_1} \qquad \text{and} \qquad \Gamma'_{\nu'\nu} = \sum_{\nu_1,\nu_2} \Gamma_{\nu\nu_1;\nu'\nu_2} w_{\nu_2} \; . \tag{2.50}$$

It is important now to notice that time reversal does not invoke $\Gamma'_{\nu\nu'} = \Gamma'_{\nu'\nu}$, but we find that at equilibrium

$$\Gamma'_{\nu'\nu} = \sum_{\nu_1,\nu_2} \Gamma_{\nu\nu_1;\nu'\nu_2} \frac{e^{-\epsilon_{\nu_2}/k_B T}}{Z} = e^{-(\epsilon_\nu - \epsilon_{\nu'})/k_B T} \sum_{\nu_1,\nu_2} \Gamma_{\nu\nu_1;\nu'\nu_2} \frac{e^{-\epsilon_{\nu_1}/k_B T}}{Z} = e^{-(\epsilon_\nu - \epsilon_{\nu'})/k_B T} \Gamma'_{\nu\nu'} \; . \tag{2.51}$$

Thus detailed balance implies here different transition rates for the two directions of processes depending on the energy difference between the two micro-states, $\epsilon_\nu - \epsilon_{\nu'}$, and the temperature T,

$$\frac{w_\nu}{w_{\nu'}} = \frac{\Gamma'_{\nu'\nu}}{\Gamma'_{\nu\nu'}} = e^{-(\epsilon_\nu - \epsilon_{\nu'})/k_B T} \; . \tag{2.52}$$

The degrees of freedom of each unit fluctuate in the environment (heat reservoir) of all the other units which can interact with it.

2.2.3 Irreversible effect and increase of entropy

Although thermodynamics shows its strength in the description of reversible processes, it provides also important insights into the irreversible changes. An example is the Gay-Lussac experiment. However, the nature of the irreversible process and in particular its time evolution is not part of thermodynamics usually. A process which considers the evolution of a system from a non-equilibrium to its equilibrium state is irreversible generally.

Here we examine briefly some aspects of this evolution by analyzing a closed system consisting of two subsystems, 1 and 2, which are initially independent (inhibited equilibrium) but then at $t = 0$ become very weakly coupled. Both systems shall have the same type of units with their micro-states which they occupy with probabilities $w_\nu^{(1)} = N_\nu^{(1)}/N_1$ and $w_\nu^{(2)} = N_\nu^{(2)}/N_2$, respectively ($N_i$ is the number of units in system i). Each subsystem is assumed to be in its equilibrium internally which can be sustained approximatively as long as the coupling between them is weak.

The number of configurations for the combined system is given by the product of the number of configurations of the subsystems:

$$W = W_1 W_2 \qquad \text{with} \qquad W_i = \frac{N_i!}{N_1^{(i)}! N_2^{(i)}! \cdots N_z^{(i)}!} \tag{2.53}$$

which then leads to the entropy

$$\begin{aligned} S &= -k_B \ln W = -k_B \ln W_1 - k_B \ln W_2 = S_1 + S_2 \\ &= -N_1 k_B \sum_\nu w_\nu^{(1)} \ln w_\nu^{(1)} - N_2 k_B \sum_\nu w_\nu^{(2)} \ln w_\nu^{(2)} \ . \end{aligned} \tag{2.54}$$

In the complete system the internal energy is conserved, $U = U_1 + U_2$. We define U_{01} and U_{02} as the internal energy of the two subsystems, if they are in equilibrium with each other. The non-equilibrium situation is then parametrized for the energy by the deviation $\tilde{U}$: $U_1 = U_{01} + \tilde{U}$ and $U_2 = U_{02} - \tilde{U}$. The entropy satisfies at equilibrium

$$\begin{aligned} & S(\tilde{U}) = S_1(U_{01} + \tilde{U}) + S_2(U_{02} - \tilde{U}) \\ & \qquad \text{with} \quad 0 = \frac{\partial S}{\partial \tilde{U}} = \frac{\partial S_1}{\partial U_1} - \frac{\partial S_2}{\partial U_2} = \frac{1}{T_1} - \frac{1}{T_2} \qquad \Rightarrow T_1 = T_2 = T_0 \ , \end{aligned} \tag{2.55}$$

for $\tilde{U} = 0$. Thus, we find for small $\tilde{U}$,

$$\begin{aligned} \frac{1}{T_1} &= \left.\frac{\partial S_1}{\partial U_1}\right|_{U_1 = U_{01} + \tilde{U}} = \frac{1}{T_0} + \tilde{U} \left.\frac{\partial^2 S_1}{\partial U_1^2}\right|_{U_1 = U_{01}} \ , \\ \frac{1}{T_2} &= \left.\frac{\partial S_2}{\partial U_2}\right|_{U_2 = U_{02} - \tilde{U}} = \frac{1}{T_0} - \tilde{U} \left.\frac{\partial^2 S_2}{\partial U_2^2}\right|_{U_2 = U_{02}} \ . \end{aligned} \tag{2.56}$$

Using (2.55) we then expand

$$S(\tilde{U}) = S_1(U_{01}) + S_2(U_{02}) + \frac{\tilde{U}^2}{2} \left(\left.\frac{\partial^2 S_1}{\partial U_1^2}\right|_{U_{01}} + \left.\frac{\partial^2 S_2}{\partial U_2^2}\right|_{U_{02}} \right) + O(\tilde{U}^3) \ . \tag{2.57}$$

so that the time evolution of the entropy is given by

$$\frac{dS}{dt} = \tilde{U} \frac{d\tilde{U}}{dt} \left(\left.\frac{\partial^2 S_1}{\partial U_1^2}\right|_{U_{01}} + \left.\frac{\partial^2 S_2}{\partial U_2^2}\right|_{U_{02}} \right) = \frac{d\tilde{U}}{dt} \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \ . \tag{2.58}$$

The derivative $d\tilde{U}/dt$ corresponds to the energy flow from one subsystem to the other. We express $d\tilde{U}/dt$ in terms of the distribution functions of one of the subsystems, say system 1.

$$\frac{d\tilde{U}}{dt} = N_1 \sum_\nu \epsilon_\nu \frac{dw_\nu^{(1)}}{dt} = N_1 \sum_\nu \epsilon_\nu \frac{\partial w_\nu^{(1)}}{\partial T_1} \frac{dT_1}{dt} \approx N_1 \frac{dT_1}{dt} \sum_\nu \epsilon_\nu \frac{\partial}{\partial T_1} \frac{e^{-\epsilon_\nu / k_B T_1}}{Z_1} \ . \tag{2.59}$$

The derivative yields

$$\frac{d\tilde{U}}{dt} \approx N_1 \frac{dT_1}{dt} \left[\sum_\nu \epsilon_\nu^2 w_\nu^{(1)} - \left\{ \sum_\nu \epsilon_\nu w_\nu^{(1)} \right\}^2 \right] \frac{1}{k_B T_1^2} = \frac{N_1}{k_B T_1^2} \frac{dT_1}{dt} \left\{ \langle \epsilon^2 \rangle - \langle \epsilon \rangle^2 \right\}_1 = C_1 \frac{dT_1}{dt} \ , \tag{2.60}$$

where C_1 denotes the specific heat of subsystem 1, since

$$\frac{d\tilde{U}}{dt} = \frac{dT_1}{dt} \frac{\partial \tilde{U}}{\partial T_1} = C_1 \frac{dT_1}{dt} \ , \tag{2.61}$$

and we have derived the relation

$$k_B T_1^2 C_1 = N_1 \left\{ \langle \epsilon^2 \rangle - \langle \epsilon \rangle^2 \right\}_1 \ ;. \tag{2.62}$$

which connects the specific heat with the fluctuations of the energy value. Note that the internal energy of subsystem 1 is not conserved due to the coupling to subsystem 2.

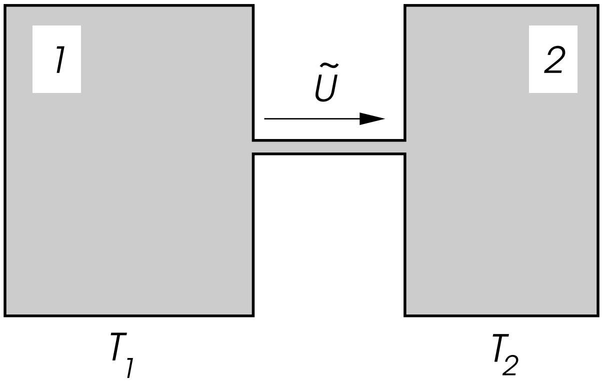


Fig. 2.2: A closed system consisting of two weakly coupled subsystems, 1 and 2, with initial temperatures T_1 and T_2 respectively, evolves towards equilibrium through energy transfer from the warmer to the colder subsystem.

In which direction does the energy flow? Because we know that $dS/dt \geq 0$, from (2.58) follows

$$0 \leq \frac{dS}{dt} = C_1 \frac{dT_1}{dt} \frac{T_2 - T_1}{T_1 T_2} \ , \tag{2.63}$$

which yields for $T_1 > T_2$ that

$$\frac{dT_1}{dt} < 0 \quad \Rightarrow \quad \frac{d\tilde{U}}{dt} < 0 \ . \tag{2.64}$$

This means that the energy flow goes from the warmer to the colder subsystem, reducing the temperature of the warmer system and increasing that of the colder subsystem as can be seen by the analogous argument. The flow stops when $T_1 = T_2 = T_0$.

Let us now consider the situation again by means of the master equation. We may approximate the equation of $w_\nu^{(a)}$ by $(a, b = 1, 2)$,

$$\frac{dw_{\nu_1}^{(a)}}{dt} \approx \sum_{\nu_1', \nu_2, \nu_2'} \sum_b \left[\Gamma^{(ab)}_{\nu_1, \nu_2; \nu_1' \nu_2'} w_{\nu_1'}^{(a)} w_{\nu_2'}^{(b)} - \Gamma^{(ab)}_{\nu_1', \nu_2'; \nu_1 \nu_2} w_{\nu_1}^{(a)} w_{\nu_2}^{(b)} \right] \tag{2.65}$$

which describes the evolution of the probability distribution transferring in general energy between the two subsystems, as only the overall energy is conserved but not in each subsystem separately,

$$\epsilon_{\nu_1} + \epsilon_{\nu_2} = \epsilon_{\nu_1'} + \epsilon_{\nu_2'} \ . \tag{2.66}$$

All processes inside each subsystem are considered at equilibrium and give by the condition of detailed balance no contribution. This remains true as long as the subsystems are very large and the coupling between them weak. For symmetry reasons

$$\Gamma^{(12)}_{\nu_1\nu_2;\nu_1'\nu_2'} = \Gamma^{(12)}_{\nu_1'\nu_2';\nu_1\nu_2} = \Gamma^{(12)}_{\nu_2\nu_1;\nu_1'\nu_2'} = \Gamma^{(12)}_{\nu_1\nu_2;\nu_2'\nu_1'} > 0 \; . \tag{2.67}$$

We use now this equation and estimate $d\tilde{U}/dt$

$$\begin{aligned} \frac{d\tilde{U}}{dt} &= \sum_{\nu_1,\nu_1',\nu_2,\nu_2'} \epsilon_{\nu_1} \left[\Gamma^{(12)}_{\nu_1,\nu_2;\nu_1'\nu_2'} w^{(1)}_{\nu_1'} w^{(2)}_{\nu_2'} - \Gamma^{(12)}_{\nu_1',\nu_2';\nu_1\nu_2} w^{(1)}_{\nu_1} w^{(2)}_{\nu_2} \right] \\ &= \frac{1}{2} \sum_{\nu_1,\nu_1',\nu_2,\nu_2'} (\epsilon_{\nu_1} - \epsilon_{\nu_1'}) \Gamma^{(12)}_{\nu_1\nu_2;\nu_1'\nu_2'} \frac{e^{-\epsilon_{\nu_1'}/k_BT_1 - \epsilon_{\nu_2'}/k_BT_2} - e^{-\epsilon_{\nu_1}/k_BT_1 - \epsilon_{\nu_2}/k_BT_2}}{Z_1 Z_2} \; . \end{aligned} \tag{2.68}$$

Using (2.66) and assuming that T_1 and T_2 are close to each other, we obtain

$$\frac{d\tilde{U}}{dt} \approx \frac{1}{2Z^2} \sum_{\nu_1,\nu_1',\nu_2,\nu_2'} \Gamma^{(12)}_{\nu_1\nu_2;\nu_1'\nu_2'} e^{-(\epsilon_{\nu_1}+\epsilon_{\nu_2})/k_BT_2} (\epsilon_{\nu_1} - \epsilon_{\nu_1'})^2 \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \approx K(T_2 - T_1) \tag{2.69}$$

with $K > 0$. The energy flow is proportional to the difference of the temperatures and stops when the two subsystems have the same temperature. It is also here obvious that the spontaneous energy transfer goes from the warmer to the colder reservoir. This corresponds to Fick's law as $d\tilde{U}/dt$ can be considered as a heat current (energy current). The described heat transport is an irreversible process during which the entropy increases, leading from the less likely system configuration of an inhibited equilibrium to the most likely of complete equilibrium with the entropy

$$S = N \sum_{\nu} w_\nu \ln w_\nu \qquad \text{with} \qquad w_\nu = \frac{N_\nu}{N} \; , \tag{2.70}$$

i.e. the probability distribution is uniform for the whole system with $N = N_1 + N_2$ units.

2.3 Boltzmann's transport theory

We now consider one of the most well-known concrete examples of a kinetic theory a system of particles which can collide with each other. This system is used to study the problem of an atomic gas within a statistical framework and is also very useful to treat various transport properties.

2.3.1 Statistical formulation

For the purpose of a statistical description we introduce here the distribution function $f(\vec{r}, \vec{p}, t)$ whose argument is a position $(\vec{r}, \vec{p})$ in the particle phase space Υ and time t. This is analogous to the previously discussed w_ν. This distribution function is defined in the following way:

$$f(\vec{r}, \vec{p}, t) d^3r \; d^3p = \begin{array}{l} \text{number of particles in the small (infinitesimal) phase space} \\ \text{volume } d^3r d^3p \text{ around the position } (\vec{r}, \vec{p}) \text{ at time } t, \end{array} \tag{2.71}$$

where it is important that d^3rd^3p is large enough to contain many particles, but very small compared to the total phase space volume (smaller than the characteristic length scales on which the distribution function changes). An alternative and equivalent definition for large N is

$$f(\vec{r}, \vec{p}, t) d^3r \; d^3p = \begin{array}{l} N \times \text{ probability that a particle is found in the} \\ \text{volume } d^3r d^3p \text{ around the position } (\vec{r}, \vec{p}) \text{ at the time } t. \end{array} \tag{2.72}$$

Based on this definition the following relations hold:

$$N = \int_\Upsilon f(\vec{r},\vec{p},t)\; d^3r\; d^3p\;, \quad n_r(\vec{r},t) = \int_\Upsilon f(\vec{r},\vec{p},t)\; d^3p\;, \quad n_p(\vec{p},t) = \int_\Upsilon f(\vec{r},\vec{p},t)\; d^3r\;, \tag{2.73}$$

where $n_r(\vec{r},t)$ and $n_p(\vec{p},t)$ are the particle density at the point $\vec{r}$ and the density of particles with momentum $\vec{p}$, respectively, at time t.

Now we consider the temporal evolution of $f(\vec{r},\vec{p},t)$. The volume d^3rd^3p at $(\vec{r},\vec{p})$ in Υ-space moves after the infinitesimal time step δt to $d^3r'd^3p'$ at $(\vec{r}\,',\vec{p}\,') = (\vec{r}+\delta\vec{r},\vec{p}+\delta\vec{p})$ where $\delta\vec{r} = \vec{v}\delta t$ and $\delta\vec{p} = \vec{F}\delta t$. Here, $\vec{v} = \vec{p}/m$ is the particle velocity (m: particle mass) and $\vec{F} = \vec{F}(\vec{r})$ is the force acting on the particles at $\vec{r}$.

If there is no scattering of the particles, then the number of particles in the time-evolving volume is conserved:

$$f(\vec{r}+\vec{v}\delta t,\vec{p}+\vec{F}\delta t,t+\delta t)\; d^3r'\; d^3p' = f(\vec{r},\vec{p},t)\; d^3r\; d^3p \tag{2.74}$$

Since the phase space volume is conserved during the temporal evolution ($d^3r\; d^3p = d^3r'\; d^3p'$), the distribution function remains unchanged,

$$f(\vec{r}+\vec{v}\delta t,\vec{p}+\vec{F}\delta t,t+\delta t) = f(\vec{r},\vec{p},t) \tag{2.75}$$

If we now include collisions of the particles among each other, then the number of particles is not conserved anymore. A particle in the small volume $d^3r\; d^3p$ scattering with any other particle during the time step δt will be removed from this volume. On the other hand, other particles may be placed into this volume through scattering processes. Thus, the more general equation is given by

$$f(\vec{r}+\vec{v}\delta t,\vec{p}+\vec{F}\delta t,t+\delta t) = f(\vec{r},\vec{p},t) + \left(\frac{\partial f}{\partial t}\right)_{coll} \delta t \tag{2.76}$$

which can be expanded leading to

$$\left(\frac{\partial}{\partial t} + \vec{v}\cdot\vec{\nabla}_{\vec{r}} + \vec{F}\cdot\vec{\nabla}_{\vec{p}}\right) f(\vec{r},\vec{p},t) = \mathcal{D}f(\vec{r},\vec{p},t) = \left(\frac{\partial f}{\partial t}\right)_{coll}\;. \tag{2.77}$$

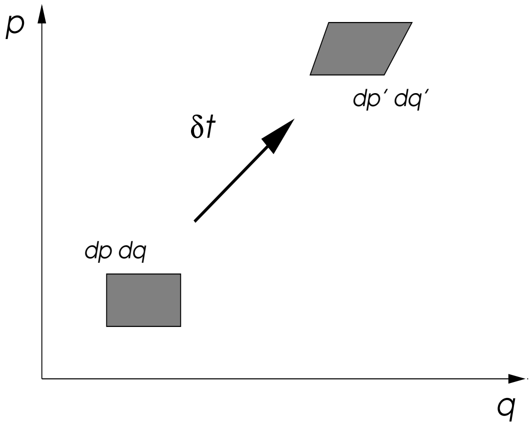


Fig. 2.3: Temporal evolution of the phase space volume $d^3p\; d^3q$.

The derivative $\mathcal{D}$ considers the change of f with the "flow" in the motion in phase space. The right-hand-side is the "collision integral" taking the effect of collisions on the evolution of f into account. Here many-body aspects appear.

2.3.2 Collision integral

In a two-particle collision, $(\vec{p}_1, \vec{p}_2) \rightarrow (\vec{p}_1', \vec{p}_2')$ we find the total momentum and energy as conserved quantities (elastic collision),

$$\vec{p}_1 + \vec{p}_2 = \vec{p}_1' + \vec{p}_2' \quad \text{and} \quad E_1 + E_2 = E_1' + E_2' \ , \tag{2.78}$$

with $E = \vec{p}^2/2m$. Of course, also the number of particles is conserved.
As indicated above there are two types of collision processes to consider, the ones which decrease and those which increase the number of particles inside the volume $d^3r\ d^3p$ at $(\vec{r}, \vec{p})$. We introduce the transition probability in a collision as

$$W(\vec{p}_1, \vec{p}_2; \vec{p}_1', \vec{p}_2') = \frac{d\sigma}{d\Omega}(\vec{p}_1, \vec{p}_2; \vec{p}_1', \vec{p}_2')\delta(\vec{p}_1 + \vec{p}_2 - \vec{p}_1' - \vec{p}_2')\delta(E_1 + E_2 - E_1' - E_2') \ . \tag{2.79}$$

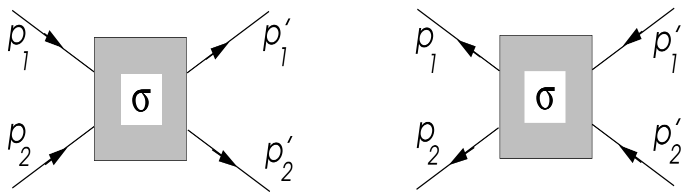


Fig. 2.4: Collision work in both time directions and the matrix elements satisfy a number of symmetry relations.

The following symmetries apply to W. The collisions are symmetric under spatial inversion

$$W(\vec{p}_1, \vec{p}_2; \vec{p}_1', \vec{p}_2') = W(-\vec{p}_1, -\vec{p}_2; -\vec{p}_1', -\vec{p}_2') \tag{2.80}$$

and time reversal invariant,

$$W(\vec{p}_1, \vec{p}_2; \vec{p}_1', \vec{p}_2') = W(-\vec{p}_1', -\vec{p}_2'; -\vec{p}_1, -\vec{p}_2) \ . \tag{2.81}$$

Moreover, the exchange of particles does not change W:

$$W(\vec{p}_1, \vec{p}_2; \vec{p}_1', \vec{p}_2') = W(\vec{p}_2, \vec{p}_1; \vec{p}_2', \vec{p}_1') \ . \tag{2.82}$$

Now we identify $\vec{p}_1$ with $\vec{p}$ in the volume $d^3r d^3p$ and write the collision integral as

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = \int_\Upsilon d^3p_2\ d^3p_1'\ d^3p_2'\ W(\vec{p}, \vec{p}_2; \vec{p}_1', \vec{p}_2') \left\{F(\vec{r}, t; \vec{p}_1', \vec{p}_2') - F(\vec{r}, t; (\vec{p}, \vec{p}_2)\right\} \ , \tag{2.83}$$

where $F(\vec{r}, t; \vec{p}_1, \vec{p}_2)$ is a two-particle correlation function giving the probability to find at the time t two particles with momenta $\vec{p}_1$ and $\vec{p}_2$, in the volume d^3r around point $\vec{r}$. The first term describes the scattering of a particle into the volume $d^3r\ d^3p$ from two particles with momenta $\vec{p}_1'$ and $\vec{p}_2'$ outside of $d^3r\ d^3p$. The second term stands for the opposite collision event.
Generally, the correlation functions $F(\vec{r}, t; \vec{p}_1, \vec{p}_2)$ are difficult to handle. Therefore we follow Boltzmann and assume "molecular chaos", i.e. the momenta of the particles are completely uncorrelated. This leads to

$$F(\vec{r}, t; \vec{p}_1, \vec{p}_2) \approx f(\vec{r}, \vec{p}_1, t) f(\vec{r}, \vec{p}_2, t) \ , \tag{2.84}$$

from which follows the well-known Boltzmann form of the collision integral

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = \int_\Upsilon d^3p_2\ d^3p_1'\ d^3p_2'\ W(\vec{p}, \vec{p}_2; \vec{p}_1', \vec{p}_2') \left\{f(\vec{r}, \vec{p}_1', t) f(\vec{r}, \vec{p}_2', t) - f(\vec{r}, \vec{p}, t) f(\vec{r}, \vec{p}_2, t)\right\} \ . \tag{2.85}$$

This equation (2.85) is known as the *Boltzmann equation* and is the analog to the previously discussed master equation for a closed system.

2.3.3 Collision conserved quantities

As mentioned above there are various quantities which are conserved in a two-particle collision: particle number, momentum and energy. We denote these quantities as $\varphi(\vec{p})$ and the corresponding density as

$$\rho_\varphi(\vec{r},t) = \int_\Upsilon d^3p\ \varphi(\vec{p}) f(\vec{r},\vec{p},t) \tag{2.86}$$

with

$$\begin{aligned} \varphi(\vec{p}) &= 1 && \text{particle number} \\ \varphi(\vec{p}) &= \vec{a}\cdot\vec{p} && \text{momentum} \\ \varphi(\vec{p}) &= \frac{\vec{p}^2}{2m} + V(\vec{r}) && \text{energy ,} \end{aligned} \tag{2.87}$$

where $\vec{a}$ is an arbitrary constant vector. We now claim that for any collision-invariant quantity $\varphi(\vec{p})$ the following relation holds:

$$\left(\frac{\partial \rho_\varphi}{\partial t}\right)_{coll} = \int_\Upsilon d^3p\ \varphi(\vec{p}) \left(\frac{\partial f}{\partial t}\right)_{coll} = 0\ . \tag{2.88}$$

For the proof of this statement we use the symmetries of $W(\vec{p},\vec{p}_2;\vec{p}_1',\vec{p}_2')$. We can then write

$$\begin{aligned} \left(\frac{\partial \rho_\varphi}{\partial t}\right)_{coll} &= \int_\Upsilon d^3p\ d^3p_2\ d^3p_1'\ d^3p_2'\ W(\vec{p},\vec{p}_2;\vec{p}_1',\vec{p}_2') \\ &\qquad \times\frac{1}{4}\left\{ \begin{array}{l} \varphi(\vec{p})\left\{f(\vec{r},\vec{p}_1',t)f(\vec{r},\vec{p}_2',t) - f(\vec{r},\vec{p},t)f(\vec{r},\vec{p}_2,t)\right\} + \\ \varphi(\vec{p}_2)\left\{f(\vec{r},\vec{p}_2',t)f(\vec{r},\vec{p}_1',t) - f(\vec{r},\vec{p}_2,t)f(\vec{r},\vec{p},t)\right\} + \\ \varphi(\vec{p}_1')\left\{f(\vec{r},\vec{p},t)f(\vec{r},\vec{p}_2,t) - f(\vec{r},\vec{p}_1',t)f(\vec{r},\vec{p}_2',t)\right\} + \\ \varphi(\vec{p}_2')\left\{f(\vec{r},\vec{p}_2,t)f(\vec{r},\vec{p},t) - f(\vec{r},\vec{p}_2',t)f(\vec{r},\vec{p}_1',t)\right\} \end{array} \right. \\ &= \int_\Upsilon d^3p\ d^3p_2\ d^3p_1'\ d^3p_2'\ W(\vec{p},\vec{p}_2;\vec{p}_1',\vec{p}_2') \\ &\qquad \times\frac{1}{4}\underbrace{\left\{\varphi(\vec{p})+\varphi(\vec{p}_2)-\varphi(\vec{p}_1')-\varphi(\vec{p}_2')\right\}}_{=0}\left\{f(\vec{r},\vec{p}_1',t)f(\vec{r},\vec{p}_2',t) - f(\vec{r},\vec{p},t)f(\vec{r},\vec{p}_2,t)\right\} \\ &= 0\ . \end{aligned} \tag{2.89}$$

2.3.4 Boltzmann's H-theorem

For simplicity we now consider a homogeneous system without external forces. Then the distribution function does not depend on $\vec{r}$. Thus, the Boltzmann equation is reduced to the form,

$$\frac{\partial f(\vec{p},t)}{\partial t} = \left(\frac{\partial f}{\partial t}\right)_{coll}\ . \tag{2.90}$$

We define the functional

$$H(t) = -\int_\Upsilon d^3p\ f(\vec{p},t)\ln f(\vec{p},t)\ , \tag{2.91}$$

whose time derivative is

$$\frac{dH(t)}{dt} = -\int d^3p \frac{\partial f(\vec{p},t)}{\partial t} \left[1 + \ln f(\vec{p},t)\right] \ . \tag{2.92}$$

The following theorem applies for $H(t)$.

Boltzmann's H-theorem: For $f(\vec{p},t)$ satisfying Boltzmann's equation, generally the inequality

$$\frac{dH}{dt} \geq 0 \tag{2.93}$$

holds and the equal-sign applies, if the system is in equilibrium.
This theorem can be proven in an analogous way as in section 2.2.2 taking (2.90) and (2.92) to write (under the assumption of molecular chaos).

$$\frac{dH}{dt} = -\int_{\Upsilon} d^3p \; d^3p_2 \; d^3p_1' \; d^3p_2' \; W(\vec{p},\vec{p}_2;\vec{p}_1',\vec{p}_2') \left\{ f(\vec{p}_1',t) f(\vec{p}_2',t) - f(\vec{p},t) f(\vec{p}_2,t) \right\} \left[1 + \ln f(\vec{p},t)\right] \ . \tag{2.94}$$

We introduce the short notation for $f(\vec{p}_1,t) = f_1$ etc.. Using the same trick as in (2.89) we can rewrite this equation as

$$\frac{dH}{dt} = -\int_{\Upsilon} d^3p \; d^3p_2 \; d^3p_1' \; d^3p_2' \; W(\vec{p},\vec{p}_2;\vec{p}_1',\vec{p}_2') \underbrace{(f_1' f_2' - f f_2) \left[\ln(f f_2) - \ln(f_1' f_2')\right]}_{\leq 0} \geq 0 \ . \tag{2.95}$$

This statement is true under all choices of f. Moreover, $dH/dt = 0$ is only realized, if the integrand is identical zero:

$$\frac{dH}{dt} = 0 \quad \Leftrightarrow \quad f(\vec{p},t) f(\vec{p}_2,t) = f(\vec{p}_1',t) f(\vec{p}_2',t) \tag{2.96}$$

where the collision process is $(\vec{p},\vec{p}_1) \leftrightarrow (\vec{p}_1',\vec{p}_2')$. Under this condition f is not time-dependent anymore and corresponds to the equilibrium situation. Any solution of the Boltzmann equation describing the evolution of a system towards equilibrium satisfies the H-theorem.
Thus, as in the previous discussions based on the master equations, the Boltzmann equation also describes a system with an apparently "preferred" time direction. This seems again at odds with the time reversal symmetry incorporated in the theory. This behavior relies again on the trend of the system to spend most of the time in the macroscopic state with the largest number of possible microscopic realizations. This discrimination of macroscopic states is better the larger the system. From this point of view also Poincaré's recurrence theorem, which states that any mechanical system would after a certain time approach any phase space configuration arbitrarily close to an initial configuration, would not be a real objection. However, in reality for a macroscopically large number of particles the recurrence time is long and would in practice exceed the age of the universe. Therefore, the H-theorem is not rigorously true for the real system, as there are fluctuations. But a spectacular violation would in practice be highly unlikely.

2.4 Maxwell-Boltzmann-Distribution

2.4.1 Equilibrium distribution function

In section 2.2.1 we have derived the equilibrium distribution by searching the maximum of the H-function under constraints such as conserved energy and number of degrees of freedom. Here, we use the concept of collision-invariants. Since the equilibrium is indicated by the time-invariance of H, we identify with a density $\rho_\varphi(t)$ in (2.86), such that $\ln f(\vec{p})$ would correspond to a collision-invariant quantity leading to $dH/dt = 0$. Thus we express $\ln f(\vec{p})$ by collision-invariants:

$$\ln f_0(\vec{p}) = A + \vec{B} \cdot \vec{p} - \frac{1}{\theta} E \ . \tag{2.97}$$

Consider the situation without potential $V(\vec{r})$, where $E = \vec{p}^2/2m$. We rewrite

$$A + \vec{B}\cdot\vec{p} - \frac{1}{\theta}\frac{\vec{p}^2}{2m} = A' - \frac{1}{\theta}\frac{(\vec{p}-\vec{p}_0)^2}{2m} = \ln a - \frac{1}{\theta}\frac{(\vec{p}-\vec{p}_0)^2}{2m} = \ln f_0(\vec{p}) \ . \tag{2.98}$$

Thus we obtain

$$f_0(\vec{p}) = a e^{-(\vec{p}-\vec{p}_0)^2/2m\theta} \ . \tag{2.99}$$

The prefactor is determined by the condition

$$\begin{aligned} n = \frac{N}{V} = \int d^3p\ f_0(\vec{p}) = a\int d^3p\ e^{-(\vec{p}-\vec{p}_0)^2/2m\theta} = a\int d^3p'\ e^{-\vec{p}'^2/2m\theta} = a\,(2\pi m\theta)^{3/2} \\ \Rightarrow \quad a = \frac{n}{(2\pi m\theta)^{3/2}} \ , \end{aligned} \tag{2.100}$$

where we assume that $C < 0$, since otherwise the integral would not converge. The mean momentum $\langle\vec{p}\rangle$ can be determined as the average

$$\langle\vec{p}\rangle = \int d^3p\ \vec{p} f_0(\vec{p}) = a\int d^3p'(\vec{p}' + \vec{p}_0)e^{-\vec{p}'^2/2m\theta} = n\vec{p}_0 \ . \tag{2.101}$$

Thus, $\vec{p}_0$ represents a net drift of the particles which we can set to zero by a proper choice of the reference frame. The energy is given by

$$\langle E\rangle = \left\langle \frac{\vec{p}^2}{2m}\right\rangle = a\int d^3p\ \frac{\vec{p}^2}{2m}e^{-\vec{p}^2/2m\theta} = \frac{3}{2}n\theta \ . \tag{2.102}$$

We will see below that $\theta = k_BT = 1/\beta$ identifying $\langle E\rangle$ as the internal energy of a dilute gas of structureless particles: $\frac{U}{V} = \langle E\rangle = \frac{3}{2}nk_BT$ and

$$f_0(\vec{p}) = n\left(\frac{\beta}{2\pi m}\right)^{3/2} e^{-\beta\vec{p}^2/2m} \ . \tag{2.103}$$

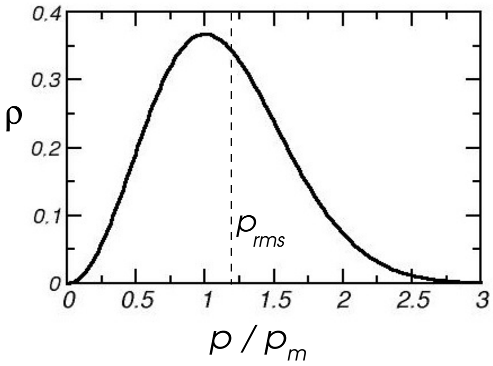


Fig. 2.5: Momentum distribution of an ideal gas in equilibrium. Maximal (p_m) and root-mean-square momentum (p_{rms}) are not identical.

Some quantities of interest are the most probable momentum $\bar{p}$ and the mean square momentum. In three-dimensional space the effective momentum distribution can be defined

$$\rho(p) = a\vec{p}^2 e^{-\beta\vec{p}^2/2m} \quad \Rightarrow \quad \left.\frac{d\rho(p)}{dp}\right|_{p=p_m} = 0 \quad \Rightarrow \quad p_m = \sqrt{\frac{2m}{\beta}} \ , \tag{2.104}$$

if we write

$$\langle g(p)\rangle = \int dp\rho(p)g(p) = \int d^3p f_0(\vec{p})g(p) \ . \tag{2.105}$$

On the other hand, the root mean square of the momentum, p_{rms} is defined by

$$p_{rms}^2 = \frac{1}{n}\int d^3p\ p^2 f_0(p) = \frac{3m}{\beta} \ . \tag{2.106}$$

Note that $p_{rms} \neq p_m$.

2.4.2 Relation to equilibrium thermodynamics - dilute gas

Let us find the connection to thermodynamics. First we define the pressure which is given by the average force per area transfered to a wall. Let the z-axis be the normal direction on a wall with specular reflection. A particle reflected at this wall reverses its momentum component p_z, loosing the momentum $2p_z$ to the wall. The density of electrons being reflected with this momentum per time unit is given by

$$\frac{p_z}{m} f_0(\vec{p}) d^3p \ , \qquad p_z > 0 \ . \tag{2.107}$$

Using $\langle \vec{p}^2 \rangle = \langle p_z^2 \rangle / 3$, we define the pressure as

$$P = \int_{p_z>0} d^3p \; 2p_z \frac{p_z}{m} f_0(\vec{p}) = \int d^3p \; p_z \frac{p_z}{m} f_0(\vec{p}) = \frac{2a}{3} \int d^3p \frac{p^2}{2m} e^{-\frac{p^2}{2m\theta}} = \frac{2}{3} \langle E \rangle = n\theta \ , \tag{2.108}$$

which is consistent with the dilute (ideal) gas, $P = nk_BT$, if $\theta = k_BT = 1/\beta$. We now consider the functional H at equilibrium,

$$H[f_0] = -\int d^3p \; f_0(\vec{p}) \left[\ln a - \frac{p^2}{2m}\beta \right] = \beta \langle E \rangle - n \ln \left[n \left(\frac{\beta}{2\pi m} \right)^{3/2} \right] \ . \tag{2.109}$$

A further step towards thermodynamics is now to identify H with the entropy, which is also suggested by the H-theorem, reminding to Clausius' inequality. The connection is given by

$$S = k_B V H \ . \tag{2.110}$$

We would like first to look at the variation of S for fixed N and V. Thus, we use the relation

$$\beta \langle E \rangle = \beta \frac{U}{V} = \frac{3}{2} \frac{N}{V} \qquad \text{and} \qquad \beta = \frac{3N}{2U} \ . \tag{2.111}$$

From this we obtain

$$S = \frac{3}{2} k_B N \ln U - k_B N \left[\ln \frac{N}{V} + \frac{3}{2} \ln N - \frac{3}{2} \right] + const. \ . \tag{2.112}$$

This expression is used to obtain,

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V,N} = \frac{3}{2} \frac{k_B N}{U} \qquad \Rightarrow \qquad \begin{cases} \beta = \dfrac{1}{k_B T} \\[2ex] \langle E \rangle = \dfrac{U}{V} = \dfrac{3}{2} n k_B T \\[2ex] P = \dfrac{N}{V} k_B T \ . \end{cases} \tag{2.113}$$

Another relation is obtained by fixing N and U:

$$\frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{U,V} = k_B \frac{N}{V} \qquad \Rightarrow \qquad PV = N k_B T \ , \tag{2.114}$$

as already given above through the relation (2.108). This corresponds to the thermodynamic equation of state of an ideal gas.

As mentioned the second law formulated as $dS \geq 0$ is here equivalent to Boltzmann's H-theorem: $dH/dt \geq 0$.

2.4.3 Local equilibrium state

So far we have concentrated on the homogeneous situation. We extend our discussion to the case where a conservative force $\vec{F}(\vec{r}) = -\vec{\nabla}V(\vec{r})$ is present, originating from the potential $V(\vec{r})$. In equilibrium the distribution function is not explicitly time-dependent and the collision integral vanishes, leaving a reduced Boltzmann equation:

$$\frac{\partial f}{\partial t} = 0 = \left(\frac{\partial f}{\partial t}\right)_{coll} \qquad \Rightarrow \qquad \left(\frac{\vec{p}}{m}\cdot\vec{\nabla}_r + \vec{F}\cdot\vec{\nabla}_p\right) f(\vec{r},\vec{p}) = 0 \; . \tag{2.115}$$

This equation is easy to solve with a separation ansatz restricting all space dependence to the particle density $n(\vec{r})$,

$$f_0(\vec{r},\vec{p}) = \frac{n(\vec{r})}{(2\pi m k_B T)^{3/2}} e^{-\frac{p^2}{2mk_BT}} \quad \text{with} \quad n(\vec{r}) = n_0 e^{-\frac{V(\vec{r})}{k_BT}} \; . \tag{2.116}$$

Gas trapped in a harmonic potential: We assume that the potential has the form $V(\vec{r}) = a|\vec{r}|^2$. Hence the distribution function for the gas at given temperature T is given by

$$f_0(\vec{r},\vec{p}) = \frac{n_0}{(2\pi m k_B T)^{3/2}} e^{-\frac{p^2}{2mk_BT}} e^{-\frac{ar^2}{k_BT}} \; , \tag{2.117}$$

with n_0 the gas density at $r = 0$. For the density and the pressure we find,

$$n(r) = n_0 e^{-\frac{ar^2}{k_BT}} \qquad \text{and} \qquad P(r) = n(r)k_BT = k_BTn_0e^{-\frac{ar^2}{k_BT}} \; . \tag{2.118}$$

The total number of particles is given by

$$N = \int d^3r \; n(r) = n_0\left(\frac{\pi k_BT}{a}\right)^{3/2} \qquad \Rightarrow \qquad n_0(T) = N\left(\frac{a}{\pi k_BT}\right)^{3/2} \tag{2.119}$$

Interesting is also the internal energy

$$U = \langle E\rangle = \langle\frac{p^2}{2m} + ar^2\rangle = \int d^3r\left\{\frac{3}{2}n(r)k_BT + n(r)ar^2\right\} = \frac{3}{2}Nk_BT + \frac{3}{2}Nk_BT = 3Nk_BT \tag{2.120}$$

and the specific heat is given by $C = dU/dT = 3Nk_B$. The volume may be defined as

$$V = \left\langle\frac{4\pi}{3}r^3\right\rangle = \frac{4N}{3\sqrt{\pi}}\left(\frac{k_BT}{a}\right)^{3/2} \; , \tag{2.121}$$

which yields a thermal expansion coefficient

$$\alpha = \frac{1}{V}\frac{\partial V}{\partial T} = \frac{3}{2T} \; . \tag{2.122}$$

Obviously, $C(T)$ and $\alpha(T)$ do not follow the third law of thermodynamics. The reason is that we consider the gas within the framework of classical physics. To account for the third law we need quantum mechanics.

2.5 Fermions and Bosons

So far we have restricted our view to classical particles following classical mechanics. If we extend our discussion to quantum systems, gases of identical quantum particles, we have to deal with Fermions and Bosons. Also in this case it is possible to formulate the Boltzmann equation, where the collision integral has to be modified in the following way:

$$\begin{aligned}\left(\frac{\partial f}{\partial t}\right)_{coll} = \int d^3p_1\, d^3p_1' d^3p_2'\, W \,\{&f(\vec{r},\vec{p}_1',t)f(\vec{r},\vec{p}_2',t)[1-f(\vec{r},\vec{p},t)][1-f(\vec{r},\vec{p}_1,t)] \\ &- f(\vec{r},\vec{p},t)f(\vec{r},\vec{p}_1,t)[1-f(\vec{r},\vec{p}_1',t)][1-f(\vec{r},\vec{p}_2',t)]\}\end{aligned} \tag{2.123}$$

for Fermions where we have to include the probability that the target states after the collisions are not occupied. For Bosons, in contrast ,we have the collision integral,

$$\begin{aligned}\left(\frac{\partial f}{\partial t}\right)_{coll} = \int d^3p_1\, d^3p_1' d^3p_2'\, W \,\{&f(\vec{r},\vec{p}_1',t)f(\vec{r},\vec{p}_2',t)[1+f(\vec{r},\vec{p},t)][1+f(\vec{r},\vec{p}_1,t)] \\ &- f(\vec{r},\vec{p},t)f(\vec{r},\vec{p}_1,t)[1+f(\vec{r},\vec{p}_1',t)][1+f(\vec{r},\vec{p}_2',t)]\}\ .\end{aligned} \tag{2.124}$$

This form is often interpreted in the way that Bosons like to go where other Bosons are. The H-function is also modified

$$\begin{aligned}H &= -\int d^3p\ \{f(\vec{r},\vec{p},t)\ln f(\vec{r},\vec{p},t) + (1-f(\vec{r},\vec{p},t))\ln(1-f(\vec{r},\vec{p},t))\} \qquad \text{Fermions} \\ H &= -\int d^3p\ \{f(\vec{r},\vec{p},t)\ln f(\vec{r},\vec{p},t) - (1+f(\vec{r},\vec{p},t))\ln(1+f(\vec{r},\vec{p},t))\} \qquad \text{Bosons}\end{aligned} \tag{2.125}$$

Analogous to the classical case the equilibrium condition corresponds to $dH/dt = 0$. The H-function takes again a maximum for the equilibrium distribution. We tackle this problem as we did previously taking the conservation laws a conditions into account,

$$H[f] \quad \rightarrow \quad H[f] + \lambda \left\{ N - \int d^3r d^3q f(\vec{r},\vec{q}) \right\} + \frac{1}{\theta}\left\{ E - \int d^3r d^3p \epsilon_{\vec{p}} f(\vec{r},\vec{p}) \right\} \tag{2.126}$$

We now maximize this functional with respect to f which we assume to be independent of $\vec{r}$.

$$\frac{\delta H}{\delta f(\vec{p})} = \ln \frac{f(\vec{p})}{1 \mp f(\vec{p})} - \lambda - \frac{\epsilon_{\vec{p}}}{\theta} = 0 \tag{2.127}$$

This equation is quickly solved and leads for both the Fermions (upper sign) and the Bosons (lower sign) to the well known distribution functions, if we identify $\theta = k_BT$ and $\lambda = -\mu/k_BT$:

$$\begin{aligned}f(\vec{p}) &= \frac{1}{e^{(\epsilon_{\vec{p}}-\mu)/k_BT}+1} \quad \text{Fermi-Dirac distribution} \\ f(\vec{p}) &= \frac{1}{e^{(\epsilon_{\vec{p}}-\mu)/k_BT}-1} \quad \text{Bose-Einstein distribution}\ .\end{aligned} \tag{2.128}$$

Note that for large energy $E \gg \mu$ both distributions approach the Maxwell-Boltzmann distribution.

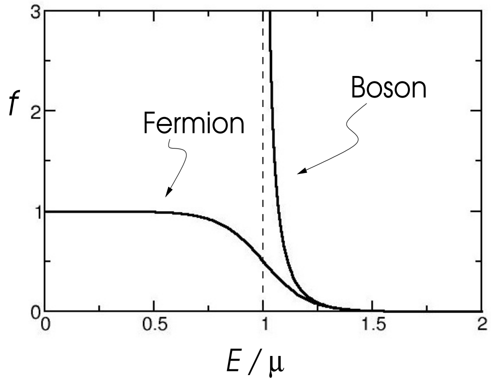


Fig. 2.6: *Bose-Einstein- and Fermi-Dirac-distribution at finite temperature.*

2.6 Transport properties

Although we have so far mainly concentrated on the equilibrium situation, the Boltzmann equations have actually been formulated to discuss systems away from equilibrium. Such non-equilibrium physics is particularly important in problems considering transport of matter, energy or other quantities, where irreversible behavior is involved.

2.6.1 Relaxation time approximation

Boltzmann's equation is complicated to solve in general as it represents a complex integro-differential equation. Here we will consider one of the most simple and useful approximations for the collision integral which leads to an easily tractable Boltzmann equation - the relaxation time approximation.
First we analyze the average time between collisions. For this purpose we consider the rate for a particle to suffer a collision,

$$\begin{aligned} Z &= \frac{2}{n_0}\int d^3p_1\, d^3p_2\, d^3p_1{}'\, d^3p_2{}'\, W(\vec{p}_1,\vec{p}_2;\vec{p}_1',\vec{p}_2')f(\vec{r},\vec{p}_1,t)f(\vec{r},\vec{p}_2,t) \\ &= \frac{2}{n_0}\int d^3p_1\, d^3p_2\, \sigma\,|\vec{v}_1-\vec{v}_2|\, f(\vec{r},\vec{p}_1,t)f(\vec{r},\vec{p}_2,t) \end{aligned} \tag{2.129}$$

where σ is total cross section. To estimate Z we use the equilibrium distribution function,

$$f(\vec{r},\vec{p}_1,t)f(\vec{r},\vec{p}_2,t) = n_0^2\left(\frac{\beta}{2\pi m}\right)^3 e^{-\beta m(\vec{V}^2+\vec{u}^2/4)} \tag{2.130}$$

with $\vec{V}=\frac{1}{2}(\vec{v}_1+\vec{v}_2)$ as center of mass velocity and $\vec{u}=\vec{v}_1-\vec{v}_2$ as relative velocity ($d^3v_1d^3v_2 = d^3Vd^3u$). In this way we obtain

$$Z = 4n_0\sigma\sqrt{\frac{k_BT}{\pi m}} = \frac{2\sqrt{2}}{\sqrt{\pi}}n_0\sigma\bar{v} = \frac{1}{\tau} \tag{2.131}$$

where we assumed that the cross section σ does not depend on the momenta and $m\bar{v}=\bar{p}=\sqrt{2mk_BT}$ (2.104), corresponds to the most probable momentum. We may then define the mean free path, the average distance for a particle between two collisions:

$$\ell = \bar{v}\tau = \sqrt{\frac{\pi}{8}}\frac{1}{n_0\sigma} \tag{2.132}$$

This discussion is valid if the system is not too far from equilibrium. Otherwise it is not acceptable to use the Maxwell-Boltzmann distribution to estimate τ.

We now look at the collision integral assuming that the system deviates only slightly from equilibrium. Thus the distribution function can be written as

$$f = f_0 + g = f_0(1+\phi) \qquad \text{with} \qquad \phi \ll 1\,. \tag{2.133}$$

Inserting this into the Boltzmann equation leads to

$$\begin{aligned} \frac{D}{Dt}f(\vec{r},\vec{p},t) &= f_0\frac{D\phi}{Dt} \\ &= -f_0(\vec{r},\vec{p})\int d^3p_1\, d^3p_1'\, d^3p_2' W(\vec{p},\vec{p}_1;\vec{p}_1',\vec{p}_2')f_0(\vec{r},\vec{p}_1) \\ &\qquad \times \underbrace{\left\{\phi(\vec{r},\vec{p},t)+\phi(\vec{r},\vec{p}_1,t)-\phi(\vec{r},\vec{p}_1',t)-\phi(\vec{r},\vec{p}_2',t)\right\}}_{=\Delta\phi} \end{aligned} \tag{2.134}$$

where we used the equilibrium condition $f_0(\vec{r},\vec{p},t)f_0(\vec{r},\vec{p}_1,t) = f_0(\vec{r},\vec{p}_1',t)f_0(\vec{r},\vec{p}_2',t)$. This defines a linear problem of the general form:

$$\frac{\partial \phi}{\partial t} = \mathcal{L}\phi \tag{2.135}$$

with the operator $\mathcal{L}$ defined through (2.134), where we assume moreover homogeneity and the absence of external forces for simplicity. Taking now a separation ansatz: $\phi(\vec{p},t) = a(t)b(\vec{p})$ we find the equations:

$$\frac{da}{dt} = \lambda a \qquad \text{and} \qquad \mathcal{L}b = \lambda b \qquad \Rightarrow \qquad \phi(\vec{p},t) = e^{\lambda t} g_\lambda(\vec{p}) \ . \tag{2.136}$$

We introduce a scalar product,

$$(b_1, b_2) = \int d^3p f_0(\vec{p}) b_1(\vec{p}) b_2(\vec{p}) = (b_2, b_1) \quad \text{with} \quad (b, b) \geq 0 \ . \tag{2.137}$$

Using like earlier the symmetries of $W(\vec{p},\vec{p}_1;\vec{p}_1',\vec{p}_2')$ we find that

$$(b_1, \mathcal{L}b_2) = (\mathcal{L}b_1, b_2) \qquad \Rightarrow \qquad \mathcal{L} : \text{"Hermitian"} \tag{2.138}$$

such that λ is real. In the same way we can show that for arbitrary $b(\vec{p})$,

$$(b, \mathcal{L}b) \leq 0 \ , \tag{2.139}$$

i.e. the eigenvalues λ are non-positive, such that the deviations from equilibrium generally decay in time.
Still the linearized problem is a complicated problem. A dramatic simplification occurs with the so-called relaxation time approximation:

$$\left(\frac{\partial f}{\partial t}\right)_{coll} = -\frac{f(\vec{r},\vec{p},t) - f_0(\vec{r},\vec{p})}{\tau} = -\frac{g(\vec{r},\vec{p},t)}{\tau} \ . \tag{2.140}$$

In order to be appropriate for a collision integral we require

$$\int d^3p \varphi(\vec{p}) g(\vec{r},\vec{p},t) = 0 \tag{2.141}$$

for any collision invariant $\varphi(\vec{p})$. The resulting Boltzmann equation has then the form

$$\frac{Df_0}{Dt} \approx -\frac{g}{\tau} \tag{2.142}$$

to linear order in τ which we consider as a small parameter to expand in,

$$g = \tau g_1 + \tau^2 g_2 + \cdots \tag{2.143}$$

where we only will keep the lowest order.

2.7 Electrical conductivity of an electron gas

We address now a topic of solid state, the electrical conductivity of electrons in the presence of a voltage, i.e. an electrical field $\vec{E}$. For this purpose we consider the steady state of electron motion driven by a uniform electrical field, using the relaxation time approximation. The corresponding Boltzmann equation reads,

$$-e\vec{E} \cdot \vec{\nabla}_p f(\vec{p}) = -\frac{f(\vec{p}) - f_0(\vec{p})}{\tau} = -\frac{g(\vec{p})}{\tau} \ . \tag{2.144}$$

This equation is easily solved to lowest order in τ:

$$g(\vec{p}) = \tau e\vec{E} \cdot \vec{\nabla}_p f_0(\vec{p}) \ , \tag{2.145}$$

and allows us to calculate the (uniform) electrical current

$$\vec{j} = -e\int d^3p\ \vec{v}f(\vec{p}) = -e\int d^3p\ \vec{v}g(\vec{p}) = -e^2\int d^3p\ \tau\vec{v}\left(\vec{E}\cdot\vec{\nabla}_p f_0(\vec{p})\right) \ . \tag{2.146}$$

Note that we use $\int d^3p\ \vec{p}\ f_0 = 0$. Now we use the relation

$$\vec{\nabla}_p f_0 = \vec{\nabla}_p \epsilon(\vec{p})\frac{\partial f_0}{\partial \epsilon} = \vec{v}\frac{\partial f_0}{\partial \epsilon} \ . \tag{2.147}$$

where we use for the clarity of the notation for the energy $\epsilon(\vec{p}) = \vec{p}^2/2m$. If the system is isotropic, $\vec{j} \parallel \vec{E}$ such that

$$j = -e^2\tau E\int d^3p\frac{v^2}{3}\frac{\partial f_0(\epsilon)}{\partial \epsilon} = \sigma E \ . \tag{2.148}$$

<u>*Electrons as classical particles:*</u> We take for f_0 the Maxwell-Boltzmann-distribution

$$f_0(\epsilon) = n\left(\frac{\beta}{2\pi m}\right)^{3/2} e^{-\beta\epsilon} \qquad \Rightarrow \qquad \frac{\partial f_0}{\partial \epsilon} = -\beta f_0(\epsilon) \ . \tag{2.149}$$

Thus we find

$$\sigma = \frac{2e^2\tau}{3mk_BT}\underbrace{\int d^3p\frac{\vec{p}^2}{2m}f_0(\epsilon)}_{=\langle E\rangle = \frac{3}{2}nk_BT} = \frac{ne^2\tau}{m} \ . \tag{2.150}$$

This corresponds to the famous Drude conductivity. We used here the approximation that τ is not momentum or energy dependent. This assumption does not generally hold.

<u>*Electrons as Fermions:*</u> In reality electrons are Fermion and we should use the Fermi-Dirac distribution for the equilibrium:

$$f_0(\epsilon) = \frac{2}{h^3}\frac{1}{e^{\beta(\epsilon-\mu)}+1} \qquad \Rightarrow \qquad \frac{\partial f_0(\epsilon)}{\partial \epsilon} = \frac{1}{2h^3k_BT\cosh^2[\beta(\epsilon-\mu)/2]} \tag{2.151}$$

where h is the Planck constant.[2] Note that this function is concentrated around the Fermi

---

[2] Integrals for the Fermi-Dirac distribution: Free electrons are described by plane wave functions

$$\psi_{\vec{k}}(\vec{r}) = \frac{1}{\sqrt{V}}e^{i\vec{k}\cdot\vec{r}} \quad \text{with} \quad \vec{p} = \hbar\vec{k} \quad \text{and} \quad \epsilon = \frac{\vec{p}^2}{2m} = \frac{\hbar^2\vec{k}^2}{2m} \tag{2.152}$$

Via periodic boundary conditions in a cube of edge length L we quantize the wavevector $\vec{k}$:

$$\psi_{\vec{k}}(x,y,z) = \psi_{\vec{k}}(x+L,y,z) = \psi_{\vec{k}}(x,y+L,z) = \psi_{\vec{k}}(x,y,z+L) \qquad \Rightarrow \quad \vec{k} = \frac{2\pi}{L}(n_x,n_y,n_z) = \frac{2\pi}{L}\vec{n} \tag{2.153}$$

with $n_{x,y,z} = 0, \pm1, \pm2, \ldots$. Summing up the states to the Fermi energy means then summing over all occupied states with quantum numbers $\vec{n}$. The total number of particles is given by

$$N = \sum_{\vec{n}}^{occ.} 2 = 2\frac{L^3}{(2\pi)^3}\int_{|\vec{k}|\le k_F} d^3k = \frac{2V}{(2\pi)^3}\frac{4\pi}{3}k_F^3 \qquad \Rightarrow \quad k_F = \{3\pi^2 n\}^{1/3} \tag{2.154}$$

with $n = N/V = N/L^3$ and the factor 2 is due to the two spin states of the electrons. From this point of view the normalization of the Fermi-Dirac distribution is given by the condition

$$f_0(\epsilon) = 2\frac{1}{h^3}\frac{1}{e^{\beta(\epsilon-\mu)}+1} \qquad \Rightarrow \qquad n = \int d^3p f_0(\epsilon(\vec{p})) = 4\pi\int dp\ p^2 f_0(\epsilon) = 4\pi\sqrt{2m^3}\int d\epsilon\epsilon^{1/2}\ f_0(\epsilon) \tag{2.155}$$

with $h = 2\pi\hbar$. It is usual to introduce here the density of states $g(\epsilon)$,

$$n = \int d\epsilon\ g(\epsilon)f_0(\epsilon) \qquad \text{with} \qquad g(\epsilon) = \frac{1}{4\pi^2}\left(\frac{2m}{\hbar^2}\right)^{3/2}\epsilon^{1/2} \ . \tag{2.156}$$

energy (chemical potential) at low temperatures ($k_B T \ll \mu$). Then we calculate the current

$$j = -e^2 \tau E \int d^3p \, \frac{2}{3m} \epsilon \frac{\partial f_0}{\partial \epsilon} = -e^2 \tau E \frac{2}{3m} 4\pi \sqrt{2m^3} \int d\epsilon \, \epsilon^{3/2} \frac{\partial f_0}{\partial \epsilon} = \frac{ne^2\tau}{m} E \, . \tag{2.157}$$

This yields the same expression for the conductivity as in the classical case: the Drude conductivity. The condition on τ is less stringent. τ essentially needs to be constant around the Fermi energy, since the first integrand in (2.157) is concentrated on the Fermi surface (Fig. 2.7).

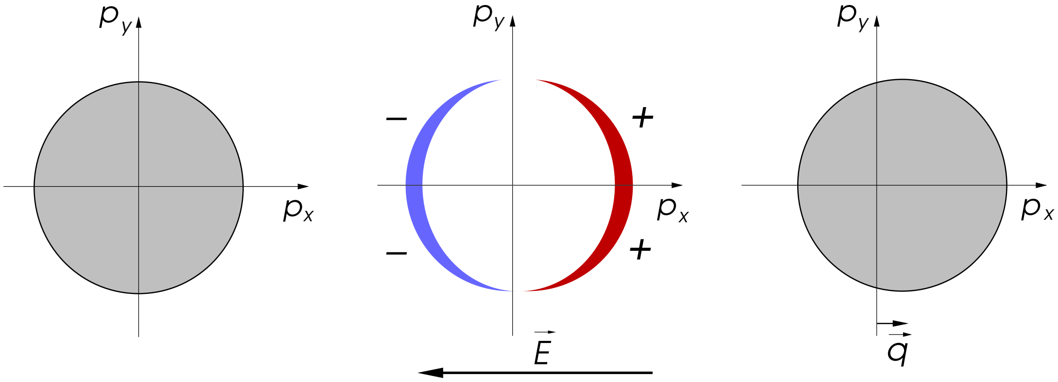


Fig. 2.7: *Schematic picture of the changed momentum distribution functions in the steady state of electrons in an electric field. Left panel: equilibrium Fermi sea* $f_0(\vec{p})$*; middle panel: change of the distribution* $g(\vec{p})$*; right panel: sum* $f = f_0 + g$ *leads to a Fermi distribution shifted by a momentum* $\vec{q} \propto \vec{E}$ *.*

Chapter 3

Classical statistical mechanics

Statistical physics deals with the equilibrium properties of matter and provides the microscopic understanding and basis for thermodynamics. Unlike the kinetic Boltzmann theory of the previous chapter the present chapter does not consider non-equilibrium situations and develops a new practical approach to equilibrium state of macroscopic systems. Time is not a variable anymore and measurements which may be considered as time averages of certain quantities, are translated as averages over a large ensemble of identical systems in different possible microscopic states under the same external parameters.

3.1 Gibbsian concept of ensembles

We consider the state of a gas of N classical particles, given by $3N$ canonical coordinates $q_1, \ldots, q_{3N}$ and by the corresponding $3N$ conjugate momenta $p_1, \ldots, p_{3N}$. These define a $6N$-dimensional space Γ, where each point in Γ represents a state of the microscopic system. Considering the whole system of N particles under certain macroscopic conditions, given external parameter such as temperature, pressure, volume, internal energy, ... , we find that an infinite number of states in Γ are compatible with the same external condition, and would not be distinguishable by macroscopic measurements.

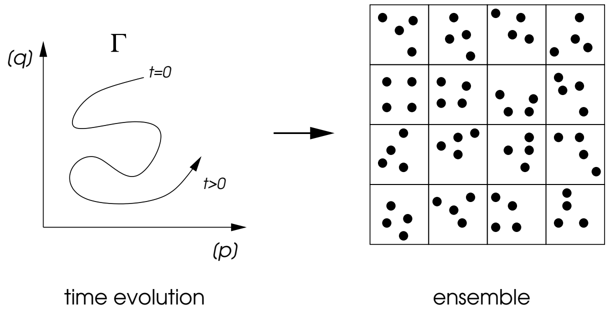


Fig. 3.1: *Time averages are replaced by averages over an ensemble of systems in different microscopic states, but with the same macroscopic conditions.*

If we would like to calculate a certain macroscopic quantity we could perform temporal average of the microscopic variables over a very long evolution time. Based on the states in Γ this is, however, not a practical method as it involves the discussion of the temporal evolution. Gibbs introduced the concept of ensembles to circumvent this problem. Taking a large (infinite) number of systems under identical macroscopic conditions, we can find macroscopic quantities by averaging over an ensemble of states. That this scheme is equivalent to a temporal average is a hypothesis, as it is assumed that the time evolution of the system would lead to all possible states also represented in the ensemble. This is the so-called *ergodicity hypothesis*.[1] The set of

---

[1] Ergodicity hypothesis: We consider the supspace defined by $\mathcal{H}(p,q) = E$ in Γ. Due to energy conservation a point in this subspace, evolving in time, does not leave this subspace. The ergodicity hypothesis states, that a

states for given macroscopic parameters is then represented by a distribution of points in the space Γ. This distribution is generally continuous for a gas of particles.

For the calculation of averages we introduce the density function $\rho(p,q)$ providing the measure of the density of points in Γ-space ((p,q) stands for the whole state $(p_1,\ldots,p_{3N};q_1,\ldots q_{3N})$. Then

$$\rho(p,q)d^{3N}pd^{3N}q \tag{3.1}$$

gives the number of representative points contained in the small volume $d^{3N}pd^{3N}q$ in Γ, very analogous to the distribution function in Boltzmann theory. This can now be used to calculate averages of any desired quantity which can be expressed in the variable (p,q), $A(p,q)$:

$$\langle A\rangle = \frac{\int dpdqA(p,q)\rho(p.q)}{\int dpdq\rho(p.q)} \; . \tag{3.2}$$

We will use from now on the short notation $dpdq$ for $d^{3N}pd^{3N}q$ where it is not misleading.

3.1.1 The Liouville Theorem

The dynamics of the system of N particles shall be described by a Hamiltonian $\mathcal{H}(p,q)$ which yields the equation of motion in the Hamiltonian formulation of classical mechanics.

$$\dot{p}_i = -\frac{\partial\mathcal{H}}{\partial q_i} \quad \text{and} \quad \dot{q}_i = \frac{\partial\mathcal{H}}{\partial p_i} \qquad (i=1,\ldots,3N)\; . \tag{3.3}$$

This equation describes the motion of the points in Γ. If $\mathcal{H}$ does not depend on time derivatives of p_i and/or q_i, then the equations of motion are time reversal invariant. They determine the evolution of any point in Γ uniquely.

Now let us consider the points in the Γ space and note that their number does not change in time, as if they form a fluid. Thus, they satisfy the continuity equation

$$\frac{\partial\rho}{\partial t} + \vec{\nabla}\cdot(\rho\vec{v}) = 0\; . \tag{3.4}$$

where $\vec{v} = (\dot{p}_1,\ldots,\dot{p}_{3N};\dot{q}_1,\ldots,\dot{q}_{3N})$ and $\vec{\nabla} = (\partial/\partial p_1,\ldots,\partial/\partial p_{3N};\partial/\partial q_1,\ldots,\partial/\partial q_{3N})$. Introducing the generalized substantial derivative we can rewrite this equation as

$$\frac{D\rho}{Dt} + \rho\vec{\nabla}\cdot\vec{v} = 0\; . \tag{3.5}$$

The divergence of the "velocity" $\vec{v}$ is

$$\vec{\nabla}\cdot\vec{v} = \sum_{i=1}^{3N}\left\{\frac{\partial\dot{q}_i}{\partial q_i} + \frac{\partial\dot{p}_i}{\partial p_i}\right\} = \sum_{i=1}^{3N}\underbrace{\left\{\frac{\partial}{\partial q_i}\frac{\partial\mathcal{H}}{\partial p_i} - \frac{\partial}{\partial p_i}\frac{\partial\mathcal{H}}{\partial q_i}\right\}}_{=0} = 0 \tag{3.6}$$

This means that the points in Γ space evolve like an incompressible fluid. This property is known as *Liouville's theorem.* We may write this also in the form

$$0 = \frac{D\rho}{Dt} = \frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left\{\dot{q}_i\frac{\partial\rho}{\partial q_i} + \dot{p}_i\frac{\partial\rho}{\partial p_i}\right\} = \frac{\partial\rho}{\partial t} + \sum_{i=1}^{3N}\left\{\frac{\partial\mathcal{H}}{\partial p_i}\frac{\partial\rho}{\partial q_i} - \frac{\partial\mathcal{H}}{\partial q_i}\frac{\partial\rho}{\partial p_i}\right\}\; . \tag{3.7}$$

---

point on $\mathcal{H}=E$ can reach in its temporal evolution every point of this subspace (Boltzmann, 1887). This is not rigorously valid, however. Ehrenfest's version of this hypothesis (1911) states: The trajectory of a point comes any other point in the same subspace arbitrarily close in the course of time. Of course, it is possible to find (singular) counter examples, for which only a tiny fraction of the subspace is accessed.

Using Poisson brackets[2] this equation reads,

$$\frac{\partial \rho}{\partial t} = \{\mathcal{H}, \rho\} \ . \tag{3.9}$$

3.1.2 Equilibrium system

A satisfactory representation of a system in equilibrium requires for an ensemble that the density function does not depend on time t, i.e. $\partial\rho/\partial t = 0$. From Liouville's theorem we get the condition that

$$0 = \sum_{i=1}^{3N} \left\{ \dot{q}_i \frac{\partial \rho}{\partial q_i} + \dot{p}_i \frac{\partial \rho}{\partial p_i} \right\} = \vec{v} \cdot \vec{\nabla} \rho = \{\mathcal{H}, \rho\} \ . \tag{3.10}$$

A general way to satisfy this is to take a density function which depends only on quantities conserved during the motion, such as energy or particle number. Then the system would evolve within a subspace where ρ is constant.

We may use this feature of ρ now to consider averages of certain quantities, based on the above mentioned equivalence between the temporal and ensemble averages. Defining the temporal average of $A(p, q)$ as

$$\langle A \rangle = \lim_{T\to\infty} \frac{1}{T} \int_0^T A(p(t), q(t)) dt \tag{3.11}$$

for any starting point $((p(t = 0), q(t = 0))$ in the space Γ and $(p(t), q(t))$ obeying the equation of motion (3.3). The hypothesis of ergodicity, even in its restricted sense, implies that this average can be taken as an average of an ensemble of an infinite number of different microscopic states (points in Γ-space). As the evolution of (p, q) conserves the energy, this leads naturally to consider an ensemble of states of a fixed energy (internal energy). We call such ensembles *microcanonical*. Although the microcanonical ensemble is not so frequently used in practice than other ensembles which we will consider later, it is very useful for illustrative purposes. It describes an isolated closed system with no energy exchange with the environment.

We postulate that in equilibrium any state of a macroscopic system satisfying the the external conditions appears with equal probability. This is in line with our experience of detailed balance in the previous chapter. In our microcanonical description with fixed energy, number of particles N and volume V, we postulate

$$\rho(p, q) = \begin{cases} const. & E \leq \mathcal{H}(p, q) \leq E + \delta E \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.12}$$

where δE is small. The average value of A is then given by (3.2). The validity of this approach is based on the assumption of small mean square fluctuations (standard deviations)

$$\frac{\langle \{A - \langle A \rangle\}^2 \rangle}{\langle A \rangle^2} \ll 1 \ . \tag{3.13}$$

Such fluctuations should be suppressed by the order N^{-1} as we will see below.

---

[2] The Poisson bracket is defined as

$$\{u, v\} = \sum_i \left\{ \frac{\partial u}{\partial q_i} \frac{\partial v}{\partial p_i} - \frac{\partial u}{\partial p_i} \frac{\partial v}{\partial q_i} \right\} = -\{v, u\} \ . \tag{3.8}$$

3.2 Microcanonical ensemble

We consider a macroscopic system of N particles in a volume V which is isolated and closed. The microcanonical ensemble for a given energy E consists of all systems of this kind, whose energy lies in the range $[E, E + \delta E]$. First we define the phase volume

$$\Phi(E) = \int_{\mathcal{H}(p,q) \leq E} dpdq \;, \tag{3.14}$$

which contains all point in Γ space with energy lower than or equal to E. From this we obtain the volume of the microcanonical ensemble as

$$\omega(E) = \Phi(E + \delta E) - \Phi(E) = \frac{d\Phi(E)}{dE} \delta E \;. \tag{3.15}$$

Thus we now can renormalize $\rho(p, q)$ with the condition,

$$1 = \int dp\, dq\; \rho(p,q) = \frac{1}{\omega(E)} \int_{E \leq \mathcal{H}(p,q) \leq E + \delta E} dp\, dq \tag{3.16}$$

such that

$$\rho(p,q) = \begin{cases} \dfrac{1}{\omega(E)} & E \leq \mathcal{H}(p,q) \leq E + \delta E \\ & \\ 0 & \text{otherwise} \end{cases} \;. \tag{3.17}$$

As postulated $\rho(p, q)$ is constant in the energy range $[E, E + \delta E]$.

3.2.1 Entropy

We use $\omega(E)$ to define the entropy

$$S(E, V, N) = k_B \ln \omega(E) \tag{3.18}$$

This agrees very well with our earlier definition based on the H-function in Eq.(2.14), $S = k_B \ln W$ where W is the number of configurations for given macroscopic distributions of microstates. Analogously this definition satisfies the general properties of the entropy as we introduced them earlier. As earlier we can consider $\omega(E)$ or S as a measure of the imprecision of our knowledge of the state of the system. The more states are available in microcanonical ensemble, the less we know in which state the system is at a given time and the larger is the volume ω and the entropy.

We consider a composite system consisting of two subsystems

$$\mathcal{H}(p,q) = \mathcal{H}_1(p_1, q_1) + \mathcal{H}_2(p_2, q_2) \;, \tag{3.19}$$

with (N_1, V_1) and (N_2, V_2), resp., for the corresponding particle number and volume. Each of the two systems is characterized by $\omega_1(E_1)$ and $\omega_2(E_2)$, respectively. The volume $\omega(E)$ of the microcanonical ensemble is the product of the subsystems under the conditions

$$E = E_1 + E_2 \;, \quad N = N_1 + N_2 \quad \text{and} \quad V = V_1 + V_2; . \tag{3.20}$$

As the individual values (E_1, N_1, V_1), (E_2, N_2, V_2) can fluctuate, we find

$$\omega(E) = \sum_{0 \leq E' \leq E} \omega_1(E')\, \omega_2(E - E') \tag{3.21}$$

where we assume a "discrete" mesh of equally spaced E'-values of mesh spacing δE ($\ll E$). We claim that this sum is approximated well by a single value E'_0 giving rise to a sharp maximum among the summands; ($E_1 = E'_0$ and $E_2 = E - E'_0$).[3] The largest term is obtained by

$$\left.\frac{\partial \omega_1(E')\,\omega_2(E-E')}{\partial E'}\right|_{E'=E'_0} = 0 \tag{3.24}$$

such that

$$\begin{aligned} 0 &= \left\{ \frac{1}{\omega_1(E')}\frac{\partial \omega_1(E')}{\partial E'} + \frac{1}{\omega_2(E-E')}\frac{\partial \omega_2(E-E')}{\partial E'} \right\}_{E'=E'_0} \\ &= \left.\frac{\partial \ln \omega_1(E_1)}{\partial E_1}\right|_{E_1=E'_0} - \left.\frac{\partial \ln \omega_2(E_2)}{\partial E_2}\right|_{E_2=E-E'_0} . \end{aligned} \tag{3.25}$$

From this we obtain with $\bar{E}_1 = E'_0$ and $\bar{E}_2 = E - E'_0$

$$\left.\frac{\partial S_1(E_1)}{\partial E_1}\right|_{E_1=\bar{E}_1} = \left.\frac{\partial S_2(E_2)}{\partial E_1}\right|_{E_2=\bar{E}_2} \tag{3.26}$$

which can be considered as the equilibrium condition. Note that this is equivalent to the statement that the equilibrium state corresponds to the macrostate with the largest number of microscopic realizations.

Identifying E as the internal energy U we define the temperature

$$\frac{\partial S}{\partial U} = \frac{1}{T} \qquad \Rightarrow \qquad \frac{1}{T_1} = \frac{1}{T_2} , \tag{3.27}$$

leading to T as an equilibrium state variable. We have ignored the other variables V, N which we will consider later.

Let us assume that there is some impediment which forces the two subsystems to specific variables such that each subsystem independently is at equilibrium, but not the total system. For example, a big vessel of gas may be separated by a wall into two subvessels. Then we find for the corresponding ω of the combined system,

$$\tilde{\omega}(E,V,N) = \omega_1(E_1,V_1,N_1)\omega_2(E_2,V_2,N_2) \qquad \text{with} \quad \left\{ \begin{array}{l} E = E_1 + E_2 \\ V = V_1 + V_2 \\ N = N_1 + N_2 \end{array} \right. \tag{3.28}$$

such that the entropy

$$\tilde{S}(E,V,N) = S_1(E_1,V_1,N_1) + S_2(E_2,V_2,N_2) \leq S(E,V,N) , \tag{3.29}$$

i.e. the entropy is concave. This means also the equilibrium is obtained by the maximal entropy, which is a consequence of the second law of thermodynamics. In terms of the volume of the microcanonical volume ω the equilibrium state assumes among all volumes the maximal one, the most likely one in terms of probability.

---

[3] Note that $\ln \omega_i \propto N_i$ and $E_i \propto N_i$ ($i = 1, 2$) as both quantities are extensive. We then find quickly the following bounds:

$$\omega_1(E'_0)\omega_2(E-E'_0) \leq \omega(E) \leq \frac{E}{\delta E}\omega_1(E'_0)\omega_2(E-E'_0) . \tag{3.22}$$

Note that $E/\delta E$ is the number of summands. Thus, we obtain for the entropy

$$k_B \ln \omega_1(E'_0)\omega_2(E-E'_0) \leq S(E) \leq k_B \ln \omega_1(E'_0)\omega_2(E-E'_0) + k_B \ln \frac{E}{\delta E} . \tag{3.23}$$

It is important to see that the last term only scales with system size as $\ln N$ (keeping δE fixed), while the first term scales with N. Therefore in the very large-N limit the last term is irrelevant. Which proves that $S(E)$ is given by the maximal term in the sum (3.21).

3.2.2 Relation to thermodynamics

We have with (3.18) a definition of the entropy analogous to the expression of $\mathcal{H}$ in (2.109,2.110) which for the variables $E(=U)$, V and N is a thermodynamic potential and allows us to calculate a variety of state variables and relations,

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN \; . \tag{3.30}$$

This allows also to calculate the equation of state,

$$p = T\left(\frac{\partial S}{\partial V}\right)_{E,N} \tag{3.31}$$

and to determine other thermodynamic potentials.
The derivative with respect to N yields finally the chemical potential

$$\mu = -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \; , \tag{3.32}$$

the energy to add a particle to the system.

3.3 Discussion of ideal systems

3.3.1 Classical ideal gas

We consider a classical gas of N independent mono-atomic particles in the fixed volume V, which is closed and isolated. The Hamiltonian is simply given by

$$\mathcal{H}(p,q) = \mathcal{H}(p) = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} \; . \tag{3.33}$$

Hence we obtain the volume

$$\Phi(E) = \int_{\mathcal{H}(p)\leq E} dpdq = V^N \int_{H(p)\leq E} dp \tag{3.34}$$

This p-integral corresponds to the volume of a sphere of Radius R in $3N$-dimensional space.[4] The integral is straightforward,

$$\Phi(E) = V^N C_{3N} \left(2mE\right)^{3N/2} \qquad \text{with} \qquad C_n = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)} \tag{3.37}$$

---

[4] Volume of sphere in ν-dimensional space: $\mathcal{V}(R) = CR^\nu$. The volume enclosed in a thin shell of width δR at the Radius R is given by

$$v_{shell} = \mathcal{V}(R) - \mathcal{V}(R-\delta R) = CR^\nu \left[1 - \left(1 - \frac{\delta R}{R}\right)^\nu\right] \tag{3.35}$$

with $\delta R \ll R$. In the limit ν very large, $\nu\delta R \gg R$, we find

$$v_{shell} \approx \mathcal{V}(R) \; . \tag{3.36}$$

Thus, the main part of the volume in the sphere is concentrated at the outermost shell.

where C_{3N} is the proper prefactor for the volume of an $3N$-dimensional sphere.[5] This leads to

$$\omega(E) = \frac{\partial \Phi(E)}{\partial E} \delta E = C_{3N} V^N \frac{3N}{2} (2mE)^{3N/2-1} \delta E \, . \tag{3.40}$$

Remarkably, for very large N ($\sim 10^{23}$) we find that the following definitions for the entropy are identical up to terms of order $\ln N$ and constants:

$$S_\omega = k_B \ln \omega(E,V,N) \qquad \text{and} \qquad S_\Phi = k_B \ln \Phi(E,V,N) \tag{3.41}$$

leading to

$$\begin{aligned} S_\omega &= k_B \ln(V^N C_{3N}) + k_B \left(\frac{3N}{2} - 1 \right) \ln(2mE) + k_B \ln \left(\frac{3N}{2} \delta E \right) \\ &= k_B \ln(V^N C_{3N}) + k_B \frac{3N}{2} \ln(2mE) + O(\ln N) \approx S_\Phi + O(\ln N) \, . \end{aligned} \tag{3.42}$$

Since we can drop terms of order $\ln N$ for the extensive entropy, we will continue, for convenience, using S_Φ instead of S_ω,

$$S(E,V,N) = N k_B \ln \left\{ V (2m\pi E)^{3/2} \right\} - \frac{3N}{2} k_B \ln \frac{3N}{2} + \frac{3N}{2} k_B \tag{3.43}$$

where we used Stirling's formula

$$\ln n! \approx n \ln n - n + \frac{1}{2} \ln(2\pi n) \qquad \text{for} \quad n \to \infty \, , \tag{3.44}$$

and neglected all terms of order $\ln N$. We then rewrite

$$S(E,V,N) = N k_B \ln \left\{ V \left(\frac{4\pi m E}{3N} \right)^{3/2} \right\} + \frac{3}{2} N k_B \, . \tag{3.45}$$

This equation may now be solved for E so that we obtain the internal energy as a thermodynamic potential

$$U(S,V,N) = E = \frac{3N}{4\pi m V^{2/3}} \exp \left\{ \frac{2S}{3Nk_B} - 1 \right\} \, . \tag{3.46}$$

The thermodynamic quantities are obtained by derivation: the temperature

$$T = \left(\frac{\partial U}{\partial S} \right)_{V,N} = \frac{2U}{3Nk_B} \qquad \Rightarrow \qquad U = \frac{3}{2} N k_B T \, , \tag{3.47}$$

the pressure

$$p = -\left(\frac{\partial U}{\partial V} \right)_{S,N} = \frac{2}{3} \frac{U}{V} = \frac{N k_B T}{V} \quad \Rightarrow \quad pV = N k_B T \tag{3.48}$$

and the chemical potential

$$\mu = -\left(\frac{\partial U}{\partial N} \right)_{S,V} = -\frac{U}{N} \left(1 - \frac{2}{3} \frac{S}{N k_B} \right) = k_B T \ln \left\{ V (2\pi m k_B T)^{3/2} \right\} \, . \tag{3.49}$$

Through further derivatives it is possible to obtain various response functions, as previously discussed in chapter 1. The ideal gas is readily described by means of the microcanonical ensemble.

---

[5] Prefactor C_n: Use the n-dimensional Gaussian integral in Cartesian coordinates

$$I = \int_{-\infty}^{+\infty} dx_1 \cdots \int_{-\infty}^{+\infty} dx_n e^{-(x_1^2 + \cdots + x_n^2)} = \left(\int_{-\infty}^{+\infty} dx e^{-x^2} \right)^n = \pi^{n/2} \, . \tag{3.38}$$

The same integral in spherical coordinates is given by

$$I = n C_n \int_0^\infty dr r^{n-1} e^{-r^2} = \frac{n}{2} C_n \int_0^\infty dt \, t^{\frac{n}{2}-1} e^{-t} = \frac{n}{2} C_n \Gamma\left(\frac{n}{2} \right) = C_n \Gamma\left(\frac{n}{2} + 1 \right) \tag{3.39}$$

such that we obtain C_n given in (3.37). Note, $\Gamma(n+1) = n!$ for $n \geq 0$ as an integer.

3.3.2 Ideal paramagnet

Consider a system of N independent magnetic moments. Each moment has only two directions $+m$ and $-m$ (Ising moments: two-state units). Now apply a magnetic field H in which these moments are either parallel or antiparallel, as described by the Hamiltonian

$$\mathcal{H}(m_i) = -\sum_{i=1}^{N} H m_i \qquad \text{with} \qquad m_i = \pm m \, , \tag{3.50}$$

so that the two states have different energy. In contrast to the gas these degrees of freedom are discrete. We determine now the volume $\omega(E)$ in the space of $(m_1, \ldots, m_N)$. For this purpose we define the magnetization of the systems as

$$M = \sum_i m_i = nm \qquad \Rightarrow \qquad \mathcal{H}(m_i) = -HM = -Hnm \, . \tag{3.51}$$

Fixed energy means a fixed number n, i.e. $\frac{1}{2}(N+n)$ moments are positively and $\frac{1}{2}(N-n)$ negatively oriented ($H > 0$). The number of configuration with fixed n are then given by

$$\omega(M) = \frac{N!}{\left[\frac{1}{2}(N+n)\right]! \left[\frac{1}{2}(N-n)\right]!} \, . \tag{3.52}$$

We can define the probability for the magnetization M by dividing $\omega(M)$ with the total number of configurations, 2^N. Let us now evaluate $\ln\omega$ using Stirling's formula,

$$\begin{aligned} \ln\omega(M) &= N\ln N - N + \frac{1}{2}\ln(2\pi N) - \frac{N}{2}\left(1+\frac{n}{N}\right)\ln\frac{N}{2}\left(1+\frac{n}{N}\right) \\ &\quad -\frac{N}{2}\left(1-\frac{n}{N}\right)\ln\frac{N}{2}\left(1-\frac{n}{N}\right) + \frac{N}{2}\left(1+\frac{n}{N}\right) + \frac{N}{2}\left(1-\frac{n}{N}\right) \\ &\quad -\frac{1}{2}\ln\left\{\pi^2 N^2\left(1-\frac{n^2}{N^2}\right)\right\} \, . \end{aligned} \tag{3.53}$$

Neglecting terms of order $\ln N$ we obtain

$$\ln\omega(M) = -\frac{N}{2}\left\{\left(1+\frac{n}{N}\right)\ln\left(1+\frac{n}{N}\right) + \left(1-\frac{n}{N}\right)\ln\left(1-\frac{n}{N}\right)\right\} + N\ln 2 \, , \tag{3.54}$$

which can be used to express the entropy,

$$\begin{aligned} S(E,H) &= k_B \ln\omega(E,H) \\ &= Nk_B\ln 2 - \frac{Nk_B}{2}\left\{\left(1+\frac{E}{HmN}\right)\ln\left(1+\frac{E}{HmN}\right) + \left(1-\frac{E}{HmN}\right)\ln\left(1-\frac{E}{HmN}\right)\right\} \end{aligned} \tag{3.55}$$

with $E = -nHm$ and introducing H as external parameter (state variable). From this we obtain

$$\begin{aligned} \frac{1}{T} &= \left(\frac{\partial S}{\partial E}\right)_H = -\frac{k_B}{2Hm}\ln\left(\frac{HmN+E}{HmN-E}\right) \\ \Rightarrow \qquad & E(T,H) = U(T,H) = -NHm\tanh\left(\frac{Hm}{k_B T}\right) \, . \end{aligned} \tag{3.56}$$

Using the following differential we determine the magnetization

$$\begin{aligned} dS &= \frac{1}{T}dU + \frac{M}{T}dH - \frac{\mu}{T}dN \\ \Rightarrow \qquad M &= T\left(\frac{\partial S}{\partial H}\right)_E = \frac{k_B T E}{H^2 m}\ln\left(\frac{HmN+E}{HmN-E}\right) = -\frac{E}{H} = Nm\tanh\left(\frac{Hm}{k_B T}\right) \, . \end{aligned} \tag{3.57}$$

This is the thermodynamic equation of state of this system. In the limit $Hm \ll k_B T$ we find the so-called Curie-type behavior of independent magnetic moments,

$$M = N\frac{Hm^2}{k_B T} \tag{3.58}$$

and in the opposite limit the magnetization tends to saturate,

$$M = Nm\left(1 - 2e^{-2Hm/k_B T}\right) . \tag{3.59}$$

The specific heat is given by

$$C_N = \left(\frac{\partial E}{\partial T}\right)_{H,N} = \frac{N(Hm)^2}{k_B T^2}\cosh^{-2}\left(\frac{Hm}{k_B T}\right) . \tag{3.60}$$

A further response function is the differential magnetic susceptibility at finite magnetic field H,

$$\chi'(T) = \left(\frac{\partial M}{\partial H}\right) = \frac{Nm^2}{k_B T}\cosh^{-2}\left(\frac{Hm}{k_B T}\right) \tag{3.61}$$

which corresponds to the "linear response". For the limit of $H \to 0$ we obtain,

$$\chi(T) = \frac{Nm^2}{k_B T} , \tag{3.62}$$

corresponding to the famous Curie-susceptibility of independent magnetic moments.

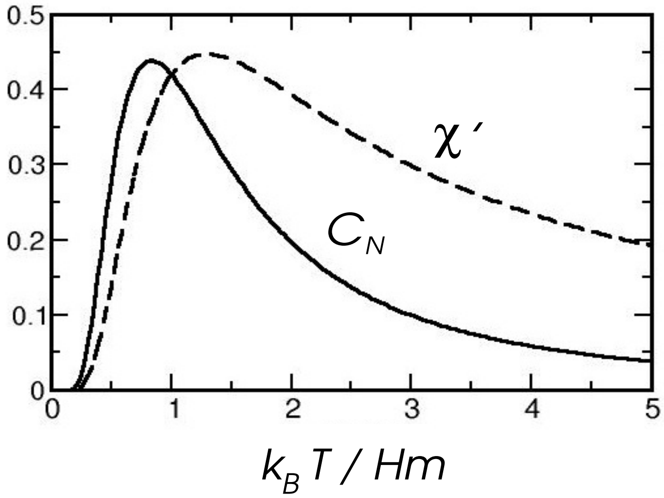


Fig. 3.2: *Specific heat and differential susceptibility (magnitude in arbitrary units): Both quantities show a maximum around $mH \sim k_B T$. At low temperature both are exponentially suppressed $\sim T^{-\alpha}e^{-2Hm/k_B T}$, which is indicating a freezing of the degrees of freedom, i.e. at $T = 0$ all magnetic moments are pointing in the same direction.*

3.4 Canonical ensemble

We change to a macroscopic system for which we control the temperature by connecting it to a very large heat reservoir. The system together with the reservoir forms a closed system of given total energy. Therefore we consider two subsystems, system 1 describing our system and system 2 being the heat reservoir:

$$\mathcal{H}(p,q) = \mathcal{H}_1(p_1,q_1) + \mathcal{H}_2(p_2,q_2) . \tag{3.63}$$

The heat reservoir is much larger than system 1: $N_2 \gg N_1$, such that energy transfer between the two subsystems would be too small to change the temperature of the reservoir. Within the microcanonical scheme we determine the phase space of the combined system in the energy range

$$E \leq E_1 + E_2 \leq E + \delta E \ . \tag{3.64}$$

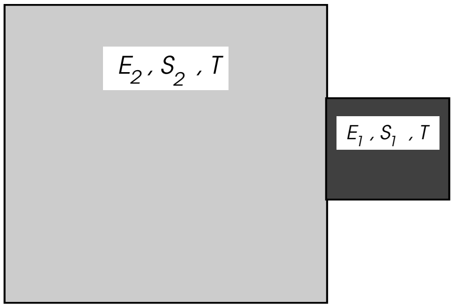


Fig. 3.3: *The two systems 1 and 2 are coupled thermally. System 2 acts as a huge reservoir fixing the temperature of the system 1 which we want to investigate.*

Therefore the volume of the microcanonical ensemble of the total system is

$$\omega(E) = \sum_{0 \leq E_1 \leq E} \omega_1(E_1)\omega_2(E - E_1) \tag{3.65}$$

Analogous to our previous discussion in section 3.2.1, there is one value $\bar{E}_1 = E'_0$ ($\bar{E}_2 = E - E'_0$) which provides the by far dominant contribution. In addition here $\bar{E}_2 \gg \bar{E}_1$ is valid. The corresponding volumes in Γ-space are $\omega_1(\bar{E}_1)$ and $\omega_2(\bar{E}_2)$ and $\omega(E) \approx \omega_1(\bar{E}_1)\omega_2(E - \bar{E}_1)$. Due to this simple product form we can determine the density function $\rho(p_1, q_1)$ of the system 1 by the integral (integrating over the degrees of freedom of system 2),

$$\rho(p_1, q_1) = \int_{\bar{E}_2 \leq \mathcal{H}_2(p_2,q_2) \leq \bar{E}_2 + \delta E} dp_2 dq_2 \rho(p, q) \propto \omega_2(E - \bar{E}_1) \tag{3.66}$$

where we take advantage of the fact the $\rho(p, q)$ is constant within the energy range $E \leq E_1 + E_2 \leq E + \delta E$, such that the integral is indeed proportional to the volume of the microcanonical ensemble of the subsystem 2 (reservoir). Using the assumption that $\bar{E}_2 \approx E \gg \bar{E}_1$ we may expand in $\bar{E}_1$:

$$k_B \ln \omega_2(E - \bar{E}_1) = S_2(E - \bar{E}_1) = S_2(E) - \bar{E}_1 \left. \frac{\partial S_2(\bar{E}_2)}{\partial E_2} \right|_{\bar{E}_2 = E} + \cdots = S_2(E) - \frac{\bar{E}_1}{T} + \cdots \tag{3.67}$$

from which we derive

$$\omega_2(E - \bar{E}_1) = e^{S_2(E)/k_B} e^{-\bar{E}_1/k_B T} \ . \tag{3.68}$$

Here T is the temperature of both systems which are in equilibrium. Therefore we identify the probability $\rho(p_1, q_1)$ for the small subsystem as

$$\rho(p_1, q_1) = \frac{1}{Z} e^{-\mathcal{H}_1(p_1,q_1)/k_B T} \ . \tag{3.69}$$

where we introduced the *partition function* Z

$$Z = \int dp_1 \, dq_1 e^{-\beta \mathcal{H}_1(p_1,q_1)} \qquad \text{with} \quad \beta = \frac{1}{k_B T} \ , \tag{3.70}$$

which, up to prefactors, corresponds to the volume of the ensemble of system 1, called *canonical ensemble*.[6]

---

[6] Note that there is, rigorously speaking, the constraint $E_1 < E$. However, in the approximation ignoring this constraint is good enough, as the main contribution is from the valid range.

3.4.1 Thermodynamics

The connection to thermodynamics is given by the relation

$$Z = e^{-\beta F(T,V,N)} \;, \tag{3.71}$$

where $F(T,V,N)$ is the Helmholtz free energy, a thermodynamical potential. Note, F is an extensive quantity, because obviously scaling the system by a factor λ would yield Z^λ. Moreover,

$$F = U - TS \tag{3.72}$$

with

$$U = \langle \mathcal{H} \rangle \qquad \text{and} \qquad S = -\left(\frac{\partial F}{\partial T} \right)_{V,N} \;. \tag{3.73}$$

This can be proven using the equation,

$$1 = \int dp\; dq e^{\beta (F - \mathcal{H})} \;, \tag{3.74}$$

which through differentiation with respect to β on both sides gives,

$$\begin{aligned} 0 &= \int dp\; dq e^{\beta(F-\mathcal{H})} \left\{ F + \beta \left(\frac{\partial F}{\partial \beta} \right)_{V,N} - \mathcal{H} \right\} \\ \Rightarrow \quad & F(T,V,N) - U(T,V,N) - T \left(\frac{\partial F}{\partial T} \right)_{V,N} = 0 \;. \end{aligned} \tag{3.75}$$

Using this formulation for the free energy we obtain for the pressure

$$p = -\left(\frac{\partial F}{\partial V} \right)_{T,N} \;. \tag{3.76}$$

which in the case of a gas leads to the thermodynamic equation of state.
The internal energy is obtained easily from the partition function in the following way:

$$U(T,V,N) = \langle \mathcal{H} \rangle = \frac{1}{Z} \int dp\; dq\; \mathcal{H} e^{-\beta \mathcal{H}} = -\frac{1}{Z} \frac{\partial Z}{\partial \beta} = -\frac{\partial}{\partial \beta} \ln Z \;. \tag{3.77}$$

This is the caloric equation of state.

3.4.2 Equipartition law

We now consider a set of special average values which will lead us to the so-called equipartition law, the equal distribution of energy on equivalent degrees of freedom. We examine the mean value,

$$\begin{aligned} \left\langle q_i \frac{\partial \mathcal{H}}{\partial q_j} \right\rangle &= \frac{1}{Z} \int dp\; dq q_i \frac{\partial \mathcal{H}}{\partial q_j} e^{-\beta \mathcal{H}} = -\frac{1}{Z\beta} \int dp\; dq q_i \frac{\partial}{\partial q_j} e^{-\beta \mathcal{H}} \\ &= -\frac{1}{Z\beta} \underbrace{\int' dp\; d'q\; q_i e^{-\beta \mathcal{H}}}_{=0} + \frac{\delta_{ij}}{Z\beta} \int dp dq e^{-\beta \mathcal{H}} = \delta_{ij} k_B T \;, \end{aligned} \tag{3.78}$$

where we used integration by parts leading to the boundary terms in the q_j-coordinate (expressed by $\int' d'q...$), which we assume to vanish. Analogously we find for the momentum

$$\left\langle p_i \frac{\partial \mathcal{H}}{\partial p_j} \right\rangle = \delta_{ij} k_B T \;. \tag{3.79}$$

If the Hamiltonian is separable into a p-dependent kinetic energy and a q-dependent potential energy part and, moreover, if the following scaling behavior is valid

$$\mathcal{H}(p,q) = E_{kin}(p) + V(q) \quad \text{with} \quad E_{kin}(\lambda p) = \lambda^2 E_{kin}(p) \quad \text{and} \quad V(\lambda q) = \lambda^\alpha V(q) \tag{3.80}$$

then we can use the above relations and find for mono-atomic particles

$$\langle E_{kin} \rangle = \frac{3N}{2} k_B T \qquad \text{and} \qquad \langle V \rangle = \frac{3N}{\alpha} k_B T \; . \tag{3.81}$$

The total energy is given by the sum of the two contributions.

Diatomic molecules: An interesting aspect appears in the case of multi-atomic molecules, as such they have internal degrees of freedom and a binding potential. We consider here a system of N atoms forming diatomic molecules. Molecules form due to a binding two-particle interaction: $V(q) = \sum_{i<j} V_{ij}(q)$ with $V_{ij}(q) = v(|\vec{q}_i - \vec{q}_j|)$. A good approximation for the binding potential is the Lenard-Jones potential composed of an attractive and a "hard-core" repulsive part,

$$v(r) = 4\epsilon \left\{ \left(\frac{\tilde{r}}{r} \right)^{12} - \left(\frac{\tilde{r}}{r} \right)^{6} \right\} , \tag{3.82}$$

with ϵ as the potential depth and $r_0 = 2^{1/6}\tilde{r}$ as the minimal point (Fig. 3.4). Around the minimum the potential can be approximated up to a constant by a harmonic potential,

$$v(r) \approx A(r - r_0)^2 \qquad \text{with} \qquad A = 36\epsilon / 2^{1/3} \tilde{r}^2 \; . \tag{3.83}$$

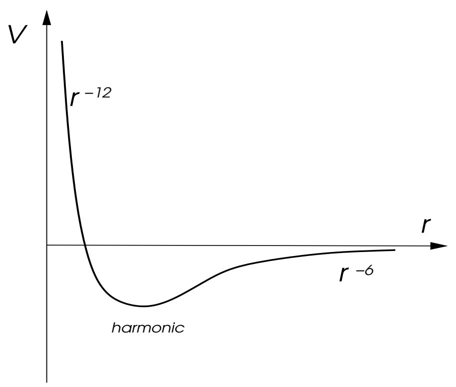


Fig. 3.4: *Lenard-Jones potential.*

The part of the potential energy is obtained from the following partition function of the molecule (for $\epsilon \gg k_B T$):

$$Z_v = \int d^3 r e^{-\beta v(r)} = 4\pi \int_0^\infty dr\, r^2 e^{-\beta A (r-r_0)^2} \approx 4\pi r_0^2 \int_{-\infty}^{+\infty} dy\, e^{-Ay^2} = 4\pi r_0^2 \sqrt{\frac{\pi}{A\beta}} , \tag{3.84}$$

which then yields,

$$\langle v \rangle = -\frac{\partial}{\partial \beta} \ln Z_v = \frac{k_B T}{2} \tag{3.85}$$

Together with the kinetic energy part we obtain

$$U = \langle E_{kin} \rangle + \langle V \rangle = \frac{3N}{2} k_B T + \frac{N}{2} \langle v \rangle \approx \frac{7N}{4} k_B T = \frac{N}{2} \left\{ \frac{5}{2} k_B T + \underbrace{k_B T}_{\text{harmonic oscillator}} \right\} . \tag{3.86}$$

This result is interesting in the following sense. The specific heat of the dissociated molecules at high temperature is lower than the one of the diatomic molecule, although there are more independent particles at high temperature:

$$\frac{C_V}{N} = \frac{1}{N}\left(\frac{\partial U}{\partial T}\right)_{V,N} = \begin{cases} \dfrac{3k_B}{2} & k_BT \gg \epsilon \\ \dfrac{7k_B}{4} & k_BT \ll \epsilon \end{cases} . \tag{3.87}$$

This result is independent of the sharpness of the harmonic potential. However, if we turn to the limit of the rigid molecular bond, i.e. no harmonic potential (or infinitely sharp), then the situation changes. The motion of the two atoms is constraint by the rigidity condition: $(\vec{p}_i - \vec{p}_j)\cdot(\vec{q}_i - \vec{q}_j) = 0$. Then the internal energy is given by

$$U = \frac{5N}{4}k_BT \qquad \Rightarrow \qquad \frac{C_V}{N} = \frac{5k_B}{4} . \tag{3.88}$$

The change to this value is singular within classical statistical mechanics. Only quantum mechanics is able to describe the "freezing out" of the vibrational degree of freedom. In fact within the quantum mechanical description we find even a further lowering when the rotational degrees of freedom freeze out such that eventually $U = 3k_BT/4$. Further temperature lowering then requires to consider the system as a quantum liquid of Bosons or Fermions.

3.4.3 Fluctuation of the energy and the equivalence of microcanonial and canonical ensembles

The internal energy is given as the average of the Hamiltonian $U = \langle \mathcal{H} \rangle$. Therefore the following relation holds:

$$\int dp\, dq\ [U - \mathcal{H}]\, e^{\beta(F-\mathcal{H})} = 0 . \tag{3.89}$$

Taking the derivative of this equation with respect to β we obtain

$$0 = \frac{\partial U}{\partial \beta} + \int dp\, dq\, (U - \mathcal{H})\left[F - T\frac{\partial F}{\partial T} - \mathcal{H}\right] e^{\beta(F-\mathcal{H})} = \frac{\partial U}{\partial \beta} + \langle (U-\mathcal{H})^2 \rangle . \tag{3.90}$$

This leads to the relation for the fluctuations of the energy around its average value U,

$$\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2 = \langle (U-\mathcal{H})^2 \rangle = -\frac{\partial U}{\partial \beta} = k_BT^2\frac{\partial U}{\partial T} = k_BT^2 C_V , \tag{3.91}$$

as we had seen in section 2.2.3. Because C_V is an extensive quantity and therefore proportional to N, it follows that

$$\frac{\langle \mathcal{H}^2 \rangle - \langle \mathcal{H} \rangle^2}{\langle \mathcal{H} \rangle^2} \propto \frac{1}{N} \tag{3.92}$$

which is a sign of the equivalence of microcanonical and canonical ensembles. In the thermodynamic limit $N \to \infty$ the fluctuations of the energy vanish compared to the energy itself.

We now consider the partition function

$$Z = \int dp\, dq\, e^{-\beta \mathcal{H}(p,q)} = \int_0^\infty dE\ \omega(E) e^{-\beta E} = \int_0^\infty dE\ e^{-\beta E + \ln \omega(E)} = \int_0^\infty dE\ e^{\beta(TS(E)-E)} \tag{3.93}$$

where the entropy $S(E)$ is defined according to the microcanonical ensemble. The maximum of the integrand at $E = E_0$ is defined by the condition

$$T\left.\frac{\partial S}{\partial E}\right|_{E=E_0} = 1 \qquad \text{and} \qquad \left.\frac{\partial^2 S}{\partial E^2}\right|_{E=E_0} < 0 . \tag{3.94}$$

Note that

$$\left.\frac{\partial^2 S}{\partial E^2}\right|_{E=E_0} = \left.\frac{\partial}{\partial E}\frac{1}{T}\right|_{E=E_0} = -\frac{1}{T^2}\left.\frac{\partial T}{\partial E}\right|_{E=E_0} = -\frac{1}{T^2 C_V} < 0 \,. \tag{3.95}$$

If we now expand the exponent in the integrand, we obtain

$$TS(E) - E = TS(E_0) - E_0 - \frac{1}{2TC_V}(E-E_0)^2 + \cdots \tag{3.96}$$

such that with $U = E_0$,

$$\begin{aligned} Z &\approx e^{\beta(TS-U)} \int_0^\infty dE\; e^{-(E-U)^2/2k_B T^2 C_V} = e^{\beta(TS-U)}\sqrt{2\pi k_B T^2 C_V} \\ \Rightarrow \quad F &\approx U - TS - \frac{1}{2}k_B T \ln\{2\pi k_B T C_V\} = U - TS + O(\ln N) \,. \end{aligned} \tag{3.97}$$

Since the free energy is extensive, the term of order $\ln N$ is in the large-N limit irrelevant.

3.4.4 Ideal gas in canonical ensemble

Consider a gas of N particles without external potential and mutual interactions described by the Hamiltonian

$$\mathcal{H}(p) = \sum_{i=1}^{N} \frac{\vec{p}_i^{\,2}}{2m} \,. \tag{3.98}$$

The partition function is given by

$$Z = \prod_{i=1}^{N} \int d^3p_i\, d^3q_i\; e^{-\vec{p}_i^{\,2}/2mk_BT} = \left\{ \int d^3p\, d^3q\; e^{-\vec{p}^{\,2}/2mk_BT} \right\}^N = V^N \{2\pi m k_B T\}^{3N/2} \,. \tag{3.99}$$

From this we obtain the free energy and the internal energy

$$\begin{aligned} F(T,V,N) &= -k_BT\ln Z = -\frac{3N}{2}k_BT\ln(2\pi m k_B T) - Nk_BT\ln V \,, \\ U(T,V,N) &= -\frac{\partial}{\partial\beta}\ln Z = \frac{3N}{2}k_BT \,. \end{aligned} \tag{3.100}$$

The entropy is given by

$$S(T,V,N) = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = Nk_B\ln(VT^{3/2}) + \frac{3N}{2}k_B + S_0 \tag{3.101}$$

and the pressure by

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \frac{Nk_BT}{V} \tag{3.102}$$

which corresponds to the thermodynamic equation of state.

An important aspect for the ideal system is the fact that the partion function has a product form because each particle is described independently. In this way it leads to an extensive free energy and internal energy.

3.4.5 Ideal paramagnet

As a next example we consider again an ideal paramagnet. Unlike the example discussed in the microcanonical ensemble we use here a magnetic moment $\vec{m}$ which can point in any direction (full rotational symmetry) and is represented as a vector of fixed length m. The externally applied magnetic field couples in the usual Zeeman form,

$$\mathcal{H}(\vec{m}_i) = -\sum_{i=1}^{N} \vec{m}_i \cdot \vec{H} = -\sum_{i=1}^{N} mH\cos\theta_i \tag{3.103}$$

where θ_i denotes the angle between magnetic field and moment (magnetic field defines z-axis). It is easy to calculate now the partion function

$$Z = \prod_{i=1}^{N} \int d\Omega_i \, e^{-\beta mH\cos\theta_i} = \left\{ 2\pi \int_{-1}^{+1} dx \, e^{-\beta mHx} \right\}^N = \left\{ \frac{4\pi}{mH\beta} \sinh(\beta mH) \right\}^N = Z_m^N . \tag{3.104}$$

This again shows the product form. The thermodynamics is again obtained via the free energy and the internal energy,

$$\begin{aligned} F(T,H,N) &= -k_BT\ln Z = Nk_BT\ln(\beta mH) - Nk_BT\ln\sinh(\beta mH) \\ U(T,H,N) &= Nk_BT - NmH\coth(\beta mH) \end{aligned} \tag{3.105}$$

From this we obtain

$$C_H(T,H,N) = Nk_B - N\frac{(mH)^2}{k_BT^2}\frac{1}{\sinh^2(mH/k_BT)} . \tag{3.106}$$

which yields the limiting behaviors

$$C_H = \begin{cases} \dfrac{Nk_B}{3}\left(\dfrac{mH}{k_BT}\right)^2 & k_BT \gg mH \\ Nk_B\left(1 - 4\left(\dfrac{mH}{k_BT}\right)^2 e^{-2mH/k_BT}\right) & k_BT \ll mH \end{cases} \tag{3.107}$$

Analogous to our previous ideal paramagnet, the magnetic field enters as a characteristic energy scale separating two regimes of thermodynamics.

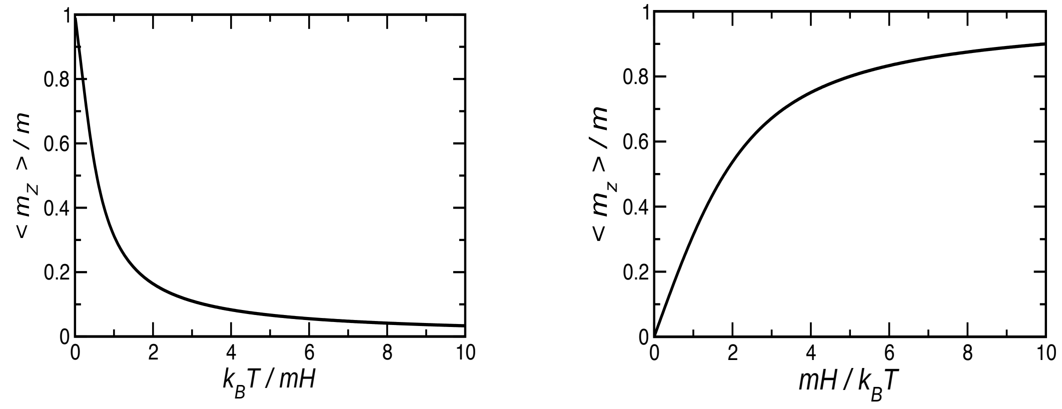


Fig.3.5: Mean magnetization per moment: left panel, as a function of temperature and fixed magnetic field, with full saturation at $T = 0$; right panel, as a function of magnetic field and fixed temperature, with full saturation at $mH \gg k_BT$.

The magnetization can be calculated as

$$\langle \vec{m} \rangle = \frac{1}{Z_m} \int d\Omega \, \vec{m}(\phi,\theta) e^{-\beta \vec{m}\cdot\vec{H}} \tag{3.108}$$

with $\vec{m}(\phi,\theta) = m(\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$ leading to $\langle m_x \rangle = \langle m_y \rangle = 0$ and

$$\langle m_z \rangle = m \coth\left(\frac{mH}{k_B T}\right) - \frac{k_B T}{H} \tag{3.109}$$

which corresponds to

$$\langle \vec{m} \rangle = -\frac{1}{N}\left(\frac{\partial F}{\partial \vec{H}}\right)_{T,N} . \tag{3.110}$$

The second derivative with respect to $\vec{H}$ yields the magnetic susceptibility as a response function (in the limit of $\vec{H} \to 0$ it is the linear response)[7]

$$\begin{aligned} \chi_{zz}(T) &= -\left(\frac{\partial^2 F}{\partial H_z^2}\right)_{T,N} = \frac{Nm^2}{k_B T}\left[\left(\frac{k_B T}{mH}\right)^2 - \frac{1}{\sinh^2(mH/k_B T)}\right] \\ &= \begin{cases} N\dfrac{m^2}{3k_B T} & k_B T \gg mH \\ N\dfrac{m^2}{k_B T}\left\{\left(\dfrac{k_B T}{mH}\right)^2 - 4e^{-2mH/k_B T}\right\} & k_B T \ll mH \end{cases} \end{aligned} \tag{3.112}$$

It is interesting to study the fluctuations of the magnetization,

$$\langle m_z^2 \rangle - \langle m_z \rangle^2 = m^2 \left[\left(\frac{k_B T}{mH}\right)^2 - \frac{1}{\sinh^2(mH/k_B T)}\right] = \frac{k_B T}{N}\chi_{zz} . \tag{3.113}$$

Note that here the zero-temperature limit of the specific heat does not vanish, but is Nk_B and corresponds to the contribution of the potential energy of a two-dimensional harmonic oscillator (the fluctuating magnetic moment nearly aligned with magnetic field feels a harmonic potential). Analogous to the fluctuations of the energy discussed above we find that the fluctuations of the magnetization are connected with a response function, the susceptibility.[8] This is a simplified version of the so-called fluctuation-dissipation theorem which will encounter later again.

---

[7] Generally we define

$$N\langle m_\nu \rangle = -\left(\frac{\partial F}{\partial H_\nu}\right)_{T,N} \quad \text{and} \quad \chi_{\nu\nu'} = -\left(\frac{\partial^2 F}{\partial H_\nu \partial H_{\nu'}}\right)_{T,N} . \tag{3.111}$$

[8] Generally we find the relation for the fluctuations of the magnetization $\vec{M} = \sum_i \vec{m}_i$ from the following argument:

$$0 = \int d\Omega_1 \cdots d\Omega_N (\langle \vec{M} \rangle - \vec{M}) e^{\beta(F-\mathcal{H})} \tag{3.114}$$

taking the derivative with respect to the field leads to

$$0 = \frac{\partial \langle M_z \rangle}{\partial H} + \beta \int d\Omega_1 \cdots d\Omega_N (\langle M_z \rangle - M_z)^2 e^{\beta(F-\mathcal{H})} = \chi_{zz} - \beta\langle(\langle M_z \rangle - M_z)^2\rangle . \tag{3.115}$$

This yields the convenient relation,

$$\chi_{zz} = \frac{1}{k_B T}\left\{\langle M_z^2 \rangle - \langle M_z \rangle^2\right\} . \tag{3.116}$$

3.4.6 More advanced example - classical spin chain

We now go beyond the ideal paramagnetic system by considering a chain of spins (magnetic moments) $\{\vec{s}_i\}$ which interact with each other via nearest-neighbor coupling in the following way:

$$\mathcal{H}(\vec{s}_i) = J \sum_{i=1}^{N} \vec{s}_i \cdot \vec{s}_{i+1} \tag{3.117}$$

where J is the coupling strength. For $J < 0$ the spins tend to align parallel (ferromagnetic) and for $J > 0$ they tend to be antiparallel, alternating (antiferromagnetic). This chain is open and occupied regularly by $N+1$ spins (N bonds).

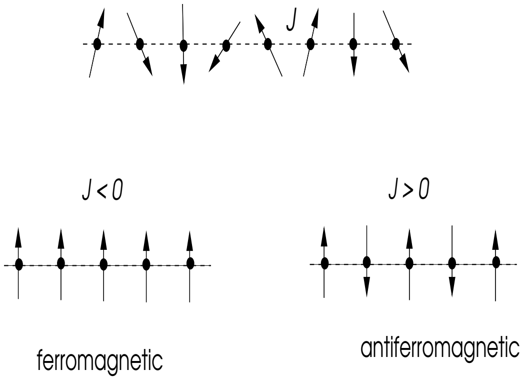


Fig.3.6: Spin chain: Coupling J between spins of arbitrary orientation. $J < 0$: ferromagnetic alignment prefered; $J > 0$: antiferromagnetic alignment prefered.

Interestingly the topology of the coupling allows to calculate the partition function exactly. We consider the spins as vectors of fixed length S and write the partition function as

$$Z = \int d\Omega_1 \cdots d\Omega_{N+1} \, e^{-\beta J \sum_i \vec{s}_i \cdot \vec{s}_{i+1}} , \tag{3.118}$$

which decays into product form, if we consider the reference polar axis ("z-axis") for the spin $\vec{s}_i$ given by the direction of the spin $\vec{s}_{i+1}$ (θ_i is defined as angle between $\vec{s}_i$ and $\vec{s}_{i+1}$). Therefore may write

$$\begin{aligned} Z &= \prod_{i=1}^{N} \left\{ \int e^{-\beta J S^2 \cos\theta_i} d\Omega_i \right\} \int d\Omega_{N+1} = 4\pi \left\{ 2\pi \int_{-1}^{+1} dx e^{-\beta J S^2 x} \right\}^N \\ &= 4\pi \left\{ 4\pi \frac{\sinh(\beta J S^2)}{\beta J S^2} \right\}^N \end{aligned} \tag{3.119}$$

which looks similar to the expression obtained for the ideal paramagnet. We may actually consider each spin being subject to the Zeeman field induced via the coupling by the neighboring spins (mean field idea). It is interesting to see that Z does not depend on the sign of J. Thus, the thermodynamics is the same for a ferromagnetic and an antiferromagnetic chain.

Easily we can determine the free and the internal energy,

$$\begin{aligned} F(T,N) &= N k_B T \ln(\beta J S^2) - N k_B T \ln\left\{\sinh(\beta J S^2)\right\} - N k_B T \ln 4\pi \\ U(T,N) &= N k_B T - N J S^2 \coth(\beta J S^2) \end{aligned} \tag{3.120}$$

which leads to the specific heat

$$C_V = N k_B \left\{ 1 - \left(\frac{\beta J S^2}{\sinh(\beta J S^2)} \right)^2 \right\} \tag{3.121}$$

with a similar form as for the ideal paramagnet.

Correlation function: Now we turn to a new aspect which did not encounter in the ideal paramagnet. The question of the mutual influence of spins is not meaningful in the ideal paramagnet, since the spins are completely independent. Moreover, there was no topology which would tell us the "distance" between two spins. The same is true for the positions of particles in an ideal gas. For our spin chain, however, the question of how the orientation of spin $\vec{s}_i$ influences the orientation of spin $\vec{s}_{i+l}$ makes sense, since we have a clear topology and coupling among the spins. We define the correlation function

$$\begin{aligned}\Gamma_l &= \langle \vec{s}_i \cdot \vec{s}_{i+l} \rangle = 3\langle s_i^z s_{i+l}^z \rangle \\ &= \frac{3}{Z} \int d\Omega_1 \cdots d\Omega_{N+1} S^2 \cos\Theta_i \cos\Theta_{i+l} e^{-\beta J S^2 \sum_i \cos\theta_i} \,.\end{aligned} \tag{3.122}$$

where Θ_i is the angle of spin i with respect to the general z-axis (e.g. along the chain). We can perform all integrations for the spin $\vec{s}_j$ with $j < i$ and $j > i + l$ which decay into the same product form as discussed above. The remaining integrals are now

$$\begin{aligned}\Gamma_l = 3S^2 \left(\frac{\beta J S^2}{4\pi \sinh(\beta J S^2)} \right)^l \int d\Omega_i \cos\Theta_i e^{-\beta J S^2 \cos\theta_i} \\ \times \int d\Omega_{i+1} e^{-\beta J S^2 \cos\theta_{i+1}} \cdots \int d\Omega_{i+l} \cos\Theta_{i+l}\end{aligned} \tag{3.123}$$

Taking again the spin s_{i+1} as a reference of the spin s_i we find the relation[9]

$$\cos\Theta_i = \cos\Theta_{i+1} \cos\theta_i + \sin\Theta_i \sin\theta_i \cos\phi_i' ; . \tag{3.126}$$

Inserting this we notice that the averages $\langle \cos\theta_i \rangle \neq 0$ and $\langle \sin\theta_i \cos\phi_i \rangle = 0$. Thus

$$\begin{aligned}\Gamma_l &= 3S^2 \left(\frac{\beta J S^2}{4\pi \sinh(\beta J S^2)} \right)^l \int d\Omega_i \cos\theta_i e^{-\beta J S^2 \cos\theta_i} \\ &\quad \times \int d\Omega_{i+1} \cos\Theta_{i+1} e^{-\beta J S^2 \cos\theta_{i+1}} \cdots \int d\Omega_{i+l} \cos\Theta_{i+l} \\ &= \Gamma_{l-1} \frac{\beta J S^2}{4\pi \sinh(\beta J S^2)} \int d\Omega_i \cos\theta_i e^{-\beta J S^2 \cos\theta_i} = \Gamma_{l-1} u(\beta J S^2)\end{aligned} \tag{3.127}$$

with

$$u(x) = \frac{1}{x} - \coth x \,. \tag{3.128}$$

If we take into account that $\Gamma_{l=0} = S^2$ we find

$$\Gamma_l = S^2 \left[u(\beta J S^2) \right]^l = S^2 e^{-l/\xi} (-sign(J))^l \quad \text{with} \quad \xi^{-1} = -\ln\{|u(\beta J S^2)|\} \,. \tag{3.129}$$

---

[9] Consider $\vec{s}_1$ and $\vec{s}_2$ and take $\hat{y}' \perp \vec{z}, \vec{s}_2$:

$$\vec{s}_1 = \vec{s}_2 \cos\theta_1 + (\hat{y}' \times \vec{s}_2) \sin\theta_1 \cos\phi_1 + \hat{y}' \sin\theta_1 \sin\phi_1 \tag{3.124}$$

and multiply by $\hat{z}$:

$$\underbrace{\hat{z} \cdot \vec{s}_1}_{= S \cos\Theta_1} = \cos\theta_1 \underbrace{\hat{z} \cdot \vec{s}_2}_{= S \cos\Theta_2} + \vec{y}' \cdot \underbrace{(\vec{s}_2 \times \hat{z})}_{= \hat{y} S \sin\Theta_2} \sin\theta_1 \cos\phi_1 \,. \tag{3.125}$$

Note that $0 < u < 1$ for all finite temperatures and couplings. The correlation function decays exponentially for all finite temperature with a correlation length ξ:

$$\xi(T) = \begin{cases} [\ln(k_B T/|J|S^2)]^{-1} & k_B T \gg |J|S^2 \\ \\ \dfrac{|J|S^2}{k_B T} & k_B T \ll |J|S^2 \end{cases} \tag{3.130}$$

For large temperature the correlation length shrinks rather slowly to zero and in the zero-temperature limit it diverges indicating an infinitely extending correlation. Indeed we find for $T = 0$ that $u \to 1$ such that

$$\Gamma_l(T = 0) = S^2(-\mathrm{sign}J)^l \tag{3.131}$$

for all l. This means that the spins order either ferromagnetically for $J < 0$ or antiferromagnetically for $J > 0$ at $T = 0$. However, for any finite temperature $\Gamma_l \to 0$ for $l \to \infty$.

Finally we want to study the susceptibility assuming a coupling to the magnetic field of the form

$$\mathcal{H}' = \mathcal{H} - g \sum_{i=1}^{N+1} \vec{s}_i \cdot \vec{H} \ ., \tag{3.132}$$

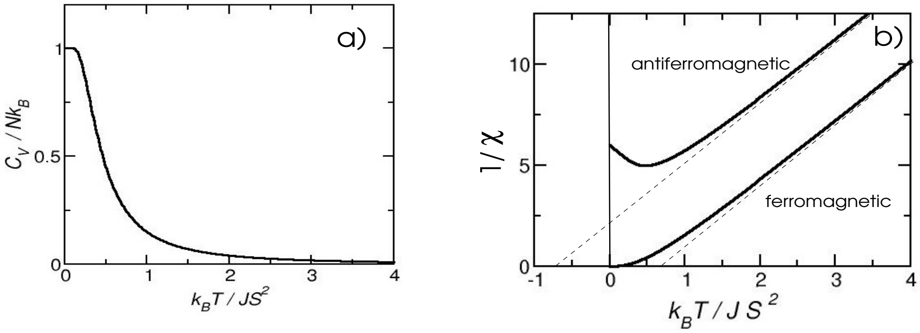


Fig.3.7: Thermodynamic properties of spin chain: a) Specific heat: saturates at low temperature to Nk_B like the potential energy of a two-dimensional harmonic oscillator; b) Susceptibility plotted inversely: ferromagnetic case show divergence at $T = 0$ and antiferromagnetic case approaches non-monotonically a finite value at $T = 0$. Extrapolations from high-temperature region cut horizontal axis at positive (ferromagnetic) and negative (antiferromagnetic) axis, indicating effective energy scales for the corresponding correlations.

since it is impossible to use the above scheme to calculate the partition function with a finite magnetic field coupling to the magnetic moments. Therefore, we will use the relation derived earlier (3.116) based on the fluctuations of the magnetization, which does not require the knowledge of the partition function at finite fields.

$$\chi_{zz} = \frac{g^2}{k_B T} \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} \left\{ \langle s_i^z s_j^z \rangle - \langle s_i^z \rangle \langle s_j^z \rangle \right\} \ . \tag{3.133}$$

The second term on the right hand side vanishes, since $\langle s_i^z \rangle = 0$ for all i, as can be easily shown. For the first term we can use our result for the correlation function. In the limit of very large N we obtain

$$\begin{aligned} \chi_{zz} &= \frac{g^2}{3k_B T} \sum_{i=1}^{N+1} \sum_{j=1}^{N+1} \Gamma_{|i-j|} \approx \frac{g^2}{3k_B T} \sum_{i=1}^{N} \left(\Gamma_{l=0} + 2S^2 \sum_{l=1}^{\infty} u^l \right) \\ &= \frac{g^2 N}{3k_B T} S^2 \left(1 + \frac{2u}{1-u} \right) = \frac{g^2 S^2 N}{3k_B T} \frac{1 + u(\beta J S^2)}{1 - u(\beta J S^2)} \ . \end{aligned} \tag{3.134}$$

This susceptibility at high temperature ($k_B T \gg JS^2$) follows the Curie behavior

$$\chi_{zz} = \frac{g^2 S^2 N}{3 k_B T} \tag{3.135}$$

irrespective of the sign of J. For the antiferromagnetic coupling χ_{zz} goes through a maximum at $k_B T_{max} \approx 0.238 JS^2$ to a finite zero-temperature value. On the other hand, for the ferromagnetic case χ_{zz} diverges at low temperatures

$$\chi_{zz} = N \frac{g^2 S^2}{3} \frac{JS^2}{(k_B T)^2} \ , \tag{3.136}$$

more strongly than the Curie $1/T$ behavior.
Let us now also look at the first lowest order correction in the high-temperature limit,

$$\frac{1}{\chi_{zz}} \approx \frac{3 k_B T}{g^2 S^2 N} \left(1 + \frac{2JS^2}{3 k_B T} \right) = \frac{3 k_B}{g^2 S^2 N} (T - \Theta_W) \quad \Rightarrow \quad \chi_{zz} = \frac{C}{T - \Theta_W} \tag{3.137}$$

where $\Theta_W = -2JS^2 / 3k_B$ defines the so-called Weiss temperature. Thus the extrapolation from the high-temperature limit of $1/\chi_{zz}$ allows from the intersection on the temperature axis to determine Θ_W and consequently the sign and magnitude of the J (see dashed lines in Fig. 3.6).

3.5 Grand canonical ensemble

We consider now a new situation by allowing beside the heat exchange also the exchange of matter of our system with a very large reservoir. Thus we take the system 1 with N_1 particles in a volume V_1 coupled to the reservoir 2 with N_2 particles in the volume V_2 with

$$N_1 \ll N_2 \qquad \text{and} \qquad V_1 \ll V_2 \ , \tag{3.138}$$

and $N = N_1 + N_2$ and $V = V_1 + V_2$ fixed. The Hamiltonian can be decomposed into two parts

$$\mathcal{H}(p, q, N) = \mathcal{H}(p_1, q_1, N_1) + \mathcal{H}(p_2, q_2, N_2) \tag{3.139}$$

such that the corresponding partion function for given temperature (everything is coupled to an even larger heat reservoir) is given by

$$Z_N(V, T) = \frac{1}{N!} \int dp \, dq \, e^{-\beta \mathcal{H}(p,q)} \ . \tag{3.140}$$

For counting reasons we introduce here the factor $1/N!$ taking into account that all possible commutation of the particles give the same states (distinguishable classical particles). Now we segregate into the subsystems fixing the volumes and particle numbers ($N_2 = N - N_1$),

$$\begin{aligned} Z_N &= \frac{1}{N!} \sum_{N_1=0}^{N} \frac{N!}{N_1! N_2!} \int dp_1 \, dp_2 \int_{V_1} dq_1 \int_{V_2} dq_2 e^{-\beta \mathcal{H}(p_1,q_1,N_1) + \mathcal{H}(p_2,q_2,N_2)} \\ &= \sum_{N_1=0}^{N} \frac{1}{N_1!} \int_{V_1} dp_1 \, dq_1 \, e^{-\beta \mathcal{H}(p_1,q_1,N_1)} \quad \frac{1}{N_2!} \int_{V_2} dp_2 \, dq_2 \, e^{-\beta \mathcal{H}(p_2,q_2,N_2)} \end{aligned} \tag{3.141}$$

Note that the distribution of the particles into the two subsystems is not fixed yielding the combinatorial factor of $N!/N_1!N_2!$ (number of configurations with fixed N_1 and N_2 by permuting the particles in each subsystem). From this we define the probability $\rho(p_1, q_1, N_1)$ that we can find N_1 particles in the volume V_1 at the space coordinates (p_1, q_1),

$$\rho(p_1, q_1, N_1) = \frac{e^{-\beta \mathcal{H}(p_1,q_1,N_1)}}{Z_N N_1! N_2!} \int_{V_2} dp_2 \, dq_2 \, e^{-\beta \mathcal{H}(p_2,q_2,N_2)} \tag{3.142}$$

which is renormalized as

$$\sum_{N_1=0}^{N} \int_{V_1} dp_1 dq_1 \rho(p_1, q_1, N_1) = 1 \tag{3.143}$$

We may write

$$\rho(p_1, q_1, N_1) = \frac{Z_{N_2}}{Z_N} \frac{1}{N_1!} e^{-\beta \mathcal{H}(p_1, q_1, N_1)} , \tag{3.144}$$

where we now use the relation

$$\frac{Z_{N_2}(V_2, T)}{Z_N(V, T)} = e^{-\beta\{F(T, V-V_1, N-N_1) - F(T, V, N)\}} \tag{3.145}$$

with

$$F(T, V - V_1, N - N_1) - F(T, V, N) \approx -\left(\frac{\partial F}{\partial V}\right)_{T,N} V_1 - \left(\frac{\partial F}{\partial N}\right)_{T,V} N_1 = -\mu N_1 + pV_1 . \tag{3.146}$$

Thus we define

$$z = e^{\beta\mu} \tag{3.147}$$

which we call *fugacity*, and we write

$$\rho(p, q, N) = \frac{z^N}{N!} e^{-\beta\{pV + \mathcal{H}(p,q,N)\}} . \tag{3.148}$$

μ is the chemical potential as introduced earlier. We now introduce the grand partition function

$$\mathcal{Z}(T, V, z) = \sum_{N=0}^{\infty} z^N Z_N(V, T) , \tag{3.149}$$

which incorporates all important information of a system of fixed volume, temperature and chemical potential.

3.5.1 Relation to thermodynamics

We use now (3.148) and integrate both sides

$$1 = e^{-\beta pV} \sum_{N=0}^{\infty} \frac{z^N}{N!} \int dp\, dq\, e^{-\beta \mathcal{H}(p,q)} = e^{-\beta pV} \mathcal{Z}(T, V, z) \tag{3.150}$$

which leads to

$$\Omega(T, V, \mu) = -pV = -k_B T \ln \mathcal{Z}(T, V, z) , \tag{3.151}$$

the grand potential:

$$d\Omega = -SdT - pdV - Nd\mu . \tag{3.152}$$

The average value of N is then given by

$$\langle N \rangle = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = k_B T \frac{\partial}{\partial \mu} \ln \mathcal{Z} = z \frac{\partial}{\partial z} \ln \mathcal{Z} = \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} N z^N Z_N . \tag{3.153}$$

It is also convenient to derive again the internal energy

$$U = -\frac{\partial}{\partial \beta} \ln \mathcal{Z} \qquad \Rightarrow \qquad C_V = \left(\frac{\partial U}{\partial T}\right)_{V,\mu} . \tag{3.154}$$

We consider now the fluctuations of the particle number. The following relation holds

$$\langle N^2\rangle - \langle N\rangle^2 = z\frac{\partial}{\partial z}z\frac{\partial}{\partial z}\ln \mathcal{Z} = z\frac{\partial}{\partial z}\frac{1}{\mathcal{Z}}\sum_N N z^N Z_N = \frac{1}{\mathcal{Z}}\sum_N N^2 z^N Z_N - \left\{\frac{1}{\mathcal{Z}}\sum_N N z^N Z_N\right\}^2 \tag{3.155}$$

from which we derive

$$\langle N^2\rangle - \langle N\rangle^2 = k_B T\frac{\partial^2}{\partial\mu^2}\Omega(T,V,\mu) = k_B T\frac{\partial\langle N\rangle}{\partial\mu} \ . \tag{3.156}$$

We now relate the right hand side with the isothermal compressibility. Introducing the specific volume $v = V/N$ we consider $\mu = \mu(v,T)$ (note that μ is not extensive). Moreover we use the Gibbs-Duhem relation (1.72),

$$d\mu = v\ dp - \frac{S}{N}\ dT \qquad \Rightarrow \qquad \left(\frac{\partial\mu}{\partial v}\right)_T = v\left(\frac{\partial p}{\partial v}\right)_T \ . \tag{3.157}$$

Since v depends on both V and N, the following derivatives are useful

$$\begin{aligned}\left(\frac{\partial}{\partial v}\right)_{V,T} &= \left(\frac{\partial N}{\partial v}\right)_{V,T}\left(\frac{\partial}{\partial N}\right)_{V,T} = -\frac{N^2}{V}\left(\frac{\partial}{\partial N}\right)_{V,T} \ , \\ \left(\frac{\partial}{\partial v}\right)_{N,T} &= \left(\frac{\partial V}{\partial v}\right)_{N,T}\left(\frac{\partial}{\partial V}\right)_{N,T} = N\left(\frac{\partial}{\partial V}\right)_{N,T} \ .\end{aligned} \tag{3.158}$$

From (3.157) then conclude

$$-\frac{N^2}{V}\left(\frac{\partial\mu}{\partial N}\right)_{V,T} = V\left(\frac{\partial p}{\partial V}\right)_{N,T} \qquad \Rightarrow \qquad \frac{1}{N}\left(\frac{\partial N}{\partial\mu}\right)_{V,T} = -\frac{1}{vV}\left(\frac{\partial V}{\partial p}\right)_{N,T} = \frac{\kappa_T}{v} \ . \tag{3.159}$$

Consequently the fluctuations of the particle number is connected with the response function, κ_T, the isothermal compressibility:

$$\kappa_T = \frac{v}{\langle N\rangle k_B T}\left\{\langle N^2\rangle - \langle N\rangle^2\right\} \ . \tag{3.160}$$

3.5.2 Ideal gas

For the ideal gas it is easy to calculated the grand partition function,

$$\mathcal{Z}(T,V,z) = \sum_{N=0}^{\infty} z^N Z_N(T,V) = \sum_{N=0}^{\infty}\frac{z^N}{N!}V^N(2\pi m k_B T)^{3N/2} = \exp\left\{zV(2\pi m k_B T)^{3/2}\right\} \tag{3.161}$$

We can also derive the probability P_N of finding the system with N particles. The average value is given by

$$\langle N\rangle = z\frac{\partial}{\partial z}zV(2\pi m k_B T)^{3/2} = zV(2\pi m k_B T)^{3/2} \tag{3.162}$$

From this we conclude

$$P_N = e^{-\langle N\rangle}\frac{\langle N\rangle^N}{N!} \tag{3.163}$$

which is strongly peaked at $N = \langle N\rangle$. The fluctuations are given by

$$\langle N^2\rangle - \langle N\rangle^2 = z\frac{\partial\langle N\rangle}{\partial z} = \langle N\rangle \qquad \Rightarrow \qquad \kappa_T = \frac{v}{k_B T} = \frac{1}{p} \tag{3.164}$$

The grand potential is given by

$$\Omega(T,V,\mu) = -k_B T e^{\beta\mu}V(2\pi m k_B T)^{3/2} = -k_B T\langle N\rangle = -pV \tag{3.165}$$

The chemical potential is obtained as

$$\mu = k_B T\ \ln\left(\frac{\langle N\rangle(2\pi m k_B T)^{-3/2}}{V}\right) \ . \tag{3.166}$$

3.5.3 Chemical potential in an external field

In order to get a better understanding of the role of the chemical potential we now consider an ideal gas in the gravitational field, i.e. the particles are subject to the potential $\phi(h) = mgh$. We take the chemical potential μ_0 as a constant. Then we write for the chemical potential,

$$\mu = \mu_0 + mgh + k_B T \ \ln\left\{n(2\pi m k_B T)^{-3/2}\right\} \tag{3.167}$$

where we define $n = \langle N \rangle / V$ as the local number density of particles. In equilibrium the temperature and the chemical potential shall be constant. We may determine μ by the condition that the at $h = 0$ the density is $n = n_0$,

$$\mu(T) = \mu_0 + k_B T \ \ln\left\{n_0(2\pi m k_B T)^{-3/2}\right\} \quad \Rightarrow \quad mgh = k_B T \ \ln\left(\frac{n_0}{n}\right) \tag{3.168}$$

We can now solve this equation for $n = n(h)$:

$$n(h) = n_0 e^{-\beta mgh} \ . \tag{3.169}$$

and with the (local) equation of state

$$p(h) = n(h) k_B T = p_0 e^{-\beta mgh} \tag{3.170}$$

This is the famous barometer formula.

3.5.4 Paramagnetic ideal gas

As a further example we consider an ideal gas of particles with a magnetic moment which only can point in two direction, parallel or antiparallel to an applied magnetic field. Thus, the Hamiltonian is given by

$$\mathcal{H}(p, q, H) = \sum_i \frac{\vec{p}_i^{\,2}}{2m} - \sum_i m_i H \ , \tag{3.171}$$

where $m_i = \pm\tilde{m}$. Having now an ideal gas of two types of particles, moment parallel (+) and antiparallel (−) to H, we treat this problem in the grand canonical ensemble language. We introduce the corresponding particle numbers N_+ and N_-. The grand canonical partition function is easily calculated as

$$\begin{aligned} \mathcal{Z} &= \sum_N \frac{1}{N!} \sum_{N_+, N_-}^{N_+ + N_- = N} \frac{N!}{N_+! \, N_-!} \left\{V a z_+ e^{\beta H \tilde{m}}\right\}^{N_+} \left\{V a z_- e^{-\beta H \tilde{m}}\right\}^{N_-} \\ &= \sum_{N_+} \frac{1}{N_+!} \left\{V a z_+ e^{\beta H \tilde{m}}\right\}^{N_+} \sum_{N_-} \frac{1}{N_-!} \left\{V a z_- e^{-\beta H \tilde{m}}\right\}^{N_-} \\ &= \exp\left\{z_+ e^{\beta \tilde{m} H} V a\right\} \exp\left\{z_- e^{-\beta \tilde{m} H} V a\right\} \end{aligned} \tag{3.172}$$

with $a = (2\pi m k_B T)^{3/2}$. The average numbers of up and down magnetic moments are then

$$N_\pm = z_\pm \frac{\partial}{\partial z_\pm} \ln \mathcal{Z} = z_\pm e^{\pm \beta \tilde{m} H} V a \tag{3.173}$$

From this we derive the chemical potential, $z_\pm = e^{\beta \mu_\pm}$,

$$\mu_\pm = k_B T \ln\left(\frac{n_\pm}{a}\right) \pm \tilde{m} H \ , \tag{3.174}$$

where $n_\pm = N_\pm/V$. In equilibrium $\mu_+ = \mu_- = \mu$ otherwise the system would shift the distribution of moments, $N_+ \leftrightarrow N_-$. This leads to the ratio

$$\frac{N_+}{N_-} = \frac{n_+}{n_-} = e^{2\beta\tilde{m}H} \; . \tag{3.175}$$

The partial pressures defined by

$$p_\pm = \frac{N_\pm k_B T}{V} \qquad \Rightarrow \qquad p = p_+ + p_- \tag{3.176}$$

add up to the total pressure. The number of particles and the magnetization are given by

$$N = N_+ + N_- = 2Vae^{\beta\mu}\cosh(\beta\tilde{m}H) \qquad \text{and} \qquad M = \tilde{m}(N_+ - N_-) = 2\tilde{m}Vae^{\beta\mu}\sinh(\beta\tilde{m}H) \tag{3.177}$$

Under the condition that there is an external reservoir of particles at the chemical potential μ, we can define the magnetization per particle in the system,

$$\frac{M}{N} = \tilde{m}\tanh(\beta\tilde{m}H) \; , \tag{3.178}$$

as we had obtained earlier for the ideal paramagnet.

Chapter 4

Quantum Statistical Physics

Quantum statistical physics opens the door to new phenomena and also provides an understanding of the third law of thermodynamics which we had found to be often violated within classical statistical physics. Moreover, we will find that for ideal quantum gases the concept of indistinguishable particles leads to the classification of particles into two categories: Bosons and Fermions.

4.1 Basis of quantum statistics

Every state of a quantum mechanical system can be described by a superposition of stationary states $|\psi_n\rangle$ of the Hamiltonian $\mathcal{H}$, which form a complete orthonormalized basis,

$$|\Psi\rangle = \sum_n c_n |\psi_n\rangle \tag{4.1}$$

where the complex coefficients c_n represent the wavefunction whose time dependence is determined by the Hamiltonian $\mathcal{H}$. For the stationary states the following holds:

$$\mathcal{H}|\psi_n\rangle = \varepsilon_n |\psi_n\rangle \qquad \text{with} \quad \langle \psi_n | \psi_{n'} \rangle = \delta_{nn'} \tag{4.2}$$

The renormalization of the state $|\Psi\rangle$ requires,

$$1 = \langle \Psi | \Psi \rangle = \sum_{n,n'} c_n^* c_{n'} \langle \psi_n | \psi_{n'} \rangle = \sum_n |c_n|^2 \ . \tag{4.3}$$

The square modulus $|c_n|^2$ denotes the probability to find the state $|\psi_n\rangle$ in $|\Psi\rangle$.
Observable quantities are represented by Hermitian operators $\mathcal{A}$ and their expectation value is

$$\langle \mathcal{A} \rangle = \langle \Psi | \mathcal{A} | \Psi \rangle = \sum_{n,n'} c_n^* c_m \langle \psi_n | \mathcal{A} | \psi_n' \rangle \ . \tag{4.4}$$

This expectation value is generally time dependent: $c_n = c_n(t)$.
Considering now a measurement of a macroscopic observable, which corresponds to an average (indicated here by over-bar) over a time much longer than the microscopic time scales (time between collisions, inverse energy spacing, etc.).

$$\overline{\langle \mathcal{A} \rangle} = \sum_{n,n'} \overline{c_n^* c_{n'}} \langle \psi_n | \mathcal{A} | \psi_{n'} \rangle \tag{4.5}$$

Let us now restrict to eigenstates in a certain narrow energy range: $E \leq \epsilon_n \leq E + \Delta$ (microcanonical point of view). The quantum statistical physics is based on two key postulates:

- Equal probability:

$$\overline{c_n^* c_n} = \begin{cases} r\ , & E \leq \epsilon_n \leq E + \Delta\ , \\ \\ 0\ , & \text{otherwise}\ . \end{cases} \tag{4.6}$$

  where r is a real constant.

- Random phase: if $n \neq n'$ then

$$\overline{c_n^* c_{n'}} = 0\ , \tag{4.7}$$

  i.e. the phase of the wavefunction is a random variable. Such a behavior requires an at least weak coupling to the environment in order to randomize the phases.

As a consequence average expectation values are given by

$$\overline{\langle \mathcal{A} \rangle} = \sum_n |c_n|^2 \langle \psi_n | \mathcal{A} | \psi_n \rangle \tag{4.8}$$

which corresponds to considering the state $|\Psi\rangle$ as an incoherent superposition of the eigenstates (pure versus mixed state). We may consider the drop out of interference terms as an effect of averaging over a long time whereby the phases depend on time. Alternatively, we average over an ensemble of many different states which different phases.

4.2 Density matrix

Mixed states or incoherent superpositions of states are conveniently described by density matrices $\hat{\rho}$. We define first the density matrix as an operator through its matrix elements

$$\langle \psi_n | \hat{\rho} | \psi_n' \rangle = |c_n|^2 \delta_{nn'} \tag{4.9}$$

In the basis of stationary states the density matrix is diagonal. Therefore we may formally write the density matrix in spectral form as

$$\hat{\rho} = \sum_n |\psi_n\rangle |c_n|^2 \langle \psi_n|\ . \tag{4.10}$$

We may also express the average value of $\mathcal{A}$ as a trace of the form,

$$\langle \mathcal{A} \rangle = \frac{\sum_n \langle \psi_n | \mathcal{A} \hat{\rho} | \psi_n \rangle}{\sum_n \langle \psi_n | \hat{\rho} | \psi_n \rangle} = \frac{tr(\mathcal{A}\hat{\rho})}{tr\hat{\rho}} \tag{4.11}$$

Note that the final expression is independent of the basis $\{\psi_n\rangle\}$, as known from standard linear algebra. The cyclic permutation of matrices (operators) does not change the trace: $tr(ABC) = tr(BCA)$. Thus the basis transformation U yields $A' = UAU^{-1}$, such that

$$tr(A') = tr(UAU^{-1}) = tr(AU^{-1}U) = tr(A)\ . \tag{4.12}$$

In this sense the density matrix is a convenient tool to perform ensemble averages as it contains basis-free informations. In order to describe equilibrium properties $\hat{\rho}$ should not depend on time t and commutes with the Hamiltonian:

$$i\hbar \frac{\partial \hat{\rho}}{\partial t} = [\mathcal{H}, \hat{\rho}] = 0\ . \tag{4.13}$$

It is obvious that the density matrix defined in (4.10) commutes with $\mathcal{H}$. Compare this with Liouville's theorem of classical mechanics.

4.3 Ensembles in quantum statistics

We now formulate the quantum statistical physics in terms of the three types of ensembles which we introduced for the classical statistical physics.

4.3.1 Microcanonical ensemble

The microcanonical ensemble is suitable for closed systems which do not exchange energy and particles with the environment. We take the density matrix in its diagonal form $\rho_{nn'} = \delta_{nn'}|c_n|^2$. In this ensemble we consider an incoherent state which consist of the equally distributed set of states within a certain narrow energy range between E and $E + \delta E$. Thus,

$$|c_n|^2 = \begin{cases} 1 & E \leq \epsilon_n \leq E + \delta E \\ \\ 0 & \text{otherwise} \end{cases} \tag{4.14}$$

The energies ϵ_n are the eigenenergies of the stationary states $|\psi_n\rangle$ for the Hamiltonian $\mathcal{H}$. Formally we then write the density tensor as

$$\hat{\rho} = \sum_{E \leq \epsilon_n \leq E+\delta E} |\psi_n\rangle\langle\psi_n| \qquad \text{with} \qquad tr\hat{\rho} = \sum_n \rho_{nn} = \omega(E) \tag{4.15}$$

Here $\omega(E)$ is the number of quantum states $|\psi_n\rangle$ with energy in the given energy range. Analogous to the classical case we use now $\omega(E)$ to define the entropy

$$S(E, V) = k_B \ln \omega(E) \; . \tag{4.16}$$

From this we may derive the thermodynamics. Note that $\omega(E)$ is connected with the density of states $d\Phi(E)/dE$ of the quantum system.

$$\Phi(E) = \sum_n \Theta(E - \epsilon_n) \qquad \Rightarrow \qquad \omega(E) = \frac{d\Phi(E)}{dE}\delta E \; . \tag{4.17}$$

with $\Theta(x)$ as the step function

$$\Theta(x) = \begin{cases} 1 & x > 0 \\ 0 & x \leq 0 \end{cases} \; . \tag{4.18}$$

4.3.2 Canonical ensemble

For a system which is in thermal equilibrium with a reservoir of given temperature T we use the canoncial ensemble formulation. Analogous to the classical statistical physics we define the density matrix here as

$$\rho(p, q) = e^{-\beta\mathcal{H}(p,q)} \qquad \rightarrow \qquad \rho_{nn'} = \delta_{nn'}e^{-\beta\epsilon_n} \; . \tag{4.19}$$

and we define the partition function as

$$Z = \sum_n e^{-\beta\epsilon_n} = tr\hat{\rho} = e^{-\beta F(T,V,N)} \tag{4.20}$$

where F is the Helmholtz free energy. The density matrix can be written as an operator,

$$\hat{\rho} = \sum_n |\psi_n\rangle e^{-\beta\epsilon_n}\langle\psi_n| = e^{-\beta\mathcal{H}} \underbrace{\sum_n |\psi_n\rangle\langle\psi_n|}_{=1} = e^{-\beta\mathcal{H}} \; . \tag{4.21}$$

Thus the partition function can also be expressed as

$$Z = tre^{-\beta\mathcal{H}} \tag{4.22}$$

and average values of observables are given by

$$\langle \mathcal{A} \rangle = \frac{tr(\mathcal{A} e^{-\beta\mathcal{H}})}{tr e^{-\beta\mathcal{H}}} = \frac{1}{Z} tr \left(\mathcal{A} e^{-\beta\mathcal{H}} \right) . \tag{4.23}$$

The connection to thermodynamics is given via the Helmholtz free energy,

$$F(T,V,N) = -k_B T \ln Z . \tag{4.24}$$

4.3.3 Grandcanonical ensemble

Now we connect the system in addition to the heat reservoir also to the particle reservoir of given chemical potential μ. In the same spirit as for the canonical ensemble we use the analogy to the classical systems. The density matrix now reads,

$$\hat{\rho} = e^{-\beta(\mathcal{H} - \mu \hat{N})} \tag{4.25}$$

with N the particle number operator and the grandcanonical partition function is then

$$\mathcal{Z} = tr e^{-\beta(\mathcal{H} - \mu N)} = \sum_N z^N Z_N \tag{4.26}$$

with $z = e^{\beta\mu}$ as the fugacity, leading to the grand potential

$$\Omega(T,V,\mu) = -k_B T \ln \mathcal{Z} . \tag{4.27}$$

4.3.4 Ideal gas in grandcanonical ensemble

A gas of independent particles in quantum mechanics has the additional key feature that indistinguishable particles are categorized in Fermions (particles with half-integer spins) and Bosons (particles with integer spins). The former follow the Pauli principle: we cannot find two Fermions in the same quantum state. For Bosons no such restriction exists. In terms of the many-particle wave function the Fermions are described by a completely antisymmetric and the symmetric wavefunction under exchange of particle variables. The free particles in quantum mechanics are represented by plane waves

$$|\psi_{\vec{p}}\rangle = \frac{1}{\sqrt{V}} e^{i\vec{p}\cdot\vec{r}/\hbar} \qquad \text{with} \quad \epsilon_{\vec{p}} = \frac{\vec{p}^2}{2m} . \tag{4.28}$$

The grand partition function for a gas of such particles is given by

$$\mathcal{Z} = \sum_{\{n_{\vec{p}}\}} g_{\{n_{\vec{p}}\}} e^{-\beta(E-\mu N)_{\{n_{\vec{p}}\}}} , \tag{4.29}$$

with

$$E = \sum_{\vec{p}} \epsilon_{\vec{p}} n_{\vec{p}} \qquad \text{and} \qquad N = \sum_{\vec{p}} n_{\vec{p}} . \tag{4.30}$$

Here $n_{\vec{p}}$ is the number of particles per state, the occupation number, whereby the sum $\sum_{\{n_{\vec{p}}\}}$ runs over all allowed configurations of occupations. The factor $g_{\{n_{\vec{p}}\}}$ is given by

$$g_{\{n_{\vec{p}}\}} = \begin{cases} 1 & \text{indistinguishable particles (Fermions and Bosons)} \;, \\[2ex] \displaystyle\prod_{\vec{p}} \frac{1}{n_{\vec{p}}!} & \text{classical particles (Boltzmann)} \;. \end{cases} \tag{4.31}$$

For Fermions we find that for each state $|\psi_{\vec{p}}\rangle$ the occupation number is $n_{\vec{p}} = 0, 1$ and for Bosons $n_{\vec{p}} = 0, 1, 2, \ldots$. Thus, calculating the partition function we obtain for Fermions/Bosons,

$$\begin{aligned}
\mathcal{Z} &= \sum_{n_{\vec{p}_1}, n_{\vec{p}_2}, \ldots} \left[\left\{ z e^{-\beta\epsilon_{\vec{p}_1}} \right\}^{n_{\vec{p}_1}} \left\{ z e^{-\beta\epsilon_{\vec{p}_2}} \right\}^{n_{\vec{p}_2}} \cdots \right] \\
&= \prod_{\vec{p}} \sum_{n_{\vec{p}}} \left(z e^{-\beta\epsilon_{\vec{p}}} \right)^{n_{\vec{p}}} = \begin{cases} \prod_{\vec{p}} \left(1 + z e^{-\beta\epsilon_{\vec{p}}} \right) & \text{Fermions} \\ \prod_{\vec{p}} \dfrac{1}{1 - z e^{-\beta\epsilon_{\vec{p}}}} & \text{Bosons} \end{cases}
\end{aligned} \tag{4.32}$$

with $z = e^{\beta\mu}$ as the fugacity.
From the partition function we arrive at the equation of state

$$\frac{pV}{k_B T} = -\beta\Omega(T, V, \mu) = \begin{cases} \sum_{\vec{p}} \ln\left(1 + z e^{-\beta\epsilon_{\vec{p}}} \right) & \text{Fermions ,} \\ -\sum_{\vec{p}} \ln\left(1 - z e^{-\beta\epsilon_{\vec{p}}} \right) & \text{Bosons ,} \end{cases} \tag{4.33}$$

and the particle number

$$N = z \frac{\partial}{\partial z} \ln \mathcal{Z} = \begin{cases} \sum_{\vec{p}} \dfrac{1}{e^{\beta\epsilon_{\vec{p}}} z^{-1} + 1} & \text{Fermions ,} \\ \sum_{\vec{p}} \dfrac{1}{e^{\beta\epsilon_{\vec{p}}} z^{-1} - 1} & \text{Bosons .} \end{cases} \tag{4.34}$$

These equations can be used to determine the chemical potential, if we fix N. The occupation number of a state is also easily obtained,

$$\begin{aligned}
\langle n_{\vec{p}} \rangle &= \frac{1}{\mathcal{Z}} \sum_{N=0}^{\infty} z^N \sum_{\{n_{\vec{p}}\}}^{\sum_{\vec{p}} n_{\vec{p}} = N} n_{\vec{p}} e^{-\beta \sum_{\vec{p}} \epsilon_{\vec{p}} n_{\vec{p}}} = -k_B T \frac{\partial}{\partial \epsilon_{\vec{p}}} \ln \mathcal{Z} \\
&= \begin{cases} \dfrac{1}{z^{-1} e^{\beta\epsilon_{\vec{p}}} + 1} & \text{Fermions ,} \\ \dfrac{1}{z^{-1} e^{\beta\epsilon_{\vec{p}}} - 1} & \text{Bosons .} \end{cases}
\end{aligned} \tag{4.35}$$

These correspond to the Fermi-Dirac and the Bose-Einstein distributions, respectively.
Eventually we compare the grand potential and the internal energy. Using (4.33) we obtain

$$\begin{aligned}
\Omega &= -k_B T \sum_{\vec{p}} \ln\left(1 \pm z e^{\beta\epsilon_{\vec{p}}} \right) = -k_B T \int d\epsilon \; g(\epsilon) \; \ln\left(1 \pm z e^{-\beta\epsilon} \right) \\
&= -V k_B T \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\infty} d\epsilon \; \epsilon^{1/2} \ln(1 \pm z e^{-\beta\epsilon}) \\
&= -\frac{2}{3} V \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^{\infty} d\epsilon \frac{\epsilon^{3/2}}{z^{-1} e^{\beta\epsilon} \pm 1} \; ,
\end{aligned} \tag{4.36}$$

where the plus (minus) sign stands for the Fermions (Bosons). For the second equality we performed an integration by parts. We have also introduced the density of states

$$g(\epsilon) = \sum_{\vec{p}} \delta(\epsilon - \epsilon_{\vec{p}}) = V \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \epsilon^{1/2} \, , \tag{4.37}$$

for the energy integration. Now consider the internal energy

$$U = \sum_{\vec{p}} \langle n_{\vec{p}} \rangle \epsilon_{\vec{p}} = V \frac{1}{4\pi^2} \left(\frac{2m}{\hbar^2} \right)^{3/2} \int_0^\infty d\epsilon \, \frac{\epsilon^{3/2}}{z^{-1} e^{\beta\epsilon} \pm 1} \, . \tag{4.38}$$

Thus from $\Omega = -pV$ we obtain the general relation

$$U = \frac{3}{2} pV \, . \tag{4.39}$$

Note that we did not so far include the spin s of the particles in our considerations. This gives an additional factor $2s+1$ to Ω and U.

4.4 Properties of Fermi gas

We consider now Fermions of spin s, i.e. there are $2s+1$ different species. The equation of state and the equation for the particle number of the Fermi gas are given by

$$\begin{aligned} \frac{p}{k_B T} &= \frac{4\pi}{h^3} (2s+1) \int_0^\infty dp \, p^2 \ln\left(1 + z e^{-\beta \epsilon_{\vec{p}}}\right) \, , \\ \frac{1}{v} = \frac{N}{V} &= \frac{4\pi}{h^3} (2s+1) \int_0^\infty dp \, p^2 \frac{1}{z^{-1} e^{\beta\epsilon_{\vec{p}}} + 1} \, . \end{aligned} \tag{4.40}$$

where we used the relation $\sum_{\vec{p}} = \frac{V}{h^3} \int d^3p$. We rewrite these equations introducing special functions of z,

$$\begin{aligned} \frac{p}{k_B T} &= \frac{2s+1}{\lambda^3} f_{5/2}(z) \\ \frac{1}{v} &= \frac{2s+1}{\lambda^3} f_{3/2}(z) \end{aligned} \tag{4.41}$$

where $\lambda = h/\sqrt{2\pi m k_B T}$ is the thermal wavelength. These functions are defined as[1]

$$\begin{aligned} f_{5/2}(z) &= \frac{4}{\sqrt{\pi}} \int_0^\infty dx \, x^2 \ln(1 + z e^{-x^2}) = -\sum_{l=1}^\infty (-1)^l \frac{z^l}{l^{5/2}} \\ f_{3/2}(z) &= z \frac{\partial}{\partial z} f_{5/2}(z) = -\sum_{l=1}^\infty (-1)^l \frac{z^l}{l^{3/2}} \end{aligned} \tag{4.43}$$

---

[1] Expansion of $f_{5/2}(z)$:

$$\frac{4}{\sqrt{\pi}} \int_0^\infty dx \, x^2 \ln(1 + z e^{-x^2}) = -\frac{4}{\sqrt{\pi}} \int_0^\infty dx \, \sum_{l=1}^\infty x^2 (-1)^i \frac{z^l e^{-l x^2}}{l} = -\sum_{l=1}^\infty (-1)^l \frac{z^l}{l^{5/2}} \, . \tag{4.42}$$

For the following it will also be important to consider the asymptotic behavior of these functions. For $z \ll 1$ we may use the expansion just given. For $z \gg 1$ we find[2]

$$\begin{aligned} f_{5/2}(z) &\approx \frac{8}{15\sqrt{\pi}}(\ln z)^{5/2}\left[1+\frac{5\pi^2}{8(\ln z)^2}+\cdots\right] \\ f_{3/2}(z) &= \frac{4}{3\sqrt{\pi}}(\ln z)^{3/2}\left[1+\frac{\pi^2}{8(\ln z)^2}+\cdots\right] \end{aligned} \tag{4.47}$$

4.4.1 High-temperature and low-density limit

Low density (high temperature) implies $\lambda^3 \ll v = V/N$ and $z = e^{\beta\mu} \ll 1$. We can use now (4.41)

$$\frac{\lambda^3}{v} = (2s+1)\left\{z - \frac{z^2}{2^{3/2}} + \cdots\right\} \qquad \Rightarrow \qquad z = \frac{\lambda^3}{v}\frac{1}{2s+1} + \frac{1}{2^{3/2}(2s+1)^2}\left(\frac{\lambda^3}{v}\right)^2 + \cdots \ . \tag{4.48}$$

This allows us to replace z in the equation of state,

$$\frac{p}{k_BT} \approx \frac{2s+1}{\lambda^3}\left\{z - \frac{z^2}{2^{5/2}}\right\} \approx \frac{N}{V}\left\{1 + \frac{N}{2^{5/2}}\frac{\lambda^3}{V}\right\} \tag{4.49}$$

The second term represents the quantum correction to the classical limit of the ideal gas. This allows us also to calculate the isothermal compressibility,

$$\kappa_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_{T,N} = \frac{V}{Nk_BT}\frac{1}{1+\frac{\lambda^3 N}{2^{3/2}V}} \tag{4.50}$$

The quantum correction decreases the compressibility, as a consequence of Pauli's exclusion principle, i.e. Fermions avoid each other. The occupation numbers is approximatively given by

$$\langle n_{\vec{p}} \rangle \approx \frac{\lambda^3}{v}\frac{1}{2s+1}e^{-\beta\epsilon_{\vec{p}}} = \frac{Nh^3}{V(2\pi m k_BT)^{3/2}}e^{-\beta\epsilon_{\vec{p}}} \tag{4.51}$$

corresponding to the Maxwell-Boltzmann distribution. With the general relation (4.39) we obtain immediately the internal energy and the specific heat

$$U \approx \frac{3}{2}Nk_BT\left\{1 + \frac{N}{2^{5/2}}\frac{\lambda^3}{V}\right\} \quad \text{and} \quad C_V = \frac{3}{2}Nk_B\left\{1 - \frac{N}{2^{7/2}}\frac{\lambda^3}{V}\right\} , \tag{4.52}$$

including the first quantum corrections.

---

[2] Derivation of the large-z-expansion: Use $\nu = \mu/k_BT = \ln z$ and rewrite

$$f_{5/2}(z) = \frac{2}{\sqrt{\pi}}\int_0^\infty dy\, y^{1/2}\ln(1+e^{\nu-y}) = \frac{8}{15\sqrt{\pi}}\int_{-\nu}^\infty dy'(y'+\nu)^{5/2}\frac{e^{y'}}{(1+e^{y'})^2} , \tag{4.44}$$

where the second equation is obtained by twice integrating by parts. More over we extend the lower integration boundary to $-\infty$, at the expense of a small error of order $e^{-\nu} = 1/z$. We now expand $(y+\nu)^{5/2}$ assuming ν large and find

$$\begin{aligned} f_{5/2}(z) &= \frac{8}{15\sqrt{\pi}}\int_{-\infty}^{+\infty} dy'\left[\nu^{5/2} + \frac{5}{2}\nu^{3/2}y' + \frac{15}{8}\nu^{1/2}y^2 + \cdots\right]\frac{e^{y'}}{(1+e^{y'})^2} \\ &= \frac{8}{15\sqrt{\pi}}\left[I_0\nu^{5/2} + \frac{5}{2}I_1\nu^{3/2} + \frac{15}{8}I_2\nu^{1/2} + \cdots\right] \end{aligned} \tag{4.45}$$

where

$$I_n = \int_{-\infty}^{+\infty} dy\frac{y^n e^y}{(1+e^y)^2} \qquad \Rightarrow \qquad I_0 = 1, I_1 = 0, I_2 = \frac{\pi^2}{3}, \ldots \ . \tag{4.46}$$

Note that all $I_n = 0$ for n odd. It is easy to obtain $f_{3/2}(z)$ by taking the derivative.

4.4.2 Low-temperature and high-density limit: degenerate Fermi gas

At low temperature we reach the "quantum limit" $\lambda^3 \gg v = V/N$, which implies large fugacity z. First we consider the zero-temperature situation. The occupation numbers follow a step function

$$\langle n_{\vec{p}} \rangle = \Theta(\epsilon_{\vec{p}} - \mu) = \begin{cases} 1\,, & \epsilon_{\vec{p}} < \mu\,, \\ & \\ 0\,, & \epsilon_{\vec{p}} > \mu\,. \end{cases} \qquad (4.53)$$

The Fermions occupy states within a sphere in momentum space, the Fermi sphere. The particle density n is

$$n = \frac{N}{V} = \frac{2s+1}{h^3} \int d^3p \langle n_{\vec{p}} \rangle = \frac{2s+1}{h^3} \frac{4\pi}{3} p_F^3 = \frac{(2s+1)k_F^3}{6\pi^2} \qquad (4.54)$$

where p_F is the Fermi momentum ($\epsilon_{\vec{p}_F} = \mu(T=0) = \epsilon_F$), isotropic, and $k_F = p_F/\hbar$ is the Fermi wavevector. The groundstate energy is

$$U_0 = \frac{2s+1}{h^3} V \int d^3p \epsilon_{\vec{p}} \langle n_{\vec{p}} \rangle = \frac{3}{5} N \epsilon_F \qquad (4.55)$$

where ϵ_F denotes the Fermi energy. The zero-point pressure is obtained through (4.33),

$$p_0 = \frac{2}{3}\frac{U}{V} = \frac{2}{5}\frac{N}{V}\epsilon_F\,. \qquad (4.56)$$

In contrast to the classical ideal gas, a Fermi gas has finite zero-point pressure which is again a consequence of the Pauli principle and is responsible for the stability of metals, neutron stars etc.

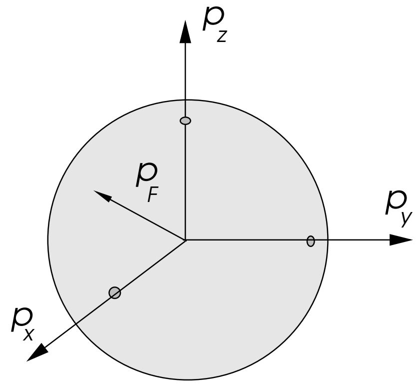


Fig.4.1: Fermi sphere of occupied single particle states. Fermi radius p_F.

Next we turn to finite temperatures for which the occupation number broadens the step at p_F. We use now (4.47, 4.54) to obtain the relation

$$\left(\frac{\epsilon_F}{k_B T}\right)^{3/2} = \frac{3}{4} \frac{\sqrt{\pi}\lambda^3}{2s+1} \frac{N}{V} = \left(\frac{\mu}{k_B T}\right)^{3/2} + \frac{\pi^2}{8} \left(\frac{\mu}{k_B T}\right)^{-1/2} + \cdots\,, \qquad (4.57)$$

which at constant density $n = N/V$ can be solved for the chemical potential,

$$\mu(T) = \epsilon_F \left(1 - \frac{\pi^2}{12} \left(\frac{k_B T}{\epsilon_F}\right)^2 + \cdots\right)\,, \qquad (4.58)$$

and in very much the same way we obtain for the pressure,

$$p(T) = p_0 \left(1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\epsilon_F}\right)^2 + \cdots\right)\,. \qquad (4.59)$$

Again we derive the internal energy from the relation (4.33)

$$U = \frac{3}{2}pV = \frac{3}{5}N\epsilon_F \left(1 + \frac{5\pi^2}{12} \left(\frac{k_B T}{\epsilon_F} \right)^2 + \cdots \right) , \tag{4.60}$$

which also leads to the specific heat for fixed N

$$C_N = k_B N \frac{\pi^2}{2} \frac{k_B T}{\epsilon_F} + \cdots . \tag{4.61}$$

This is the famous linear temperature dependence of the specific heat, which can be observed well for the electrons in simple metals. Obviously now the third law of thermodynamics is satisfied: $C_N \overset{T \to 0}{\to} 0$. Also the entropy goes to zero linearly in T.

4.5 Bose gas

There are two situations for Bosons: (1) a system with well-defined particle number, e.g. bosonic atoms, ^{4}He, ... ; (2) Bosons which results as modes of harmonic oscillators (no fixed particle number), e.g. photons, phonons, magnons, etc..

4.5.1 Bosonic atoms

We consider Bosons without spin ($S = 0$) for which ^{4}He is a good example. Analogous to the Fermions we introduce functions of z to express the equation of state and the particle number,

$$\begin{aligned} \frac{p}{k_B T} &= \frac{1}{\lambda^3} g_{5/2}(z) = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \frac{z^l}{l^{5/2}} \\ \frac{1}{v} = \frac{N}{V} &= \frac{1}{\lambda^3} g_{3/2}(z) = \frac{1}{\lambda^3} \sum_{l=1}^{\infty} \frac{z^l}{l^{3/2}} . \end{aligned} \tag{4.62}$$

For small z both functions grow linearly from zero and $g_{3/2}(z)$ has a divergent derivative for $z \to 1$. We concentrate on the range $0 < z \leq 1$, such that $\mu(T) \leq 0$. For $z = 1$ we obtain

$$g_{3/2}(1) = \sum_l \frac{1}{l^{3/2}} = \zeta(3/2) \approx 2.612 \qquad \text{and} \quad g_{5/2}(1) = \sum_l \frac{1}{l^{5/2}} = \zeta(5/2) \approx 1.342 \tag{4.63}$$

where $\zeta(x)$ is Riemann's ζ-function.

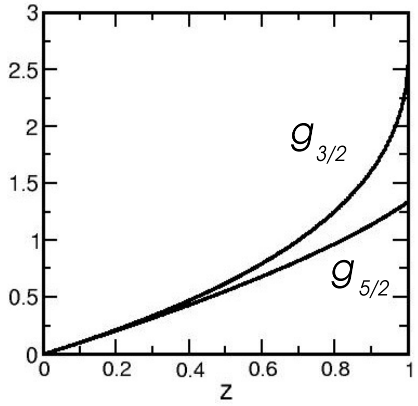


Fig.4.2: Functions $g_{3/2}(z)$ and $g_{5/2}(z)$.

4.5.2 High-temperature and low-density limit

It is easy to see that like the Fermions the Bosons behave in this limit as a classical ideal gas. An intriguing aspect occurs, however, in the quantum corrections. For the pressure we find

$$p(T) = \frac{N}{V} k_B T \left\{ 1 - \frac{N}{2^{5/2}} \frac{\lambda^3}{V} + \cdots \right\} . \tag{4.64}$$

The quantum correction reduces the classical ideal gas pressure and yields the compressibility

$$\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_{T,N} = \frac{V}{N k_B T} \frac{1}{1 - \frac{\lambda^3 N}{2^{3/2} V}} \tag{4.65}$$

In contrast to the Fermions where the quantum nature of the particles diminishes the compressibility, here the compressibility is enhanced. Actually, in this approximation the compressibility even diverges, if

$$2^{3/2} = \frac{N}{V} \lambda^3 \tag{4.66}$$

i.e. at low enough temperature or high enough density. We will now see that this indeed indicates an instability of the Bose gas.

4.5.3 Low-temperature and high-density limit: Bose-Einstein condensation

Let us now consider equation (4.62) carefully, The function $g_{3/2}(z)$ is monotonically increasing with z. If T is lowered, $\lambda \propto T^{-1/2}$ increases, such that z has to increase too in order to satisfy (4.62). Therefore μ approaches the singular point at 0 ($z = 1$). The critical point is determined by

$$g_{3/2}(1) = \zeta(3/2) = \frac{N}{V} \lambda^3 \qquad \Rightarrow \qquad \begin{cases} T_c = \dfrac{h^2}{2\pi k_B m [\zeta(3/2) V/N]^{2/3}} , \\ V_c = \dfrac{N}{\zeta(3/2)} \dfrac{h^3}{(2\pi m k_B T)^{3/2}} . \end{cases} \tag{4.67}$$

This defines a critical temperature T_c and critical volume V_c below which a new state of the Bose gas occurs. Note that this equation is qualitatively very similar to (4.66) and even quantitatively not so far ($\zeta(3/2) \approx 2.612 \leftrightarrow 2^{3/2} \approx 2.85$) . The question arises what happens for $T < T_c$ or $V < V_c$. Actually the problem occurring in (4.62) and (4.67) arises in the step

$$N = \sum_{\vec{p}} \frac{1}{e^{\beta(\epsilon_{\vec{p}} - \mu)} - 1} \rightarrow \frac{V}{h^3} \int d^3p \frac{1}{e^{\beta(\epsilon_{\vec{p}} - \mu)} - 1} . \tag{4.68}$$

The integral does not count the occupation of the state $\vec{p} = 0$, because the momentum distribution function entering the integral,

$$\rho(p) = \frac{p^2}{e^{\beta(\epsilon_{\vec{p}} - \mu)} - 1} \qquad \Rightarrow \qquad \rho(0) = 0 . \tag{4.69}$$

This is fine as long as the occupation of the $\vec{p} = 0$-state (single-particle groundstate) is vanishingly small compared to N. However, for $T < T_c$ ($V < V_c$) the occupation becomes macroscopic, $\langle n_{\vec{p}=0} \rangle / N > 0$ and we cannot neglect this contribution in the calculation of N. Thus the correct density is

$$\frac{N}{V} = \frac{1}{\lambda^3} g_{3/2}(z) + n_0(T) = n_n(T) + n_0(T) \tag{4.70}$$

with $n_0(T)$ denoting the density of Bosons in the single-particle groundstate ($\vec{p} = 0$). These particles form a condensate, the *Bose-Einstein condensate*. What happens at T_c is a phase transition. We encounter here a "two-fluid" system for which the total particle density split into a condensed fraction n_0 and a normal fraction n_n. From (4.70) we find the temperature dependence of n_0,

$$n_0(T) = \frac{N}{V}\left[1 - \left(\frac{T}{T_c}\right)^{3/2}\right] . \tag{4.71}$$

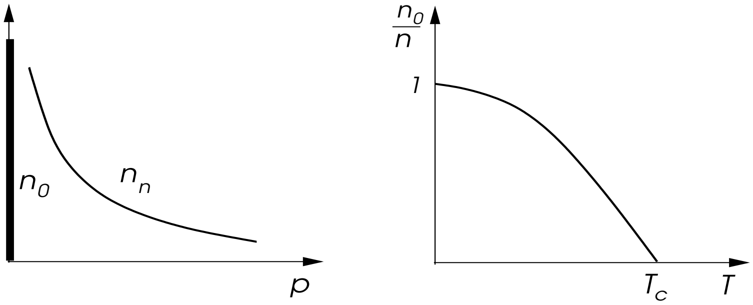


Fig.4.3: Occupation: Left panel: A macroscopic fraction of particle occupy the momentum $p = 0$-state for $T < T_c$. Right panel: Temperature dependence of the condensate fraction.

Next we also determine the equation of state,

$$p = \begin{cases} \dfrac{k_B T}{\lambda^3} g_{5/2}(z) , & V > V_c \\[2ex] \dfrac{k_B T}{\lambda^3} g_{5/2}(1) , & V < V_c \end{cases} \tag{4.72}$$

We now consider the compressibility for $V > V_c$. For this purpose we first determine

$$\frac{\partial V}{\partial z} = -N\lambda^3 \frac{g'_{3/2}(z)}{g_{3/2}(z)^2} , \tag{4.73}$$

and consider

$$\frac{\partial p}{\partial V} = \frac{k_B T}{\lambda^3} g'_{5/2}(z) \frac{\partial z}{\partial V} \qquad \Rightarrow \qquad \kappa_T = \frac{N\lambda^6}{k_B T g_{3/2}(z)^2} \frac{g'_{3/2}(z)}{g'_{5/2}(z)} . \tag{4.74}$$

where we use the notation $g'_n(z) = dg_n(z)/dz$. As anticipated earlier the compressibility diverges at the transition $V \to V_c$ (or $T \to T_c$), since the derivative $g'_{3/2}(z) \to \infty$ for $z \to 1$. In the condensed phase the pressure is independent of V as is obvious from (4.72). Therefore the condensed phase is infinitely compressible, i.e. it does not resist to compression.

Some further thermodynamic quantities can be derived. First we consider the entropy S from the grand canonical potential

$$S(T,V,\mu) = -\left(\frac{\partial \Omega}{\partial T}\right)_{V,\mu} = \left(\frac{\partial pV}{\partial T}\right)_{V,\mu} = \begin{cases} N k_B \left(\dfrac{5v}{2\lambda^3} g_{5/2}(z) - \ln z\right) , & T > T_c , \\[2ex] N k_B \dfrac{5}{2} \dfrac{g_{5/2}(1)}{g_{3/2}(1)} \left(\dfrac{T}{T_c}\right)^{3/2} , & T < T_c . \end{cases} \tag{4.75}$$

where we used (4.62).[3] For the specific heat at fixed particle number N we find from the internal energy $U = \frac{3}{2}pV$,

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{V,N} = \begin{cases} Nk_B\left(\frac{15v}{4\lambda^3}g_{5/2}(z) - \frac{9}{4}\frac{g_{3/2}(z)}{g_{1/2}(z)}\right) \ , & T > T_c \ , \\ Nk_B\frac{15}{4}\frac{g_{5/2}(1)}{g_{3/2}(1)}\left(\frac{T}{T_c}\right)^{3/2} \ , & T < T_c \ . \end{cases} \tag{4.79}$$

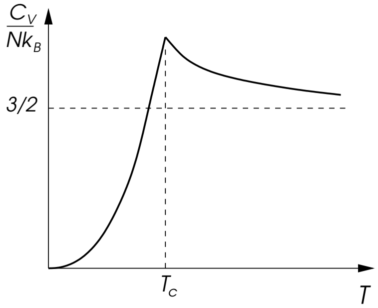


Fig.4.4: Specific heat: C_V has a cusp at the transition and vanishes as $T^{3/2}$ towards zero-temperature. In the high-temperature limit C_v approaches $3Nk_B/2$ which corresponds to the equipartition law of a mono-atomic gas.

In accordance with the third law of thermodynamics both the entropy and the specific heat go to zero in the zero-temperature limit. The entropy for $T < T_c$ can be viewed as

$$\frac{S}{N} = s\left(\frac{T}{T_c}\right)^{3/2} = \frac{n_n(T)}{n}s \qquad \text{with} \quad s = \frac{5}{2}k_B\frac{g_{5/2}(1)}{g_{3/2}(1)} \tag{4.80}$$

where s is the entropy per normal particle (specific entropy), i.e. a non-vanishing contribution to the entropy is only provided by the normal fraction (two-fluid model). The specific heat has a cusp at $T = T_c$.

---

[3] Calculation of the temperature derivatives: (1) Fixed chemical potential:

$$\frac{\partial}{\partial T}\frac{Vk_BT}{\lambda^3}g_{5/2}(z) = \frac{5Vk_B}{\lambda^3}g_{5/2}(z) + \underbrace{\frac{Vk_BT}{\lambda^3}\frac{g_{3/2}(z)}{z}\frac{\partial z}{\partial T}}_{= -k_B\frac{V}{\lambda^3}g_{3/2}(z)\beta\mu = -Nk_B\ln z} \tag{4.76}$$

where we used $\frac{\partial}{\partial z}g_{5/2} = g_{3/2}/z$.
(2) Fixed particle number: we use

$$g_{3/2}(z) = \frac{N}{V}\lambda^3 \quad \Rightarrow \quad \frac{dg_{3/2}}{dT} = \frac{g_{1/2}(z)}{z}\frac{dz}{dT} = -\frac{3}{2}\frac{N}{V}\frac{\lambda^3}{T} \tag{4.77}$$

which leads to the relation

$$\frac{dg_{5/2}}{dT} = \frac{g_{3/2}(z)}{z}\frac{dz}{dT} = -\frac{9}{4}\frac{g_{3/2}(z)}{g_{1/2}(z)}\frac{N\lambda^3}{V} \ . \tag{4.78}$$

This leads to the expression for the specific heat.

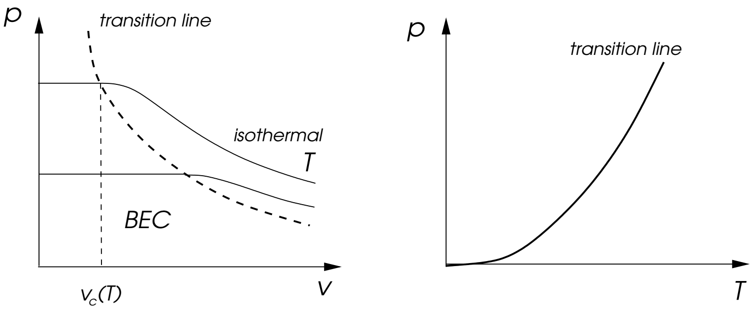


Fig.4.5: Phasediagrams; Left panel: p-v-diagram; the isothermal lines reach the transition line with zero-slope, i.e. the compressibility diverges. Right panel: p-T-diagram; the condensed phase corresponds to the transition line, there is no accessible space above the transition line.

We consider now the phase diagram of various state variable.
(1) p-v-diagram: phase transition line

$$p_0 v^{5/3} = \frac{h^2}{2\pi m} \frac{g_{5/2}(1)}{[g_{3/2}(1)]^{5/3}} \ . \tag{4.81}$$

(2) p-T-diagram: phase transition line

$$p_0 = \frac{k_B T}{\lambda^3} g_{5/2}(1) \propto T^{5/2} \tag{4.82}$$

which is the vapor pressure (constant for $T < T_c$). We use this line to determine the latent heat l per particle via the Clausius-Clapeyron relation,

$$\frac{dp_0}{dT} = \frac{l}{T\Delta v} \quad \text{with} \qquad l = T\Delta s \tag{4.83}$$

The condensate takes no specific volume compared to the normal fraction. Thus, $\Delta v = v_c$. Therefore we obtain

$$l = T v_c \frac{dp_0}{dT} = T v_c \frac{5}{2} \frac{k_B g_{5/2}(1)}{\lambda^3} = T \frac{5}{2} k_B \frac{g_{5/2}(1)}{g_{3/2}(1)} \tag{4.84}$$

where we used the relation $\lambda^3 = v_c g_{3/2}(1)$. Note that this is consistent with our result on the specific entropy s. The condensed phase has no specific entropy such that $\Delta s = s$ and, consequently, $l = T\Delta s$ using (4.80).

Examples of the Bose-Einstein condensates is the quantum fluid ^{4}He which shows a condensation below $T_c \approx 2.18K$ into a superfluid phase. The superfluid phase is characterized by a frictionless flow, which is a signature of the "rigidity" of the condensate.

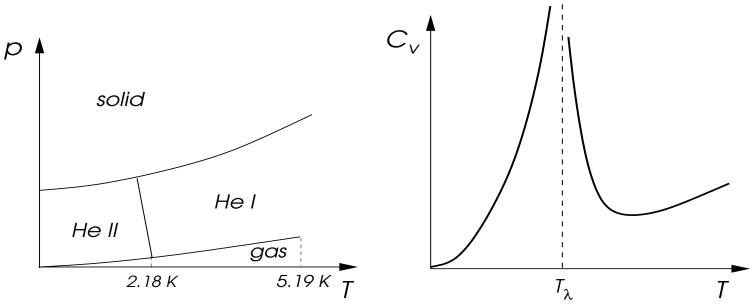


Fig.4.6: ^{4}He: Left panel: Schematic phase diagram below 2.18 K the superfluid phase HeII appears and the phase He I is a normal quantum fluid. At 5.19 K a critical end point is located. At high pressure Helium becomes a solid. Right panel: Specific heat shows a singularity at the transition point. Due to the characteristic shape this transition is called "λ-transition".

A further very modern example are ultracold atoms in optical traps, e.g. ^{87} Rb (37 electrons + 87 nucleons = 124 Fermions → Boson). For 2000 atoms in the trap the critical temperature is 170 nK.

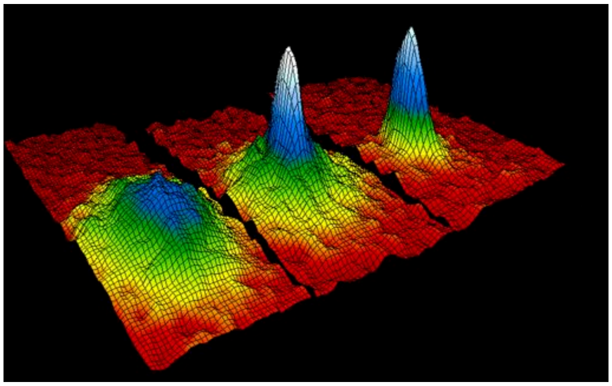


Fig.4.7: Velocity distribution of Rb-atoms: Left panel: $T > T_c$; middle panel: $T \sim T_c$; right panel $T \ll T_c$. A peak at the center develops corresponding to the fraction of particles with "zero-velocity" and at the same time the fraction of particles with finite velocity shrinks. (Source: http://www.nist.gov/public_affairs/gallery/bosein.htm)

4.6 Photons and phonons

We consider now classes of Bose gases whose particle number is not conserved. They are derived as normal modes of harmonic oscillators. Thus we first consider the statistical physics of the harmonic oscillator. The most simple example is the one-dimensional harmonic oscillator whose spectrum is given by

$$\epsilon_n = \hbar\omega\left(n+\frac{1}{2}\right) \qquad \text{with} \qquad n = 0,1,2,\dots \tag{4.85}$$

and the eigenstates $|n\rangle$. The quantum number n is considered as the occupation number of the oscillator mode. We analyze this within the canonical ensemble formulation with a given temperature T. The partition function reads

$$Z = tr e^{-\beta\mathcal{H}} = \sum_{n=0}^{\infty}\langle n|e^{-\beta\mathcal{H}}|n\rangle = \sum_{n=0}^{\infty} e^{-\beta\epsilon_n} = e^{-\beta\hbar\omega/2}\sum_{n=0}^{\infty} e^{-\beta\hbar\omega n} = \frac{e^{-\beta\hbar\omega/2}}{1-e^{-\beta\hbar\omega}} \,. \tag{4.86}$$

The internal energy is obtained through

$$U = -\frac{\partial \ln Z}{\partial\beta} = \frac{1}{2}\hbar\omega + \frac{\hbar\omega}{e^{\beta\hbar\omega}-1} \,. \tag{4.87}$$

The specific heat is

$$C = \frac{dU}{dT} = k_B\left(\frac{\hbar\omega}{2k_BT}\right)^2 \frac{1}{\sinh^2(\beta\hbar\omega/2)} \,, \tag{4.88}$$

with the limiting properties

$$C = \begin{cases} k_B & k_BT \gg \hbar\omega \\ \\ k_B\left(\dfrac{\hbar\omega}{2k_BT}\right)^2 e^{-\beta\hbar\omega} & k_BT \ll \hbar\omega \,. \end{cases} \tag{4.89}$$

In the large-temperature limit the specific heat approaches the equipartition law of a one-dimensional classical harmonic oscillator. The mean occupation number is given by

$$\langle n \rangle = \frac{1}{Z} \sum_{n=0}^{\infty} n e^{-\beta \epsilon_n} = \frac{1}{e^{\beta \hbar \omega} - 1} \ . \tag{4.90}$$

This corresponds to the Bose-Einstein distribution function. Thus we interpret n as a number of Bosons occupying the mode ω.

4.6.1 Blackbody radiation - photons

Electromagnetic radiation in a cavity is a good example of a discrete set of independent harmonic oscillators. Consider a cubic cavity of edge length L. The wave equation is expressed in terms of the vector potential

$$\left(\frac{1}{c^2} \frac{\partial^2}{\partial t^2} - \vec{\nabla}^2 \right) \vec{A} = 0 \tag{4.91}$$

and the electric and magnetic field are then

$$\vec{E} = -\frac{1}{c} \frac{\partial \vec{A}}{\partial t} \qquad \text{and} \qquad \vec{B} = \vec{\nabla} \times \vec{A} \ , \tag{4.92}$$

where we used the Coulomb gauge $\vec{\nabla} \cdot \vec{A} = 0$ and $\phi = 0$. This can be solved by a plane wave:

$$\vec{A}(\vec{r}, t) = \frac{1}{\sqrt{V}} \sum_{\vec{k}, \lambda} \left\{ A_{\vec{k}\lambda} \vec{e}_{\vec{k}\lambda} e^{i\vec{k}\cdot\vec{r} - i\omega t} + A^*_{\vec{k}\lambda} \vec{e}^{\,*}_{\vec{k}\lambda} e^{-i\vec{k}\cdot\vec{r} + i\omega t} \right\} \qquad \text{with} \qquad \left\{ \begin{array}{l} \omega = \omega_{\vec{k}} = c|\vec{k}| \ , \\ \\ \vec{e}_{\vec{k}\lambda} \cdot \vec{k} = 0 \ , \end{array} \right. \tag{4.93}$$

i.e. a linear dispersion relation and a transverse polarization $\vec{e}_{\vec{k}\lambda}$ (λ: polarization index). Assuming for simplicity periodic boundary conditions in the cube we obtain the quantization of the wavevector,

$$\vec{k} = \frac{2\pi}{L} (n_x, n_y, n_z) \qquad \text{with} \qquad n_i = 0, \pm 1, \pm 2, \ldots \tag{4.94}$$

Each of the parameter set $(\vec{k}, \lambda)$ denotes a mode representing an independent harmonic oscillator. These oscillators can be quantized again in the standard way.[4] The states of a mode differ by energy quanta $\hbar \omega_{\vec{k}}$. The occupation number $n_{\vec{k}\lambda}$ is interpreted as the number of *photons* in this mode.

The partition function is then derived from that of a harmonic oscillator

$$Z = \prod_{\vec{k}, \lambda} \frac{e^{-\beta \hbar \omega_{\vec{k}}/2}}{1 - e^{-\beta \hbar \omega_{\vec{k}}}} = \prod_{\vec{k}} \left(\frac{e^{-\beta \hbar \omega_{\vec{k}}/2}}{1 - e^{-\beta \hbar \omega_{\vec{k}}}} \right)^2 \tag{4.98}$$

---

[4] Canonical quantization of the radiation field: Introduce the variables

$$Q_{\vec{k}\lambda} = \frac{1}{\sqrt{4\pi c}} \left(A_{\vec{k}\lambda} + A^*_{\vec{k}\lambda} \right) \quad \text{and} \quad P_{\vec{k}\lambda} = \frac{i\omega_{\vec{k}}}{\sqrt{4\pi c}} \left(A_{\vec{k}\lambda} - A^*_{\vec{k}\lambda} \right) \tag{4.95}$$

which leads to the following expression for the Hamiltonian

$$\mathcal{H} = \int d^3 r \frac{\vec{E}^2 + \vec{B}^2}{8\pi} = \sum_{\vec{k}, \lambda} \frac{\omega_{\vec{k}}}{2\pi c} \left| A_{\vec{k}\lambda} \right|^2 = \frac{1}{2} \sum_{\vec{k}, \lambda} \left(P^2_{\vec{k}\lambda} + \omega^2_{\vec{k}} Q^2_{\vec{k}\lambda} \right) \ . \tag{4.96}$$

This is the Hamiltonian of a harmonic oscillator for each mode which we can quantize and obtain the new form

$$\mathcal{H} = \sum_{\vec{k}, \lambda} \hbar \omega_{\vec{k}} \left(a^\dagger_{\vec{k}\lambda} a_{\vec{k}\lambda} + \frac{1}{2} \right) = \sum_{\vec{k}, \lambda} \hbar \omega_{\vec{k}} \left(n_{\vec{k}\lambda} + \frac{1}{2} \right) \tag{4.97}$$

where $A_{\vec{k}\lambda} \to a_{\vec{k}\lambda}$ annihilates and $A^*_{\vec{k}\lambda} \to a^\dagger_{\vec{k}\lambda}$ creates a photon in the mode $(\vec{k}, \lambda)$.

where the exponent 2 originates from the two polarization directions. The internal energy follows from

$$U(T) = -\frac{\partial \ln Z}{\partial \beta} = 2\sum_{\vec{k}} \frac{\hbar\omega_{\vec{k}}}{e^{\beta\hbar\omega_{\vec{k}}}-1} = \int d\omega D(\omega)\frac{\hbar\omega}{e^{\beta\hbar\omega}-1} = V\int d\omega u(\omega,T) \tag{4.99}$$

where we have neglected the zero point motion term (irrelevant constant). The density of modes in (4.99) is denoted as

$$D(\omega) = \sum_{\vec{k},\lambda} \delta(\omega - \omega_{\vec{k}}) = \frac{2V}{(2\pi)^3} 4\pi \int dk\, k^2 \delta(\omega - ck) = V\frac{\omega^2}{\pi^2 c^3} \tag{4.100}$$

which leads to the spectral energy density

$$u(\omega,T) = \frac{\omega^2}{\pi^2 c^3}\frac{\hbar\omega}{e^{\beta\hbar\omega}-1} , \tag{4.101}$$

which is the famous Planck formula. There are two limits

$$u(\omega,T) \approx \begin{cases} k_B T \dfrac{\omega^2}{\pi^2 c^3} & \hbar\omega \ll k_B T \quad \text{Rayleigh-Jeans-law} \\ \\ \dfrac{\hbar\omega^3}{\pi^2 c^3} e^{-\beta\hbar\omega} & \hbar\omega \gg k_B T \quad \text{Wien's law} \end{cases} \tag{4.102}$$

whereby the Rayleigh-Jeans law corresponds to the classical limit. The maximum for given T follows Wien's displacement law,

$$\hbar\omega_0 = 2.82 k_B T . \tag{4.103}$$

The total internal energy density leads to the Stefan-Boltzmann law

$$\frac{U}{V} = \int d\omega u(\omega,T) = \frac{\pi^2}{15}\frac{(k_B T)^4}{(\hbar c)^3} \propto T^4 . \tag{4.104}$$

The energy current density of a blackbody is defined as

$$\frac{U}{V}c = \frac{\text{energy}}{\text{area} \cdot \text{time}} \tag{4.105}$$

Thus the emission power of electromagnetic radiation per unit area for the surface of a blackbody is defined by

$$P_{em} = \frac{U}{V}c\frac{1}{4\pi}\int d\Omega_{\vec{k}} \frac{\vec{k}\cdot\vec{n}}{|\vec{k}|} = \frac{U}{V}c\frac{1}{4\pi}\int d\Omega_{\vec{k}} \cos\theta = \frac{Uc}{4V} = \frac{\pi^2}{60}\frac{(k_B T)^4}{\hbar^3 c^2} = \sigma T^4 \tag{4.106}$$

where for the current density the component perpendicular to the surface counts ($\vec{n}$: surface normal vector).

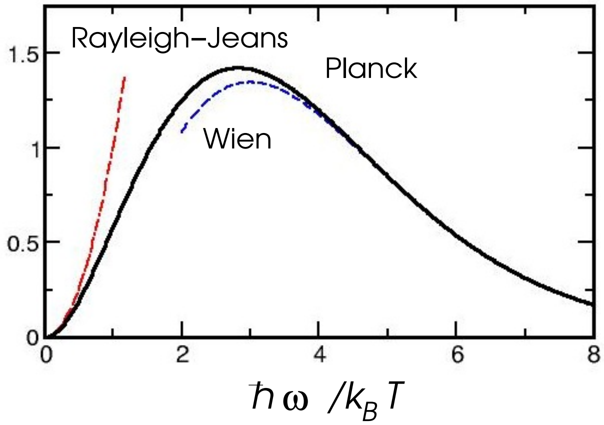


Fig.4.8: Spectral density of black body radiation.

This blackbody radiation plays an important role for the energy budget of the earth. The sun can be considered a blackbody emitting an energy current at the temperature of $T \approx 6000K$. This delivers an energy supply of $1.37kW/m^2$ to the earth. The earth, on the other hand, has to emit radiation back to the universe in order not to heat up arbitrarily. The earth is not a black body but a "gray" body as it is strongly reflecting in certain parts of the spectrum. A further example of blackbody radiation is the cosmic background radiation at a temperature 2.73 K which originates from the big bang.

4.6.2 Phonons in a solid

We consider Debye's theory of the lattice vibration and their influence on the thermodynamics of a solid. A solid consists of atoms which form a lattice. They interact with each other through a harmonic potential. Let us assume that the solid consists of N_A atoms arranged in a cube of edge length L, i.e. there are $3N_A$ degrees of freedom of motion. For our purpose it is convenient and sufficient to approximate this solid as a homogenous isotropic elastic medium whose vibration are described by the following equations of motion:

$$\begin{aligned} &\frac{1}{c_l^2}\frac{\partial^2 \vec{u}}{\partial t^2} - \vec{\nabla}(\vec{\nabla}\cdot\vec{u}) = 0 \quad \text{longitudinal sound mode ,} \\ &\frac{1}{c_t^2}\frac{\partial^2 \vec{u}}{\partial t^2} - \vec{\nabla}^2\vec{u} = 0 \qquad\quad \text{transversal sound mode .} \end{aligned} \tag{4.107}$$

There are two independent transversal ($\vec{k}\cdot\vec{u} = 0$) and one longitudinal ($\vec{k}\times\vec{u} = 0$) sound mode. These equation can be solved by plane waves and yield linear dispersion analogous to the electromagnetic waves:

$$\omega^{(l)}_{\vec{k}} = c_l|\vec{k}| \qquad \text{and} \qquad \omega^{(t)}_{\vec{k}} = c_t|\vec{k}| \; . \tag{4.108}$$

The density of states is obtained analogously using periodic boundary conditions for the waves,

$$D(\omega) = \frac{V\omega^2}{2\pi^2}\left(\frac{1}{c_l^3} + \frac{2}{c_t^3}\right) . \tag{4.109}$$

A difference occurs due to the finite number of degrees of freedom. In the end we get $3N_A$ modes. Thus there must be a maximal value of ω and $|\vec{k}|$. We take the sum

$$3N_A = \sum_{|\vec{k}|\le k_D} 3 = \frac{3V}{(2\pi)^3}4\pi\int_0^{k_D} dk\, k^2 = \frac{Vk_D^3}{2\pi^2} \qquad \Rightarrow \qquad k_D = \left(\frac{6\pi^2 N_A}{V}\right)^{1/3} \tag{4.110}$$

and define in this way the Debye wave vector k_D and the Debye frequency $\omega_D = c_{eff}k_D$ where

$$\frac{3}{c_{eff}^3} = \left(\frac{1}{c_l^3} + \frac{2}{c_t^3}\right) . \tag{4.111}$$

The internal energy is obtained again in the same way as for the electromagnetic radiation apart from the limit on the frequency integration:

$$\frac{U(T)}{V} = \int_0^{\omega_D} d\omega\; u(\omega, T) \; . \tag{4.112}$$

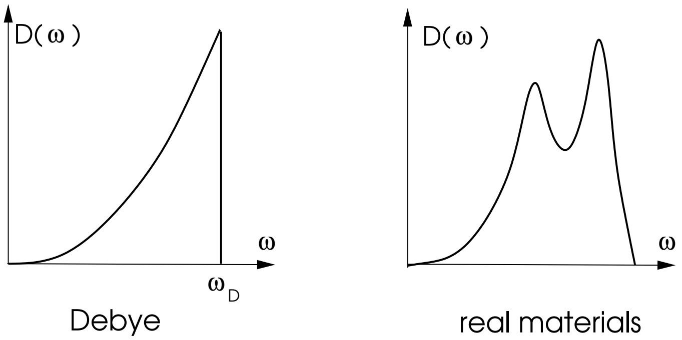


Fig.4.9: Density of states of phonons. Left panel: Debye model; right panel: more realistic spectrum. Note that the low frequency part in both cases follows an ω^2 law and leads to the T^3-behavior of the specific heat at low temperature. This is a consequence of the linear dispersion which is almost independent of the lattice structure and coupling.

We consider first the limit of small temperatures $k_BT \ll k_B\theta_D = \hbar\omega_D$ (θ_D: Debye temperature). The internal energy is given by

$$U(T) = V\frac{(k_BT)^4}{(2\pi\hbar)^3}\frac{3}{c_{eff}^3}4\pi\underbrace{\int_0^\infty dx\frac{x^3}{e^x-1}}_{=\pi^4/15} = V\frac{\pi^2(k_BT)^4}{10\hbar^3c_{eff}^3} = \frac{3\pi^4k_BT}{5}\left(\frac{T}{\theta_D}\right)^3 N_A \tag{4.113}$$

and correspondingly the low-temperature specific heat is

$$C_V = \frac{12\pi^4}{5}N_Ak_B\left(\frac{T}{\theta_D}\right)^3 , \tag{4.114}$$

the famous Debye law. On the other hand, at high temperatures ($T \gg \theta_D$) we use

$$\frac{1}{e^{\beta\hbar\omega}-1} \approx \frac{1}{\beta\hbar\omega} - \frac{1}{2} + \frac{\beta\hbar\omega}{12} . \tag{4.115}$$

This yields for the internal energy

$$\begin{aligned} U(T) &= \frac{3V}{2\pi^2c_{eff}^3}\int_0^{\omega_D} d\omega \left(\omega^2k_BT - \frac{\hbar\omega^3}{2} + \frac{\hbar^2\omega^4}{12k_BT}\right) + \cdots \\ &= 3N_Ak_BT\left\{1 - \frac{\hbar\omega_D}{k_BT} + \frac{1}{20}\left(\frac{\hbar\omega_D}{k_BT}\right)^2\right\} + \cdots \end{aligned} \tag{4.116}$$

and leads to the specific heat

$$C_V = 3N_Ak_B\left\{1 - \frac{1}{20}\frac{\theta_D^2}{T^2}\right\} + \cdots . \tag{4.117}$$

In the high-temperature limit the specific heat approaches the value of the equipartition law for $3N_A$ harmonic oscillators.

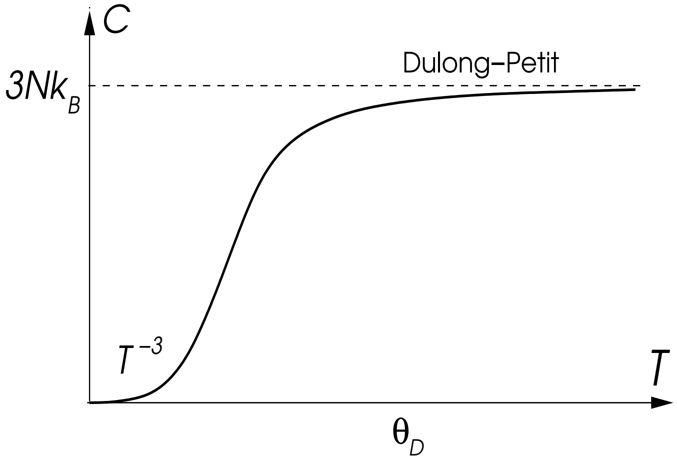


Fig.4.10: Specific heat: Low-temperature behavior follows a T^{-3}-law. At high-temperature the universal Dulong-Petit law recovers, which is a classical result of the equipartition law for particles in a harmonic potential.

The Debye temperature lies around room temperature usually. However, there also notable exception such as lead (Pb) with $\theta_D = 88K$ or diamond with $\theta_D = 1860K$.

4.6.3 Diatomic molecules

We now return to the problem of the diatomic molecules made of N atoms which we had already analyzed in the framework of classical statistical physics (Sect.3.4.2). Quantum mechanical aspects appear in the low-temperature regime $k_B T \ll \epsilon$ (ϵ: depth of the Lenard-Jones potential). Under this condition we consider again the Lenard-Jones potential around its minimum as a harmonic potential. Ignoring for the moment the motion of the center of mass, the Hamiltonian can be restricted to the relative coordinates r and relative momenta $\vec{p}$,

$$\mathcal{H} = \frac{\vec{p}^2}{2m^*} + v(r) = \frac{\vec{p}^2}{2m^*} + A(r - r_0)^2 \approx \frac{p_r^2}{2m^*} + \frac{\vec{L}^2}{2m^* r_0^2} + A(r - r_0)^2 \tag{4.118}$$

where the reduced mass is given by $m^* = m/2$ and p_r is radial momentum along the axis of the molecule bond. The motion separates into radial vibrational part which corresponds to a one-dimensional harmonic oscillator and the rotational motion. The two independent spectra are given by

$$E_n^{vib} = \hbar\omega \left(n + \frac{1}{2} \right) \qquad \text{and} \qquad E_l^{rot} = \frac{\hbar^2 l(l+1)}{2m^* r_0^2} \tag{4.119}$$

with $\omega = \sqrt{2A/m^*}$.

Let us now analyze the partition function of the different degrees of freedom. The translational degree of freedom (center of mass motion) corresponds to the motion of a particle of mass $2m$ in three dimensions. At low temperatures this has to be treated as bosonic quantum liquid in the case that the atoms are identical (not only chemically but also as an isotope, same mass m),

$$Z_{trans} = \prod_{\vec{p}} \frac{1}{1 - z^{-1} e^{-\beta \vec{p}^2/4m}} \; . \tag{4.120}$$

Next we have the vibrational part and the rotational part which we want to consider more carefully,

$$Z_{vib} = \left(\frac{e^{-\beta\hbar\omega/2}}{1 - e^{-\beta\hbar\omega}} \right)^{N/2} \tag{4.121}$$

and

$$Z_{rot} = \left(\sum_{l=0}^{\infty} (2l+1) e^{-\beta l(l+1)/I_{rot}} \right)^{N/2} \tag{4.122}$$

where $I_{rot} = mr_0^2/\hbar^2$. Note that per angular momentum quantum number l there are $2l+1$ degenerate states. Since we know the behavior of Z_{trans} and Z_{vib} already from the previous sections, we address now only the rotational part. The partition function cannot be expressed in closed form. Thus we examine only the limiting behaviors. As a characteristic energy scale we take $k_B\theta_{rot} = 2/I_{rot}$. For $T \gg \theta_{rot}$

$$\begin{aligned} Z_{rot} &\approx \left(\int_0^{\infty} dl \; (2l+1) e^{-\beta l(l+1)/I_{rot}} \right)^{N/2} = \left(k_B \int_0^{\infty} dl \; \frac{d}{dl} e^{-\beta l(l+1)/I_{rot}} \right)^{N/2} \\ &= (I_{rot} k_B T)^{N/2} = \left(2 \frac{T}{\theta_{rot}} \right)^{N/2} . \end{aligned} \tag{4.123}$$

and for $T \ll \theta_{rot}$,

$$Z_{rot} \approx \left(1 + 3e^{-\beta 2/I_{rot}} + \cdots \right)^{N/2} . \tag{4.124}$$

There is a hierarchy of temperatures in this system, $T_c \ll \theta_{rot} \ll \theta_{vib} \ll T_{dis}$, where T_c is the critical temperature for the Bose-Einstein condensation of the molecules, $k_B\theta_{vib} = \hbar\omega$ and $k_B T_{dis}$ is the molecule dissociation temperature ($\sim$ binding energy). We consider in the different regimes the behavior of the specific heat per molecule, $C(T) = dU/dT$:

$$\frac{2C(T)}{N} = \begin{cases} \frac{3}{2}k_B + 3k_B \left(\frac{\theta_{rot}}{T} \right)^2 e^{-\theta_{rot}/T} & T_c \ll T \ll \theta_{rot} \\ \frac{3}{2}k_B + k_B + k_B \left(\frac{\theta_{vib}}{2T} \right)^2 e^{-\theta_{vib}/T} & \theta_{rot} \ll T \ll \theta_{vib} \\ \frac{3}{2}k_B + k_B + k_B & \theta_{vib} \ll T \ll T_{dis} \\ 3k_B & T_{dis} \ll T \end{cases} \tag{4.125}$$

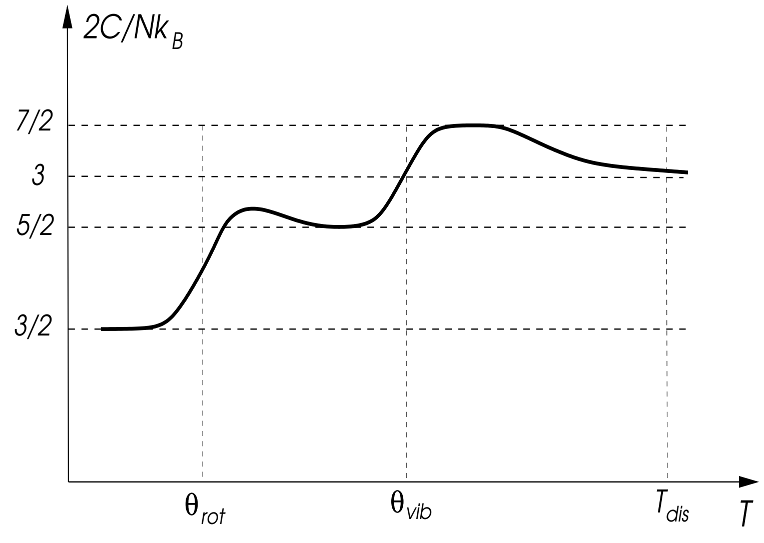


Fig. 4.11: Schematic behavior of specific heat of a diatomic molecule.

Note that due to the vibrational modes the specific heat goes through maxima when molecules are formed. Also the rotational modes are responsible for a non-monotonic behavior. Upon lowering temperature it looses gradually in magnitude due to the quantum mechanical freezing out of degrees of freedom.
For the hydrogen molecule H$_2$ the energy scales are $T_{dis} \sim 50000K$, $\theta_{vib} \sim 2000K$ and $\theta_{rot} \sim 85K$. [5] There is no Bose-Einstein condensation for H$_2$, because it solidifies at $14K$.

---

[5] Note that due to the fact that the nuclear spins of the H-atom is 1/2, there is a subtlety about the degeneracies. Actually, as identical Fermions their wave function has to be totally antisymmetric under exchange. If the two spins form a spin singlet then the orbital part has to have even parity, i.e. $l = 0, 2, 4, \dots$ (called "para-hydrogen"), while for the spin triplet configuration odd parity is required, i.e. $l = 1, 3, 5, \dots$ (called "ortho-hydrogen"). As a consequence in the partition function (4.122) the summands of odd l should have a prefactor 3. This does not affect the specific heat at temperatures $T \gg \theta_{rot}$ which is the classical limit of the rotator. But there is a (minor) correction at lower temperature such that

$$\frac{2C}{N} \approx \frac{3}{2} k_B T + 9 k_B \left(\frac{\theta_{rot}}{T} \right)^2 e^{-\theta_{rot}/T} \tag{4.126}$$

for $T_c \ll T \ll \theta_{rot}$.

Chapter 5

Phase transitions

Phase transitions in macroscopic systems are ubiquitous in nature and represent a highly important topics in statistical physics and thermodynamics. Phase transitions define a change of state of a system upon changing external parameters. In many cases this kind of change is obvious, e.g. transition between liquid and gas or between paramagnetic and ferromagnetic phase, and in most cases it is accompanied by anomalies in measurable macroscopic quantities.

In the previous chapter we have seen a phase transition, the Bose-Einstein condensation. This transition is special in the sense that it occurs for non-interacting particles. Generally, phase transitions require an interaction favoring an ordered phase. Then the phase transition occurs as a competition between the internal energy (or enthalpy) which is lowered by the order and the entropy which at finite temperature favors disorder. The relevant thermodynamic potentials to describe phase transitions are the Helmholtz free energy $F(T, V, N)$ and the Gibbs free energy $G(T, p, N)$,

$$F = U - TS \qquad \text{and} \qquad G = H - TS \, . \tag{5.1}$$

These potentials show anomalies (singularities) at the phase transition.

5.1 Ehrenfest classification of phase transitions

The type of singularity in the thermodynamic potential defines the order of the phase transition. According to Ehrenfest classification we call a phase transition occurring at a critical temperature T_c (different phase for $T > T_c$ and $T < T_c$) to be of n^{th} order, if the following properties hold:

$$\left.\left(\frac{\partial^m G}{\partial T^m}\right)_p\right|_{T=T_{c+}} = \left.\left(\frac{\partial^m G}{\partial T^m}\right)_p\right|_{T=T_{c-}} \qquad \text{and} \qquad \left(\frac{\partial^m G}{\partial p^m}\right)_{T=T_{c+}} = \left(\frac{\partial^m G}{\partial p^m}\right)_{T=T_{c-}} \tag{5.2}$$

for $m \leq n-1$, and

$$\left.\left(\frac{\partial^n G}{\partial T^n}\right)_p\right|_{T=T_{c+}} \neq \left.\left(\frac{\partial^n G}{\partial T^n}\right)_p\right|_{T=T_{c-}} \qquad \text{and} \qquad \left(\frac{\partial^n G}{\partial p^n}\right)_{T=T_{c+}} \neq \left(\frac{\partial^n G}{\partial p^n}\right)_{T=T_{c-}} \tag{5.3}$$

The same definition is used for the free energy. In practice this classification is rarely used beyond $n = 2$.

<u>$n = 1$:</u> A discontinuity is found in the entropy and in the volume:

$$S = -\left(\frac{\partial G}{\partial T}\right)_p \qquad \text{and} \qquad V = \left(\frac{\partial G}{\partial p}\right)_T \tag{5.4}$$

The discontinuity of the entropy is experimentally the *latent heat*. The change in volume is connected with the difference in density of the substance. A well-known example is the transition between liquid and gas phase, for which the former is much denser than the latter and accordingly takes a much smaller volume.

$\underline{n = 2}$: The discontinuities are not in the first derivatives but in the second derivatives of the Helmholtz free energy or Gibbs free energy, i.e. in the response functions. Some such quantities are the specific heat, the compressibility or the thermal expansion coefficient:

$$C_p = -T\left(\frac{\partial^2 G}{\partial T^2}\right)_p , \kappa_T = -\frac{1}{V}\left(\frac{\partial^2 G}{\partial p^2}\right)_T , \alpha = \frac{1}{V}\left(\frac{\partial^2 G}{\partial T \partial p}\right) \tag{5.5}$$

As we will see later, second order phase transitions are usually connected with spontaneous symmetry breaking and can be associated with the continuous growth of an order parameter. Such transitions show also interesting fluctuation features which lead to the so-called critical phenomena and universal behavior at the phase transition.

<u>Ehrenfest relations:</u> Interesting relations between various discontinuities at the phase transition exist. They are generally known at Ehrenfest relations. We consider first a first-order transition such as the gas-liquid transition. The phase boundary line in the p-T-phase diagram describes the relation between temperature and vapor pressure in the case of liquid-gas transition. For the differentials of the free enthalpy in the two phases the following equality holds:

$$dG_l = dG_g \qquad \Rightarrow \qquad -S_l dT + V_l dp = -S_g dT + V_g dp \; . \tag{5.6}$$

This allows us to get from the vapor pressure curve ($p(T)$ at the phase boundary in the p-T-plane) the relation

$$\frac{dp}{dT} = \frac{S_g - S_l}{V_g - V_l} = \frac{L}{T \Delta V} \tag{5.7}$$

where $L = T(S_g - S_l)$ is the *latent heat* and $\Delta V = V_g - V_l$ is the change of the volume. This relation is known as the *Clausius-Clapeyron* equation.
If the transition is of second order then the both the entropy and the volume are continuous through the transition between two phase A and B:

$$S_A(T,p) = S_B(T,p) \quad \text{and} \quad V_A(T,p) = V_B(T,p) \; , \tag{5.8}$$

which yields the relations through the equality of their differentials,

$$\begin{aligned} \left(\frac{\partial S_A}{\partial T}\right)_p dT + \left(\frac{\partial S_A}{\partial p}\right)_T dp &= \left(\frac{\partial S_B}{\partial T}\right)_p dT + \left(\frac{\partial S_B}{\partial p}\right)_T dp \; , \\ \left(\frac{\partial V_A}{\partial T}\right)_p dT + \left(\frac{\partial V_A}{\partial p}\right)_T dp &= \left(\frac{\partial V_B}{\partial T}\right)_p dT + \left(\frac{\partial V_B}{\partial p}\right)_T dp \; . \end{aligned} \tag{5.9}$$

We now use the Maxwell relation

$$\left(\frac{\partial S}{\partial p}\right)_T = -\left(\frac{\partial V}{\partial T}\right)_p = -V\alpha \tag{5.10}$$

and obtain

$$\frac{dp}{dT} = -\frac{\left(\frac{\partial S_B}{\partial T}\right)_p - \left(\frac{\partial S_A}{\partial T}\right)_p}{\left(\frac{\partial S_B}{\partial p}\right)_T - \left(\frac{\partial S_A}{\partial p}\right)_T} = \frac{\Delta C_p}{TV\Delta\alpha} \tag{5.11}$$

and analogously

$$\frac{dp}{dT} = -\frac{\left(\frac{\partial V_B}{\partial T}\right)_p - \left(\frac{\partial V_A}{\partial T}\right)_p}{\left(\frac{\partial V_B}{\partial p}\right)_T - \left(\frac{\partial V_A}{\partial p}\right)_T} = \frac{\Delta\alpha}{\Delta\kappa_T} \; . \tag{5.12}$$

Various other relations exist and are of experimental importance.

5.2 Phase transition in the Ising model

The Ising model is the simplest model of a magnetic system. Like for our model of the ideal paramagnet we consider magnetic moments or spins with two states, $s_i = \pm s$ (Ising spins). Sitting on a lattice they interact with their nearest neighbors (analogous to the spin chain in the Chapter 3). We write the model Hamiltonian as

$$\mathcal{H} = -J \sum_{\langle i,j \rangle} s_i s_j - \sum_i s_i H \ . \tag{5.13}$$

The sum $\sum_{\langle i,j \rangle}$ denotes summation over nearest neighbors on the lattice, counting each bond only once. J is the coupling constant which we assume positive. The second term corresponds to a Zeeman term due to an external magnetic field. The Ising spins are classical variables, unlike quantum spins $\vec{s}$ whose different components do not commute with each other. Ising spins represent only one component of a quantum spin.

The interaction favors the parallel alignment of all spins such that this model describes a ferromagnetic system. The ferromagnetic phase is characterized by a finite uniform mean value $\langle s_i \rangle = m$, the magnetization, even in the absence of an external magnetic field.

5.2.1 Mean field approximation

The mean field approximation is based on the view that a given spin in the lattice coupling to surrounding spins feels an average field which acts analogous to a Zeeman field. Thus for the spin s_i we may define

$$\mathcal{H} = \sum_i \mathcal{H}_i = \frac{1}{2} \sum_i s_i \left\{ -J \sum_{j \in \Lambda_i} s_j - 2H \right\} = -\sum_i s_i h_i \tag{5.14}$$

where Λ_i denotes the set of neighbors of site i (factor 1/2 due to double counting). The most simple form of the mean field approximation is to replace the field of the neighboring by their mean value, i.e. $h_i = h_{\text{eff}}$ which is independent of i.

Let us now tackle the problem in a more systematic way. We rewrite the spin for each site

$$s_i = \langle s_i \rangle + (s_i - \langle s_i \rangle) = m + (s_i - m) = m + \delta s_i \tag{5.15}$$

and insert it into the Hamiltonian, where we approximate $\langle s_i \rangle = m$ uniformly.

$$\begin{aligned} \mathcal{H} &= -J \sum_{\langle i,j \rangle} \{m + (s_i - m)\}\{m + (s_j - m)\} - \sum_i s_i H \\ &= -J \sum_{\langle i,j \rangle} \left\{ m^2 + m(s_i - m) + m(s_j - m) + \delta s_i \delta s_j \right\} - \sum_i s_i H \\ &= -J \sum_i \left(z m s_i - \frac{z}{2} m^2 \right) - \sum_i s_i H - J \sum_{\langle i,j \rangle} \delta s_i \delta s_j \ . \end{aligned} \tag{5.16}$$

Here z is the number of nearest neighbors (for a hypercubic lattice in d dimensions $z = 2d$). In the mean field approximation we neglect the last term assuming that it is small. This means that the fluctuations around the mean value would be small,

$$E_{ij} = \frac{\langle \delta s_i \delta s_j \rangle}{\langle s_i \rangle \langle s_j \rangle} = \frac{\langle \delta s_i \delta s_j \rangle}{m^2} \ll 1 \ , \tag{5.17}$$

to guarantee the validity of this approximation. We will see later that this condition is not satisfied very near the phase transition and that its violation is the basis for so-called critical phenomena. We now write the mean field Hamiltonian

$$\mathcal{H}_{mf} = -\sum_i s_i h_{\text{eff}} + NJ\frac{z}{2}m^2 \qquad \text{with} \qquad h_{\text{eff}} = Jzm + H \ , \tag{5.18}$$

which has the form of an ideal paramagnet in a magnetic field h_{eff}. It is easy to calculate the partition function and the free energy as a function of the parameter m,

$$Z(T, m, H) = e^{-\beta Jzm^2 N/2} \left\{2\cosh(\beta s h_{\text{eff}})\right\}^N \tag{5.19}$$

and

$$F(T, H, m) = -k_B T \ln Z = NJ\frac{z}{2}m^2 - Nk_B T \ln\left\{2\cosh(\beta s h_{\text{eff}})\right\} \ . \tag{5.20}$$

To find the equilibrium condition we have now to find the minimum of F for given T and H. To this end we minimize F with respect to m as the only free variable,

$$0 = \frac{\partial F}{\partial m} = NJzm - NJzs\tanh(\beta s h_{\text{eff}}) \ . \tag{5.21}$$

This equation is equivalent to the self-consistence equation for the mean value of s_i:

$$m = \langle s_i \rangle = \frac{\sum_{s_i=\pm s} s_i e^{\beta s_i h_{\text{eff}}}}{\sum_{s_i=\pm s} e^{\beta s_i h_{\text{eff}}}} = s\tanh(\beta s h_{\text{eff}}) \tag{5.22}$$

This is a non-linear equation whose solution determines m and eventually through the free energy all thermodynamic properties.

5.2.2 Instability of the paramagnetic phase

The disordered phase above a certain critical temperature T_c is called paramagnetic phase. For this phase we consider first the magnetic susceptibility $\chi(T)$ at zero magnetic field, which is obtained from

$$\chi(T) = N\frac{d\langle s_i\rangle}{dH}\bigg|_{H=0} = -\frac{d^2F}{dH^2}\bigg|_{H=0} = -\frac{d}{dH}\left\{\frac{\partial F}{\partial H} + \frac{\partial F}{\partial m}\frac{\partial m}{\partial H}\right\}\bigg|_{H=0} = -\frac{d}{dH}\frac{\partial F}{\partial H}\bigg|_{H=0} \tag{5.23}$$

where we used the equilibrium condition (5.21). Thus we obtain

$$\begin{aligned} \chi(T) &= Ns\frac{d}{dH}\tanh\left[\beta(Jzsm(H) + sH)\right]\bigg|_{H=0} \\ &= \frac{Ns}{k_BT}\left\{Jzs\left.\frac{dm}{dH}\right|_{H=0} + s\right\} \\ &= \frac{s}{k_BT}Jzs\chi(T) + \frac{Ns^2}{k_BT} \ . \end{aligned} \tag{5.24}$$

where we used that for a paramagnet $m(H=0) = 0$. This leads to the susceptibility

$$\chi(T) = \frac{Ns^2}{k_BT - Jzs^2} \tag{5.25}$$

which is modified compared to that of the ideal paramagnet. If $k_BT \to Jzs^2$ from above $\chi(T)$ is singular. We define this as the critical temperature

$$T_c = \frac{Jzs^2}{k_B} \ . \tag{5.26}$$

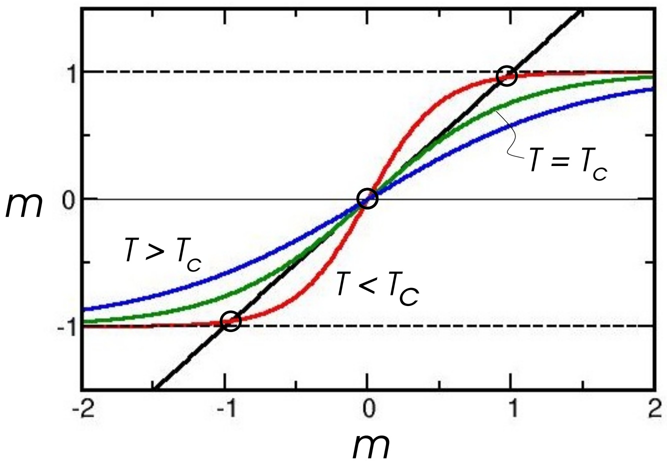


Fig. 5.1: Graphical solution of the self-consistence equation (5.22). The crossing points of the straight line and the step-like function gives the solution. There is only one solution at $m = 0$ for $T \geq T_c$ and three solutions for $T < T_c$.

As the system approaches $T = T_c$ it becomes more and more easy to polarize its spin by a tiny magnetic field. This indicates an instability of the system which we now analyze in terms of the self-consistence equation (5.22) in the absence of a magnetic field. Looking at Fig. 5.1 we find that indeed the critical temperature T_c plays an important role in separating two types of solutions of equations (5.22). For $T \geq T_c$ there is one single solution at $m = 0$ and for $T < T_c$ there are three solutions including $m = 0, m(T), -m(T)$. The physically relevant solution is then the one with finite value of m, as we will show below. It is obvious that below $T = T_c$ the mean field m grows continuously from zero to a finite value. In order to see which of the solutions is a minimum of the free energy we expand F in m assuming that m and H are small.

$$F(T,H,m) \approx NJz\left[\frac{m^2}{2} - \frac{k_BT}{Jz}\left\{\frac{(\beta s h_{\text{eff}})^2}{2} - \frac{(\beta s h_{\text{eff}})^4}{12}\right\}\right] - Nk_BT\ln 2 \tag{5.27}$$

For $H = 0$ we find

$$\begin{aligned} F(T,H,m) &\approx F_0(T) + NJz\left[\left(1-\frac{T_c}{T}\right)\frac{m^2}{2} + \frac{1}{12s^2}\left(\frac{T_c}{T}\right)^3 m^4\right] \\ &\approx F_0(T) + NJz\left[\left(\frac{T}{T_c}-1\right)\frac{m^2}{2} + \frac{m^4}{12s^2}\right] \end{aligned} \tag{5.28}$$

where for the last step we took into account that our expansion is only valid for $T \approx T_c$. Moreover, $F_0 = -k_BT\ln 2$. This form of the free energy expansion is the famous Landau theory of a continuous phase transition.

It is obvious that for $T > T_c$ the minimum lies at $m = 0$. For $T < T_c$ the coefficient of the m^2-term (2^{nd}-order) changes sign and a finite value of m minimizes F (see Fig.5.2). The minimization leads to

$$m(T) = \begin{cases} \pm s\sqrt{3\tau} & T < T_c \\ 0 & T \geq T_c \end{cases} \tag{5.29}$$

with $\tau = 1 - T/T_c$ as a short-hand notation. There are two degenerate minima and the system chooses spontaneously one of the two (spontaneous symmetry breaking).

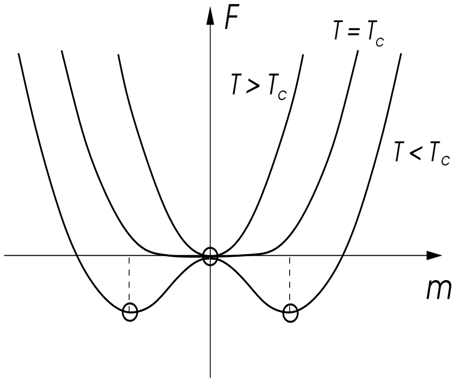


Fig. 5.2: Landau free energy: $T > T_c$: 2^{nd}-order term is positive and minimum of F at $m = 0$; $T = T_c$, 2^{nd} vanishes and free energy minimum at $m = 0$ becomes very shallow; $T < T_c$: 2^{nd}-order term is negative and minimum of F is at finite value of m, bounded by the 4^{th}-order term.

Next we analyze the behavior of the free energy and other thermodynamic quantities around the phase transition. The temperature dependence of the free energy and the entropy is given by

$$F(T) = F_0(T) - \frac{3Nk_BT_c\tau^2}{4}\Theta(\tau) \qquad \text{and} \qquad S(T) = -\frac{\partial F(T)}{\partial T} = Nk_B \ln 2 - \frac{3Nk_B\tau}{2}\Theta(\tau) \; , \tag{5.30}$$

and eventually we obtain for the specific heat,

$$\frac{C}{T} = \frac{\partial S}{\partial T} = \frac{3Nk_B}{2T_c}\Theta(\tau) + C_0 \tag{5.31}$$

where C_0 is zero in the present approximation for $H = 0$. While the free energy and the entropy are continuous through the transition, the specific heat shows a jump indicating the release of entropy through ordering.

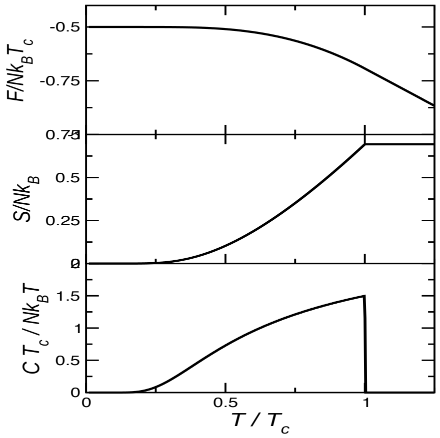


Fig. 5.3: Thermodynamic quantities within mean field theory. Free energy, entropy and specific heat.

Within mean field approximation the region close to T_c is described by the Landau expansion. However, taking the solution of the complete mean field equations leads to the thermodynamic behavior for the whole temperature range as shown in Fig. 5.3. Note that in this mean field approximation the entropy is $Nk_B \ln 2$ in the paramagnetic phase, the maximal value the entropy can reach.

5.2.3 Phase diagram

So far we have concentrated on the situation without magnetic field. In this case the phase transition goes to one of two degenerate ordered phases. Either the moments order to $m = +|m|$ or $m = -|m|$. An applied magnetic field lifts the degeneracy by introducing a bias for one of the two states. The order with m parallel to the field is prefered energetically. In a finite field the transition turns into a crossover, since there is already a moment m for temperatures above T_c. This is also reflected in the thermodynamic properties which show broad features around T_c and not anymore the sharp transition, e.g. the specific heat is turned into a broadened anomaly.

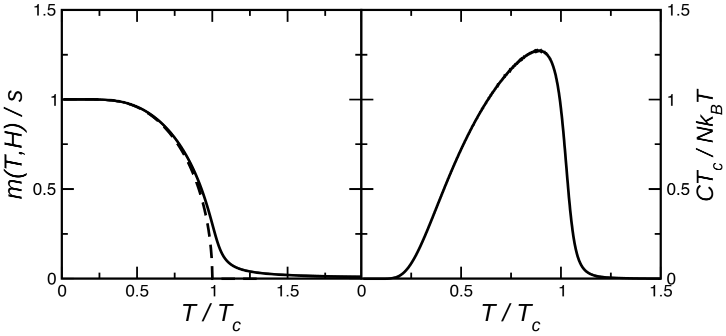


Fig. 5.4: Ising system in a finite magnetic field: (left panel) Magnetization as a function of temperature in a fixed magnetic field (solid line) and in zero field (dashed line); (right panel) specific heat for a fixed magnetic field. In a magnetic field no sharp transition exists.

Next we turn to the behavior of the magnetization m as a function of the magnetic field and temperature (illustrated in Fig.5.5 and 5.6). At $H = 0$ going from high to low temperatures the slope of $m(H)|_{H=0}$ is linear and diverges as we approach T_c. This reflects the diverging susceptibility as a critical behavior.

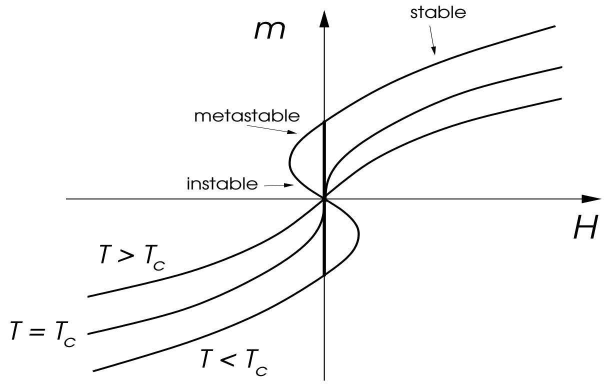


Fig. 5.5: Magnetization as a function of magnetic field for different temperatures.

For all temperatures $T > T_c$ $m(H)$ is a single-valued function in Fig.5.5. Below T_c, however, $m(H)$ is triply valued as a solution of the self-consistence equation. The part with $dm/dH > 0$ is stable or metastable representing local minima of the free energy. The part of $dm/dH < 0$ corresponds to a local maximum of the free energy and is unstable. Considering for $T < T_c$ the magnetic field running from negative to positive values, we find for $H < 0$ the state with $m < 0$ has lowest free energy. At $H = 0$ there is a degeneracy between $+|m|$ and $-|m|$ as both minima have the same energy and for $H > m > 0$ is the lowest energy phase. Thus we encounter a level-crossing of the free energy for the two states at $H = 0$ and magnetization jumps from negative to positive direction. This represents a first-order transition, since the free energy as a

function of H is singular, i.e. it has a sharp slope change (a discontinuity in the first derivative of F with respect to H).[1]
This may be visualized also in a three-dimensional graph in m, H and T (Fig. 5.6). The shaded region appearing for $T < T_c$ is a coexistence region of two degenerate phases. As m is changed in this region, the fraction of the two degenerate finite magnetization phases is changed continuously following a Maxwell construction.

$$m(q) = q|m(H=0,T)| + (1-q)\{-|m(H=0,T)|\} = (2q-1)|m(H=0,T)| \tag{5.32}$$

where q can change continuously with $0 \leq q \leq 1$. This may be viewed as domains of the two states changing in size.

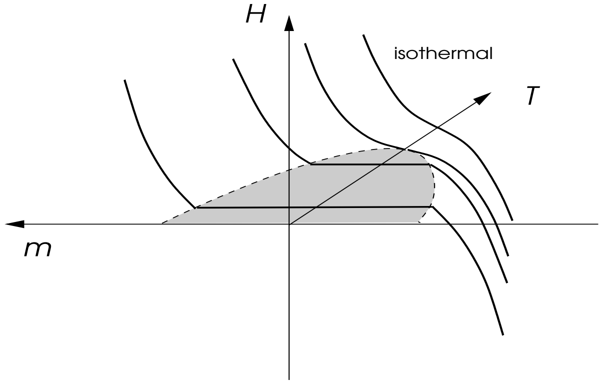


Fig. 5.6: Phase diagram in (H, m, T). The shaded region represents a coexistence region.

In the H-T-plane this shaded region is simply a line for the first order transition of a discontinuously changing magnetization.

5.2.4 Hubbard-Stratonovich transformation

We analyze the mean field approximation from a different point of view, using the Hubbard-Stratonovich transformation. The partition function of the Ising model can be rewritten by introducing an auxiliary field ϕ_i:

$$\begin{aligned} Z &= \sum_{\{s_i\}} e^{-\frac{\beta}{2}\sum_{i,j} J_{ij} s_i s_j + \beta \sum_i s_i H_i} \\ &= \frac{1}{(2\pi k_B T)^{N/2}\sqrt{detJ}} \int_{-\infty}^{+\infty} \left(\prod_{i'} d\phi_{i'}\right) \; e^{\frac{\beta}{2}\sum_{i,j}(J^{-1})_{ij}(\phi_i - H_i)(\phi_j - H_j)} \prod_i \sum_{s_i = \pm s} e^{\beta \phi_i s_i} \\ &= \frac{1}{(2\pi k_B T)^{N/2}\sqrt{detJ}} \int_{-\infty}^{+\infty} \left(\prod_{i'} d\phi_{i'}\right) \; e^{\frac{\beta}{2}\sum_{i,j}(J^{-1})_{ij}(\phi_i - H_i)(\phi_j - H_j) + \sum_i \ln[2\cosh(\beta s \phi_i)]} \end{aligned} \tag{5.33}$$

where we use the $N \times N$-matrix

$$J_{ij} = \begin{cases} -J & (i,j) \text{ nearest neighbors} \\ 0 & \text{otherwise} \end{cases} \tag{5.34}$$

---

[1] Note that in reality ramping H through 0 does not necessarily lead to a sharp transition. Flipping all magnetic moments at once is not a trivial matter. Thus the system may stay trapped in the metastable free energy minimum for a certain range of positive H. The transition may happen through the nucleation of reversed magnetic bubbles (domains) and the subsequent expansion of the reversed region by means of domain wall motion. (Similar to the nucleation and expansion of droplets in undercooled vapor.) There is a certain field where the metastable minimum becomes a saddle point and thus unstable for infinitesimally higher fields. Varying the magnetic field periodically can lead to hysteretic loops.

and $(J^{-1})_{ij}$ is the inverse of J_{ij}. We take the magnetic field as site dependent, which will be convenient later. We used the identity

$$\begin{aligned} &\int_{-\infty}^{+\infty} d\phi\, e^{-\frac{\phi^2}{2a}+s\phi} = e^{\frac{a}{2}s^2}\int_{-\infty}^{+\infty} d\phi\, e^{-\frac{1}{2a}(\phi-sa)^2} = \sqrt{2\pi a}e^{\frac{a}{2}s^2} \\ \Rightarrow \quad &\int_{-\infty}^{+\infty}\left(\prod_i d\phi_i\right) e^{-\frac{1}{2}\sum_{i,j}\phi_i(A^{-1})_{ij}\phi_j+\sum_i \phi_i s_i} = (2\pi)^{N/2}\sqrt{detA}\; e^{\frac{1}{2}\sum_{ij}s_iA_{ij}s_j} \end{aligned} \tag{5.35}$$

with A being a positive definite $N\times N$-matrix. This exact rewriting of the partition function is called Hubbard-Stratonovich transformation. We replaced here the discontinuous variable s_i by a continuous field ϕ_i.
We introduce the potential $S(\phi_i, H_i)$ and write

$$Z = C\int_{-\infty}^{+\infty}\left(\prod_{i'} d\phi_{i'}\right) e^{-\beta S(\phi_i,H_i)} = e^{-\beta F} \tag{5.36}$$

with $C = 1/(2\pi k_BT)^{N/2}\sqrt{detJ}$ and

$$S(\phi_i,H_i) = -\frac{1}{2}\sum_{i,j}(J^{-1})_{ij}(\phi_i-H_i)(\phi_j-H_j) - \frac{1}{\beta}\sum_i \ln[2\cosh(\beta s\phi_i)] \tag{5.37}$$

From this we can derive the mean field result by the saddle point approximation (sometimes also called method of steepest decent) which is based on the idea that one set replaces the auxiliary field by the value $\bar{\phi}_i$ which dominates the integral for the partition function. [2] This is analogous to discussions we had earlier when we tested the equivalence of different ensembles. This method relies on the fact that the fluctuations of the field ϕ_i are small - something which has to be tested. Therefore we look for the maximum of S now and approximate Z then by

$$Z \approx Ce^{-\beta S(\bar{\phi}_i,H_i)} \qquad \text{with} \quad 0 = \left.\frac{\partial S}{\partial \phi_i}\right|_{\phi_i=\bar{\phi}_i} = -\sum_j (J^{-1})_{ij}(\bar{\phi}_j - H_j) - s\tanh(\beta s\bar{\phi}_i) \tag{5.41}$$

which leads to the saddle-point equation

$$\bar{\phi}_i = H_i - s\sum_j J_{ij}\tanh(\beta s\bar{\phi}_j) \tag{5.42}$$

For $H_i = 0$ the saddle point is given by the uniform solution $\bar{\phi}_i = \bar{\phi}$, satisfying

$$\bar{\phi} = -s\sum_j J_{ij}\tanh(\beta s\bar{\phi}) = Jzs\tanh(\beta s\bar{\phi})\ . \tag{5.43}$$

---

[2] *Method of steepest decent:* We consider the integral

$$I = \int_a^b e^{Ng(x)}dx \tag{5.38}$$

and $N \gg 1$ and $g(x)$ a function with a unique maximum at $x = \bar{x}$ in the interval $[a,b]$, i.e.

$$g(x) = g(\bar{x}) + g'(\bar{x})(x-\bar{x}) + \frac{1}{2}g''(\bar{x})(x-\bar{x})^2 + \cdots\ . \tag{5.39}$$

For the maximum, the conditions $g'(\bar{x}) = 0$ and $g''(\bar{x}) < 0$ holds. We approximate the integral

$$I \approx e^{Ng(\bar{x})}\int_a^b e^{-N|g''(\bar{x})|(x-\bar{x})^2} \approx e^{Ng(\bar{x})}\int_{-\infty}^{+\infty} e^{-N|g''(\bar{x})|(x-\bar{x})^2} = \left(\frac{2\pi}{N|g''(\bar{x})|}\right)^{1/2} e^{Ng(\bar{x})} \tag{5.40}$$

which is exact in the limit $N \to \infty$. Considering $\ln I$ we find that $\ln I \approx Ng(\bar{x}) + O(\ln N)$.

This yields the same critical temperature for the onset of a finite solution for $\bar{\phi}$ as the mean field solution. The relation to m of the mean field approximation is given by the condition

$$\langle s_i \rangle = -k_B T \frac{\partial \ln Z}{\partial H_i} = -\frac{dS(\bar{\phi}_i, H_i)}{dH_i} = -\frac{\partial S}{\partial H_i} = -\sum_j (J^{-1})_{ij}(\bar{\phi}_i - H_i) = s \tanh(\beta s \bar{\phi}_i) \qquad (5.44)$$

such that

$$m = s \tanh(\beta s \bar{\phi}) \qquad \Rightarrow \qquad \bar{\phi} = Jzm \ . \qquad (5.45)$$

The discussion of the thermodynamic properties are in this formulation analogous to the ordinary mean field treatment and give exactly the same behavior. We are, however, now in a position to go beyond the mean field level and, in particular, to test the validity of the mean field approximation.

5.2.5 Correlation function and susceptibility

We consider first the correlation function

$$\Gamma_{ij} = \langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle = -(k_B T)^2 \frac{\partial \ln Z}{\partial H_i \partial H_j} \approx -k_B T \frac{d^2 S(\bar{\phi}_i, H_i)}{dH_i dH_j} \ . \qquad (5.46)$$

With (5.44) we obtain

$$\beta \Gamma_{ij} = \frac{d}{dH_j} s \tanh(\beta s \bar{\phi}_i) \qquad (5.47)$$

which if inverted yields

$$k_B T \Gamma^{-1}_{ij} = \frac{k_B T}{s^2} \cosh^2(\beta s \bar{\phi}_i) \frac{dH_j}{d\bar{\phi}_i} = \frac{k_B T}{s^2} \cosh^2(\beta s \bar{\phi}_i) \left\{ \delta_{ij} + \frac{\beta s^2 J_{ij}}{\cosh^2(\beta s \bar{\phi})} \right\} \qquad (5.48)$$

where we use (5.42)

$$H_j = \bar{\phi}_j + s \sum_{j'} J_{jj'} \tanh(\beta s \bar{\phi}_{j'}) \qquad \Rightarrow \qquad \frac{dH_j}{d\bar{\phi}_i} = \delta_{ij} + \frac{\beta s^2 J_{ij}}{\cosh^2(\beta s \bar{\phi}_i)} \ . \qquad (5.49)$$

The following Fourier-transformations lead to a simple form for (5.48),[3]

$$J_{ij} = \int \frac{d^3q}{(2\pi)^3} \, J(\vec{q}) e^{i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} \ , \quad \Gamma_{ij} = \int \frac{d^3q}{(2\pi)^3} \, \Gamma(\vec{q}) e^{i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} \ , \quad \delta_{ij} = \int \frac{d^3q}{(2\pi)^3} \, e^{i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} \qquad (5.54)$$

---

[3] For the inversion of Γ_{ij} it is important to realize that $\Gamma_{ij} = \Gamma(\vec{r}_i - \vec{r}_j)$ is translation invariant. We use now the linear equation

$$a_i = \sum_j \Gamma^{-1}_{ij} b_j \qquad \Rightarrow \qquad \sum_j \Gamma_{ij} a_j = b_i \ . \qquad (5.50)$$

We perform now the Fourier transform of the second equation,

$$\sum_i \sum_j \Gamma_{ij} a_j e^{-i\vec{q}\cdot\vec{r}_i} = \sum_i \sum_j \Gamma_{ij} a_j e^{-i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} e^{-i\vec{q}\cdot\vec{r}_j} = \Gamma(\vec{q}) a(\vec{q}) = \sum_i b_i e^{-\vec{q}\cdot\vec{r}_i} = b(\vec{q}) \ . \qquad (5.51)$$

On the other hand, we find

$$\sum_i \sum_j \Gamma^{-1}_{ij} b_j e^{-i\vec{q}\cdot\vec{r}_i} = \sum_i \sum_j \Gamma^{-1}_{ij} b_j e^{-i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} e^{-i\vec{q}\cdot\vec{r}_j} = \Gamma^{-1}(\vec{q}) b(\vec{q}) = \sum_i a_i e^{-\vec{q}\cdot\vec{r}_i} = a(\vec{q}) \ , \qquad (5.52)$$

leading to the relation

$$\Gamma^{-1}(\vec{q}) = \frac{1}{\Gamma(\vec{q})} \ , \qquad (5.53)$$

which is a convenient way to handle the inverted matrix Γ^{-1}_{ij}.

with

$$\Gamma(\vec{q}) = \frac{k_B T \Gamma_0}{1 + \Gamma_0 J(\vec{q})} \qquad \text{with} \qquad \Gamma_0 = \frac{\beta s^2}{\cosh^2(\beta s \bar{\phi})} = \beta(s^2 - m^2) \;, \tag{5.55}$$

using (5.45). On a d-dimensional hypercubic lattice with only nearest-neighbor coupling we obtain for $J(\vec{q})$,

$$J(\vec{q}) = \frac{1}{N} \sum_{i,j} J_{ij} e^{-i\vec{q}\cdot(\vec{r}_i - \vec{r}_j)} = -2J \sum_{\alpha=1}^{d} \cos q_\alpha a \tag{5.56}$$

with the lattice constant a. If we focus on the long-distance correlations only, we can restrict ourselves to the small $\vec{q}$ range and we expand $J(\vec{q})$ as

$$J(\vec{q}) \approx -Jz + Jq^2a^2 \qquad \Rightarrow \qquad \Gamma(\vec{q}) \approx \frac{k_B T}{\frac{k_B T}{s^2 - m^2} - Jz + Jq^2 a^2} \approx \frac{k_B T s^2}{k_B(T - T_c) + J s^2 q^2 a^2 + k_B T m^2 / s^2} \tag{5.57}$$

where for the last approximation we assumed $m \ll s$ as is the case in ordered phase close to T_c and for $T > T_c$.

First let us use this result to determine the susceptibility. For this purpose we take the earlier derived connection of the susceptibility with the fluctuations.

$$\chi = \beta \sum_{i,j} \{\langle s_i s_j \rangle - \langle s_i \rangle \langle s_j \rangle\} = \beta \sum_{i,j} \Gamma_{ij} = N\beta \Gamma(\vec{q} = 0) = \frac{N s^2}{k_B(T - T_c) + k_B T m^2 / s^2} \tag{5.58}$$

We use now the earlier mean field result for $m^2 = 3s^2\tau$ in (5.29) and obtain for the regime very close to T_c the behavior,

$$\chi(T) = \begin{cases} \dfrac{N s^2}{k_B(T - T_c)} & T > T_c \;, \\[2ex] \dfrac{N s^2}{2k_B |T - T_c|} & T < T_c \;, \end{cases} \tag{5.59}$$

showing that the susceptibility is singular approaching T_c from both sides of the phase transition with the same exponent for $|\tau|$.

Now return to the correlation function in the paramagnetic phase ($T > T_c$) and represent the correlation function in real space, where we restrict ourselves to the long-distance behavior $r \to \infty$. Thus we have the Fourier-transform $\Gamma(\vec{q})$. For simplicity we consider only the three-dimensional case,

$$\begin{aligned} \Gamma_{\vec{r}} &= \int \frac{d^3 q}{(2\pi)^3} \Gamma(\vec{q}) e^{i\vec{q}\cdot\vec{r}} = \frac{A}{4\pi^2} \int_0^\infty dq\, q^2 \int d\theta\ \sin\theta\, \frac{e^{iqr\cos\theta}}{1 + \xi^2 q^2} = \frac{A}{4\pi^2 i r} \int_0^\infty dq\, q \frac{e^{iqr} - e^{-iqr}}{1 + \xi^2 q^2} \\ &= \frac{A}{4\pi^2 i r} \int_{-\infty}^{+\infty} dq\, q \frac{e^{iqr}}{1 + \xi^2 q^2} = \frac{A}{4\pi} \frac{e^{-r/\xi}}{r\xi^2} = \frac{k_B T}{4\pi J} \frac{e^{-r/\xi}}{r} \end{aligned} \tag{5.60}$$

where we used residue calculation for the last integral and introduced the parametrization,

$$A = \frac{s^2}{1 - T_c/T} = \frac{k_B T \xi^2}{J a^2} \qquad \text{and} \qquad \xi^2 = \frac{J s^2 a^2}{k_B(T - T_c)} \;. \tag{5.61}$$

The general form of the correlation function for other dimensions d is

$$\Gamma_{\vec{r}} \propto \frac{e^{-r/\xi}}{r^{(d-1)/2}} \tag{5.62}$$

if $T > T_c$. In all cases there is a correlation length which diverges as we approach $T \to T_{c+}$. At T_c we find

$$\Gamma_{\vec{r}} = \frac{k_B T}{J} \int \frac{d^d q}{(2\pi)^3} \frac{e^{i\vec{q}\cdot\vec{r}}}{q^2} \propto \begin{cases} \ln r & d = 2 \\ \\ r^{2-d} & d \geq 3 \end{cases} \tag{5.63}$$

This suggests that for $T \to T_{c+}$ the correlation function should rather behave as

$$\Gamma_{\vec{r}} \propto \begin{cases} \ln r \; e^{-r/\xi} & d = 2 \\ \\ \dfrac{e^{-r/\xi}}{r^{d-2}} & d \geq 3 \end{cases} \tag{5.64}$$

which we will encounter later in the context of scaling again.

5.3 Ginzburg-Landau theory

We have used the Landau expansion of the free energy above to discuss phase transitions in the vicinity of the critical temperature where m was small. This method can be extended to a highly convenient scheme which allows us to discuss phase transition more generally, in particular, those of second order. Landau's concept is based on symmetry and spontaneous symmetry breaking. The disordered high-temperature phase has a certain symmetry which is characterized by a group $\mathcal{G}$ of all symmetry operations leaving the system invariant. At the second order phase transition a form of order appears reducing this symmetry such that the low-temperature phase obeys a symmetry group $\mathcal{G}'$ which is a subgroup of $\mathcal{G}$. This change of symmetry is called *spontaneous symmetry breaking*. This symmetry lowering is described by the appearance of an *order parameter*, a quantity which does not possess all symmetries of the group $\mathcal{G}$.

A further important aspect emerges when long-length scale variations of the order parameter are taken into account. This can be easily incorporated in the Ginzburg-Landau theory and allows to discuss spatial variations of the ordered phase as well as fluctuations.

5.3.1 Ginzburg-Landau theory for the Ising model

For the Ising model of the previous section we can identify m as the order parameter. The order parameter m is not invariant under time reversal symmetry $\hat{K}$,

$$\hat{K} m = -m \; . \tag{5.65}$$

The two states with positive and negative m are degenerate. The relevant symmetry group above the phase transition is

$$\mathcal{G} = G \times \mathcal{K} \tag{5.66}$$

with G as the space group of the lattice (simple cubic) and $\mathcal{K}$, the group $\{E, \hat{K}\}$ (E denotes the identity operation). As for the space group we consider the magnetic moment here detached from the crystal lattice rotations such that G remains untouched through the transition so that the corresponding subgroup is

$$\mathcal{G}' = G \subset \mathcal{G} \tag{5.67}$$

The degeneracy of the ordered phase is given by the order of $\mathcal{G}/\mathcal{G}'$ which is 2 in our case.

The Ginzburg-Landau free energy functional has in d dimensions the general form

$$\begin{aligned} F[m; H, T] &= F_0(H,T) + \int d^d r \left\{ \frac{A}{2} m(\vec{r})^2 + \frac{B}{4} m(\vec{r})^4 - H(\vec{r}) m(\vec{r}) + \frac{\kappa}{2} [\vec{\nabla} m(\vec{r})]^2 \right\} \\ &= F_0(H,T) + \int d^d r \; f(m, \vec{\nabla} m; HT) \end{aligned} \tag{5.68}$$

where we choose the coefficients according to the expansion done in (5.28) as

$$A = \frac{Jz}{a^d}\left(\frac{T}{T_c} - 1\right) = -Jz\tau \quad \text{and} \quad B = \frac{Jz}{3s^2 a^d} \; . \tag{5.69}$$

Here a is the lattice constant. We have introduced the spatial continuum limit for the order parameter m which is based on the procedure of *coarse graining*. We take a block of sites with the volume L_b^d with L_b much larger than the lattice constant a and define

$$m(\vec{r}) = \frac{1}{N_b} \sum_{i \in \Lambda_b(\vec{r})} \langle s_i \rangle \qquad \text{with} \qquad N_b = \frac{a^d}{L_b^d} \tag{5.70}$$

and $\Lambda_b(\vec{r})$ is the set of sites in the block around the center position $\vec{r}$. Here we assume that $\langle s_i \rangle$ is changing slowly in space on the length scale L_b.

Under this condition we can now also determine κ from the model Hamiltonian using the following consideration. The variational equation of the free energy functional is given by

$$0 = \frac{\delta F}{\delta m} = \frac{\partial f}{\partial m} - \vec{\nabla} \cdot \frac{\partial f}{\partial \vec{\nabla} m} = -\kappa \vec{\nabla}^2 m + Am + Bm^3 - H \tag{5.71}$$

Let us compare this equation with the saddle-point equation (5.42,5.43) assuming $H = 0$. It is sufficient to keep only the terms of first order in m. We then take the equations (5.42,5.45) and expand the self-consistence equation to linear order

$$\begin{aligned} \bar{\phi}_i &= \bar{\phi}(\vec{r}_i) \approx -\beta s^2 \sum_j J_{ij} \bar{\phi}(\vec{r}_j) = \beta s^2 \sum_{\{\vec{a}\}_{n.n.}} J \bar{\phi}(\vec{r}_i + \vec{a}) \\ &= \beta s^2 J \left[z\bar{\phi}(\vec{r}_i) + \sum_{\{\vec{a}\}_{n.n.}} \vec{a} \cdot \vec{\nabla} \bar{\phi}(\vec{r}_i) + \frac{1}{2} \sum_{\{\vec{a}\}_{n.n.}} \sum_{\mu,\nu = x,y\ldots} a_\mu a_\nu \frac{\partial^2}{\partial r_\mu \partial r_\nu} \bar{\phi}(\vec{r}_i) \right] \; . \end{aligned} \tag{5.72}$$

The sum $\sum_{\{\vec{a}\}_{n.n.}}$ runs over nearest-neighbor sites. Note that the second term in the bracket $[\ldots]$ vanishes due to symmetry. Now using the coarse graining procedure we may replace $\bar{\phi}(\vec{r}_i) = Jzm(\vec{r})$ and obtain

$$0 = Jz\left(\frac{T}{T_c} - 1\right) m(\vec{r}) - Ja^2 \vec{\nabla}^2 m(\vec{r}) \; , \tag{5.73}$$

and the comparison of coefficients leads to

$$\kappa = Ja^{2-d} \; . \tag{5.74}$$

We may rewrite the equation (5.73) as

$$0 = m - \xi^2 \vec{\nabla}^2 m \qquad \text{with} \qquad \xi^2 = \frac{k_B T_c}{z k_B (T - T_c)} = \frac{J s^2 a^2}{k_B (T - T_c)} \tag{5.75}$$

where we introduced the length ξ which is exactly equal to the correlation length for $T > T_c$ in (5.61).

5.3.2 Critical exponents

Close to the phase transition at T_c various quantities have a specific temperature or field dependence which follows powerlaws in $\tau = 1 - T/T_c$ with characteristic exponents, so-called critical exponents. We introduce here the exponents relevant for a magnetic system like the Ising model. The specific heat C and the susceptibility χ follow the behavior

$$C(T) \propto |\tau|^{-\alpha} \qquad \text{and} \qquad \chi(T) \propto |\tau|^{-\gamma} \tag{5.76}$$

for both $\tau > 0$ and $\tau < 0$. Also the coherence length displays a powerlaw

$$\xi(T) \propto |\tau|^{-\nu} \, . \tag{5.77}$$

For $\tau > 0$ (ordered phase) the magnetization grows as

$$m(T) \propto |\tau|^{\beta} \, . \tag{5.78}$$

At $T = T_c$ ($\tau = 0$) the magnetization has the field dependence

$$m \propto H^{1/\delta} \tag{5.79}$$

and the correlation function has a powerlaw dependence on the distance r

$$\Gamma_{\vec{r}} \propto \frac{1}{r^{d-2+\eta}} \, . \tag{5.80}$$

These exponents are not completely independent but are related by means of so-called scaling laws:

- Rushbrooke scaling: $\alpha + 2\beta + \gamma = 2$
- Widom scaling: $\gamma = \beta(\delta - 1)$
- Fisher scaling: $\gamma = (2 - \eta)\nu$
- Josephson scaling: $\nu d = 2 - \alpha$

We do not derive all relations, but restrict to Fisher's result in order to show the basic idea. We consider the correlation function for $\tau > 0$ but very close to T_c. Then using (5.80) we assume that we can write $\Gamma_{\vec{r}}$ as

$$\Gamma_{\vec{r}} \propto \frac{1}{r^{d-2+\eta}} g(r/\xi) \, . \tag{5.81}$$

According to our previous discussion the susceptibility is given by the integral of $\Gamma_{\vec{r}}$ over space

$$\chi \propto \int d^d r \, \Gamma_{\vec{r}} \propto \int d^d r \, \frac{1}{r^{d-2+\eta}} g(r/\xi) \propto \xi^{2-\eta} \int d^d y \frac{1}{y^{d-2+\eta}} g(y) \propto |\tau|^{-\nu(2-\eta)} \propto |\tau|^{-\gamma} \tag{5.82}$$

which leads to $\gamma = \nu(2 - \eta)$.
Let us now determine the exponents within mean field theory. The only one we have not determined so far is δ. Using the Ginzburg-Landau equations for $\tau = 0$ leads to

$$Bm^3 = H \qquad \Rightarrow \qquad \delta = 3 \tag{5.83}$$

Thus the list of exponents is

$$\alpha = 0 \, , \quad \beta = \frac{1}{2} \, , \quad \gamma = 1 \quad \delta = 3 \, , \quad \nu = \frac{1}{2} \quad \eta = 0 \tag{5.84}$$

These exponents satisfy the scaling relations apart from the Josephson scaling which depends on the dimension d.
The critical exponents arise from the specific fluctuation (critical) behavior around a second-order phase transition. They are determined by dimension, structure of order parameter and coupling topology, and are consequently identical for equivalent phase transitions. Therefore, the critical exponents incorporate *universal* properties.

5.3.3 Range of validity of the mean field theory - Ginzburg criterion

In equation (5.17) we gave a condition for the validity of the mean field approximation. The fluctuations around the mean field should be small. We formulate this statement more precisely here. In previous chapters we have found that for large systems the fluctuations are very small compared to the mean value (e.g. energy, particle number, ...). Looking at the fluctuations of the magnetic moments, the question arises what is the "system size". The natural length scale for fluctuations here is the correlation length which we know from the correlation function Γ_{ij}. Thus, the relevant "system size" corresponds to the volume $V_\xi = \xi^d$. Looking at the ratio

$$E_{ij} = \frac{\langle (s_i - \langle s_i \rangle)(s_j - \langle s_j \rangle) \rangle}{\langle s_i \rangle \langle s_j \rangle} \qquad \Rightarrow \qquad E_{GL} = \frac{\sum_{j \in V_\xi} \Gamma_{0j}}{\sum_{j \in V_\xi} \langle s_0 \rangle \langle s_j \rangle} \ . \tag{5.85}$$

We first look at the numerator and use the fact that $\Gamma_{\vec{r}}$ decays on the length ξ quickly. Thus we use

$$\int_{V_\xi} d^d r \ \Gamma_{\vec{r}} = f k_B T \frac{\chi(T)}{N} \ . \tag{5.86}$$

where the factor f is independent of ξ and temperature and gives the fraction to which integral limited to V_ξ corresponds to the susceptibility. We can estimate it by using the scaling form of $\Gamma_{\vec{r}}$,

$$\int_0^\xi dr \frac{r^{d-1}}{r^{d-2}} g(r/\xi) = f \int_0^\infty dr \frac{r^{d-1}}{r^{d-2}} g(r/\xi) \qquad \Rightarrow \qquad \xi^2 \int_0^1 dx \ x g(x) = f \xi^2 \int_0^\infty dx \ x g(x) \ , \tag{5.87}$$

which leads indeed to f independent of ξ and of order one. Next we consider the denominator of E_{GL} and we restrict to $T < T_c$ so that we immediately find

$$\sum_{j \in V_\xi} \langle s_0 \rangle \langle s_j \rangle = m(T)^2 N_\xi \tag{5.88}$$

with $N_\xi = (\xi/a)^d$ as the number of lattice sites within the volume V_ξ. The criterion for the mean field theory to be applicable is then given by

$$f k_B T_c \frac{\chi(T)}{N} \ll \left(\frac{\xi}{a} \right)^d m(T)^2 \quad \Rightarrow \quad \frac{f s^2}{\tau} \ll 3 s^2 \tau \left(\frac{\xi_0}{a} \right)^d \tau^{-d/2} \tag{5.89}$$

with ξ_0 as the correlation length at zero-temperature. From these we obtain the range of validity

$$\tau^{\frac{4-d}{2}} \gg \frac{f}{3} \left(\frac{a}{\xi_0} \right)^d \tag{5.90}$$

which excludes obviously a narrow range of temperature close to T_c for dimensions $d \leq 3$. For dimensions $d > 4$ there is no limitations and the mean field theory including the Ginzburg-Landau theory is valid also close to T_c. The temperature region

$$\Delta T = T_c - T \sim T_c \left\{ \frac{f}{3} \left(\frac{a}{\xi_0} \right)^d \right\}^{\frac{2}{4-d}} \tag{5.91}$$

is called the critical region and grows with shrinking ξ_0 as the number of spins in the correlation volume V_ξ becomes smaller.

One important result is the fact that the dimension $d_c = 4$ has been identified as an upper critical dimension. For $d > d_c$ the mean field exponents are valid. This does, however, not mean that the mean field approximation delivers the exact thermodynamics.

5.4 Self-consistent field approximation

Fluctuation effects affect not only the critical exponents, they also alter the condition for the instability. The mean field approximation overestimates the tendency to order by ignoring fluctuations. The functional integral formulation based on the Hubbard-Stratonovich transformations allows to tackle the discussion of fluctuation effects in a systematic way. Since the Ginzburg-Landau free energy is a good approximation to the saddle point solution close to the phase transition, it shall be the basis for our discussion. A good insight into some aspects of the fluctuation effects is obtained by the so-called self-consistent field approximation which is used to analyze the instability condition. We consider the fluctuations around the mean order parameter value by looking at the partion function

$$Z = \int \mathcal{D}m(\vec{r}) e^{-\beta F(m;T,H)} \tag{5.92}$$

which represent a functional integral in the order parameter $m(\vec{r})$. Our aim in the following will be to examine the instability condition of the paramagnetic state. Including thermal fluctuations we find that the mean square of the order parameter is finite above the transition temperature, while the simple mean value vanishes, i.e. $\langle m(\vec{r})^2\rangle \neq 0$ while $\langle m(\vec{r})\rangle = 0$.

5.4.1 Renormalization of the critical temperature

In order to avoid difficulties with the functional integral due to the fourth-order terms in F, we expand F to second order around its mean value in the paramagnetic phase ($\langle m(\vec{r})\rangle = 0$),

$$F = F_0 + F_L + \frac{1}{2}\int d^d r\, d^d r'\, m(\vec{r}) \left\langle \frac{\delta^2 F}{\delta m(\vec{r})\delta m(\vec{r}\,')} \right\rangle m(\vec{r}\,') \tag{5.93}$$

where we use

$$\left\langle \frac{\delta^2 F}{\delta m(\vec{r})\delta m(\vec{r}\,')} \right\rangle = \left\{ A + 3B\langle m(\vec{r})^2\rangle - \kappa \vec{\nabla}^2 \right\} \delta(\vec{r}-\vec{r}\,')\ . \tag{5.94}$$

F_L is the term in zeroth order. which is $F_L = 0$ for $T > T_c$ where $m = 0$. This leads to the so-called Gaussian approximation. The mean value $\langle m(\vec{r})^2\rangle = \langle m^2\rangle$ is taken to be uniform.
It is now convenient to represent the order parameter fluctuations in Fourier space,

$$m(\vec{r}) = \frac{1}{\sqrt{L^d}} \sum_{\vec{q}} m_{\vec{q}}\, e^{i\vec{q}\cdot\vec{r}} \qquad \text{with} \qquad m^*_{\vec{q}} = m_{-\vec{q}}\ , \tag{5.95}$$

in a hypercube of edge length L to introduce periodic boundary conditions for $m(\vec{r})$. Then F reads

$$F = F_0 + \frac{1}{2}\sum_{\vec{q}} \left\{A + 3B\langle m^2\rangle + \kappa q^2\right\} m_{\vec{q}} m_{-\vec{q}} = F_0 + \sum_{\vec{q}} G^{-1}(\vec{q}) m_{\vec{q}} m_{-\vec{q}} \tag{5.96}$$

Now we may rewrite the partition function in this approximation as

$$\begin{aligned} Z' &= Z_0 \prod_{\vec{q}} \int dm_{\vec{q}}\, dm_{-\vec{q}}\, \exp\left\{-\beta G^{-1}(\vec{q}) m_{\vec{q}} m_{-\vec{q}}\right\} \\ &= Z_0 \prod_{\vec{q}} \int dm'_{\vec{q}}\, dm''_{\vec{q}}\, \exp\left\{-\beta G^{-1}(\vec{q})({m'_{\vec{q}}}^2 + {m''_{-\vec{q}}}^2)\right\} \end{aligned} \tag{5.97}$$

where we used the parametrization $m_{\pm\vec{q}} = m'_{\vec{q}} \pm i m''_{\vec{q}}$. Now determine the mean value

$$\begin{aligned} \langle m^2\rangle &= \frac{1}{L^d}\int d^d r\, \langle m(\vec{r})^2\rangle = \frac{1}{L^d}\sum_{\vec{q}} \langle m_{\vec{q}} m_{-\vec{q}}\rangle \\ &= \frac{1}{L^d}\sum_{\vec{q}} \langle {m'_{\vec{q}}}^2 + {m''_{\vec{q}}}^2\rangle = \frac{k_B T}{L^d}\sum_{\vec{q}} G(\vec{q}) = \frac{1}{L^d}\sum_{\vec{q}} \frac{k_B T}{A + 3B\langle m^2\rangle + \kappa q^2} \end{aligned} \tag{5.98}$$

which represents a self-consistent equation for $\langle m^2 \rangle$. Note that the $G(\vec{q})$ corresponds to a renormalized correlation function, analogous to $\Gamma_{\vec{q}}$ in (5.57).

On the other hand, we can determine the susceptibility using the fluctuations,

$$\begin{aligned} \chi(T) &= \beta \frac{1}{L^d} \int d^d r \, d^d r' \left\{ \langle m(\vec{r}) m(\vec{r}\,') \rangle - \langle m(\vec{r}) \rangle \langle m(\vec{r}\,') \rangle \right\} \\ &= \beta \langle m^2_{\vec{q}=0} \rangle = G(\vec{q}=0) = \frac{1}{A + 3B\langle m^2 \rangle} \end{aligned} \tag{5.99}$$

The fact that the susceptibility diverges at the phase transition can now be used to determine the instability temperature T_c^* which is different from the "bare" T_c defining the zero of A. Thus we analyze the equation

$$\chi^{-1} = [A + 3B\langle m^2 \rangle] = A + \frac{3Bk_BT}{L^d} \sum_{\vec{q}} \frac{1}{\chi^{-1} + \kappa q^2} \, . \tag{5.100}$$

where $\chi^{-1} = 0$ determines T_c^*. For the sum over $\vec{q}$ we have to be careful because the form we have chosen for the expansion in small powers of q is not valid at large q. We correct for this by introducing an upper cutoff Λ for q, which is of the order a^{-1}, the inverse lattice constant. Then going to integral form $L^{-d} \sum_{\vec{q}} \to \int d^d q / (2\pi)^d$ and setting $\chi^{-1} = 0$ we reach at equation

$$A_c = Jza^{-d} \left(\frac{T_c^*}{T_c} - 1 \right) = -\frac{Jza^{-d}}{s^2} \frac{C_d k_B T_c^*}{(2\pi)^d} \int_0^\Lambda dq \, \frac{q^{d-1}}{\kappa q^2} \tag{5.101}$$

where C_d is the volume of a unit-sphere in d dimensions as defined int (3.37). Thus the renormalized transition temperature is

$$T_c^* = \frac{T_c}{1 + \dfrac{C_d z}{(2\pi)^d} \dfrac{(\Lambda a)^{d-2}}{d-2}} \quad < T_c \, . \tag{5.102}$$

Therefore the transition temperature is reduced by the fluctuations. The dimension $d = d_L = 2$ appears as lower critical dimension. For dimension $d < d_L$ the integral diverges at the lower boundary (infrared divergence) such that no finite transition temperature is possible. The dimension d_L is more subtle. Below we will comment from a different point of view on this result.

5.4.2 Renormalized critical exponents

Now we turn to the behavior of the susceptibility at the new critical temperature. For this purpose we rewrite (5.100) with the help of (5.101) and obtain for $T \to T_{c+}^*$

$$\begin{aligned} \chi^{-1} &= (A - A_c) + \frac{3BC_d}{(2\pi)^d} \int_0^\Lambda dq \, q^{d-1} \left[\frac{k_B T}{\chi^{-1} + \kappa q^2} - \frac{k_B T_c^*}{\kappa q^2} \right] \\ &\approx (A - A_c) - \frac{3BC_d k_B T_c^*}{(2\pi)^d \kappa} \int_0^\Lambda dq \frac{q^{d-3}}{1 + \chi \kappa q^2} \\ &= (A - A_c) - \frac{3BC_d k_B T_c^*}{(2\pi)^d \kappa} \{\kappa \chi\}^{(2-d)/2} \int_0^{\Lambda \sqrt{\kappa \chi}} dx \frac{x^{d-3}}{1 + x^2} \, . \end{aligned} \tag{5.103}$$

Note that $\kappa\chi = \xi^2$. We distinguish now two cases. First choose the dimension larger than the upper critical dimension $d > d_c = 4$ and obtain

$$\begin{aligned}\chi^{-1} &\approx (A - A_c) - \frac{3BC_d k_B T_c^*}{(2\pi)^d \kappa} \{\kappa\chi\}^{(2-d)/2} \frac{\{\Lambda(\kappa\chi)^{1/2}\}^{d-4}}{d-4} \\ &= (A - A_c) - \frac{3BC_d k_B T_c^*}{(2\pi)^d \kappa^2} \frac{\Lambda^{d-4}}{d-4} \chi^{-1} = \frac{k_B}{a^d s^2}(T - T_c^*) - \frac{C_d z^2}{2(2\pi)^d} \frac{T_c^*}{T_c} \frac{(\Lambda a)^{d-4}}{d-4} \chi^{-1} \end{aligned} \tag{5.104}$$

$$\Rightarrow \qquad \chi(T) = \frac{a^d s^2}{k_B(T - T_c^*)} \left\{ 1 + \frac{C_d z^2}{2(2\pi)^d} \frac{T_c^*}{T_c} \frac{(\Lambda a)^{d-4}}{d-4} \right\}^{-1} \propto |T - T_c^*|^{-1} .$$

We observe that this corresponds to the behavior found in mean field calculation. Indeed the above the critical dimension d_c the mean field exponents are correct, and the only effect of fluctuations is to renormalize certain quantities such as the critical temperature.

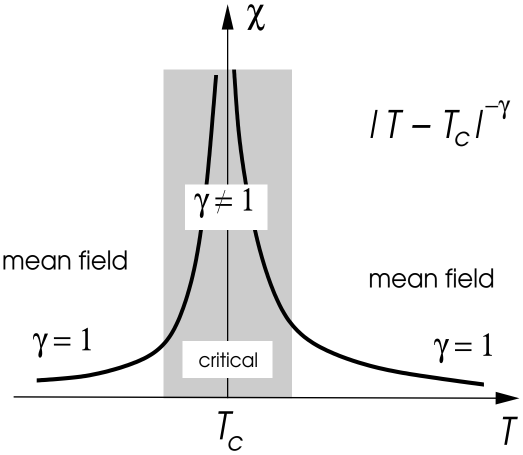


Fig. 5.7: Critical region.

Next we consider $d < d_c = 4$. In this case the integral over x in (5.103) converges and the upper bound does not enter in the limit $T \to T_c^*$. The integral only depends on d and is given by $K_d = \Gamma[(d-2)/2]\Gamma[(4-d)/2]$. Therefore, we obtain

$$\chi^{-1} = (A - A_c) - \frac{3BC_d k_B T_c^*}{(2\pi)^d \kappa^{d/2}} K_d \chi^{(2-d)/2} \tag{5.105}$$

This equation has two regimes. For large χ the second term on the right hand side is dominating over χ^{-1}. On the other hand, if χ is small, χ^{-1} is dominant. The first case applies for T very close to T_c^* such that we arrive at the behavior

$$\chi(T) \propto |T - T_c^*|^{-\gamma} \qquad \text{with} \quad \gamma = \frac{2}{d-2} . \tag{5.106}$$

Away from T_c^* the second case holds and we return back to the usual mean field behavior,

$$\chi(T) \propto |T - T_c^*|^{-1} . \tag{5.107}$$

The crossover happens roughly at the boundary defined by the Ginzburg criterion. It is now also obvious that the critical exponent γ depends on the dimension. The critical dimensions $d_L = 2$ and $d_c = 4$ need a special care in this discussion as the expression we have obtained do not straightforwardly apply to them.

We can now also use this result to determine the exponent of the correlation length, ν. We may use $\xi^2 = \kappa\chi$ such that we find immediately

$$\nu = \frac{\gamma}{2} \; . \tag{5.108}$$

5.5 Long-range order - Peierls' argument

We have seen in the previous section that the critical temperature is reduced from its mean field value by quantum fluctuations. The lower the dimension (the coordination number) the more severe thermal fluctuations act against order. We consider here the case of the Ising model in one and two dimensional order. In one dimension no magnetic order is possible at any finite temperature, while in two dimensions there is a finite-temperature phase transition.

5.5.1 Absence of finite-temperature phase transition in the 1D Ising model

We consider the groundstate of the ferromagnetic Ising chain. All spins are aligned. The lowest energy excitation is a domain wall for which the spins point up on the left and down on the right of the domain wall. The energy for such a wall is $2Js^2$. With N spins on the Ising chain, there would be $N-1$ positions possible for this wall and would yield an entropy $S = k_B \ln(N-1)$ accounting for the uncertainty of the configuration of the state. Therefore a simple estimate of the free energy for the presence of a domain wall leads to

$$\Delta F = 2Js^2 - k_B T \ln(N-1) \; . \tag{5.109}$$

Taking now the thermodynamic limit $N \to \infty$ would yield for any finite temperature a negative ΔF such that the ordered phase (without domain wall) would be unfavorable.

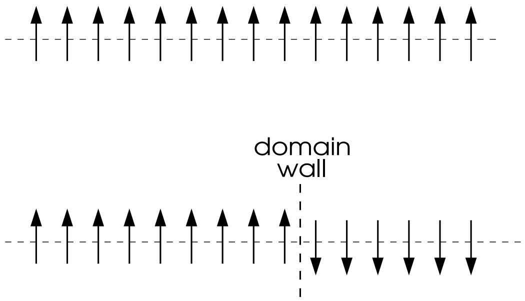


Fig. 5.8: Domain wall as the lowest excitation.

In the one-dimensional Ising chain long-range order exists only at zero temperature analogous to the classical spin chain discussed earlier.

5.5.2 Long-range order in the 2D Ising model

The argument given by Peierls to proof long-range order for the 2D Ising model is also based on domains. We consider a square lattice and assume that all spins at the boundary are pointing up. This requirement is not relevant, if the thermodynamic limit is considered, i.e. the number of site or spins $N \to \infty$. However, it implies that all domain walls on the lattice are closed. The energy cost for a domain wall of the length L compared to the completely aligned configuration is

$$E(L) = 2Js^2 L \; . \tag{5.110}$$

Now choose a site j somewhere in the interior of the system. We define $P_\pm$ as the probability that this spin s_j is $\pm s$. Thus the mean value of the spin is

$$\langle s_j \rangle = s(P_+ - P_-) \tag{5.111}$$

where in the case of long range order $\langle s_j \rangle > 0$ due to the boundary condition. Moreover the correlation function satisfies for $|i-j| \to \infty$,

$$\langle s_i s_j \rangle \to \langle s_i \rangle \langle s_j \rangle = s^2 (P_+ - P_-)^2 \tag{5.112}$$

which is finite for the long-range ordered system, if $P_+ \neq P_-$. Thus our proof has to address this point.
In case $s_j = -s$ we find an odd number of domain walls encircling the site j. The relative probability that the shortest domain wall around j (closest to site j) has the length L is given by

$$P_{L,j} = \frac{e^{-KL}}{Z} \sum_{L'} e^{-KL'} \tag{5.113}$$

where the sum $\sum_{L'}$ runs over all other domain walls and $K = 2\beta J s^2$. If we consider $s_j = +s$ this shortest domain wall is absent. Thus we may write

$$P_+ = \frac{1}{Z} \sum_{L'} e^{-KL'} \qquad \text{and} \qquad P_- = \sum_L g(L) e^{-KL} P_+ \ . \tag{5.114}$$

Here $g(L)$ is the number of domain wall shapes of length L around j.

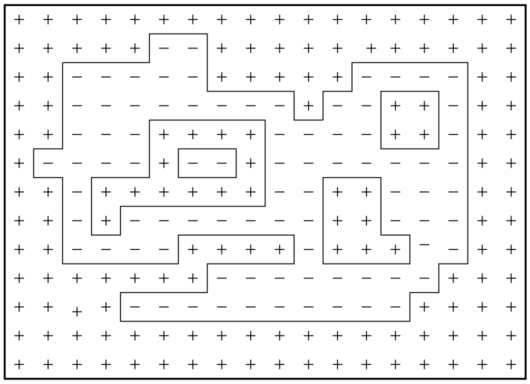


Fig. 5.9: Domain wall configurations.

From topological point of view the domain wall should not cut each other. If we ignore this condition and in addition the requirement that the domain wall is closed, we can derive a simple upper bound for $g(L)$,

$$g(L) < 4 \times 3^{L-1} \times \frac{1}{L} \times \frac{1}{2} \times \left(\frac{L}{4}\right)^2 = \frac{L^2}{24L} e^{L \ln 3} \ . \tag{5.115}$$

which can be understood as a walk of length L through the lattice. From the starting point, we may start in 4 directions and every following step has 3 possible directions. On the contour there L equivalent starting points (assuming close contour) and 2 directions. Then there are maximally $(L/4)^2$ sites enclosed by the domain and the starting point may be shift on any of them (for given circumference the square has the largest area among all rectangles). Therefore we can write

$$P_- < \sum_{L \geq 4} \frac{L^2}{24L} e^{L(\ln 3 - K)} P_+ < \sum_{L=4,6,8,\ldots} \frac{L^2}{96} e^{L(\ln 3 - K)} \tag{5.116}$$

since $P_+ < 1$. For sufficiently low temperature $\ln 3 - K = \ln 3 - 2Js^2\beta < 0$ so that the sum converges very well. Making the temperature small enough also the condition $P_- < \frac{1}{2}$ can be reached. With the condition $P_+ + P_- = 1$ follows that

$$P_+ - P_- > 0 \qquad \Rightarrow \qquad \lim_{|i-j| \to \infty} \langle s_i s_j \rangle = const. \tag{5.117}$$

which means that we have long- range order at a finite temperature.

Chapter 6

Linear Response Theory

The observation of its response to a small external perturbation, such as an external field, provides much information about a macroscopic sytem. If the perturbation is sufficiently small we can consider the response of the system in lowest order only, as a so-called linear response. We already are familiar with this from magnetism where the magnetization is proportional to the applied magnetic field. This susceptibility as a linear response function incorporates important information about the macroscopic system when we take dynamical aspects into account. Thus, we will extend our discussion to the dynamical linear response function looking at the response to spatial and time dependent perturbations. If we knew all stationary states of a macroscopic quantum system with many degrees of freedom we would in principle be ready to calculate any desired quantity. As we mentioned earlier this full information is hard to store and is also unnecessary in view of our real experimental interests. The linear response functions are an efficient way to provide in a condensed form the most important and relevant information of the system accessible in an experiment. The linear response function is one element of quantum field theory of solid state physics. We will introduce it here on an elementary level.

6.1 Linear Response function

Some of the well-known examples of linear response functions are the dielectric and the magnetic susceptibilities which describe the dielectric and magnetic response to an external field, respectively.

$$\begin{aligned} \vec{P}(\vec{r},t) &= \int d^3r' \int dt' \; \chi_e(\vec{r}-\vec{r}\,',t-t')\vec{E}(\vec{r}\,',t') \\ \vec{M}(\vec{r},t) &= \int d^3r' \int dt' \; \chi_m(\vec{r}-\vec{r}\,',t-t')\vec{H}(\vec{r}\,',t') \end{aligned} \tag{6.1}$$

where we consider here a homogeneous, isotropic medium. It is obvious from this form that the response functions χ describes how a field at the position $\vec{r}\,'$ at time t' influences the system at $\vec{r}$ at a later time t (causality). Causality actually requires that $\chi(\vec{r},t) = 0$ for $t < 0$. The response functions are non-local in space and time. We may get a "local" form by going to momentum-frequency space (convolution).

$$\vec{P}(\vec{q},\omega) = \chi_e(\vec{q},\omega)\vec{E}(\vec{q},\omega) \quad \text{and} \quad \vec{M}(\vec{q},\omega) = \chi_m(\vec{q},\omega)\vec{H}(\vec{q},\omega) \tag{6.2}$$

where the Fourier transformation is performed as follows,

$$f(\vec{r},t) = \frac{1}{V}\sum_{\vec{q}} \int_{-\infty}^{+\infty} \frac{d\omega}{2\pi} f(\vec{q},\omega) e^{-i(\omega t - \vec{q}\cdot\vec{r})} \; . \tag{6.3}$$

We now determine the response function for a general external field and response quantity.

6.1.1 Kubo formula - retarded Green's function

We consider a system described by the Hamiltonian $\mathcal{H}_0$ and analyze its response to an external field $h(\vec{r},t)$ which couples to the field operator $\hat{A}(\vec{r})$,

$$\mathcal{H} = \mathcal{H}_0 + \mathcal{H}'(t) = \mathcal{H}_0 + \int d^3r\ \hat{A}(\vec{r})h(\vec{r},t)e^{\eta t} \tag{6.4}$$

where $\eta = 0_+$ is a small positive parameter allowing to switch the perturbation adiabatically on. The behavior of the system can now be determined by the density matrix $\hat{\rho}(t)$. Possessing $\hat{\rho}$ we are able to calculate interesting mean values of operators, $\langle \hat{B}(t)\rangle = tr(\hat{\rho}(t)\hat{B})$. We find the density matrix by means of the equation of motion,

$$i\hbar\frac{\partial\hat{\rho}}{\partial t} = -[\hat{\rho},\mathcal{H}] = -[\hat{\rho},\mathcal{H}_0+\mathcal{H}'] \ . \tag{6.5}$$

We proceed using the concept of perturbation theory, $\hat{\rho} = \hat{\rho}_0 + \delta\hat{\rho}(t)$, with

$$\hat{\rho}_0 = \frac{1}{Z}e^{-\beta\mathcal{H}_0} \qquad \text{and} \qquad Z = tre^{-\beta\mathcal{H}_0} \tag{6.6}$$

Then we insert this separated form and truncate the equation in linear order in $\mathcal{H}'$,

$$i\hbar\frac{\partial}{\partial t}\delta\hat{\rho} = -[\delta\hat{\rho},\mathcal{H}_0] - [\hat{\rho}_0,\mathcal{H}'] + \cdots \tag{6.7}$$

We introduce now the interaction representation (time-dependent perturbation theory),

$$\delta\hat{\rho}(t) = e^{-i\mathcal{H}_0 t/\hbar}\hat{y}(t)e^{i\mathcal{H}_0 t/\hbar} \quad \Rightarrow \quad i\hbar\frac{\partial}{\partial t}\delta\hat{\rho} = -[\delta\hat{\rho},\mathcal{H}_0] + e^{-i\mathcal{H}_0 t/\hbar}\left\{i\hbar\frac{\partial\hat{y}(t)}{\partial t}\right\}e^{i\mathcal{H}_0 t/\hbar} \ . \tag{6.8}$$

Comparing Eqs. (6.7) and (6.8) and using (6.5) we arrive at the equation for $\hat{y}$,

$$i\hbar\frac{\partial\hat{y}(t)}{\partial t} = -[\hat{\rho}_0,\mathcal{H}'_{int}(t)] \qquad \text{with} \qquad \mathcal{H}'_{int}(t) = e^{i\mathcal{H}_0 t/\hbar}\mathcal{H}'e^{-i\mathcal{H}_o t/\hbar} \tag{6.9}$$

which is formally solved by

$$\hat{y}(t) = \frac{i}{\hbar}\int_{-\infty}^{t} dt'\ [\hat{\rho}_0,\mathcal{H}'_{int}(t')] \ . \tag{6.10}$$

We now look at the mean value of the observable $\hat{B}(\vec{r})$. For simplicity we assume that the expectation value of $\hat{B}$ vanishes if there is no perturbation, i.e. $\langle\hat{B}\rangle_0 = tr\{\hat{\rho}_0\hat{B}\} = 0$. We determine

$$\begin{aligned}\langle\hat{B}(\vec{r},t)\rangle &= tr\left\{\delta\hat{\rho}(\vec{r},t)\hat{B}(\vec{r})\right\} \\ &= tr\left\{\frac{i}{\hbar}e^{-i\mathcal{H}_0 t/\hbar}\int_{-\infty}^{t} dt'\ [\hat{\rho}_0,\mathcal{H}'_{int}(t')]e^{i\mathcal{H}_0 t/\hbar}\hat{B}(\vec{r})\right\} \ .\end{aligned} \tag{6.11}$$

By means of cyclic permutation which does not affect the trace we arrive at the form

$$\begin{aligned}\langle\hat{B}(\vec{r},t)\rangle &= -\frac{i}{\hbar}\int_{-\infty}^{t} dt' \int d^3r'\ tr\left\{\hat{\rho}_0[\hat{B}_{int}(\vec{r},t),\hat{A}_{int}(\vec{r}\,',t')]\right\}h(\vec{r}\,',t')e^{\eta t'} \\ &= \int dt' \int d^3r'\ \chi_{BA}(\vec{r}-\vec{r}\,',t-t')h(\vec{r}\,',t') \ ,\end{aligned} \tag{6.12}$$

which defines the response function. Notably, it is entirely determined by the properties of the unperturbed system.

Recipe for the linear response function: We arrive at the following recipe to obtain a general linear response function: From now on we denote the Hamiltonian of the (unperturbed) system

$\mathcal{H}$. Then the linear response function of the pair of operators $\hat{A}, \hat{B}$ (they are often in practice conjugate operators, $\hat{A} = \hat{B}^\dagger$) is given by

$$\chi_{BA}(\vec{r}-\vec{r}\,', t-t') = -\frac{i}{\hbar}\Theta(t-t')\langle[\hat{B}_H(\vec{r},t), \hat{A}_H(\vec{r}\,',t')]\rangle_{\mathcal{H}} \tag{6.13}$$

where $\langle\ldots\rangle_{\mathcal{H}}$ is the thermal mean value with respect to the Hamiltonian $\mathcal{H}$,

$$\langle\hat{C}\rangle_{\mathcal{H}} = \frac{tr\{\hat{C}e^{-\beta\mathcal{H}}\}}{tr\{e^{-\beta\mathcal{H}}\}} \ , \tag{6.14}$$

$\hat{A}_H(t) = e^{i\mathcal{H}t/\hbar}\hat{A}e^{-i\mathcal{H}t/\hbar}$ is the Heisenberg representation of the operator $\hat{A}$ (analog for $\hat{B}$). Note that the temporal step function $\Theta(t-t')$ ensures the causality, i.e. there is no response for the system before there is a perturbation. The form (6.13) is often called *Kubo formula* or *retarded Green's function.*

Frequently used examples are:

- *magnetic susceptibility:*

perturbation $\quad \mathcal{H}' = -\int d^3r\ \mu_B\hat{S}^z(\vec{r})h(\vec{r},t)$

conjugate magnetization $\quad \hat{M}(\vec{r}) = \mu_B\hat{S}^z(\vec{r})$

response function

$$\chi_{zz}(\vec{r}-\vec{r}\,', t-t') = \frac{i}{\hbar}\Theta(t-t')\langle[\hat{S}^z_H(\vec{r},t), \hat{S}^z_H(\vec{r}\,',t')]\rangle_{\mathcal{H}} \ . \tag{6.15}$$

- *dielectric susceptibility:*

$$\begin{aligned}
&\text{perturbation} && \mathcal{H}' = \int d^3r\ e\hat{n}(\vec{r})\phi(\vec{r},t) \\
&\text{conjugate density} && e\hat{n}(\vec{r}) \\
&\text{response function} && \chi_e(\vec{r}-\vec{r}\,', t-t') = -\frac{i}{\hbar}\Theta(t-t')\langle[\hat{n}_H(\vec{r},t), \hat{n}_H(\vec{r}\,',t')]\rangle_{\mathcal{H}} \ .
\end{aligned} \tag{6.16}$$

6.1.2 Information in the response function

The information stored in the response function can be most easily visualized by assuming that we know the complete set of stationary states of the system Hamiltonian $\mathcal{H}$: $\mathcal{H}|n\rangle = \epsilon_n|n\rangle$. For simplicity we will from now on assume that $\hat{A} = \hat{B}^\dagger$ which is the case in many practical examples, and will simplify our notation. We can then rewrite the response function χ as

$$\begin{aligned}
\chi(\vec{r}-\vec{r}\,', t-t') &= -\frac{i}{\hbar}\Theta(t-t')\sum_n \frac{e^{-\beta\epsilon_n}}{Z}\Big\{\langle n|e^{i\mathcal{H}t/\hbar}\hat{B}(\vec{r})e^{-i\mathcal{H}t/\hbar}e^{i\mathcal{H}t'/\hbar}\hat{B}(\vec{r}\,')^\dagger e^{-i\mathcal{H}t'/\hbar}|n\rangle \\
&\qquad -\langle n|e^{i\mathcal{H}t'/\hbar}\hat{B}(\vec{r}\,')^\dagger e^{-i\mathcal{H}t'/\hbar}e^{i\mathcal{H}t/\hbar}\hat{B}(\vec{r})e^{-i\mathcal{H}t/\hbar}|n\rangle\Big\} \\
&= -\frac{i}{\hbar}\Theta(t-t')\sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z}\Big\{\langle n|\hat{B}(\vec{r})|n'\rangle\langle n'|\hat{B}(\vec{r}\,')^\dagger|n\rangle e^{i(\epsilon_n-\epsilon_{n'})(t-t')/\hbar} \\
&\qquad -\langle n|\hat{B}(\vec{r}\,')^\dagger|n'\rangle\langle n'|\hat{B}(\vec{r})|n\rangle e^{i(\epsilon_{n'}-\epsilon_n)(t-t')/\hbar}\Big\} \ .
\end{aligned} \tag{6.17}$$

It is convenient to work in momentum and frequency space. Thus, we perform now the Fourier transform

$$\begin{aligned}
\chi(\vec{q},\omega) &= \int d^3\tilde{r} \int_{-\infty}^{+\infty} d\tilde{t}\ \chi(\tilde{\vec{r}},\tilde{t}) e^{i\omega\tilde{t} - i\vec{q}\cdot\tilde{\vec{r}}} \\
&= -\frac{i}{\hbar}\sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z} |\langle n|\hat{B}_{\vec{q}}|n'\rangle|^2 \int_0^\infty d\tilde{t}\ \left\{ e^{i(\epsilon_n - \epsilon_{n'} + \hbar\omega)\tilde{t}/\hbar} - e^{i(\epsilon_{n'} - \epsilon_n + \hbar\omega)\tilde{t}/\hbar} \right\} e^{-\eta\tilde{t}}
\end{aligned} \tag{6.18}$$

where we introduce

$$\hat{B}_{\vec{q}} = \int d^3\tilde{r} \hat{B}(\tilde{\vec{r}}) e^{-i\vec{q}\cdot\tilde{\vec{r}}} \ . \tag{6.19}$$

with $\hat{B}_{\vec{q}} = \hat{B}^\dagger_{-\vec{q}}$ due to the inversion symmetry ($\vec{r} \to -\vec{r}$), i.e. $\chi(\vec{r}-\vec{r}\,', t-t') = \chi(\vec{r}\,'-\vec{r}, t-t')$. Performing the time integral in (6.18) we obtain

$$\begin{aligned}
\chi(\vec{q},\omega) &= \sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z} |\langle n|\hat{B}_{\vec{q}}|n'\rangle|^2 \left\{ \frac{1}{\hbar\omega - \epsilon_{n'} + \epsilon_n + i\hbar\eta} - \frac{1}{\hbar\omega - \epsilon_n + \epsilon_{n'} + i\hbar\eta} \right\} \\
&= \int_0^\infty d\omega'\ S(\vec{q},\omega') \left\{ \frac{1}{\omega - \omega' + i\eta} - \frac{1}{\omega + \omega' + i\eta} \right\} \ .
\end{aligned} \tag{6.20}$$

In the last line we write the response function in a spectral form with $S(\vec{q},\omega)$ as the spectral function,

$$S(\vec{q},\omega) = \sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z} |\langle n|\hat{B}_{\vec{q}}|n'\rangle|^2 \delta(\hbar\omega - \epsilon_{n'} + \epsilon_n) \tag{6.21}$$

We call $S(\vec{q},\omega)$ also *dynamical structure factor*. It represents a correlation function[1] ,

$$S(\vec{r}-\vec{r}\,', t-t') = \frac{1}{h} \left\langle \hat{B}_H(\vec{r},t) \hat{A}_H(\vec{r}\,',t')^\dagger \right\rangle_{\mathcal{H}} \ , \tag{6.23}$$

and contains the spectrum of the excitations which can be coupled to by the external perturbation.

6.1.3 Analytical properties

The representation of the linear response function in (6.20) shows that $\chi_{BA}(\vec{q},\omega)$ has poles only in the lower half of the complex ω-plane. This property reflects causality ($\chi(\vec{r},t) = 0$ for $t < 0$. We separate now $\chi = \chi' + i\chi''$ in real and imaginary part and use the relation

$$\lim_{\eta\to 0_+} \frac{1}{x + i\eta} = \mathcal{P}\frac{1}{x} - i\pi\delta(x) \ . \tag{6.24}$$

---

[1] Consider the Fourier transform

$$\begin{aligned}
S(\vec{q},\omega) &= \int d^3\tilde{r} \int_{-\infty}^{+\infty} dt S(\tilde{\vec{r}},t) e^{i\omega t - i\vec{q}\cdot\tilde{\vec{r}}} \\
&= \frac{1}{Vh} \int d^3\tilde{r} d^3r' \int_{-\infty}^{+\infty} dt \sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z} \langle n|e^{i\mathcal{H}t/\hbar} \hat{B}(\tilde{\vec{r}} + \vec{r}\,') e^{-i\mathcal{H}t/\hbar}|n'\rangle\langle n'|\hat{B}(\vec{r}\,')^\dagger|n\rangle e^{-i\vec{q}\cdot(\tilde{\vec{r}}+\vec{r}\,') - \vec{q}\,'\cdot\vec{r}\,' + i\omega t} \\
&= \frac{1}{h} \int_{-\infty}^{+\infty} dt \sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z} |\langle n|\hat{B}_{\vec{q}}|n'\rangle|^2 e^{i(\epsilon_n - \epsilon_{n'} + \hbar\omega)t/\hbar} \\
&= \sum_{n.,n'} \frac{e^{-\beta\epsilon_n}}{Z} |\langle n|\hat{B}_{\vec{q}}|n'\rangle|^2 \delta(\hbar\omega - \epsilon_{n'} + \epsilon_n) \ ,
\end{aligned} \tag{6.22}$$

as gven in (6.21).

with $\mathcal{P}$ denoting the principal part. This relation leads to

$$\begin{aligned} \chi'(\vec{q},\omega) &= \int_0^\infty d\omega'\; S(\vec{q},\omega') \left\{ \mathcal{P}\frac{1}{\omega-\omega'} - \mathcal{P}\frac{1}{\omega+\omega'} \right\} \;, \\ \chi''(\vec{q},\omega) &= -\pi \left\{ S(\vec{q},\omega) - S(\vec{q},-\omega) \right\} \;. \end{aligned} \tag{6.25}$$

Therefore the imaginary part of χ corresponds to the excitation spectrum of the system. Finally, it has to be noted that $\chi(\vec{q},\omega)$ follows the Kramers-Kronig relations:[2]

$$\begin{aligned} \chi'(\vec{q},\omega) &= -\frac{1}{\pi} \int_{-\infty}^{+\infty} d\omega'\; \mathcal{P}\frac{\chi''(\vec{q},\omega')}{\omega-\omega'} \;, \\ \chi''(\vec{q},\omega) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} d\omega'\; \mathcal{P}\frac{\chi'(\vec{q},\omega')}{\omega-\omega'} \;. \end{aligned} \tag{6.27}$$

6.1.4 Fluctuation-Dissipation theorem

First we consider the aspect of dissipation incorporated in the response function. For this purpose we ignore for simplicity the spatial dependence and consider a perturbative part of the Hamiltonian which only depends on time.

$$\mathcal{H}' = h(t)\hat{B} \qquad \Rightarrow \qquad \langle \hat{B}(t) \rangle = \int_0^\infty dt'\; \chi(t-t') h(t') \tag{6.28}$$

with $\hat{B} = \hat{B}^\dagger$. We assume now a monochromatic external field,

$$h(t) = \frac{1}{2}(h_0 e^{-i\omega t} + h_0^* e^{i\omega t})$$

$$\Rightarrow \qquad \langle \hat{B}(t) \rangle = \int_0^\infty dt' \chi(t-t') \frac{1}{2}(h_0 e^{-i\omega t'} + h_0^* e^{i\omega t'}) = \frac{1}{2}\left\{ \chi(\omega)^* h_0 e^{-i\omega t} + \chi(\omega) h_0^* e^{i\omega t} \right\} \;. \tag{6.29}$$

---

[2] *Kramers-Kronig relation:* This relation results from the analytic structure of χ. Consider a contour in the upper half-plane of ω' where $\chi(\vec{q},\omega')$ has no poles due to causality.

$$\begin{aligned} &\oint_C d\omega' \frac{\chi(\vec{q},\omega')}{\omega-\omega'-i\eta'} = 0 \\ &\Rightarrow \quad \int_{-\infty}^{+\infty} d\omega'\; \chi(\vec{q},\omega')\mathcal{P}\tfrac{1}{\omega-\omega'} + i\pi\chi(\vec{q},\omega) = 0 \;. \end{aligned} \tag{6.26}$$

Separating this equation into real and imaginary part yields the Kramers-Kronig relation.

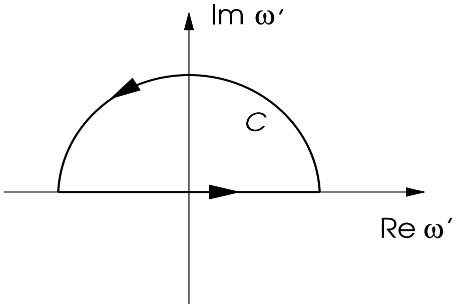


Fig. 6.1: Contour C close in the upper half of the ω′-plane.

The energy dissipation rate is determined by[3]

$$\frac{dE}{dt} = \langle \frac{\partial}{\partial t}\mathcal{H}\rangle = \langle \hat{B}\rangle \frac{\partial h}{\partial t} = \frac{i\omega}{4}[\chi(\omega)^* - \chi(\omega)]|h_0|^2 = \frac{\omega}{2}\chi''(\omega)|h_0|^2 < 0 \tag{6.31}$$

where we dropped oscillating terms with the time dependence $e^{\pm i2\omega t}$. The imaginary part of the dynamical susceptibility describes the dissipation of the system.
From the definition of the dynamical structure factor it follows that

$$S(\vec{q},-\omega) = e^{-\beta\hbar\omega} S(\vec{q},\omega) \tag{6.32}$$

because

$$\begin{aligned} S(\vec{q},-\omega) &= \sum_{n,n'} \frac{e^{-\beta\epsilon_n}}{Z} |\langle n|\hat{B}_{\vec{q}}|n'\rangle|^2 \delta(-\hbar\omega - \epsilon_{n'} + \epsilon_n) \\ &= \sum_{n,n'} \frac{e^{-\beta\epsilon_{n'}-\beta\hbar\omega}}{Z} |\langle n'|\hat{B}_{-\vec{q}}|n\rangle|^2 \delta(\hbar\omega - \epsilon_n + \epsilon_{n'}) = e^{-\beta\hbar\omega} S(\vec{q},\omega) \; . \end{aligned} \tag{6.33}$$

This is a statement of detailed balance. The transition matrix element between two states is the same whether the energy is absorbed or emitted. For emitting, however, the thermal occupation of the initial state has to be taken into account.
Using (6.25) we can derive the following relation

$$\chi''(\vec{q},\omega) = -\pi[1 - e^{-\beta\hbar\omega}] S(\vec{q},\omega) \tag{6.34}$$

which is known as the *fluctuation-dissipation theorem*. Let us consider here some consequences and find the relation to our earlier simplified formulations.

$$\int_{-\infty}^{+\infty} d\omega \; S(\vec{q},\omega) = \langle \hat{B}_{\vec{q}}(t)\hat{B}_{-\vec{q}}(t)\rangle = -\frac{1}{\pi}\int_{-\infty}^{+\infty} d\omega \frac{\chi''(\vec{q},\omega)}{1 - e^{-\beta\hbar\omega}} \tag{6.35}$$

This corresponds to the equal-time correlation function (assuming $\langle \hat{B}\rangle = 0$). Now we turn towards the static limit which we consider as the limit $k_B T \gg \hbar\omega$. Then we may approximate this equation by

$$\langle |\hat{B}_{\vec{q}}|^2\rangle \approx -\frac{k_B T}{\pi}\int_{-\infty}^{+\infty} d\omega \frac{\chi''(\vec{q},\omega)}{\omega} = k_B T \chi'(\vec{q},0) = k_B T \chi(\vec{q},0) \; . \tag{6.36}$$

This is valid, if $\chi''(\vec{q},\omega)$ essentially vanishes for frequencies comparable and larger than the temperature. This leads us now to

$$\int d^3r d^3r' \; \langle \hat{B}(\vec{r},t)\hat{B}^\dagger(\vec{r}',t)\rangle = k_B T \chi(\vec{q} = 0, \omega = 0) = k_B T \chi \; , \tag{6.37}$$

i.e. the static uniform susceptibility is related to the integration of the equal-time correlation function as we had used previously several times.

---

[3] The time-derivative of the Hamiltonian is given by

$$\frac{d\mathcal{H}}{dt} = \frac{\partial\mathcal{H}}{\partial t} + \frac{i}{\hbar}[\mathcal{H},\mathcal{H}] = \frac{\partial\mathcal{H}}{\partial t} \tag{6.30}$$

for a quantum mechanical problem. The analogous relation is obtained for classical systems.

6.2 Example - Heisenberg ferromagnet

In this section we apply the linear response function to discuss the ferromagnetic Heisenberg quantum spin model with $s = \hbar/2$. Like in the Ising model the spins reside on a hypercubic lattice and are coupled via nearest-neighbor interaction. The Heisenberg model has the full $SU(2)$ spin rotation symmetry.

$$\begin{aligned} \mathcal{H} &= -J \sum_{\langle i,j \rangle} \hat{\vec{S}}_i \cdot \hat{\vec{S}}_j = -J \sum_{\langle i,j \rangle} \left[\hat{S}_i^z \hat{S}_j^z + \frac{1}{2} \left\{ \hat{S}_i^+ \hat{S}_j^- + \hat{S}_i^- \hat{S}_j^+ \right\} \right] \\ &= -\frac{J}{V} \sum_{\vec{q}} \gamma_{\vec{q}} \hat{\vec{S}}_{\vec{q}} \cdot \hat{\vec{S}}_{-\vec{q}} = -\frac{J}{V} \sum_{\vec{q}} \gamma_{\vec{q}} \left[\hat{S}_{\vec{q}}^z \hat{S}_{-\vec{q}}^z + \frac{1}{2} \left\{ \hat{S}_{\vec{q}}^+ \hat{S}_{-\vec{q}}^- + \hat{S}_{\vec{q}}^- \hat{S}_{-\vec{q}}^+ \right\} \right] \end{aligned} \tag{6.38}$$

with $\gamma_{\vec{q}} = 2 \sum_{\alpha = x,y,\ldots} \cos(q_\alpha)$ (lattice constant $a = 1$, unit of length). The exchange coupling $-J$ is negative so that the ground state is ferromagnetic (all spins aligned). This ground state has maximal spin $\hat{\vec{S}}_{tot} = \sum_i \hat{\vec{S}}_i$ which is a symmetry operator of the Hamiltonian. In the second line we have introduced the Fourier transformed Hamiltonian with

$$\hat{\vec{S}}_i = \frac{1}{V} \sum_{\vec{q}} \hat{\vec{S}}_{\vec{q}} e^{i \vec{q} \cdot \vec{r}_i} \tag{6.39}$$

where we use again periodic boundary condition with a hypercube of edge length L. The quantum spins satisfy the following commutation relations:

$$\begin{aligned} \left[\hat{S}_i^z, \hat{S}_j^\pm \right] &= \pm \hbar \delta_{ij} \hat{S}_i^\pm \ , \quad \left[\hat{S}_i^+, \hat{S}_j^- \right] = 2\hbar \delta_{ij} \hat{S}_i^z \ , \\ \left[\hat{S}_{\vec{q}}^z, \hat{S}_{\vec{q}'}^\pm \right] &= \pm \hbar \hat{S}_{\vec{q}+\vec{q}'}^\pm \ , \quad \left[\hat{S}_{\vec{q}}^+, \hat{S}_{\vec{q}'}^- \right] = 2\hbar \hat{S}_{\vec{q}+\vec{q}'}^z \ . \end{aligned} \tag{6.40}$$

It is possible to discuss the ferromagnetic state by means of mean field approximation which is formally the same as for the Ising model leading to the same critical temperature $k_B T_c = Jzs^2 = Jd\hbar^2/2$.

6.2.1 Tyablikov decoupling approximation

Here we would like to go beyond mean field theory using the response function for the Heisenberg model. We introduce the transverse dynamical spin susceptibility,

$$\chi(\vec{q}, t - t') = -\frac{i}{\hbar} \Theta(t - t') \langle [\hat{S}_{\vec{q}}^-(t), \hat{S}_{-\vec{q}}^+(t')] \rangle \ . \tag{6.41}$$

We use now a method based on the equation of motion to find this response function.

$$\begin{aligned} i\hbar \frac{\partial}{\partial t} \chi(\vec{q}, t - t') &= \delta(t - t') \langle [\hat{S}_{\vec{q}}^-, \hat{S}_{-\vec{q}}^+] \rangle - \frac{i}{\hbar} \Theta(t - t') \langle [[\hat{S}_{\vec{q}}^-, \mathcal{H}](t), \hat{S}_{-\vec{q}}^+(t')] \rangle \\ &= -2\delta(t - t') \hbar \langle \hat{S}_{\vec{q}=0}^z \rangle \\ &\qquad + \frac{i}{\hbar} \Theta(t - t') \frac{J\hbar}{V} \sum_{\vec{q}'} (\gamma_{\vec{q}'} - \gamma_{\vec{q}+\vec{q}'}) \left\{ \left\langle \left[\hat{S}_{\vec{q}'}^z(t) \hat{S}_{\vec{q}-\vec{q}'}^-(t), \hat{S}_{-\vec{q}}^+(t') \right] \right\rangle \right. \\ &\qquad\qquad \left. + \langle \left[\hat{S}_{\vec{q}+\vec{q}'}^-(t) \hat{S}_{-\vec{q}'}^z(t), \hat{S}_{-\vec{q}}^+(t') \right] \rangle \right\} \ . \end{aligned} \tag{6.42}$$

This equation leads on the right-hand side to new retarded Green's functions containing three spin operators. Thus we would have also to set up equations of motion for them which would in turn lead to Green's functions with four spin operators and so on. This hierarchy of equation is intractable. Therefore we introduce here the approximation which truncates the hierarchy and leads to a closed equation system,

$$
\begin{aligned}
\hat{S}^z_{\vec{q}'}(t)\hat{S}^-_{\vec{q}-\vec{q}'}(t) &\to \langle \hat{S}^z_0 \rangle \hat{S}^-_{\vec{q}}(t)\delta_{0,\vec{q}'} \\
\hat{S}^-_{\vec{q}+\vec{q}'}(t)\hat{S}^z_{\vec{q}'}(t) &\to \langle \hat{S}^z_0 \rangle \hat{S}^-_{\vec{q}}(t)\delta_{0,\vec{q}'}
\end{aligned}
\tag{6.43}
$$

Replacing the corresponding terms on the right-hand side of the equation of motion and performing the Fourier transform for the time we obtain the closed equation

$$
(\omega + i\eta)\chi(\vec{q},\omega) = -2\langle \hat{S}^z_0 \rangle - 2\frac{J}{V}\langle \hat{S}^z_0 \rangle (\gamma_0 - \gamma_{\vec{q}})\chi(\vec{q},\omega) \tag{6.44}
$$

which leads to the solution

$$
\chi(\vec{q},\omega) = \frac{-2\langle \hat{S}^z_0 \rangle}{\omega + 2\frac{J}{V}\langle \hat{S}^z_0 \rangle (\gamma_0 - \gamma_{\vec{q}}) + i\eta} \, . \tag{6.45}
$$

Here we define now the mean field $\langle \hat{S}^z_0 \rangle = -Vm$ with $0 \leq m \leq \hbar/2$. We have introduced the small imaginary part $i\eta$ in order to guarantee the causality.
First let us look at the spectrum of excitations if $m \neq 0$. This is given by the imaginary part of χ.

$$
\begin{aligned}
\chi''(\vec{q},\omega) &= -2\pi V m \delta[\omega - 2Jm(\gamma_0 - \gamma_{\vec{q}})] \\
\Rightarrow \quad \omega_{\vec{q}} &= 2Jm(\gamma_0 - \gamma_{\vec{q}}) = 4Jm\sum_{\alpha}(1 - \cos q_\alpha)
\end{aligned}
\tag{6.46}
$$

This is the dispersion of a collective magnetic mode, a *magnon* or spin wave, which corresponds analogous to the phonons in a solid to a bosonic mode. This mode appears as an excitation with well-defined momentum-energy relation. For small energy the dispersion is quadratic, $\omega_{\vec{q}} = 2Jm\vec{q}^2$.

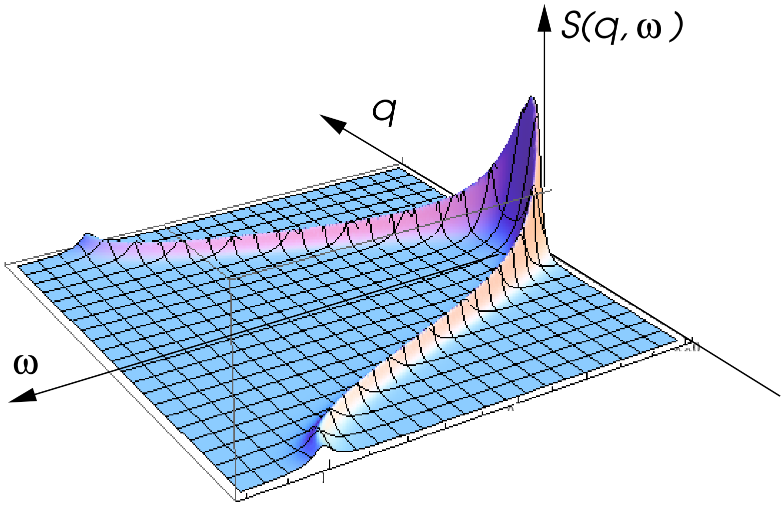


Fig. 6.2: Dynamical structure factor for the spin excitation spectrum in the ferromagnetic phase.

6.2.2 Instability condition

Now we construct the self-consistence equation for the mean field. In principle we could use the mean field approximation solution. However, here we require now that the fluctuation-dissipation theorem is satisfied in order to take the effects of fluctuations into account. Thus, we consider the dynamical structure factor,

$$S(\vec{q},\omega) = \int dt\, e^{i\omega t} \frac{1}{\hbar} \langle \hat{S}^-_{\vec{q}}(t) \hat{S}^+_{-\vec{q}}(0) \rangle \,. \tag{6.47}$$

Taking the $\vec{q}$-sum and ω integral provides an onsite equal time correlation function:

$$\frac{1}{V^2} \sum_{\vec{q}} \int d\omega\, S(\vec{q},\omega) = \frac{1}{\hbar} \langle \hat{S}^-_i(0) \hat{S}^+_i(0) \rangle = \frac{1}{\hbar} \left\{ \langle \hat{\vec{S}}^2_i \rangle - \langle \hat{S}^{z2}_i \rangle - \langle \hat{S}^z_i \rangle \hbar \right\} = \frac{\hbar}{2} + m \,. \tag{6.48}$$

Now we may replace $S(\vec{q},\omega)$ by means of the fluctuation-dissipation theorem (6.34).

$$\frac{\hbar}{2} + m = -\frac{1}{\pi} \int d\omega \frac{1}{1 - e^{-\beta\hbar\omega}} \frac{1}{V^2} \sum_{\vec{q}} \chi''(\vec{q},\omega) = \frac{1}{V} \sum_{\vec{q}} \frac{2m}{1 - e^{-\beta\hbar\omega_{\vec{q}}}} \tag{6.49}$$

The instability condition defining the critical temperature is obtained from the limit $m \to 0$, assuming it to be a second order (continuous) phase transition. Then we may expand the denominator on the right-hand side and obtain the equation

$$\frac{\hbar}{2} = \frac{k_B T_c}{J\hbar} \frac{1}{V} \sum_{\vec{q}} \frac{1}{\gamma_0 - \gamma_{\vec{q}}} \tag{6.50}$$

For the three-dimensional system a numerical solution shows $k_B T_c \approx 1.1 J\hbar^2$ which is considerably smaller than the mean field result of $1.5 J\hbar^2$. The reduction is, analogous to the self-consistent fluctuation approximation, caused by thermal fluctuations which are ignored in the ordinary mean field approximation.
In one and two dimensions we find that the integral on the right-hand side diverges at the lower boundary (infrared divergence).

$$\frac{1}{V} \sum_{\vec{q}} \frac{1}{\gamma_0 - \gamma_{\vec{q}}} \to \int \frac{d^d q}{(2\pi)^d} \frac{1}{\gamma_0 - \gamma_{\vec{q}}} \to \int_0^\Lambda dq \frac{q^{d-1}}{q^2} \to \infty \quad \text{for} \quad d \leq 2 \,. \tag{6.51}$$

Thus $k_B T_c \to 0$. This is a consequence of the spin wave fluctuations which destroy the magnetic moment. This result is also known as the Hohenberg-Mermin-Wagner theorem. It applies to systems with a continuous spin symmetry and short-range interaction. Note that Peierls' domain wall argument fails here, since the domain walls have an infinite width due to the continuous degeneracy of the magnetic ground state, i.e. any direction for the ordered moment has the same energy.

6.2.3 Low-temperature properties

How do the spin wave fluctuations influence the low-temperature behavior? Let us first consider $m(T) = \hbar/2 - \delta m(T)$. We insert this into the self-consistence equation,

$$\begin{aligned} \hbar - \delta m &= \left(\frac{\hbar}{2} - \delta m \right) \frac{2}{V} \sum_{\vec{q}} \left(1 + \sum_{n=1}^{\infty} e^{-\beta\hbar\omega_{\vec{q}} n} \right) \\ &\approx (\hbar - 2\delta m) \left\{ 1 + \sum_{n=1}^{\infty} \int \frac{d^3 q}{(2\pi)^3} e^{-3Jm\hbar\beta n q^2} \right\} = (\hbar - 2\delta m) \left\{ 1 + \frac{\zeta(3/2)(k_B T)^{3/2}}{(6\pi J\hbar^2)^{3/2}} \right\} \end{aligned} \tag{6.52}$$

which leads to

$$\delta m(T) \approx \frac{\zeta(3/2)}{(6\pi J\hbar^2)^{3/2}}(k_B T)^{3/2} \, . \tag{6.53}$$

If we compare this with the result of the ordinary mean field approximation, $\delta m \propto \exp(-T/T_c)$ we find that the spin wave fluctuations suppress the magnetic order more strongly.
Finally we consider the spin wave contribution to the low-temperature specific heat. The magnons as bosons have a dispersion $\omega_{\vec{q}} = c_s q^2$ leading to a density of states as

$$D(\omega) \propto \omega^{1/2} \, . \tag{6.54}$$

With this we obtain

$$U \propto \int d\omega \omega^{1/2} \frac{\hbar\omega}{e^{\beta\hbar\omega} - 1} \propto T^{5/2} \quad \Rightarrow \quad C \propto T^{3/2} \tag{6.55}$$

which is also a consequence of low-energy spin wave excitations.

Chapter 7

Renormalization group

While the mean field approximation describes second-order phase transitions in a very handy way, we have seen that it treats fluctuations poorly so that the critical behavior is not adequately reproduced. This is particularly true for the renormalization of the transition point and the critical exponents of quantities which are singular at the phase transition. In this context a powerful method has been developed in the late sixties and seventies which introduce a new concept to describe critical phenomena: *the renormalization group method.* This method is based on a method for the successive decimation of high energy degrees of freedom of the system with the aim to extract the most important information on the low-energy physics from a complex many-body system. In this way the properties of a phase transitions, in particular, its critical regime can be described. There are many variants of the renormalization procedure also beyond the discussion of phase transitions and much development is still going on. Here we will restrict to one special formulation which gives a good insight to the ides.

7.1 Basic method - Block spin scheme

Close to a second order phase transition fluctuations are large and hard to treat. The renormalization group method has the aim to connect the behavior of a model close to with the behavior of a related model far from the phase transition. The basic idea is to make a change of the length scale at which the system is considered. Close to the critical point the only relevant length scale is the correlation length, the length over which the degrees of freedom effectively interact. This length is diverging, if the critical point approached. The procedure rests on diluting (reducing) the number of degrees of freedom in a proper way so as to shrink effectively the length scale and to "move away" from the critical point.
In order to illustrate the method we consider again the Ising model with the degrees of freedom $s_i = \pm s$ on a regular lattice. The partition function is given by

$$Z(\vec{K}, N) = \sum_{\{s_i\}} e^{\mathcal{H}(\vec{K}, \{s_i\}, N)} \tag{7.1}$$

where we assume a generalized Hamiltonian of the form

$$\mathcal{H}(\vec{K}, \{s_i\}, N) = NK_0 + K_1 \sum_i s_i + K_2 \sum_{\langle i,j \rangle} s_i s_j + \cdots \tag{7.2}$$

with couplings among many spins, $\vec{K} = (K_0, K_1, K_2, \ldots)$. In the standard form of the Ising model we have

$$K_0 = 0 \; , \quad K_1 = H/k_B T \; , \quad K_2 = J/K_B T \; , \quad K_{n>2} = 0 \; . \tag{7.3}$$

The aim is now to reduce the number of degrees of freedom without changing the partition function. This can be done by deriving an appropriate model of the remaining degrees of freedom. We introduce $b > 1$ which denotes the change of the length scale in the system. In the decimation procedure we reduce the degrees of freedom by factor b^d. In the following we will use the so-called block spin or related formulations which gives an intuitive view of the renormalization group procedure. We separate the spin degrees of freedom into two groups $\{S_b\}$ and $\{s'\}$. The first set will be eliminated while the second set represents the remaining degrees of freedom.

$$Z(\vec{K}, N) = \sum_{\{s'\}} \sum_{\{S_b\}} e^{\mathcal{H}(\vec{K},\{S_b\},\{s'\},N)} = \sum_{\{s'\}} e^{\mathcal{H}(\vec{K}',\{s'\},Nb^{-d})} = Z(\vec{K}', Nb^{-d}) \ . \tag{7.4}$$

Thus we express the Hamiltonian in terms of new coupling constants between the remaining degrees of freedom. Looking at the reduced free energy per spin we find

$$f(\vec{K}) = -\lim_{N\to\infty} \frac{1}{N} \ln \left[Z(\vec{K}, N) \right] = -b^{-d} \lim_{N\to\infty} \frac{1}{Nb^{-d}} \ln \left[Z(\vec{K}', Nb^{-d}) \right] = b^{-d} f(\vec{K}') \ . \tag{7.5}$$

The transformation $\vec{K} \to \vec{K}'$ can be repeated in the same way

$$\vec{K}^{(n)} = R\vec{K}^{(n-1)} = R^n \vec{K} \qquad \text{with} \qquad N \to Nb^{-dn} \tag{7.6}$$

where the set of operations $\{R^n\}$ with $n = 0, 1, 2, 3, \ldots$ represents a semigroup, leading to the name *renormalization group*. It is a semigroup only, because it lacks the inversion. There is namely no unique way to undo a decimation step. The renormalization group transformations lead to a "flow" of the coupling constants $\vec{K}$.
The instabilities of the system are determined by so-called unstable *fixed points* where we find

$$R\vec{K}_c = \vec{K}_c \ , \tag{7.7}$$

i.e. the coupling constants do not change under renormalization group transformation. The region of $\vec{K}$ close to the fixed point is most interesting as we can approximate the flow of the coupling constants within a linear approximation,

$$\vec{K} = \vec{K}_c + \delta\vec{K} \qquad \text{and} \qquad R\vec{K} \approx \vec{K}_c + \Lambda\delta\vec{K} \tag{7.8}$$

with Λ being a quadratic matrix of the dimensions of $\vec{K}$ which can be infinite in principle. This matrix is generally not Hermitian and its eigenvectors $\vec{e}_i$ (unit length) are not orthogonal. Nevertheless we can express (7.8) as

$$\vec{K} = \vec{K}_c + \sum_{i=0}^{\infty} c_i \vec{e}_i \qquad \text{and} \qquad R\vec{K} = \vec{K}' = \vec{K}_c + \sum_{i=0}^{\infty} c_i b^{y_i} \vec{e}_i \tag{7.9}$$

Thus we find

$$Rc_i = c_i' = c_i b^{y_i} \ . \tag{7.10}$$

The exponents y_i characterize the eigenvectors $\vec{e}_i$. If $y_i > 0$ ($y_i < 0$) then $\vec{e}_i$ is called *relevant* (*irrelevant*). $\vec{e}_i$ with $y_i = 0$ is called *marginal*. Along relevant eigenvectors we move away from the fixed point under successive decimation (unstable fixed point).
Consider the case that only $\vec{e}_1$ is relevant, setting the external field H to zero. The phase transition (critical temperature) corresponds to the temperature T_c where c_1 vanishes (fixed point). Close to the fixed point ($\vec{K}$ close to $\vec{K}_c$ or T close to T_c) we can approximate

$$c_1 = A\left(\frac{T}{T_c} - 1\right) = -A\tau \ , \tag{7.11}$$

and then consider the renormalization group step

$$c_1' = Rc_1 = A\left(\frac{T'}{T_c} - 1\right) = -A\tau' \qquad \Rightarrow \qquad \tau' = b^{y_1}\tau \ . \tag{7.12}$$

Thus the renormalization group step corresponds effectively to a change of the temperature. Let us use this relation first for the length scale. The correlation length is changed in a renormalization group step by

$$\xi \to \xi' = \xi(T)/b = \xi(T') \tag{7.13}$$

The fixed point is characterized by $\xi = \xi/b$ so that only two situations are possible $\xi = \infty$ for the unstable fixed point and $\xi = 0$ for a stable fixed point which can be interpreted as the limit of non-interacting degrees of freedom (no correlations). In case of an unstable fixed point we obtain close to T_c,

$$\frac{|\tau|^{-\nu}}{b} = |\tau'|^{-\nu} \quad \Rightarrow \quad \tau' = b^{1/\nu}\tau \; . \tag{7.14}$$

This leads immediately to

$$\nu = \frac{1}{y_1} \; . \tag{7.15}$$

Therefore the exponent y_1 describing the change of temperature under the renormalization group step is connected directly with the critical exponent, ν of the coherence length.

We now turn to the free energy, still assuming that c_1 is the only component in the renormalization group procedure. Then the singular part of the free energy per spin close to the critical point has the following form:

$$f(\tau) = b^{-d} f(\tau') = b^{-d} f(\tau b^{y_1}) = b^{-dl} f(\tau b^{l y_1}) \; , \tag{7.16}$$

where the last equality means that we consider l renormalization group steps. We now choose l as a real number so that $\tau b^{l y_1} = 1$ and so $b^l = \tau^{-1/y_1}$. Then the free energy takes the form

$$f(\tau) = \tau^{d/y_1} f(1) \tag{7.17}$$

with the critical behavior of the specific heat $C \propto |\tau|^{-\alpha}$ we find here that

$$2 - \alpha = \frac{d}{y_1} \; . \tag{7.18}$$

Thus, from the exponent y_1 we have obtained ν and α. Equations (7.15) and (7.18) can be combined to the Josephson scaling $\nu d = 2 - \alpha$, which we had seen in Section 5.3.2. Modifications and additional exponents can be obtained by keeping the magnetic field as another coupling constant. For simplicity we ignore this here and turn to the examples.

7.2 One-dimensional Ising model

The one-dimensional Ising model can be solved exactly in a rather simple way. In the absence of a magnetic field the partition function and the reduced free energy are given by

$$Z = [2\cosh(K)]^N \qquad \text{and} \qquad f(K) = \ln[2\cosh(K)] \; , \tag{7.19}$$

respectively, where $K = J/k_B T$.
We define the partition functions as

$$Z = \sum_{\{s\}} e^{N K_0 + K_2 \sum_i s_i s_{i+1}} \; , \tag{7.20}$$

which is convenient to illustrate the way of decimation. Note that $K = K_2$ here. We now reorganize this partition function in order to identify a decimation scheme, where we sum over the spins on the even sides,

$$\begin{aligned} Z &= \sum_{\{s_{odd}\}} \sum_{\{s_{even}\}} e^{K_0 + K_2 \sum_i (s_{2i-1} + s_{2i+1}) s_{2i}} \\ &= \sum_{\{s_{odd}\}} \prod_i e^{2K_0} \left\{ e^{K_2(s_{2i-1}+s_{2i+1})} + e^{-K_2(s_{2i-1}+s_{2i+1})} \right\} = \sum_{\{s_{odd}\}} \prod_i e^{K_0' + K_2' s_{2i-1} s_{2i+1}} \end{aligned} \tag{7.21}$$

where $s = 1$. The new coupling parameter K' is related through the condition

$$e^{2K_0}\left\{e^{K_2(s_{2i-1}+s_{2i+1})}+e^{-K_2(s_{2i-1}+s_{2i+1})}\right\}=e^{K_0'+K_2's_{2i-1}s_{2i+1}} \tag{7.22}$$

In order to solve this equation we consider the configurations $(s_{2i-1}, s_{2i+1}) = (+1, +1)$ and $(+1, -1)$ which yields the equations

$$\left.\begin{array}{ll} e^{2K_0}\left\{e^{2K_2}+e^{-2K_2}\right\} & = e^{K_0'+K_2'} \\ & \\ 2e^{2K_0} & = e^{K_0'-K_2'} \end{array}\right\} \quad \Rightarrow \quad \left\{\begin{array}{ll} K_0' & = 2K_0+\dfrac{1}{2}\ln\left[4\cosh(2K_2)\right] \\ & \\ K_2' & = \dfrac{1}{2}\ln\left[\cosh(2K_2)\right] \end{array}\right. \tag{7.23}$$

Note that this decimation scheme conserves the form of the partition function, i.e. we end up with the same connectivity of spins in the exponent (nearest neighbor coupling). We will see in the next section that this is not generally the case.
There are two fixed points:

$$(K_{0c}, K_{2c}) = (\infty, \infty) \qquad \text{and} \qquad (K_{0c}, K_{2c}) = (-\ln 2, 0)\ . \tag{7.24}$$

For the coupling constant K_2 the first fixed point is unstable (the iterative renormalization leads the coupling constants away from the fixed point) and the second is stable.

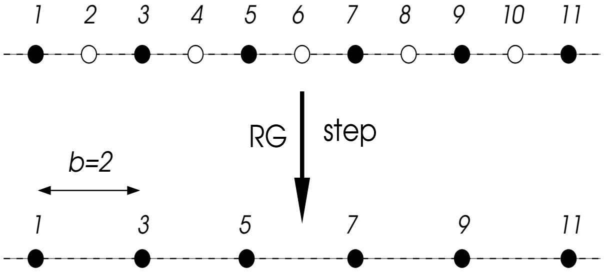


Fig.7.1: Decimation scheme: Every second spin is integrated out to reach a renormalized new system with only half the spins of the previous system.

We do not find a fixed point at a finite value of K_2 which states that no phase transition occurs, as expected in view of Peierls argument. The unstable fixed point corresponds to zero-temperature limit where the spins order in the ground state. The stable fixed point $K_{2c} = 0$ is the limit of non-interacting spins. The renormalization group treatment which had been here performed exactly, shows that there is no phase transition at a finite temperature.

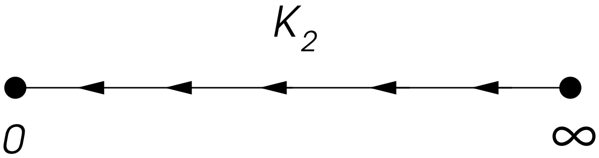


Fig.7.2: Flow of the coupling constant $K_2 = J/k_BT$ of the one-dimensional Ising model under the renormalization group scheme. For any finite coupling and temperature the flow goes towards $K = 0$ the limit of completely decoupled spins.

Therefore we find that starting at any finite value of K_2 leads us through successive application of the decimation procedure towards weaker and weaker coupling K_2. The fixed point of $K_2 = 0$ eventually corresponding to non-interacting spins is the ultimate limit of disordered spins. Through the renormalization group procedure it is possible to generate an effective model which can be solved perturbatively and obtain the essential physics of the system. We introduce $K_0 = g(K_2)$ and write

$$g(K_2') = \frac{1}{2}\left(\ln 4 + K_2'\right) \ . \tag{7.25}$$

from (7.23). We start with $K_0 = 0$ and write

$$Z_N(K_2) = \sum_{\{s\}} e^{K_2 \sum_i s_i s_{i+1}} = e^{Ng(K_2')} Z_{N/2}(K_2') = e^{Ng(K_2') + Ng(K_2'')/2} Z_{N/4}(K_2'') = \dots \ . \tag{7.26}$$

The reduced free energy

$$f(K_2) = -\frac{1}{N} \ln Z_N(K_2) \tag{7.27}$$

can now be approximated by

$$f^{(n)}(K_2) = -\sum_{n'=1}^{n} \frac{g(K_2^{(n'-1)})}{2^{n'-1}} - \frac{1}{2^n} \ln 2 \ , \tag{7.28}$$

by successively increasing n which leads to $K_2^{(n)} \to 0$. For $K_2 = 0$ we find $f = -\ln 2$. Thus knowing the weak coupling limit allows us to determine the reduced free energy at any parameter K_2 (temperature).

7.3 Two-dimensional Ising model

Also for the two-dimensional Ising model there is an exact solution, due to Onsager, which is, however, considerably more complicated than for the one-dimensional system. The partition function is given by

$$Z = \left[2\cosh(2K) e^{I}\right]^N \tag{7.29}$$

with

$$I = \int_0^\infty \frac{d\phi}{2\pi} \ln\left\{\frac{1}{2}\left[1 + (1 - \kappa^2 \sin^2\phi)^{1/2}\right]\right\} \tag{7.30}$$

and

$$\kappa = \frac{2\sinh(2K)}{\cosh^2(2K)} \ . \tag{7.31}$$

In this case there is a phase transition at finite temperature given by

$$\sinh(2K_c) = 1 \qquad \Rightarrow \qquad k_B T_c = 2.27J \quad K_c = 0.4407 \ . \tag{7.32}$$

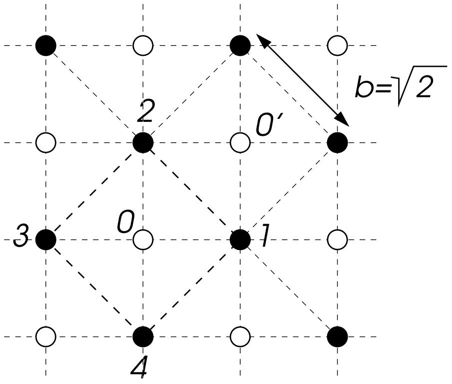


Fig. 7.3: Decimation scheme for two-dimensional Ising model: Every second site is integrate yielding an effective coupling among all surrounding spins.

Now we turn to the renormalization group treatment.[1] There are various decimation schemes we could imagine. In analogy to the one-dimensional case we divide the square lattice into two sublattices as shown in Fig. 7.3: The white sites are integrated out. We take the plaquette indicated in the figure. Spins 1, 2, 3, and 4 encircle spin 0. The latter couples through nearest neighbor interaction to the former four spins. Thus, our decimation works as follows

$$\begin{aligned} Z &= \cdots \sum_{s_1,s_2,s_3,s_4} \sum_{s_0} e^{K s_0(s_1+s_2+s_3+s_4)} \dots \\ &= \cdots \sum_{s_1,s_2,s_3,s_4} \left[e^{K(s_1+s_2+s_3+s_4)} + e^{-K(s_1+s_2+s_3+s_4)} \right] \cdots . \end{aligned} \tag{7.33}$$

We reformulate the partition function for the remaining spin degrees of freedom

$$Z = \cdots \sum_{s_1,s_2,s_3,s_4} e^{K_0'+K_1'(s_1s_2+s_2s_3+s_3s_4+s_4s_1)+K_2'(s_1s_3+s_2s_4)+K_3's_1s_2s_3s_4} \cdots . \tag{7.34}$$

Going through the space of spin configurations we find new effective interactions between the four surrounding spins with the relation[2]

$$\begin{aligned} K_0' &= \frac{1}{8}\ln\left\{\cosh^4(2K)\cosh(4K)\right\} + \ln 2 \\ K_1' &= \frac{1}{8}\ln\left\{\cosh(4K)\right\} \\ K_2' &= \frac{1}{8}\ln\left\{\cosh(4K)\right\} \\ K_3' &= \frac{1}{8}\ln\left\{\cosh(4K)\right\} - \frac{1}{2}\ln\left\{\cosh(2K)\right\} \end{aligned} \tag{7.38}$$

where K_0' is again connected with the reduced free energy, K_1' and K_2' denote nearest- and next-nearest-neighbor spin-spin interaction, and K_3' gives a four-spin interaction on the plaquette. Note that the neighboring plaquettes contribute to the nearest-neighbor interaction, e.g. summing in Fig. 7.3 over $s_{0'}$ on site $0'$, yields another interaction between s_1 and s_2. Therefore we have to modify the second equation in (7.38) by multiplying by a factor 2,

$$K_1' = \frac{1}{4}\ln\left\{\cosh(4K)\right\} . \tag{7.39}$$

---

[1] Literature: H.J. Maris and L.P. Kadanoff, Am. J. Phys. 46, 652 (1978).

[2] The renormalization group relations are obtained by looking at different spin configurations for

$$\begin{aligned} & e^{K(s_1+s_2+s_3+s_4)} + e^{-K(s_1+s_2+s_3+s_4)} \\ &\qquad = e^{K_0'+K_1'(s_1s_2+s_2s_3+s_3s_4+s_4s_1)+K_2'(s_1s_3+s_2s_4)+K_3's_1s_2s_3s_4} . \end{aligned} \tag{7.35}$$

We use now the configurations

$$(s_1,s_2,s_3,s_4) = (+,+,+,+),(+,+,+,-),(+,-,+,-),(+,+,-,-) \tag{7.36}$$

and obtain the equations,

$$\begin{aligned} e^{4K} + e^{-4K} &= e^{K_0'+4K_1'+2K_2'+K_3'} , \\ e^{2K} + e^{-2K} &= e^{K_0'-K_3'} , \\ 2 &= e^{K_0'-4K_1'+2K_2'+K_3'} , \\ 2 &= e^{K_0'-2K_2'+K_3'} , \end{aligned} \tag{7.37}$$

whose solution leads to (7.38).

Unlike in the Ising chain we end up here with a different coupling pattern than we started. More spins are coupled on a wider range. Repeating the decimation procedure would even further enlarge the interaction range and complexity. This is not a problem in principle. However, in order to have a more practical iterative procedure we have to make an approximation. We restrict ourselves to the nearest neighbor interactions which would give a well-defined iterative procedure. But simply ignoring the other couplings which additionally help to align the spins would lead to an oversimplified behavior and would actually give no phase transition. Thus we have to add the other couplings in some way to the nearest-neighbor coupling. It can be shown that the four-spin interaction is small and not important close to the transition point, and we concentrate on K_1' and K_2' only. Let us define the effective nearest-neighbor in a way to give the same ground state energy as both couplings. Each site has four nearest- and four next-nearest neighbors, which yield the energy per site for full spin alignment

$$E_0 = 2NK_1' + 2NK_2' = 2NK' \tag{7.40}$$

Thus we define the new renormalized nearest-neighbor coupling

$$K' = K_1' + K_2' = \frac{3}{8} \ln \{\cosh(4K)\} \tag{7.41}$$

which now can be iterated. We find a stable fixed points of this new equation at $K = 0$ and ∞. There is also an unstable fixed point at

$$K_c = 0.507 \tag{7.42}$$

This unstable fixed point corresponds now to a finite-temperature phase transition at $k_B T_c = 1.97J$, which is lower than the mean field result $k_B T_c = 4J$, but relatively inaccurate compared to the exact result of $2.27J$.

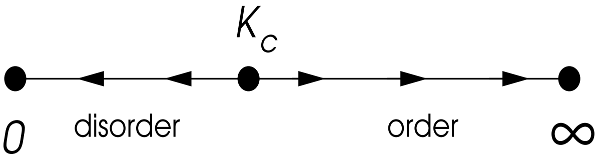


Fig. 7.4: Renormalization group flow of coupling constant; The unstable fixed point K_c represents the critical point. On the left hand side, the flow leads to the stable fixed point $K = 0$ corresponding to the uncoupled spins: disorder phase. The right hand side flows to the stable fixed point $K = \infty$, where system is ordered.

It is now interesting to consider the exponents which we had discussed above. Thus we take into account that in our decimation scheme $b = \sqrt{2}$ and calculate

$$\begin{aligned} K' - K_c &= \frac{3}{8}\left[\ln\{\cosh(4K)\} - \ln\{\cosh(4K_c)\}\right] \\ &= \frac{3}{8}(K - K_c)\,\frac{d}{dK}\ln\{\cosh(4K)\}\bigg|_{K=K_c} \\ &= (K - K_c)\frac{3}{2}\tanh(4K_c) = 1.45(K - K_c) \end{aligned} \tag{7.43}$$

$$\Rightarrow \qquad b^{y_1} = 2^{y_1/2} = 1.45 \qquad \Rightarrow \quad y_1 = 2\frac{\ln 1.45}{\ln 2} = 1.07\,.$$

From this result we obtain the critical exponents ν and α:

$$\nu = \frac{1}{y_1} = 0.93 \qquad \text{and} \qquad \alpha = 2 - \frac{d}{y_1} = 0.135\,. \tag{7.44}$$

The exact result is $\nu_{exact} = 1$ (mean field $\nu_{mf} = 1/2$) and $\alpha_{exact} = 0$ meaning that the specific heat has a logarithmic singularity not describable by an exponent.

The decimation method used here is only one among many which could be imagined and have been used. Unfortunately, for practice reasons approximations had to be made so that the results are only of qualitative value. Nevertheless these results demonstrate that non-trivial properties appear in the critical region close to the second order phase transition.

Other decimations schemes can be used. Decimations in real space are only one type of method, know under the name of block spin method. Alternatively, also decimations schemes in momentum space can be used. Depending on the problem and the purpose different methods are more useful.

Appendix A

2D Ising model: Monte Carlo method and Metropolis algorithm

By Christian Iniotakis

A.1 Monte Carlo integration

In the description of many physical scenarios, one may be confronted with an integration problem, which is not solvable analytically. Thus, the correct result of the integration has to be approximated by evaluating the function at a sufficiently large set of sample points:

$$\int dx\, f(x) \to \sum_{x_i} f(x_i). \tag{A.1}$$

There are several well-established numerical integration methods based on sample points, which are equally spaced, e.g. the midpoint or trapezoidal rules, as well as some more evolved adaptive methods[1]. These methods work fine in lower dimensions. However, if the space of integration is of higher dimension, they are practically not applicable anymore, since the number of sample points grows exponentially with the number of dimension.
In contrast to this, the main idea of Monte Carlo integration is to use sample points, which are generated randomly[2]. For low-dimensional problems, such a method is not useful, if compared to the other methods mentioned above (an example is given in the lecture). Nevertheless, Monte Carlo integration is easy to generalize and is the method of choice for high-dimensional problems.

A.2 Monte Carlo methods in thermodynamic systems

We have already seen, that the the expectation value of a quantity A in the canonical ensemble (where $\beta = 1/k_B T$ is fixed) can be written as

$$\langle A \rangle = \frac{\sum_{S_i} A(S_i) e^{-\beta H(S_i)}}{\sum_{S_i} e^{-\beta H(S_i)}}. \tag{A.2}$$

Here the sums have to be performed over all possible microstates S_i of the system and $H(S_i)$ is the energy of the state S_i. With the partition function

$$Z = \sum_{S_i} e^{-\beta H(S_i)} \tag{A.3}$$

---

[1] These methods decide, whether and where more additional points should be evaluated.

[2] like in a casino, which gives some hint about the origin of the name.

we may define

$$P(S_i) = \frac{1}{Z} e^{-\beta H(S_i)} \tag{A.4}$$

so that

$$\langle A \rangle = \sum_{S_i} P(S_i) A(S_i). \tag{A.5}$$

The expectation value for A is the sum of the values $A(S_i)$, weighted with the function P. Although this structure looks rather simple, it is generally not possible to evaluate it in a naive way. One concrete problem is the tremendous number[3] of states S_i involved. One might think about using the Monte Carlo method here, generating a random sequence of n sample states $S_{i_0}, S_{i_1}, ..., S_{i_n}$, for which we could evaluate the corresponding $A(S_{i_n})$. However, normally it is not possible to calculate the partition function Z, even with a Monte Carlo technique. Thus we do not know the corresponding weighting factors $P(S_{i_n})$. How can we manage to get the expectation value, then?

Firstly, we notice that the weighting factor P represents the probability distribution of thermodynamic equilibrium, since

$$\sum_{S_i} P(S_i) = 1. \tag{A.6}$$

Thus, a rather elegant solution of the problem would be, to generate a series of sample states for the Monte Carlo method, which contains the states S_i according to their distribution P - instead of a completely random sequence[4]. For these states, we could simply sum up all their corresponding A values to get a good approximation for the expectation value. But still the problem is, how to generate such a series, since we do not know the probability distribution P. What we may access, however, is the relative probability of two states, because the partition function drops out:

$$\frac{P(S_i)}{P(S_j)} = e^{-\beta[H(S_i) - H(S_j)]}. \tag{A.7}$$

It turns out, that an algorithm originally developed by Metropolis et al. solves the problem[5]. It is capable of generating sequences of random states, which fulfill the corresponding probability distribution, based on the knowledge of the relative probabilities only. In the following, we will illustrate the functional principle of the Metropolis algorithm for a very simple example, namely an Ising model consisting of two sites only.

A.3 Example: Metropolis algorithm for the two site Ising model

Consider the two site Ising model, where the configuration space consists of the four states $S_i = \{\uparrow\uparrow, \downarrow\downarrow, \uparrow\downarrow, \downarrow\uparrow\}$. Without an applied magnetic field, we have a ground state energy E_G and an excited state energy E_E, and both of them are twofold degenerate. The two states S_1 and S_2 belong to the ground state level, and both S_3 and S_4 to the excited one. In the following, p denotes the probability of the excited state relative to the ground state, given by

$$p = e^{-\beta(E_E - E_G)}. \tag{A.8}$$

Now, take a look at the following algorithm, which produces a sequence of states $S^{(n)}$:

1. Choose a starting state $S^{(0)}$ out of the four S_i.

---

[3] Note, that we assume a model system with a discrete configuration space here. For a continuous model, the sums should be replaced by integrals over configuration space.

[4] For a non-degenerate system at low temperatures, for example, the ground state should appear much more often in this sequence than any specific state of higher energy (if the sequence is long enough...).

[5] Cf. N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

2. Given a state $S^{(n)}$, randomly pick one of the other states and denote it $\tilde{S}$.
3. Generate the next state $S^{(n+1)}$ according to the rule:
   If the energy $H(\tilde{S}) \leq H(S^{(n)})$, set $S^{(n+1)} := \tilde{S}$.
   If the energy $H(\tilde{S}) > H(S^{(n)})$, set $S^{(n+1)} := \tilde{S}$ with probability p,
   and keep $S^{(n+1)} := S^{(n)}$ otherwise.
4. Continue with step 2.

The algorithm works quite simple, but what are its properties? We can answer this question in terms of probability. Let $p_i^{(n)}$ denote the probability, that $S^{(n)} = S_i$. The corresponding probability vector of the nth state lives in the configuration space and is given by

$$\mathbf{p}^{(n)} = (p_1^{(n)}, p_2^{(n)}, p_3^{(n)}, p_4^{(n)})^T. \tag{A.9}$$

One step of the algorithm may be represented by a probability matrix Π in the following way:

$$\mathbf{p}^{(n+1)} = \Pi \mathbf{p}^{(n)}, \tag{A.10}$$

where the matrix elements Π_{ij} are the probabilities $P_{j\to i}$ of the step getting from state S_j to S_i. It is not difficult to find, e.g., $P_{1\to 2} = \frac{1}{3} \cdot 1$, $P_{1\to 3} = P_{1\to 4} = \frac{1}{3} \cdot p$ and $P_{1\to 1} = \frac{2}{3}(1-p)$. Eventually,

$$\Pi = \frac{1}{3} \begin{pmatrix} 2(1-p) & 1 & 1 & 1 \\ 1 & 2(1-p) & 1 & 1 \\ p & p & 0 & 1 \\ p & p & 1 & 0 \end{pmatrix}. \tag{A.11}$$

Note, that Π is diagonalizable. The matrices

$$M = \begin{pmatrix} \frac{1}{p} & 0 & -2 & -1 \\ \frac{1}{p} & 0 & 0 & 1 \\ 1 & -1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \tag{A.12}$$

and

$$M^{-1} = \frac{1}{2(1+p)} \begin{pmatrix} p & p & p & p \\ 0 & 0 & -(1+p) & 1+p \\ -p & -p & 1 & 1 \\ -1 & 1+2p & -1 & -1 \end{pmatrix} \tag{A.13}$$

result in

$$M^{-1}\Pi M = \Lambda \tag{A.14}$$

with the diagonal matrix

$$\Lambda = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -\frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1-2p}{3} & 0 \\ 0 & 0 & 0 & \frac{1-2p}{3} \end{pmatrix}. \tag{A.15}$$

This is quite helpful in finding the actual probability vector after n steps. We have

$$\begin{aligned} \mathbf{p}^{(n)} &= \Pi \mathbf{p}^{(n-1)} \\ &= \Pi^n \mathbf{p}^{(0)} \\ &= M \Lambda^n M^{-1} \mathbf{p}^{(0)}. \end{aligned}$$

The nth power of the diagonal matrix Λ simply is a diagonal matrix with the elements of Λ, but each raised to the power of n. For a large number of iteration steps, $n \to \infty$, only the top-left entry of Λ^n will survive. The other eigenvalues have a modulus smaller than 1 and vanish accordingly. In this limit, we get

$$\mathbf{p}^{(\infty)} = \frac{1}{2(1+p)} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ p & p & p & p \\ p & p & p & p \end{pmatrix} \mathbf{p}^{(0)} = \frac{1}{2(1+p)} \begin{pmatrix} 1 \\ 1 \\ p \\ p \end{pmatrix}. \tag{A.16}$$

Now, we can list several remarkable properties, which are typical for a Metropolis algorithm in general: Firstly note, that $\mathbf{p}^{(\infty)}$ is independent from the actual starting vector $\mathbf{p}^{(0)}$ due to the normalization condition $\sum_i p_i^{(0)} = 1$. Furthermore, it remains unchanged by Π, since we have

$$\mathbf{p}^{(\infty)} = \Pi \mathbf{p}^{(\infty)}. \tag{A.17}$$

Finally, most important, the elements of $\mathbf{p}^{(\infty)}$ show the correct probability distribution[6]. In the practical use of such a Metropolis algorithm, one has to wait for a sufficiently large number n of first initial steps[7]. After this initial process, the algorithm effectively produces a random sequence of states according to the probability distribution of thermodynamic equilibrium. A sufficiently large number of those states can then be used for the Monte Carlo integration.
The concrete generalization of the sample Metropolis algorithm presented above to an Ising model with a larger amount of sites is straightforward. A typical way is as follows:

1. Choose a starting state $S^{(0)}$.
2. Given a state $S^{(n)}$, keep the spins at all sites, apart from one randomly picked site, where you flip the spin. Denote this new state $\tilde{S}$.
3. Generate the next state $S^{(n+1)}$ according to the rule:
   If $H(\tilde{S}) \leq H(S^{(n)})$, set $S^{(n+1)} := \tilde{S}$.
   If $H(\tilde{S}) > H(S^{(n)})$, set $S^{(n+1)} := \tilde{S}$ with prob. $p = e^{-\beta[H(\tilde{S}) - H(S^{(n)})]}$,
   and keep $S^{(n+1)} := S^{(n)}$ otherwise.
4. Continue with step 2.

The basic step 3. remains unchanged. It fulfills the so-called detailed balance property

$$P(S_i) \cdot P_{i \to j} = P(S_j) \cdot P_{j \to i}, \tag{A.18}$$

which is a sufficient condition for P being the stationary probability distribution. Thus, one can be sure to finally get the thermodynamic equilibrium distribution using the Metropolis algorithm. Note, that the algorithm given here differs from the one presented above by a minor point regarding step 2. The new state is not randomly chosen from the huge amount of all possible states, as above. Instead, the new state is generated from the old one by randomly picking a site and flipping its spin, but all the spins at other sites remain unchanged[8]. Alternatively, the local spin flip may be performed not for a randomly chosen site, but for all sites, one after the other, in a sweep. In the lecture, some sample solutions for the $2D$ Ising model, that are based on the Monte Carlo Metropolis method, will be presented. There are even more evolved techniques using cluster updates to overcome the effect of the critical slowing down.

---

[6] Actually, the Metropolis algorithm implements a Markov chain in configuration space, for which the thermodynamic equilibrium distribution is stationary.

[7] The actual number severely depends on the initial values and the circumstances.

[8] This kind of local updating allows to determine the energy difference and thus the relative probability much faster. An important aspect is to ensure the reachability of any state in a finite number of steps to keep ergodicity.

Appendix B

High-temperature expansion of the 2D Ising model: Finding the phase transition with Padé approximants

By Christian Iniotakis

In this section, we investigate the $2D$ Ising model using a specific approximation method, which allows to determine singular critical points at phase transitions in a rather elegant numerical way. Since we consider only the simplest two-dimensional cases, we may compare the final results with those of exact analytical solutions[1]. We start with the Ising Hamiltonian

$$H = -J \sum_{<i,j>} \sigma_i \sigma_j - h \sum_i \sigma_i. \tag{B.1}$$

Here, we assume a regular lattice of N sites with allowed spin values $\sigma_i = \pm 1$ at site i. The index $< i,j >$ denotes nearest neighbor connections. Due to the constants $J > 0$ and h, it is energetically favorable if a spin is parallel to its next neighbors and to the direction of the magnetic field, respectively. Obviously, the configuration space of this system contains 2^N possible states.

B.1 High-temperature expansion

For a general quantity A, the expectation value is given by

$$\langle A \rangle = \frac{\mathrm{Tr} A\, e^{-\beta H}}{\mathrm{Tr} e^{-\beta H}}, \tag{B.2}$$

where the abbreviation $\beta = 1/k_B T$ has been used. In the following, we concentrate on the average magnetization per spin/site

$$m = \langle \frac{1}{N} \sum_{i=1}^{N} \sigma_i \rangle \tag{B.3}$$

and the resulting zero-field susceptibility

$$\chi_0 = \left. \frac{\partial m}{\partial h} \right|_{h=0}. \tag{B.4}$$

---

[1] Cf., e.g., L. Onsager, Phys. Rev. **65**, 117 (1944).

The Tr-operator in the expression for the expectation value sums over all 2^N possible states of our system. Furthermore, for the concrete evaluation it is helpful to realize that

$$\mathrm{Tr}\, 1 = \mathrm{Tr}\, \sigma_i^{2n} \sigma_j^{2m} ... \sigma_k^{2l} = 2^N, \tag{B.5}$$

whereas

$$\mathrm{Tr}\, \sigma_i^{2n+1} \sigma_j^{m} ... \sigma_k^{l} = 0. \tag{B.6}$$

In words, as long as we deal with products of even powers of the σ_i only, we get 2^N. However, as soon as at least one of the powers is odd, we get zero.
Exercise 1: Using above relations, show that the reduced zero-field susceptibility

$$\chi = k_B T \chi_0 \tag{B.7}$$

is given by

$$\chi = \frac{1}{N} \frac{\mathrm{Tr}\, \sum_{i,j} \sigma_i \sigma_j e^{-\beta H_0}}{\mathrm{Tr}\, e^{-\beta H_0}} = 1 + \frac{1}{N} \frac{\mathrm{Tr}\, \sum_{i \neq j} \sigma_i \sigma_j e^{-\beta H_0}}{\mathrm{Tr}\, e^{-\beta H_0}} \tag{B.8}$$

with H_0 denoting the zero-field Hamiltonian

$$H_0 = -J \sum_{<i,j>} \sigma_i \sigma_j. \tag{B.9}$$

As a next step, we want to evaluate both the numerator and the denominator of the reduced zero-field susceptibility. For this purpose, we use the identity

$$\begin{aligned} e^{\beta J \sigma_i \sigma_j} &= \sum_{n=0}^{\infty} \frac{(\beta J)^n}{n!} \sigma_i^n \sigma_j^n \\ &= \sum_{n=0}^{\infty} \frac{(\beta J)^{2n}}{(2n)!} + \sigma_i \sigma_j \sum_{n=0}^{\infty} \frac{(\beta J)^{2n+1}}{(2n+1)!} \\ &= \cosh \beta J + \sigma_i \sigma_j \sinh \beta J \\ &= \cosh(\beta J)(1 + w \sigma_i \sigma_j) \end{aligned} \tag{B.10}$$

with the definition

$$w = \tanh \beta J. \tag{B.11}$$

Note, that $w \to 0$ for $T \to \infty$ and $w < 1$ for finite T. Thus, w is an ideal candidate (better than β itself) for a series expansion in a small parameter, starting from the high-temperature regime. Due to

$$e^{-\beta H_0} = e^{\beta J \sum_{<i,j>} \sigma_i \sigma_j} = \prod_{<i,j>} e^{\beta J \sigma_i \sigma_j}, \tag{B.12}$$

Eq. (B.10) can be plugged into the expression for the reduced zero-field susceptibility Eq. (B.8), and we find in orders of w:

$$\begin{aligned} \chi &= 1 + \frac{1}{N} \frac{\mathrm{Tr} \sum\limits_{i \neq j} \sigma_i \sigma_j \prod\limits_{<r,s>} (1 + w \sigma_r \sigma_s)}{\mathrm{Tr} \prod\limits_{<r,s>} (1 + w \sigma_r \sigma_s)} \\ &= 1 + \frac{1}{N} \frac{\mathrm{Tr} \sum\limits_{i \neq j} \sigma_i \sigma_j \left(1 + w \sum\limits_{<r,s>} \sigma_r \sigma_s + w^2 \sum\limits_{<r,s> \neq <n,m>} \sigma_r \sigma_s \sigma_n \sigma_m + ... \right)}{\mathrm{Tr} \left(1 + w \sum\limits_{<r,s>} \sigma_r \sigma_s + w^2 \sum\limits_{<r,s> \neq <n,m>} \sigma_r \sigma_s \sigma_n \sigma_m + ... \right)}. \end{aligned}$$

n	square lattice, $q = 4$	triangular lattice, $q = 6$	honeycomb lattice, $q = 3$
0	1	1	1
1	4	6	3
2	12	30	6
3	36	138	12
4	100	606	24
5	276	2586	48
6	740	10818	90
7	1972	44574	168
8	5172	181542	318
9	13492	732678	600
10	34876	2 935218	1098
11	89764	11 687202	2004
12	229628	46 296210	3696
13	585508	182 588850	6792
14	1 486308	717 395262	12270
15	3 763460	2809 372302	22140

Table B.1: Coefficients a_n of the high-temperature expansion of the reduced zero-field susceptibility χ in orders of $w = \tanh \beta J$ (cf. Exercise 2). The data for the different lattice types is according to Sykes et al.

Exercise 2: Calculate the lowest orders in the numerator and the denominator. Show, that we end up with

$$\begin{aligned} \chi &= 1 + \frac{1}{N} \frac{2^N N q w + 2^N N q (q-1) w^2 + O(w^3)}{2^N + O(w^3)} \\ &= 1 + q w + q(q-1) w^2 + O(w^3), \end{aligned}$$

where q denotes the number of nearest neighbors in the lattice.

The evaluation of higher coefficients is merely a combinatorial task. In Table B.1, we list the results published by Sykes et al.[2] in 1972.

B.2 Finding the singularity with Padé approximants

The coefficients in Table B.1 grow rapidly for higher orders. In particular, the series

$$\chi = \sum_n a_n w^n \tag{B.13}$$

does not converge for all temperatures T or values of w, respectively. This is not surprising, since we expect a singularity of the form

$$\chi \sim A(w - w_c)^{-\gamma} \tag{B.14}$$

to appear when $w \to w_c^-$ due to the phase transition. The specific values of both w_c and the critical exponent $\gamma > 0$ are unknown so far. And obviously, an "ordinary" series expansion of the Taylor-type is not sufficient to handle such a singularity appropriately. Therefore we employ the method of the so-called Padé approximants[3] in the following, which will allow us to determine the critical values w_c and γ with a high accuracy.

---

[2] Cf. M.F. Sykes, D.S. Gaunt, P.D. Roberts and J.A. Wyles, J. Phys. A **5**, 624 (1972).

[3] Cf. H. Padé's thesis of the year 1892 as well as G.A. Baker, Phys. Rev. **124**, 768 (1961).

The Padé approximant $[m,n]$ to a given function f is the ratio of a polynomial of degree m divided by a polynomial of degree n

$$[m,n] = \frac{p_0 + p_1 x + p_2 x^2 + \ldots + p_m x^m}{1 + q_1 x + q_2 x^2 + \ldots + q_n x^n} \tag{B.15}$$

such that it agrees with f up to order $m+n$:

$$f = [m,n] + O(x^{m+n+1}). \tag{B.16}$$

Note, that the coefficients p_i, q_j of the polynomials are uniquely determined, and with them also is the Padé approximant $[m,n]$. The Padé approximants $[m,0]$ just coincide with the Taylor series expansion for f up to order m and are of no use for our purpose here. However, Padé approximants of higher order in the denominator naturally exhibit poles. This feature makes them good candidates for catching the singularity at the phase transition.

Since the expected singularity is of order γ, whereas the poles of the Padé approximants are restricted to integer orders, we may introduce the following trick: Instead of χ, we use the auxiliary function $\partial_w \ln \chi$, since

$$\begin{aligned} \chi &\sim A(w-w_c)^{-\gamma} \\ \ln \chi &\sim \ln A - \gamma \ln(w - w_c) \\ \partial_w \ln \chi &\sim \frac{-\gamma}{w - w_c}. \end{aligned}$$

Thus we expect the Padé approximants for the function $\partial_w \ln \chi$ to display a singularity at the critical value w_c, representing the physical phase transition. Moreover, the value of the residue corresponding to this singularity should provide us with $-\gamma$. In Table B.2, we list the first coefficients for the auxiliary function $\partial_w \ln \chi$. Results for the concrete numerical evaluation of the Padé approximants will be presented in the lecture. Some samples can be seen in Figs. B.1, B.2 and B.3 for the square, triangular and honeycomb lattice, respectively. We get a very good estimate for the critical exponent compared to its exact value

$$\gamma = 7/4 = 1.75, \tag{B.17}$$

which is independent of the lattice type. Furthermore, the results for the critical values w_c are in excellent agreement with those of analytical solutions given in the brackets:

$$\begin{aligned} w_c^{SQU} &\approx 0.414211 \;\; (\sqrt{2} - 1 \approx 0.414214) \\ w_c^{TRI} &\approx 0.267944 \;\; (2 - \sqrt{3} \approx 0.267949) \\ w_c^{HON} &\approx 0.577348 \;\; (1/\sqrt{3} \approx 0.577350). \end{aligned}$$

The critical temperature T_c can easily be calculated from w_c. We get

$$\frac{k_B T_c}{J} \approx 2.269 \,/\, 3.641 \,/\, 1.519 \tag{B.18}$$

for the square, triangular and honeycomb lattice, respectively. Note, that the critical temperature roughly scales with the number q of nearest neighbors in the lattice, which are accordingly given by

$$q = 4 \,/\, 6 \,/\, 3. \tag{B.19}$$

This behavior indicates, that a coupling to more neighbors enhances parallel alignment of the spins, allowing the phase transition to occur already at higher temperatures.

Summarizing, employing Padé approximants revealed the phase transitions of the $2D$ Ising model on different lattice types with a high accuracy. The main work within this method lies in calculating sufficiently many coefficients of the power series of the corresponding singular quantity[4]. A generalization to, e.g., the $3D$-case is straightforward.

---

[4] Note, that the success of this method crucially depends on the concrete type of singularity to deal with.

n	square lattice, $q = 4$	triangular lattice, $q = 6$	honeycomb lattice, $q = 3$
0	4	6	3
1	8	24	3
2	28	90	9
3	48	336	15
4	164	1266	33
5	296	4752	27
6	956	17646	87
7	1760	65760	159
8	5428	245646	297
9	10568	917184	243
10	31068	3422898	795
11	62640	12773952	1503
12	179092	47674386	2499
13	369160	177928824	2355
14	1034828	664051230	7209

Table B.2: Coefficients of the high-temperature expansion of the function $\partial_w \ln \chi$ in orders of $w = \tanh \beta J$. We only give results up to order 14, generated from Table B.1.

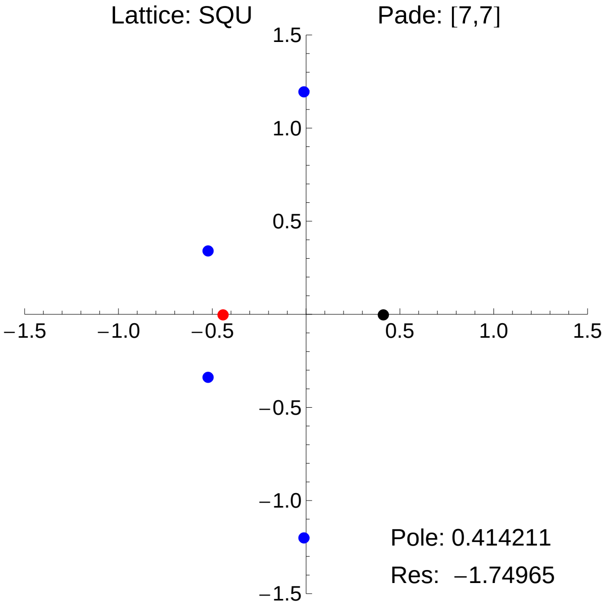


Fig. B.1: Poles of the $[7, 7]$ *Padé approximant for* $\partial_w \ln \chi$ *on the square lattice. The physical pole is at* $w_c \approx 0.414211$ *with residue* -1.7497.

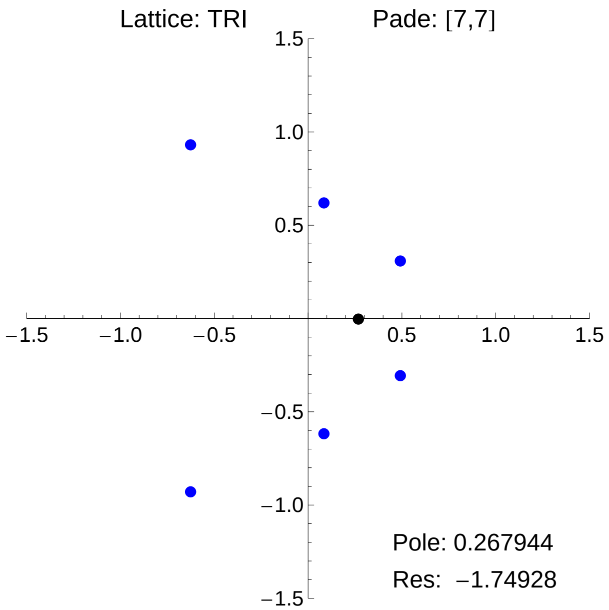


Fig. B.2: Poles of the $[7,7]$ *Padé approximant for* $\partial_w \ln \chi$ *on the triangular lattice. The physical pole is at* $w_c = 0.267944$ *with residue* -1.7493*.*

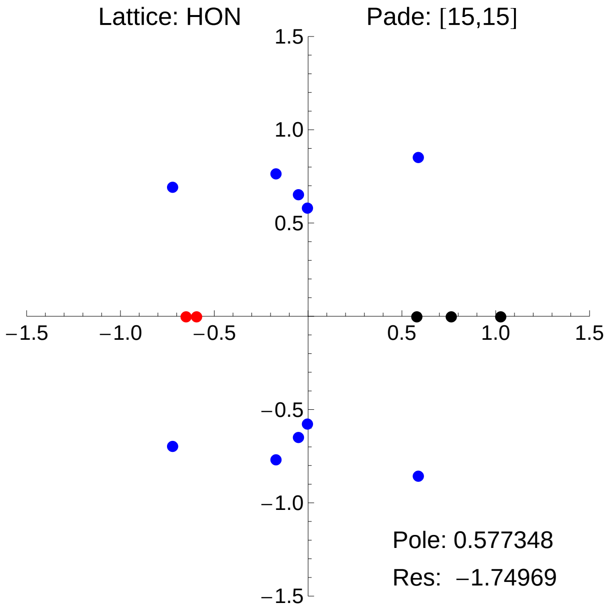


Fig. B.3: Poles of the $[15,15]$ *Padé approximant for* $\partial_w \ln \chi$ *on the honeycomb lattice. The physical pole is at* $w_c = 0.577348$ *with residue* -1.7497*.*