Lecture Notes, Statistical Mechanics (Theory F)

Jörg Schmalian

April 14, 2014

Institute for Theory of Condensed Matter (TKM) Karlsruhe Institute of Technology

Summer Semester, 2014

Contents

Preface

These lecture notes summarize the main content of the course Statistical Mechanics (Theory F), taught at the Karlsruhe Institute of Technology during the summer semester 2012 and 2014. They are based on the graduate course Statistical Mechanics taught at Iowa State University between 2003 and 2005.

Chapter 1

Introduction

Many-particle systems are characterized by a huge number of degrees of freedom. However, in essentially all cases a complete knowledge of all quantum or classical states is neither possible nor useful and necessary. For example, it is hard to determine the initial coordinates and velocities of 10^{23} Ar-atoms in a high-temperature gas state, needed to integrate Newton's equations. In addition, it is known from the investigation of classical chaos that in classical systems with many degrees of freedom the slightest change (i.e. lack of knowledge) in the initial conditions usually causes dramatic changes in the long time behavior of the individual particles are concerned. On the other hand the macroscopic properties of container of gas or a bucket of water are fairly generic and don't seem to depend on how the individual particles have been initialized. This interesting observation clearly suggests that there are principles at work, ensuring that only a few variables are needed to characterize the macroscopic properties of a macroscopic system. It is obviously worthwhile trying to identify these principles instead of going through the effort to identify all particle momenta and positions.

The tools and insights of statistical mechanics enable us to determine the macroscopic properties of many particle systems with known microscopic Hamiltonian, albeit in many cases only approximately. This bridge between the microscopic and macroscopic world is based on the concept of a "lack of knowledge" in the precise characterization of the system and therefore has a probabilistic aspect. This is indeed a lack of knowledge which, different from the probabilistic aspects of quantum mechanics, could be fixed if one we were only able to fully characterize and solve the many particle problem. For finite but large systems this is an extraordinary tough problem. It becomes truly impossible in the limit of infinitely many particles. It is this limit of large systems where statistical mechanics is extremely powerful. One way to see that the "lack of knowledge" problem is indeed more fundamental than solely laziness of the physicist is that essentially every physical system is embedded in an environment. Only complete knowledge of system *and* environment allows for a complete characterization. Even the observable part of our universe seems to behave this way, denying us

full knowledge of any given system as a matter of principle.

Chapter 2

Thermodynamics

Even though this course is about statistical mechanics, it is useful to summarize some of the key aspects of thermodynamics. Clearly these comments cannot replace a course on thermodynamics itself. Thermodynamics and statistical mechanics have a relationship which is quite special. It is well known that classical mechanics covers a set of problems which are a subset of the ones covered by quantum mechanics. Even more clearly is non-relativistic mechanics a part of relativistic mechanics. Such a statement cannot be made if one tries to relate thermodynamics and statistical mechanics. Thermodynamics makes very general statements about equilibrium states. The observation that a system in thermodynamic equilibrium does not depend on its preparation in the past for example is being beautifully formalized in terms of exact and inexact differentials. However, it also covers the energy balance and efficiency of processes which can be reversible or irreversible. Using the concept of extremely slow, so called quasi-static processes it can then make far reaching statements which only rely on the knowledge of equations of state like for example

$$pV = k_{\mathrm{B}}NT \tag{2.1}$$

in case of a dilute gas at high temperatures. On the other hand, equilibrium statistical mechanics provides us with the tools to derive such equations of state, even though it has not much to say about the actual processes, like for example in a Diesel engine. The latter may however be covered as part of the rapidly developing field of non-equilibrium statistical mechanics. The main conclusion from these considerations is that it is important to summarize some, but not necessary to summarize all aspects of thermodynamics for this course.

2.1 Equilibrium and the laws of thermodynamics

Thermodynamics is based on four laws which are in short given as:

- 0th law: Thermodynamic equilibrium exists and is characterized by a temperature

- 1st law: Energy is conserved

- 2nd law: Not all heat can be converted into work

- 3rd law: One cannot reach absolute zero temperature.

Zeroth law: A closed system reaches after long time the state of thermo-dynamic equilibrium. Here closed stands for the absence of directed energy, particle etc. flux into or out of the system, even though a statistical fluctuation of the energy, particle number etc. may occur. The equilibrium state is then characterized by a set of variables like:

- volume, V

- electric polarization, $\mathbf{P}$

- magnetization, $\mathbf{M}$

- particle numbers, N_i of particles of type i

- etc.

This implies that it is irrelevant what the previous volume, magnetization etc. of the system were. The equilibrium has no memory! If a function of variables does not depend on the way these variables have been changed it can conveniently written as a total differential like dV or dN_i etc.

   If two system are brought into contact such that energy can flow from one system to the other. Experiment tells us that after sufficiently long time they will be in equilibrium with each other. Then they are said to have the same temperature. If for example system A is in equilibrium with system B and with system C, it holds that B and C are also in equilibrium with each other. Thus, the temperature is the class index of the equivalence class of the thermodynamic equilibrium. There is obviously large arbitrariness in how to chose the tempera-ture scale. If T is a given temperature scale then any monotonous function $t\,(T)$ would equally well serve to describe thermodynamic systems. The temperature is typically measured via a *thermometer*, a device which uses changes of the system upon changes of the equilibrium state. This could for example be the volume of a liquid or the magnetization of a ferromagnet etc.

   In classical system one frequently uses the averaged kinetic energy of the particles

$$k_{\text{B}}T = \frac{2}{3}\left\langle \varepsilon_{\text{kin}} \right\rangle \tag{2.2}$$

as measure of the temperature. We will derive this later. Here we should only keep in mind that this relation is not valid within quantum mechanics, i.e. fails at low temperatures. The equivalence index interpretation given above is a much more general concept.

First law: The first law is essentially just energy conservation. The total energy is called the *internal energy U*. Below we will see that U is nothing else but the expectation value of the Hamilton operator. Changes, dU of U occur only by causing the system to do work, δW, or by changing the heat content, δQ. To do work or change heat is a process and not an equilibrium state and the amount of work depends of course on the process. Nevertheless, the sum of these two contributions is a total differential[1]

$$dU = \delta Q + \delta W \tag{2.3}$$

which is obvious once one accepts the notion of energy conservation, but which was truly innovative in the days when R. J. Mayer (1842) and Joule (1843-49) realized that heat is just another energy form.

The specific form of δW can be determined from mechanical considerations. For example we consider the work done by moving a cylinder in a container. Mechanically it holds

$$\delta W = -\mathbf{F} \cdot \mathbf{ds} \tag{2.4}$$

where $\mathbf{F}$ is the force exerted by the system and $d\mathbf{s}$ is a small distance change (here of the wall). The minus sign in δW implies that we count energy which is added to a system as positive, and energy which is subtracted from a system as negative. Considering a force perpendicular to the wall (of area A) it holds that the pressure is just

$$p = \frac{|\mathbf{F}|}{A}. \tag{2.5}$$

If we analyze the situation where one pushes the wall in a way to reduce the volume, then $\mathbf{F}$ and $\mathbf{ds}$ point in opposite directions, and and thus

$$\delta W = pAds = -pdV. \tag{2.6}$$

Of course in this case $\delta W > 0$ since $dV = -Ads < 0$. Alternatively, the wall is pushed out, then $\mathbf{F}$ and $\mathbf{ds}$ point in the same direction and

$$\delta W = -pAds = -pdV.$$

Now $dV = Ads > 0$ and $\delta W < 0$. Note, we can only consider an infinitesimal amount of work, since the pressure changes during the compression. To calculate the total compressional work one needs an equation of state $p(V)$. It is a general property of the energy added to or subtracted from a system that it is the product of an intensive quantity (pressure) and the change of an extensive state quantity (volume).

---

[1]A total differential of a function $z = f(x_i)$ with $i = 1, \cdots, n$, corresponds to $dz = \sum_i \frac{\partial f}{\partial x_i} dx_i$. It implies that $z\left(x_i^{(1)}\right) - z\left(x_i^{(2)}\right) = \int_C \sum_i \frac{\partial f}{\partial x_i} dx_i$, with contour C connecting $x_i^{(2)}$ with $x_i^{(1)}$, is independent on the contour C. In general, a differential $\sum_i F_i dx_i$ is total if $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$, which for $F_i = \frac{\partial f}{\partial x_i}$ corresponds to the interchangability of the order in which the derivatives are taken.

More generally holds

$$\delta W = -pdV + \mathbf{E}\cdot d\mathbf{P} + \mathbf{H}\cdot d\mathbf{M} + \sum_i \mu_i dN_i, \qquad (2.7)$$

where $\mathbf{E}$, $\mathbf{H}$ and μ_i are the electrical and magnetic field and the chemical potential of particles of type i. $\mathbf{P}$ is the electric polarization and $\mathbf{M}$ the magnetization. To determine electromagnetic work $\delta W_{em} = \mathbf{E}\cdot d\mathbf{P} + \mathbf{H}\cdot d\mathbf{M}$ is in fact rather subtle. As it is not really relevant for this course we only sketch the derivation and refer to the corresponding literature: J. A. Stratton, "Electromagnetic Theory", Chap. 1, McGraw-Hill, New York, (1941) or V. Heine, Proc. Cambridge Phil. Sot., Vol. 52, p. 546, (1956), see also Landau and Lifshitz, Electrodynamics of Continua.

Finally we comment on the term with chemical potential μ_i. Essentially by definition holds that μ_i is the energy needed to add one particle in equilibrium to the rest of the system, yielding the work $\mu_i dN_i$.

Second Law: This is a statement about the stability of the equilibrium state. After a closed system went from a state that was out of equilibrium (right after a rapid pressure change for example) into a state of equilibrium it would not violate energy conservation to evolve back into the initial out-of-equilibrium state. In fact such a time evolution seems plausible, given that the microscopic laws of physics are invariant under time reversal. The content of the second law is, however, that the tendency to evolve towards equilibrium can only be reversed **by changing work into heat** (i.e. the system is not closed anymore). We will discuss in some detail how this statement can be related to the properties of the microscopic equations of motion.

Historically the second law was discovered by Carnot. Lets consider the Carnot process of an ideal gas

1. Isothermal $(T = T_h = const.)$ expansion from volume $V_1 \to V_2$:

$$\frac{V_2}{V_1} = \frac{p_1}{p_2} \qquad (2.8)$$

   Since U of an ideal gas is solely kinetic energy $\sim T_h$, it holds $dU = 0$ and thus

$$\begin{aligned} \Delta Q &= -\Delta W = -\int_{V_1}^{V_2} \delta W = \int_{V_1}^{V_2} pdV \\ &= Nk_{\mathrm{B}}T_h \int_{V_1}^{V_2} \frac{dV}{V} = Nk_{\mathrm{B}}T_h \log\left(\frac{V_2}{V_1}\right) \end{aligned} \qquad (2.9)$$

2. Adiabatic $(\delta Q = 0)$ expansion from $V_2 \to V_3$ with

$$\Delta Q = 0 \qquad (2.10)$$

   The system will lower its temperature according to

$$\frac{V_3}{V_2} = \left(\frac{T_h}{T_l}\right)^{3/2}. \qquad (2.11)$$

This can be obtained by using

$$dU = CdT = -\frac{Nk_{\mathrm{B}}T}{V}dV \qquad (2.12)$$

and $C = \frac{3}{2}Nk_{\mathrm{B}}$ and integrating this equation.

3. Isothermal compression $V_3 \to V_4$ at T_l where similar to the first step:

$$\Delta Q_{3\to4} = Nk_{\mathrm{B}}T_l \log\left(\frac{V_4}{V_3}\right) \qquad (2.13)$$

4. Adiabatic compression to the initial temperature and volume, i.e.

$$\Delta Q = 0 \qquad (2.14)$$

$$\frac{V_1}{V_4} = \left(\frac{T_l}{T_h}\right)^{3/2}. \qquad (2.15)$$

As expected follows that $\Delta U_{\mathrm{tot}} = 0$, which can be obtained by using $\Delta W = -C\left(T_l - T_h\right)$ for the first adiabatic and $\Delta W = -C\left(T_h - T_l\right)$ for the second. On the other hand $\Delta Q_{\mathrm{tot}} > 0$, which implies that the system does work since $\Delta W_{\mathrm{tot}} = -\Delta Q_{\mathrm{tot}}$. As often remarked, for the efficiency (ratio of the work done by the heat absorbed) follows $\eta = \frac{|\Delta W_{\mathrm{tot}}|}{\Delta Q_{1\to2}} < 1$.

Most relevant for our considerations is however the observation:

$$\frac{\Delta Q_{1\to2}}{T_h} + \frac{\Delta Q_{3\to4}}{T_l} = Nk_B\left(\log\left(\frac{V_2}{V_1}\right) + \log\left(\frac{V_4}{V_3}\right)\right) = 0 \qquad (2.16)$$

Thus, it holds

$$\int \frac{\delta Q}{T} = 0. \qquad (2.17)$$

This implies that (at least for the ideal gas) the *entropy*

$$dS \equiv \frac{\delta Q}{T} \qquad (2.18)$$

is a total differential and thus a quantity which characterizes the state of a system. This is indeed the case in a much more general context, where we write

$$\delta Q = TdS \qquad (2.19)$$

It then follows in particular that in equilibrium the entropy S is a maximum. As we will see, the equilibrium is the least structured state possible at a given total energy. In this sense it is very tempting to interpret the maximum of the entropy in equilibrium in a way that S is a measure for the lack of structure, or disorder.

It is already now useful to comment on the microscopic, statistical interpretation of this behavior and the origin of irreversibility. In classical mechanics

a state of motion of N particles is uniquely determined by the $3N$ coordinates and $3N$ momenta (q_i, p_i) of the N particles at a certain time. The set (q_i, p_i) is also called the micro-state of the system, which of course varies with time. Each micro-state (q_i, p_i) corresponds to one point in a $6N$-dimensional space, the phase space. The set (q_i, p_i) i.e., the micro-state, can therefore be identified with a point in phase space. Let us now consider the diffusion of a gas in an initial state $(q_i(t_0), p_i(t_0))$ from a smaller into a larger volume. If one is really able to reverse all momenta in the final state $(q_i(t_f), p_i(t_f))$ and prepare a state $(q_i(t_f), -p_i(t_f))$, the process would in fact be reversed. From a statistical point of view, however, this is an event with an incredibly small probability. For there is only one point (micro-state) in phase space which leads to an exact reversal of the process, namely $(q_i(t_f), -p_i(t_f))$. The great majority of micro-states belonging to a certain macro-state, however, lead under time reversal to states which cannot be distinguished macroscopically from the final state (i.e., the equilibrium or Maxwell- Boltzmann distribution). The fundamental assumption of statistical mechanics now is that all micro-states which have the same total energy can be found with equal probability. This, however, means that the micro-state $(q_i(t_f), -p_i(t_f))$ is only one among very many other micro-states which all appear with the same probability.

As we will see, the number Ω of micro-states that is compatible with a given macroscopic observable is a quantity closely related to the entropy of this macro-state. The larger Ω, the more probable is the corresponding macro-state, and the macro-state with the largest number $\Omega_{\max}$ of possible microscopic realizations corresponds to thermodynamic equilibrium. The irreversibility that comes with the second law is essentially stating that the motion towards a state with large Ω is more likely than towards a state with smaller Ω.

Combining first and second law, follows (we ignore electrostatic effects and consider only one type of particles):

$$dU = TdS - pdV + \mu dN \tag{2.20}$$

It is easily possible to determine the zeroth law from this result and the principle of maximum entropy. To this end we consider a perfectly isolated system consisting of two subsystems. Let the total energy be $U = U_1 + U_2$ and the total entropy be

$$S = S_1(U_1) + S_2(U_2) \tag{2.21}$$

Since U is fixed, we only have one variable, say U_1 that we need to adjust to reach equilibrium. We want to determine U_1 and thus the way the energy is distributed among the subsystems from the principle of maximal entropy, i.e. from

$$
\begin{aligned}
\frac{dS}{dU_1} &= \frac{dS_1}{dU_1} + \frac{dS_2}{dU_2}\frac{dU_2}{dU_1} \\
&= \frac{dS_1}{dU_1} - \frac{dS_2}{dU_2} = 0
\end{aligned}
\tag{2.22}
$$

We perform this analysis for fixed total volume + particle number and fixed volume and particle number of the subsystems such that

$$\frac{1}{T_1} = \frac{dS_1}{dU_1}$$

that the condition of maximal entropy implies that

$$T_1 = T_2,$$

which is the zeroth law of thermodynamics.

Third law: This law was postulated by Nernst in 1906 and is closely related to quantum effects at low T. If one cools a system down it will eventually drop into the lowest quantum state. Then, there is no lack of structure and one expects $S \to 0$. This however implies that one can not change the heat content anymore if one approaches $T \to 0$, i.e. it will be increasingly harder to cool down a system the lower the temperature gets.

2.2 Thermodynamic potentials

The internal energy of a system is written (following the first and second law) as

$$dU = TdS - pdV + \mu dN \tag{2.23}$$

where we consider for the moment only one type of particles. Thus it is obviously a function

$$U(S, V, N) \tag{2.24}$$

with internal variables entropy, volume and particle number. In particular $dU = 0$ for fixed S, V, and N, i.e. $U = const$. In case one considers a physical situation where indeed these internal variables are fixed the internal energy is constant in equilibrium. Clearly, the internal energy is a natural variable to describe the system if we consider a scenario with fixed S, V, and N.

Let us, instead, look at a situation where we want to fix the temperature T instead of S. A simple trick however enables us to find another quantity

$$F = U - TS \tag{2.25}$$

which is much more convenient in such a situation. F obviously has a differential

$$dF = dU - TdS - SdT \tag{2.26}$$

which gives:

$$dF = -SdT - pdV + \mu dN. \tag{2.27}$$

This implies that

$$F = F(T, V, N) \tag{2.28}$$

i.e. the free energy has the internal variables T, V and N and it does not change ($dF = 0$) if the system has constant temperature, volume and particle number. The transformation from U to F is called a Legendre transformation.

Subsystems: There is a more physical interpretation for the origin of the Legendre transformation. To this end we consider a system with internal energy U and its environment with internal energy U_{env}. Let the entire system, consisting of subsystem and environment be closed with fixed energy

$$U_{tot} = U\left(S, V, N\right) + U_{env}\left(S_{env}, V_{env}, N_{env}\right) \tag{2.29}$$

Consider the situations where all volume and particle numbers are known and fixed and we are only concerned with the entropies. The change in energy due to heat fluxes is

$$dU_{tot} = TdS + T_{env}dS_{env}. \tag{2.30}$$

The total entropy for such a state must be fixed, i.e.

$$S_{tot} = S + S_{env} = const \tag{2.31}$$

such that

$$dS_{env} = -dS. \tag{2.32}$$

As $dU = 0$ by assumption for such a closed system in equilibrium, we have

$$0 = \left(T - T_{env}\right)dS, \tag{2.33}$$

i.e. equilibrium implies that $T = T_{env}$. It next holds

$$d\left(U + U_{env}\right) = dU + T_{env}dS_{env} = 0 \tag{2.34}$$

which yields for fixed temperature $T = T_{env}$ and with $dS_{env} = -dS$

$$d\left(U - TS\right) = dF = 0. \tag{2.35}$$

with *free energy* F. From the perspective of the subsystem (without fixed energy) is therefore the free energy the more appropriate thermodynamic potential. Thus in a situation where a system is embedded in an external bath and is rather characterized by a constant temperature. Then, U is not the most convenient *thermodynamic potential* to characterize the system.

Of course, F is not the only thermodynamic potential one can introduce this way, and the number of possible potentials is just determined by the number of internal variables. For example, in case of a constant pressure (as opposed to constant volume) one uses the *enthalpy*

$$H = U + pV \tag{2.36}$$

with

$$dH = TdS + Vdp + \mu dN. \tag{2.37}$$

If both, pressure and temperature, are given in addition to the particle number one uses the *free enthalpy*

$$G = U - TS + pV \qquad (2.38)$$

with

$$dG = -SdT + Vdp + \mu dN \qquad (2.39)$$

In all these cases we considered systems with fixed particle number. It is however often useful to be able to allow exchange with a particle bath, i.e. have a given chemical potential rather than a given particle number. The potential which is most frequently used in this context is the grand-canonical potential

$$\Omega = F - \mu N \qquad (2.40)$$

with

$$d\Omega = SdT - pdV - Nd\mu. \qquad (2.41)$$

$\Omega = \Omega(T, V, \mu)$ has now temperature, volume and chemical potential as internal variables, i.e. is at a minimum if those are the given variables of a physical system.

Maxwell relations: The statement of a total differential is very powerful and allows to establish connections between quantities that are seemingly unrelated. Consider again

$$dF = -SdT - pdV + \mu dN \qquad (2.42)$$

Now we could analyze the change of the entropy

$$\frac{\partial S(T, V, N)}{\partial V} \qquad (2.43)$$

with volume at fixed T and N or the change of the pressure

$$\frac{\partial p(T, V, N)}{\partial T} \qquad (2.44)$$

with temperature at fixed V and N. Since $S(T, V, N) = -\frac{\partial F(T,V,N)}{\partial T}$ and $p(T, V, N) = -\frac{\partial F(T,V,N)}{\partial V}$ follows

$$\begin{aligned}
\frac{\partial S(T, V, N)}{\partial V} &= -\frac{\partial^2 F(T, V, N)}{\partial V \partial T} = -\frac{\partial^2 F(T, V, N)}{\partial T \partial V} \\
&= \frac{\partial p(T, V, N)}{\partial T} \qquad (2.45)
\end{aligned}$$

Thus, two very distinct measurements will have to yield the exact same behavior. Relations between such second derivatives of thermodynamic potentials are called Maxwell relations.

On the heat capacity: The heat capacity is the change in heat $\delta Q = TdS$ that results from a change in temperature dT, i.e.

$$C(T, V, N) = T\frac{\partial S(T, V, N)}{\partial T} = -T\frac{\partial^2 F(T, V, N)}{\partial T^2} \qquad (2.46)$$

It is interesting that we can alternatively obtain this result from the internal energy, if measured not with its internal variables, but instead $U(T, V, N) = U(S(T, V, N), V, N)$:

$$
\begin{aligned}
C(T, V, N) &= \frac{\partial U(T, V, N)}{\partial T} \\
&= \frac{\partial U(S, V, N)}{\partial S} \frac{\partial S}{\partial T} = -T \frac{\partial S}{\partial T}.
\end{aligned} \tag{2.47}
$$

2.2.1 Example of a Legendre transformation

Consider a function

$$
f(x) = x^2 \tag{2.48}
$$

with

$$
df = 2x dx \tag{2.49}
$$

We would like to perform a Legendre transformation from x to the variable p such that

$$
g = f - px \tag{2.50}
$$

and would like to show that

$$
dg = df - p dx - x dp = -x dp. \tag{2.51}
$$

Obviously we need to chose $p = \frac{\partial f}{\partial x} = 2x$. Then it follows

$$
g = \left(\frac{p}{2}\right)^2 - \frac{p^2}{2} = -\frac{p^2}{4} \tag{2.52}
$$

and it follows

$$
dg = -\frac{p}{2} dp = -x dp \tag{2.53}
$$

as desired.

2.3 Gibbs Duhem relation

Finally, a very useful concept of thermodynamics is based on the fact that thermodynamic quantities of big systems can be either considered as extensive (proportional to the size of the system) of intensive (do not change as function of the size of the system).

Extensive quantities:

- volume

- particle number

- magnetization

- entropy

Intensive quantities:

- pressure

- chemical potential

- magnetic field

- temperature

Interestingly, the internal variables of the internal energy are all extensive

$$U = U\left(S, V, N_i\right). \tag{2.54}$$

Now if one increases a given thermodynamic system by a certain scale,

$$
\begin{aligned}
V &\rightarrow \lambda V \\
N_i &\rightarrow \lambda N_i \\
S &\rightarrow \lambda S
\end{aligned} \tag{2.55}
$$

we expect that the internal energy changes by just that factor $U \rightarrow \lambda U$ i.e.

$$U\left(\lambda S, \lambda V, \lambda N_i\right) = \lambda U\left(S, V, N_i\right) \tag{2.56}$$

whereas the temperature or any other intensive variable will not change

$$T\left(\lambda S, \lambda V, \lambda N_i\right) = T\left(S, V, N_i\right). \tag{2.57}$$

Using the above equation for U gives for $\lambda = 1 + \varepsilon$ and small ε:

$$
\begin{aligned}
U\left(\left(1 + \varepsilon\right) S, ...\right) &= U\left(S, V, N_i\right) + \frac{\partial U}{\partial S}\varepsilon S + \frac{\partial U}{\partial V}\varepsilon V + \sum_i \frac{\partial U}{\partial N_i}\varepsilon N_i. \\
&= U\left(S, V, N_i\right) + \varepsilon U\left(S, V, N_i\right) \tag{2.58}
\end{aligned}
$$

Using the fact that

$$
\begin{aligned}
T &= \frac{\partial U}{\partial S} \\
p &= -\frac{\partial U}{\partial V} \\
\mu_i &= \frac{\partial U}{\partial N_i} \tag{2.59}
\end{aligned}
$$

it follows

$$
\begin{aligned}
U\left(\left(1 + \varepsilon\right) S, ...\right) &= U\left(S, V, N_i\right) + \varepsilon U\left(S, V, N_i\right) \\
&= U\left(S, V, N_i\right) + \varepsilon \left(TS - pV + \sum_i \mu_i N_i\right) \tag{2.60}
\end{aligned}
$$

which then gives

$$U = TS - pV + \sum_i \mu_i N_i. \tag{2.61}$$

Since

$$dU = TdS - pdV + \mu dN \tag{2.62}$$

it follows immediately

$$0 = SdT - Vdp + \sum_i N_i d\mu_i. \tag{2.63}$$

This relationship is useful if one wants to study for example the change of the temperature as function of pressure changes etc. Another consequence of the Gibbs Duhem relations is for the other potentials like

$$F = -pV + \sum_i \mu_i N_i. \tag{2.64}$$

or

$$\Omega = -pV \tag{2.65}$$

which can be very useful.

Chapter 3

Summary of probability theory

We give a very brief summary of the key aspects of probability theory that are needed in statistical mechanics. Consider a physical observable x that takes with probability $p(x_i)$ the value x_i. In total there are N such possible values, i.e. $i = 1, \cdots, N$. The observable will with certainty take one out of the N values, i.e. the probability that x is either $p(x_1)$ or $p(x_2)$ or ... or $p(x_N)$ is:

$$\sum_{i=1}^{N} p(x_i) = 1. \tag{3.1}$$

The probability is normalized.

The mean value of x is given as

$$\langle x \rangle = \sum_{i=1}^{N} p(x_i) x_i. \tag{3.2}$$

Similarly holds for an arbitrary function $f(x)$ that

$$\langle f(x) \rangle = \sum_{i=1}^{N} p(x_i) f(x_i), \tag{3.3}$$

e.g. $f(x) = x^n$ yields the n-th moment of the distribution function

$$\langle x^n \rangle = \sum_{i=1}^{N} p(x_i) x_i^n. \tag{3.4}$$

The variance of the distribution is the mean square deviation from the averaged value:

$$\left\langle (x - \langle x \rangle)^2 \right\rangle = \left\langle x^2 \right\rangle - \langle x \rangle^2. \tag{3.5}$$

If we introduce $f_t(x) = \exp(tx)$ we obtain the characteristic function:

$$
\begin{aligned}
c(t) &= \langle f_t(x) \rangle = \sum_{i=1}^{N} p(x_i) \exp(tx_i) \\
&= \sum_{n=0}^{\infty} \frac{t^n}{n!} \sum_{i=1}^{N} p(x_i) x_i^n = \sum_{n=0}^{\infty} \frac{\langle x^n \rangle}{n!} t^n.
\end{aligned}
\tag{3.6}
$$

Thus, the Taylor expansion coefficients of the characteristic function are identical to the moments of the distribution $p(x_i)$.

Consider next two observables x and y with probability $P(x_i \& y_j)$ that x takes the value x_i and y becomes y_i. Let $p(x_i)$ the distribution function of x and $q(y_j)$ the distribution function of y. If the two observables are statistically *independent*, then, and only then, is $P(x_i \& y_j)$ equal to the product of the probabilities of the individual events:

$$
P(x_i \& y_j) = p(x_i) q(y_j) \quad \text{iff } x \text{ and } y \text{ are statistically independent.} \tag{3.7}
$$

In general (i.e. even if x and y are not independent) holds that

$$
\begin{aligned}
p(x_i) &= \sum_{j} P(x_i \& y_j), \\
q(y_j) &= \sum_{i} P(x_i \& y_j).
\end{aligned}
\tag{3.8}
$$

Thus, it follows

$$
\langle x + y \rangle = \sum_{i,j} (x_i + y_j) P(x_i \& y_j) = \langle x \rangle + \langle y \rangle. \tag{3.9}
$$

Consider now

$$
\langle xy \rangle = \sum_{i,j} x_i y_j P(x_i \& y_j) \tag{3.10}
$$

Suppose the two observables are independent, then follows

$$
\langle xy \rangle = \sum_{i,j} x_i y_j p(x_i) q(y_j) = \langle x \rangle \langle y \rangle. \tag{3.11}
$$

This suggests to analyze the covariance

$$
\begin{aligned}
C(x,y) &= \langle (x - \langle x \rangle)(y - \langle y \rangle) \rangle \\
&= \langle xy \rangle - 2 \langle x \rangle \langle y \rangle + \langle x \rangle \langle y \rangle \\
&= \langle xy \rangle - \langle x \rangle \langle y \rangle
\end{aligned}
\tag{3.12}
$$

The covariance is therefore only finite when the two observables are not independent, i.e. when they are *correlated*. Frequently, x and y do not need to be

distinct observables, but could be the same observable at different time or space arguments. Suppose $x = S(r, t)$ is a spin density, then

$$\chi(r, r'; t, t') = \langle (S(r, t) - \langle S(r, t) \rangle)(S(r', t') - \langle S(r', t') \rangle) \rangle$$

is the spin-correlation function. In systems with translations invariance holds $\chi(r, r'; t, t') = \chi(r - r'; t - t')$. If now, for example

$$\chi(r; t) = A e^{-r/\xi} e^{-t/\tau}$$

then ξ and τ are called correlation length and time. They obviously determine over how-far or how-long spins of a magnetic system are correlated.

Chapter 4

Equilibrium statistical mechanics

4.1 The maximum entropy principle

The main contents of the second law of thermodynamics was that the entropy of a closed system in equilibrium is maximal. Our central task in statistical mechanics is to relate this statement to the results of a microscopic calculation, based on the Hamiltonian, H, and the eigenvalues

$$H\psi_i = E_i\psi_i \tag{4.1}$$

of the system.

Within the ensemble theory one considers a large number of essentially identical systems and studies the statistics of such systems. The smallest contact with some environment or the smallest variation in the initial conditions or quantum preparation will cause fluctuations in the way the system behaves. Thus, it is not guaranteed in which of the states ψ_i the system might be, i.e. what energy E_i it will have (remember, the system is in thermal contact, i.e. we allow the energy to fluctuate). This is characterized by the probability p_i of the system to have energy E_i. For those who don't like the notion of ensembles, one can imagine that each system is subdivided into many macroscopic subsystems and that the fluctuations are rather spatial.

If one wants to relate the entropy with the probability one can make the following observation: Consider two identical large systems which are brought in contact. Let p_1 and p_2 be the probabilities of these systems to be in state 1 and 2 respectively. The entropy of each of the systems is S_1 and S_2. After these systems are combined it follows for the entropy as a an extensive quantity that

$$S_{\text{tot}} = S_1 + S_2 \tag{4.2}$$

and for the probability of the combined system

$$p_{\text{tot}} = p_1 p_2. \tag{4.3}$$

The last result is simply expressing the fact that these systems are independent whereas the first ones are valid for extensive quantities in thermodynamics. Here we assume that short range forces dominate and interactions between the two systems occur only on the boundaries, which are negligible for sufficiently large systems. If now the entropy S_i is a function of p_i it follows

$$S_i \sim \log p_i. \tag{4.4}$$

It is convenient to use as pre-factor the so called Boltzmann constant

$$S_i = -k_{\mathrm{B}} \log p_i \tag{4.5}$$

where

$$k_{\mathrm{B}} = 1.380658 \times 10^{-23} \mathrm{JK}^{-1} = 8.617385 \times 10^{-5} \mathrm{eVK}^{-1}. \tag{4.6}$$

The averaged entropy of each subsystems, and thus of each system itself is then given as

$$S = -k_{\mathrm{B}} \sum_{i=1}^{\mathcal{N}} p_i \log p_i. \tag{4.7}$$

Here, we have established a connection between the probability to be in a given state and the entropy S. This connection was one of the truly outstanding achievements of Ludwig Boltzmann.

Eq.4.7 helps us immediately to relate S with the degree of disorder in the system. If we know exactly in which state a system is we have:

$$p_i = \left\{ \begin{array}{ll} 1 & i = 0 \\ 0 & i \neq 0 \end{array} \right. \implies S = 0 \tag{4.8}$$

In the opposite limit, where all states are equally probable we have instead:

$$p_i = \frac{1}{\mathcal{N}} \implies S = k_{\mathrm{B}} \log \mathcal{N}. \tag{4.9}$$

Thus, if we know the state of the system with certainty, the entropy vanishes whereas in case of the complete equal distribution follows a large (a maximal entropy). Here $\mathcal{N}$ is the number of distinct state

The fact that $S = k_{\mathrm{B}} \log \mathcal{N}$ is indeed the largest allowed value of S follows from maximizing S with respect to p_i. Here we must however keep in mind that the p_i are not independent variables since

$$\sum_{i=1}^{\mathcal{N}} p_i = 1. \tag{4.10}$$

This is done by using the method of Lagrange multipliers summarized in a separate handout. One has to minimize

$$\begin{aligned} I & = S + \lambda \left(\sum_{i=1}^{\mathcal{N}} p_i - 1 \right) \\ & = -k_{\mathrm{B}} \sum_{i=1}^{\mathcal{N}} p_i \log p_i + \lambda \left(\sum_{i=1}^{\mathcal{N}} p_i - 1 \right) \end{aligned} \tag{4.11}$$

We set the derivative of I with respect to p_i equal zero:

$$\frac{\partial I}{\partial p_j} = -k_{\mathrm{B}} \left(\log p_i + 1 \right) + \lambda = 0 \tag{4.12}$$

which gives

$$p_i = \exp\left(\frac{\lambda}{k_{\mathrm{B}}} - 1 \right) = P \tag{4.13}$$

independent of i! We can now determine the Lagrange multiplier from the constraint

$$1 = \sum_{i=1}^{\mathcal{N}} p_i = \sum_{i=1}^{\mathcal{N}} P = \mathcal{N} P \tag{4.14}$$

which gives

$$p_i = \frac{1}{\mathcal{N}}, \ \forall_i. \tag{4.15}$$

Thus, one way to determine the entropy is by analyzing the number of states $\mathcal{N}(E, V, N)$ the system can take at a given energy, volume and particle number. This is expected to depend exponentially on the number of particles

$$\mathcal{N} \sim \exp\left(sN \right), \tag{4.16}$$

which makes the entropy an extensive quantity. We will perform this calculation when we investigate the so called micro-canonical ensemble, but will follow a different argumentation now.

4.2 The canonical ensemble

Eq.4.9 is, besides the normalization, an unconstraint extremum of S. In many cases however it might be appropriate to impose further conditions on the system. For example, if we allow the energy of a system to fluctuate, we may still impose that it has a given averaged energy:

$$\langle E \rangle = \sum_{i=1}^{\mathcal{N}} p_i E_i. \tag{4.17}$$

If this is the case we have to minimize

$$
\begin{aligned}
I \ &= \ S + \lambda \left(\sum_{i=1}^{\mathcal{N}} p_i - 1 \right) - k_{\mathrm{B}} \beta \left(\sum_{i=1}^{\mathcal{N}} p_i E_i - \langle E \rangle \right) \\
&= \ -k_{\mathrm{B}} \sum_{i=1}^{\mathcal{N}} p_i \log p_i + \lambda \left(\sum_{i=1}^{\mathcal{N}} p_i - 1 \right) - k_{\mathrm{B}} \beta \left(\sum_{i=1}^{\mathcal{N}} p_i E_i - \langle E \rangle \right) \ (4.18)
\end{aligned}
$$

We set the derivative of I w.r.t p_i equal zero:

$$\frac{\partial I}{\partial p_j} = -k_{\mathrm{B}} \left(\log p_i + 1 \right) + \lambda - k_{\mathrm{B}} \beta E_i = 0 \tag{4.19}$$

which gives

$$p_i = \exp\left(\frac{\lambda}{k_B} - 1 - \beta E_i\right) = \frac{1}{Z}\exp\left(-\beta E_i\right) \tag{4.20}$$

where the constant Z (or equivalently the Lagrange multiplier λ) are determined by

$$Z = \sum_i \exp\left(-\beta E_i\right) \tag{4.21}$$

which guarantees normalization of the probabilities.

The Lagrange multiplier β is now determined via

$$\langle E \rangle = \frac{1}{Z}\sum_i E_i \exp\left(-\beta E_i\right) = -\frac{\partial}{\partial \beta}\log Z. \tag{4.22}$$

This is in general some implicit equation for β given $\langle E \rangle$. However, there is a very intriguing interpretation of this Lagrange multiplier that allows us to avoid solving for $\beta\left(\langle E \rangle\right)$ and gives β its own physical meaning.

For the entropy follows ($\log p_i = -\beta E_i - \log Z$)

$$
\begin{aligned}
S &= -k_B \sum_{i=1}^{\mathcal{N}} p_i \log p_i = k_B \sum_{i=1}^{\mathcal{N}} p_i\left(\beta E_i + \log Z\right) \\
&= k_B \beta \langle E \rangle + k_B \log Z = -k_B\beta\frac{\partial}{\partial\beta}\log Z + k_B \log Z \tag{4.23}
\end{aligned}
$$

If one substitutes:

$$\beta = \frac{1}{k_B T} \tag{4.24}$$

it holds

$$S = k_B T\frac{\partial \log Z}{\partial T} + k_B \log Z = \frac{\partial\left(k_B T \log Z\right)}{\partial T} \tag{4.25}$$

Thus, there is a function

$$F = -k_B T \log Z \tag{4.26}$$

which gives

$$S = -\frac{\partial F}{\partial T} \tag{4.27}$$

and

$$\langle E \rangle = \frac{\partial}{\partial\beta}\beta F = F + \beta\frac{\partial F}{\partial\beta} = F + TS \tag{4.28}$$

Comparison with our results in thermodynamics lead after the identification of $\langle E \rangle$ with the internal energy U to:

$$
\begin{aligned}
T &: \quad \text{temperature} \\
F &: \quad \text{free energy.}
\end{aligned}
\tag{4.29}
$$

Thus, it might not even be useful to ever express the thermodynamic variables in terms of $\langle E \rangle = U$, but rather keep T.

The most outstanding results of these considerations are:

- the statistical probabilities for being in a state with energy E_i are

$$p_i \sim \exp\left(-\frac{E_i}{k_{\mathrm{B}}T}\right) \qquad (4.30)$$

- all thermodynamic properties can be obtained from the so called partition function

$$Z = \sum_i \exp\left(-\beta E_i\right) \qquad (4.31)$$

- within quantum mechanics it is useful to introduce the so called density operator

$$\rho = \frac{1}{Z}\exp\left(-\beta H\right), \qquad (4.32)$$

where

$$Z = \operatorname{tr}\exp\left(-\beta \mathrm{H}\right) \qquad (4.33)$$

ensures that $\operatorname{tr}\rho = 1$. The equivalence between these two representations can be shown by evaluating the trace with respect to the eigenstates of the Hamiltonian

$$
\begin{aligned}
Z &= \sum_n \langle n| \exp\left(-\beta H\right) |n\rangle = \sum_n \langle n|n\rangle \exp\left(-\beta E_n\right) \\
&= \sum_n \exp\left(-\beta E_n\right) \qquad (4.34)
\end{aligned}
$$

as expected.

The evaluation of this partition sum is therefore the major task of equilibrium statistical mechanics.

4.2.1 Spin $\frac{1}{2}$ particles within an external field (paramagnetism)

Consider a system of spin-$\frac{1}{2}$ particles in an external field $\mathbf{B} = (0, 0, B)$ characterized by the Hamiltonian

$$H = -g\mu_{\mathrm{B}}\sum_{i=1}^{N} \widehat{s}_{z,i}B \qquad (4.35)$$

where $\mu_{\mathrm{B}} = \frac{e\hbar}{2mc} = 9.27 \times 10^{-24}\mathrm{JT}^{-1} = 0.671\,41\ \mathrm{k_BK/T}$ is the Bohr magneton. Here the operator of the projection of the spin onto the z-axis, $\widehat{s}_{z,i}$, of the particle at site i has the two eigenvalues

$$s_{z,i} = \pm\frac{1}{2} \qquad (4.36)$$

Using for simplicity $g = 2$ gives with $S_i = 2s_{z,i} = \pm 1$, the eigenenergies are characterized by the set of variables $\{S_i\}$

$$E_{\{S_i\}} = -\mu_{\mathrm{B}} \sum_{i=1}^{N} S_i B. \tag{4.37}$$

Different states of the system can be for example

$$S_1 = 1, \ S_2 = 1, ..., \ S_N = 1 \tag{4.38}$$

which is obviously the ground state if $B > 0$ or

$$S_1 = 1, \ S_2 = -1, ..., \ S_N = -1 \tag{4.39}$$

etc. The partition function is now given as

$$
\begin{aligned}
Z &= \sum_{\{S_i\}} \exp\left(\beta\mu_{\mathrm{B}} \sum_{i=1}^{N} S_i B\right) \\
&= \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} ... \sum_{S_N=\pm 1} \exp\left(\beta\mu_{\mathrm{B}} \sum_{i=1}^{N} S_i B\right) \\
&= \sum_{S_1=\pm 1} \sum_{S_2=\pm 1} ... \sum_{S_N=\pm 1} \prod_{i=1}^{N} \exp\left(\beta\mu_{\mathrm{B}} S_i B\right) \\
&= \sum_{S_1=\pm 1} e^{\beta\mu_{\mathrm{B}} S_1 B} \sum_{S_2=\pm 1} e^{\beta\mu_{\mathrm{B}} S_2 B} ... \sum_{S_N=\pm 1} e^{\beta\mu_{\mathrm{B}} S_N B} \\
&= \left(\sum_{S=\pm 1} e^{\beta\mu_{\mathrm{B}} S B}\right)^{N} = (Z_1)^{N}
\end{aligned}
$$

where Z_1 is the partition function of only one particle. Obviously statistical mechanics of only one single particle does not make any sense. Nevertheless the concept of single particle partition functions is useful in all cases where the Hamiltonian can be written as a sum of commuting, non-interacting terms and the particles are distinguishable, i.e. for

$$H = \sum_{i=1}^{N} h\left(\mathbf{X}_i\right) \tag{4.40}$$

with wave function

$$|\Psi\rangle = \prod_{i} |n_i\rangle \tag{4.41}$$

with

$$h\left(\mathbf{X}_i\right)|n_i\rangle = \varepsilon_n |n_i\rangle. \tag{4.42}$$

It holds that $Z_N = (Z_1)^{N}$ where $Z_1 = \mathrm{tr}\exp\left(-\beta\mathrm{h}\right)$.

For our above example we can now easily evaluate

$$Z_1 = e^{\beta \mu_{\mathrm{B}} B} + e^{-\beta \mu_{\mathrm{B}} B} = 2 \cosh(\beta \mu_{\mathrm{B}} B) \qquad (4.43)$$

which gives

$$F = -N k_{\mathrm{B}} T \log \left[2 \cosh \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right) \right] \qquad (4.44)$$

For the internal energy follows

$$U = \langle E \rangle = -\frac{\partial}{\partial \beta} \log Z = -\mu_{\mathrm{B}} N B \tanh \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right) \qquad (4.45)$$

which immediately gives for the expectation value of the spin operator

$$\langle \widehat{s}_i^z \rangle = \frac{1}{2} \tanh \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right). \qquad (4.46)$$

The entropy is given as

$$S = -\frac{\partial F}{\partial T} = N k_{\mathrm{B}} \log \left[2 \cosh \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right) \right] - \frac{N \mu_{\mathrm{B}} N B}{T} \tanh \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right) \qquad (4.47)$$

which turns out to be identical to $S = \frac{U-F}{T}$.

If $B = 0$ it follows $U = 0$ and $S = N k_{\mathrm{B}} \log 2$, i.e. the number of configurations is degenerate (has equal probability) and there are 2^N such configurations.

For finite B holds

$$S(k_B T \gg \mu_B B) \to N k_B \left(\log 2 - \frac{1}{2} \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right)^2 ... \right) \qquad (4.48)$$

whereas for

$$S(k_B T \ll \mu_B B) \;\; \to \;\; N \frac{\mu_{\mathrm{B}} B}{T} + N e^{-\frac{2 \mu_{\mathrm{B}} B}{k_{\mathrm{B}} T}} - N \frac{\mu_{\mathrm{B}} B}{T} \left(1 - 2 e^{-\frac{2 \mu_{\mathrm{B}} B}{k_{\mathrm{B}} T}} \right)$$

$$\to \;\; \frac{N 2 \mu_{\mathrm{B}} B}{T} \; e^{-\frac{2 \mu_{\mathrm{B}} B}{k_{\mathrm{B}} T}} \to 0 \qquad (4.49)$$

in agreement with the third law of thermodynamics.

The Magnetization of the system is

$$M = g \mu_{\mathrm{B}} \sum_{i=1}^{N} \widehat{s}_i \qquad (4.50)$$

with expectation value

$$\langle M \rangle = N \mu_{\mathrm{B}} \tanh \left(\frac{\mu_{\mathrm{B}} B}{k_{\mathrm{B}} T} \right) = -\frac{\partial F}{\partial B}, \qquad (4.51)$$

i.e.

$$dF = -\langle M \rangle \, dB - S dT. \qquad (4.52)$$

This enables us to determine the magnetic susceptibility

$$\chi = \frac{\partial \langle M \rangle}{\partial B} = \frac{N \mu_{\mathrm{B}}^2}{k_{\mathrm{B}} T} = \frac{C}{T}. \tag{4.53}$$

which is called the Curie-law.

If we are interested in a situation where not the external field, but rather the magnetization is fixed we use instead of F the potential

$$H = F + \langle M \rangle B \tag{4.54}$$

where we only need to invert the $\langle M \rangle$-B dependence, i.e. with

$$B\left(\langle M \rangle\right) = \frac{k_{\mathrm{B}} T}{\mu_{\mathrm{B}}} \tanh^{-1}\left(\frac{\langle M \rangle}{N \mu_B}\right). \tag{4.55}$$

4.2.2   Quantum harmonic oscillator

The analysis of this problem is very similar to the investigation of ideal Bose gases that will be investigated later during the course. Lets consider a set of oscillators. The Hamiltonian of the problem is

$$H = \sum_{i=1}^{N} \left(\frac{p_i^2}{2m} + \frac{k}{2} x_i^2\right)$$

where p_i and x_i are momentum and position operators of N independent quantum Harmonic oscillators. The energy of the oscillators is

$$E = \sum_{i=1}^{N} \hbar \omega_0 \left(n_i + \frac{1}{2}\right) \tag{4.56}$$

with frequency $\omega_0 = \sqrt{k/m}$ and zero point energy $E_0 = \frac{N}{2} \hbar \omega_0$. The integers n_i determine the oscillator eigenstates of the i-th oscillator. They can take the values from 0 to infinity. The Hamiltonian is of the form $H = \sum_{i=1}^{N} h\left(p_i, x_i\right)$ and it follows for the partition function

$$Z_N = \left(Z_1\right)^N.$$

The single oscillator partition function is

$$\begin{aligned} Z_1 &= \sum_{n=0}^{\infty} e^{-\beta \hbar \omega_0 \left(n + \frac{1}{2}\right)} = e^{-\beta \hbar \omega_0 / 2} \sum_{n=0}^{\infty} e^{-\beta \hbar \omega_0 n} \\ &= \frac{e^{-\beta \hbar \omega_0 / 2}}{1 - e^{-\beta \hbar \omega_0}}. \end{aligned}$$

This yields for the partition function

$$\log Z_N = N \log Z_1 = -N \beta \hbar \omega_0 / 2 - N \log\left(1 - e^{-\beta \hbar \omega_0}\right)$$

which enables us to determine the internal energy

$$\begin{aligned}
\langle E \rangle &= -\frac{\partial}{\partial \beta} \log Z_N \\
&= N\hbar\omega_0 \left(\frac{1}{e^{\beta\hbar\omega_0} - 1} + \frac{1}{2} \right).
\end{aligned}$$

The mean value of the oscillator quantum number is then obviously given as

$$\langle n_i \rangle = \langle n \rangle = \frac{1}{e^{\beta\hbar\omega_0} - 1},$$

which tells us that for $k_B T \ll \hbar\omega_0$, the probability of excited oscillator states is exponentially small. For the entropy follows from

$$\begin{aligned}
F &= -k_B T \log Z_N \\
&= N\hbar\omega_0/2 + Nk_B T \log \left(1 - e^{-\frac{\hbar\omega_0}{k_B T}} \right) \quad (4.57)
\end{aligned}$$

that

$$\begin{aligned}
S &= -\frac{\partial F}{\partial T} \\
&= -k_B N \log \left(1 - e^{-\frac{\hbar\omega_0}{k_B T}} \right) + N \frac{\hbar\omega_0}{T} \frac{1}{e^{\frac{\hbar\omega_0}{k_B T}} - 1} \\
&= k_B \left((\langle n \rangle + 1) \log (1 + \langle n \rangle) - \langle n \rangle \log \langle n \rangle \right) \quad (4.58)
\end{aligned}$$

As $T \to 0$ (i..e. for $k_B T \ll \hbar\omega_0$) holds

$$S \simeq k_B N \frac{\hbar\omega_0}{k_B T} e^{-\frac{\hbar\omega_0}{k_B T}} \to 0 \quad (4.59)$$

in agreement with the 3rd law of thermodynamics, while for large $k_B T \gg \hbar\omega_0$ follows

$$S \simeq k_B N \log \left(\frac{k_B T}{\hbar\omega_0} \right). \quad (4.60)$$

For the heat capacity follows accordingly

$$C = T \frac{\partial S}{\partial T} = Nk_B \left(\frac{\hbar\omega_0}{k_B T} \right)^2 \frac{1}{4 \sinh^2 \left(\frac{\hbar\omega_0}{k_B T} \right)}. \quad (4.61)$$

Which vanishes at small T as

$$C \simeq Nk_B \left(\frac{\hbar\omega_0}{k_B T} \right)^2 e^{-\frac{\hbar\omega_0}{k_B T}}, \quad (4.62)$$

while it reaches a constant value

$$C \simeq Nk_B \quad (4.63)$$

as T becomes large. The last result is a special case of the equipartition theorem that will be discussed later.

4.3   Heat capacity and Maxwell relations

The knowledge of the probability distribution

$$p_i = \frac{1}{Z} \exp\left(-\beta E_i\right)$$

allows for a statistical interpretation of various thermodynamic quantities, if we use $Z = \sum_i \exp\left(-\beta E_i\right)$. As we discussed earlier, the free energy

$$F = -k_{\mathrm{B}} T \log Z \tag{4.64}$$

determines the mean (internal) energy

$$\langle E \rangle = \frac{\partial}{\partial \beta} \beta F = F + \beta \frac{\partial F}{\partial \beta} = F + TS \tag{4.65}$$

It holds indeed

$$
\begin{aligned}
\langle E \rangle &= -\frac{\partial}{\partial \beta} \log \sum_i \exp\left(-\beta E_i\right) \\
&= \frac{1}{Z} \sum_i E_i \exp\left(-\beta E_i\right) = \sum_i E_i p_i
\end{aligned}
\tag{4.66}
$$

It is then tempting to analyze the heat capacity

$$C = \frac{\partial \langle E \rangle}{\partial T} = \frac{\partial \beta}{\partial T} \frac{\partial \langle E \rangle}{\partial \beta} = -\frac{1}{k_B T^2} \frac{\partial \langle E \rangle}{\partial \beta} \tag{4.67}$$

It holds

$$
\begin{aligned}
\frac{\partial \langle E \rangle}{\partial \beta} &= \frac{\partial}{\partial \beta} \frac{1}{Z} \sum_i E_i \exp\left(-\beta E_i\right) \\
&= -\frac{1}{Z^2} \frac{\partial Z}{\partial \beta} \sum_i E_i \exp\left(-\beta E_i\right) - \frac{1}{Z} \sum_i E_i^2 \exp\left(-\beta E_i\right) \\
&= -\sum_i E_i^2 p_i + \left(\sum_i E_i p_i\right)^2 .
\end{aligned}
\tag{4.68}
$$

This yields

$$C = k_B \frac{\langle E^2 \rangle - \langle E \rangle^2}{\left(k_B T\right)^2} = k_B \frac{\left\langle \left(E - \langle E \rangle\right)^2 \right\rangle}{\left(k_B T\right)^2} \tag{4.69}$$

i.e. the heat capacity is a measure of the energy fluctuations of a system. Notice that C and $\langle E \rangle$ are extensive. Thus, we introduce the energy per particle $\varepsilon_i = E_i / N$ and the specific heat capacity $c = C/N$ and obtain

$$\left\langle \left(\varepsilon - \langle \varepsilon \rangle\right)^2 \right\rangle = \frac{c k_B T^2}{N} \tag{4.70}$$

which implies that the fluctuations of the energy per particle vanishes as $N \to \infty$.

Let X be a given extensive thermodynamic variable (e.g. volume) and x be the conjugated variable (pressure in case of volume), such that

$$dU = TdS - xdX \tag{4.71}$$

and

$$dF = -SdT - xdX \tag{4.72}$$

The Maxwell relations imply that

$$\frac{\partial S}{\partial X} = \frac{\partial x}{\partial T} \tag{4.73}$$

In the canonical formalism, the dependence on X enters through the variation of the energy eigenvalues $E_i(X)$. For example, in case of X being the volume, the energy changes if the boundary conditions are being changed as function of a variation of the volume. It is interesting to analyze the temperature dependence of x. We start with

$$
\begin{aligned}
x &= -\frac{\partial F}{\partial X} = k_B T \frac{\partial}{\partial X} \log Z \\
&= \frac{k_B T}{Z} \frac{\partial}{\partial X} \sum_i \exp(-\beta E_i) = -\frac{1}{Z} \sum_i \exp(-\beta E_i) \frac{\partial E_i}{\partial X} \\
&= -\left\langle \frac{\partial E}{\partial X} \right\rangle
\end{aligned}
\tag{4.74}
$$

The identity $\frac{\partial F}{\partial X} = \left\langle \frac{\partial E}{\partial X} \right\rangle$ is also referred to as the Hellman-Feynman theorem. We can now analyze

$$
\begin{aligned}
\frac{\partial x}{\partial T} &= -\frac{1}{k_B T^2} \frac{\partial x}{\partial \beta} = -\frac{1}{k_B T^2} \frac{1}{Z^2} \frac{\partial Z}{\partial \beta} \sum_i \exp(-\beta E_i) \frac{\partial E_i}{\partial X} \\
&\quad -\frac{1}{k_B T^2 Z} \sum_i \exp(-\beta E_i) E_i \frac{\partial E_i}{\partial X} \\
&= -\frac{\left\langle E \frac{\partial E}{\partial X} \right\rangle - \left\langle \frac{\partial E}{\partial X} \right\rangle \left\langle E \right\rangle}{k_B T^2} = -\frac{\left\langle (E - \langle E \rangle)\left(\frac{\partial E}{\partial X} - \left\langle \frac{\partial E}{\partial X} \right\rangle\right)\right\rangle}{k_B T^2}
\end{aligned}
\tag{4.75}
$$

i.e. $\frac{\partial x}{\partial T}$ determines the correlations between the energy and $\frac{\partial E}{\partial X}$. On the other hand it holds

$$S = -k_B \sum_i p_i \log p_i = -\frac{\partial F}{\partial T} \tag{4.76}$$

and we have

$$\frac{\partial S}{\partial X} = -\frac{\partial^2 F}{\partial X \partial T} = \frac{\partial x}{\partial T}. \tag{4.77}$$

Finally, it holds

$$
\begin{aligned}
\frac{\partial x}{\partial X} &= \frac{1}{Z^2}\frac{\partial Z}{\partial X}\sum_i \exp\left(-\beta E_i\right)\ \frac{\partial E_i}{\partial X} - \frac{1}{Z}\sum_i \exp\left(-\beta E_i\right)\ \frac{\partial^2 E_i}{\partial X^2} \\
&\quad + \frac{\beta}{Z}\sum_i \exp\left(-\beta E_i\right)\ \left(\frac{\partial E_i}{\partial X}\right)^2 \\
&= -\left\langle\frac{\partial^2 E}{\partial X^2}\right\rangle.
\end{aligned}
\tag{4.78}
$$

4.4   The micro-canonical ensemble

The canonical ensemble is an approach to determine the thermodynamic potentials and equation of state. In both cases we assumed fluctuations of energy and in the grand canonical case even fluctuations of the particle number. Only the expectation numbers of the energy or particle number are kept fixed.

A much more direct approach to determine thermodynamic quantities is based on the micro-canonical ensemble. In accordance with our earlier analysis of maximizing the entropy we start from

$$
S = k_{\mathrm{B}}\log\mathcal{N}\left(E\right)
\tag{4.79}
$$

where we consider an isolated system with conserved energy, i.e. take into account that the we need to determine the number of states with a given energy. If $S\left(E\right)$ is known and we identity E with the internal energy U we obtain the temperature from

$$
\frac{1}{T} = \left.\frac{\partial S\left(E\right)}{\partial E}\right|_{V,N}.
\tag{4.80}
$$

4.4.1   Quantum harmonic oscillator

Let us again consider a set of oscillators with energy

$$
E = E_0 + \sum_{i=1}^{N}\hbar\omega_0 n_i
\tag{4.81}
$$

and zero point energy $E_0 = \frac{N}{2}\hbar\omega_0$. We have to determine the number or realizations of $\{n_i\}$ of a given energy. For example $\mathcal{N}\left(E_0\right) = 1$. There is one realization ($n_i = 0$ for all i) to get $E = E_0$. There are N realization to have an energy $E = E_0 + \omega_0$, i.e. $\mathcal{N}\left(E_0 + \hbar\omega_0\right) = N$. The general case of an energy $E = E_0 + M\hbar\omega_0$ can be analyzed as follows. We consider M black balls and $N-1$ white balls. We consider a sequence of the kind

$$
b_1 b_2...b_{n_1} w b_1 b_2...b_{n_2} w....w b_1 b_2...b_{n_N}
\tag{4.82}
$$

where $b_1 b_2 ... b_{n_i}$ stands for n_i black balls not separated by a white ball. We need to find the number of ways how we can arrange the $N-1$ white balls keeping the total number of black balls fixed at M. This is obviously given by

$$\mathcal{N}(E_0 + M\hbar\omega_0) = \binom{M+N-1}{N-1} = \frac{(M+N-1)!}{(N-1)!M!} \tag{4.83}$$

This leads to the entropy

$$S = k_{\mathrm{B}}\left(\log(M+N-1)! - \log(N-1)! - M!\right) \tag{4.84}$$

For large N it holds $\log N! \simeq N(\log N - 1)$

$$S = k_{\mathrm{B}}N\log\left(\frac{N+M}{N}\right) + k_{\mathrm{B}}M\log\left(\frac{N+M}{M}\right)$$

Thus it holds

$$\beta = \frac{1}{k_B}\frac{\partial S}{\partial E} = \frac{1}{k_B}\frac{1}{\hbar\omega_0}\frac{\partial S}{\partial M} = \frac{1}{\hbar\omega_0}\log\left(1 + \frac{N}{M}\right) \tag{4.85}$$

Thus it follows:

$$M = \frac{N}{e^{\beta\hbar\omega_0} - 1} \tag{4.86}$$

Which finally gives

$$E = N\hbar\omega_0\left(\frac{1}{e^{\beta\hbar\omega_0} - 1} + \frac{1}{2}\right) \tag{4.87}$$

which is the result obtained within the canonical approach.

It is instructive to analyze the entropy without going to the large N and M limit. Then

$$S = k_{\mathrm{B}}\log\frac{\Gamma(N+M)}{\Gamma(N)\Gamma(M+1)} \tag{4.88}$$

and

$$\beta = \frac{1}{\hbar\omega_0}(\psi(N+M) - \psi(M+1)) \tag{4.89}$$

where $\psi(z) = \frac{\Gamma'(z)}{\Gamma(x)}$ is the digamma function. It holds

$$\psi(L) = \log L - \frac{1}{2L} - \frac{1}{12L^2} \tag{4.90}$$

for large L, such that

$$\beta = \frac{1}{\hbar\omega_0}\log\left(\frac{N+M}{M}\right) + \frac{1}{2}\frac{N}{(N+M)M} \tag{4.91}$$

and we recover the above result for large N, M. Here we also have an approach to make explicit the corrections to the leading term. The canonical and micro-canonical obviously differ for small system since $\beta\log\left(1 + \frac{\hbar\omega_0 N}{E-E_0}\right)$ is the exact result of the canonical approach for all system sizes.

Another way to look at this is to write the canonical partition sum as

$$Z = \sum_i e^{\beta E_i} = \frac{1}{\Delta E} \int dE \mathcal{N}(E) e^{-\beta E} = \frac{1}{\Delta E} \int dE e^{-\beta E + S(E)/k_\mathrm{B}} \qquad (4.92)$$

If E and S are large (proportional to N) the integral over energy can be estimated by the largest value of the integrand, determined by

$$-\beta + \frac{1}{k_\mathrm{B}} \frac{\partial S(E)}{\partial E} = 0. \qquad (4.93)$$

This is just the above micro-canonical behavior. At the so defined energy it follows

$$F = -k_\mathrm{B} T \log Z = E - TS(E). \qquad (4.94)$$

Again canonical and micro-canonical ensemble agree in the limit of large systems, but not in general.

Chapter 5

Ideal gases

5.1 Classical ideal gases

5.1.1 The non-relativistic classical ideal gas

Before we study the classical ideal gas within the formalism of canonical ensemble theory we summarize some of its thermodynamic properties. The two equations of state (which we assume to be determined by experiment) are

$$
\begin{aligned}
U &= \frac{3}{2} N k_{\mathrm{B}} T \\
pV &= N k_{\mathrm{B}} T
\end{aligned}
\tag{5.1}
$$

We can for example determine the entropy by starting at

$$
dU = T dS - p dV
\tag{5.2}
$$

which gives (for fixed particle number)

$$
dS = \frac{3}{2} N k_{\mathrm{B}} \frac{dT}{T} + N k_{\mathrm{B}} \frac{dV}{V}
\tag{5.3}
$$

Starting at some state T_0, V_0 with entropy S_0 we can integrate this

$$
\begin{aligned}
S(T,V) &= S_0(T,V) + \frac{3}{2} N k_{\mathrm{B}} \log \frac{T}{T_0} + N k_{\mathrm{B}} \log \frac{V}{V_0} \\
&= N k_{\mathrm{B}} \left(s_0(T,V) + \log \left[\left(\frac{T}{T_0} \right)^{3/2} \frac{V}{V_0} \right] \right).
\end{aligned}
\tag{5.4}
$$

Next we try to actually derive the above equations of state. We start from the Hamiltonian

$$
H = \sum_i \frac{p_i^2}{2m}.
\tag{5.5}
$$

and use that a classical system is characterized by the set of three dimensional momenta and positions $\{\mathbf{p}_i, \mathbf{x}_i\}$. This suggests to write the partition sum as

$$Z = \sum_{\{\mathbf{p}_i, \mathbf{x}_i\}} \exp\left(-\beta \sum_i \frac{p_i^2}{2m}\right). \tag{5.6}$$

For practical purposes it is completely sufficient to approximate the sum by an integral, i.e. to write

$$\sum_{p,x} f(p,x) = \frac{\Delta p \Delta x}{\Delta p \Delta x} \sum_{p,x} f(p,x) \simeq \int \frac{dpdx}{\Delta p \Delta x} f(p,x). \tag{5.7}$$

Here $\Delta p \Delta x$ is the smallest possible unit it makes sense to discretize momentum and energy in. Its value will only be an additive constant to the partition function, i.e. will only be an additive correction to the free energy $\sim 3NT \log \Delta p \Delta x$. The most sensible choice is certainly to use

$$\Delta p \Delta x = h \tag{5.8}$$

with Planck's quantum $h = 6.6260755 \times 10^{-34}$Js. Below, when we consider ideal quantum gases, we will perform the classical limit and demonstrate explicitly that this choice for $\Delta p \Delta x$ is the correct one. It is remarkable, that there seems to be no natural way to avoid quantum mechanics in the description of a classical statistical mechanics problem. Right away we will encounter another left-over of quantum physics when we analyze the entropy of the ideal classical gas.

With the above choice for $\Delta p \Delta x$ follows:

$$\begin{aligned} Z &= \prod_{i=1}^{N} \int \frac{d^d p_i d^d x_i}{h^d} \exp\left(-\beta \frac{p_i^2}{2m}\right) \\ &= V^N \prod_{i=1}^{N} \int \frac{d^d p_i}{h^d} \exp\left(-\beta \frac{p_i^2}{2m}\right) \\ &= \left(\frac{V}{\lambda^d}\right)^N, \end{aligned} \tag{5.9}$$

where we used

$$\int dp \exp\left(-\alpha p^2\right) = \sqrt{\frac{\pi}{\alpha}}, \tag{5.10}$$

and introduced:

$$\lambda = \sqrt{\frac{\beta h^2}{2\pi m}} \tag{5.11}$$

which is the thermal de Broglie wave length. It is the wavelength obtained via

$$k_{\mathrm{B}}T = C\frac{\hbar^2 k_\lambda^2}{2m} \text{ and } k_\lambda = \frac{2\pi}{\lambda} \tag{5.12}$$

with C some constant of order unity. If we compare both expressions, it follow $\lambda = \sqrt{\frac{Ch^2\beta}{2m}} = \sqrt{\frac{Ch^2}{2mk_\mathrm{B}T}}$, i.e. $C = 1/\pi$ is indeed simply a constant of order unity. In case of a generic ideal gas with energy momentum relation $\varepsilon(\mathbf{p})$, it holds in d-dimensions:

$$\left(\frac{h}{\lambda}\right)^d = \int d^d p \exp\left(-\beta\varepsilon(\mathbf{p})\right).$$

For the free energy follows

$$F(V,T) = -k_\mathrm{B}T \log Z = -Nk_\mathrm{B}T \log\left(\frac{V}{\lambda^3}\right) \tag{5.13}$$

Using

$$dF = -SdT - pdV \tag{5.14}$$

gives for the pressure:

$$p = -\left.\frac{\partial F}{\partial V}\right|_T = \frac{Nk_\mathrm{B}T}{V}, \tag{5.15}$$

which is the well known equation of state of the ideal gas. Next we determine the entropy

$$
\begin{aligned}
S &= -\left.\frac{\partial F}{\partial T}\right|_V = Nk_\mathrm{B}\log\left(\frac{V}{\lambda^3}\right) - 3Nk_\mathrm{B}T\frac{\partial\log\lambda}{\partial T} \\
&= Nk_\mathrm{B}\log\left(\frac{V}{\lambda^3}\right) + \frac{3}{2}Nk_\mathrm{B}
\end{aligned}
\tag{5.16}
$$

which gives

$$U = F + TS = \frac{3}{2}Nk_\mathrm{B}T. \tag{5.17}$$

Thus, we recover both equations of state, which were the starting points of our earlier thermodynamic considerations. Nevertheless, there is a discrepancy between our results obtained within statistical mechanics and what follows from thermodynamics. The entropy is not extensive but has a term of the form $N\log V$ which grows faster with the system size than linear and thus seems to overestimate the number of states.

The issue is that there are physical configurations, where (for some values for the momenta P_A and P_B)

$$p_1 = P_A \text{ and } p_2 = P_B \tag{5.18}$$

as well as

$$p_2 = P_A \text{ and } p_1 = P_B, \tag{5.19}$$

i.e. we counted them both. Assuming indistinguishable particles however, all what matters are whether there is some particle with momentum P_A and some particle with P_B, but not which particles are in those states. Thus we assumed

particles to be distinguishable. There are however $N!$ ways to relabel the particles which are all identical. We simply counted all these configurations. The proper expression for the partition sum is therefore

$$Z = \frac{1}{N!} \prod_{i=1}^{N} \int \frac{d^d p_i d^d x_i}{h^d} \exp\left(-\beta \frac{p_i^2}{2m}\right) = \frac{1}{N!} \left(\frac{V}{\lambda^d}\right)^N \qquad (5.20)$$

Using

$$\log N! = N\left(\log N - 1\right) \qquad (5.21)$$

valid for large N, gives

$$
\begin{aligned}
F &= -Nk_{\mathrm{B}}T \log\left(\frac{V}{\lambda^3}\right) + Nk_{\mathrm{B}}T\left(\log N - 1\right) \\
&= -Nk_{\mathrm{B}}T\left(1 + \log\left(\frac{V}{N\lambda^3}\right)\right).
\end{aligned}
\qquad (5.22)
$$

This result differs from the earlier one by a factor $Nk_{\mathrm{B}}T\left(\log N - 1\right)$ which is independent of the volume. Thus that the equation of state for the pressure is unchanged by this issue. However, for the entropy follows now

$$S = Nk_{\mathrm{B}} \log\left(\frac{V}{N\lambda^3}\right) + \frac{5}{2}Nk_{\mathrm{B}} \qquad (5.23)$$

Which gives again

$$U = F + TS = \frac{3}{2}Nk_{\mathrm{B}}T. \qquad (5.24)$$

The reason that the internal energy is correct is due to the fact that it can alternatively be written as an expectation value

$$U = \frac{\frac{1}{N!} \prod_{i=1}^{N} \int \frac{d^d p_i d^d x_i}{h^d} \sum_i \frac{p_i^2}{2m} \exp\left(-\beta \frac{p_i^2}{2m}\right)}{\frac{1}{N!} \prod_{i=1}^{N} \int \frac{d^d p_i d^d x_i}{h^d} \exp\left(-\beta \frac{p_i^2}{2m}\right)} \qquad (5.25)$$

and the factor $N!$ does not change the value of U. The pre-factor $N!$ is called Gibbs correction factor.

The general partition function of a classical real gas or liquid with pair potential $V\left(\mathbf{x}_i - \mathbf{x}_j\right)$ is then characterized by the partition function

$$Z = \frac{1}{N!} \int \prod_{i=1}^{N} \frac{d^d p_i d^d x_i}{h^d} \exp\left(-\beta \sum_{i=1}^{N} \frac{p_i^2}{2m} - \beta \sum_{i,j=1}^{N} V\left(\mathbf{x}_i - \mathbf{x}_j\right)\right). \qquad (5.26)$$

5.1.2 Binary classical ideal gas

We consider a classical, non-relativistic ideal gas, consisting of two distinct types of particles. There are N_A particles with mass M_A and N_B particles with mass M_B in a volume V. The partition function is now

$$Z = \frac{V^{N_A + N_B}}{\lambda_A^{3N_A} \lambda_A^{3N_B} N_A! N_B!} \tag{5.27}$$

and yields with $\log N! = N (\log N - 1)$

$$F = -N_A k_B T \left(1 + \log \left(\frac{V}{N_A \lambda_A^3}\right)\right) - N_B k_B T \left(1 + \log \left(\frac{V}{N_B \lambda_A^3}\right)\right) \tag{5.28}$$

This yields

$$p = \frac{N_A + N_B}{V} k_B T \tag{5.29}$$

for the pressure and

$$S = N_A k_B \log \left(\frac{V}{N_A \lambda_A^3}\right) + N_B k_B \log \left(\frac{V}{N_B \lambda_B^3}\right) + \frac{5}{2} (N_A + N_B) k_B \tag{5.30}$$

for the entropy. We can compare this result with the entropy of an ideal gas of $N = N_A + N_B$ particles.

$$S_0 = (N_A + N_B) k_B \log \left(\frac{V}{(N_A + N_B) \lambda^3}\right) + \frac{5}{2} N k_B \tag{5.31}$$

It follows

$$S - S_0 = N_A k_B \log \left(\frac{(N_A + N_B) \lambda^3}{N_A \lambda_A^3}\right) + N_B k_B \log \left(\frac{(N_A + N_B) \lambda^3}{N_B \lambda_B^3}\right) \tag{5.32}$$

which can be simplified to

$$S - S_0 = -N k_B \left(\phi \log \left(\frac{\lambda_A^3}{\lambda^3} \phi\right) + (1 - \phi) \log \left(\frac{\lambda_B^3}{\lambda^3} (1 - \phi)\right)\right) \tag{5.33}$$

where $\phi = \frac{N_A}{N_A + N_B}$. The additional contribution to the entropy is called mixing entropy.

5.1.3 The ultra-relativistic classical ideal gas

Our calculation for the classical, non-relativistic ideal gas can equally be applied to other classical ideal gases with an energy momentum relation different from $\varepsilon(\mathbf{p}) = \frac{p^2}{2m}$. For example in case of relativistic particles one might consider

$$\varepsilon(\mathbf{p}) = \sqrt{m^2 c^4 + c^2 p^2} \tag{5.34}$$

which in case of massless particles (photons) becomes

$$\varepsilon\left(\mathbf{p}\right) = cp. \tag{5.35}$$

Our calculation for the partition function is in many steps unchanged:

$$
\begin{aligned}
Z &= \frac{1}{N!} \prod_{i=1}^{N} \int \frac{d^3 p_i d^3 x_i}{h^d} \exp\left(-\beta c p_i\right) \\
&= \frac{1}{N!} \left(\frac{V}{h^3}\right)^N \left(\int d^3 p \exp\left(-\beta c p\right)\right)^N \tag{5.36}
\end{aligned}
$$

The remaining momentum integral can be performed easily

$$
\begin{aligned}
I &= \int d^3 p \exp\left(-\beta c p\right) = 4\pi \int_0^\infty p^2 dp \exp\left(-\beta c p\right) \\
&= \frac{4\pi}{(\beta c)^3} \int_0^\infty dx\, x^2 e^{-x} = \frac{8\pi}{(\beta c)^3}. \tag{5.37}
\end{aligned}
$$

This leads to

$$Z = \frac{1}{N!} \left(8\pi V \left(\frac{k_{\mathrm{B}} T}{hc}\right)^3\right)^N \tag{5.38}$$

where obviously the thermal de Broglie wave length of the problem is now given as

$$\lambda = \frac{hc}{k_{\mathrm{B}} T}. \tag{5.39}$$

For the free energy follows with $\log N! = N\left(\log N - 1\right)$ for large N:

$$F = -N k_{\mathrm{B}} T \left[1 + \log\left(\frac{8\pi V}{N\lambda^3}\right)\right]. \tag{5.40}$$

This allows us to determine the equation of state

$$p = -\frac{\partial F}{\partial V} = \frac{N k_{\mathrm{B}} T}{V} \tag{5.41}$$

which is identical to the result obtained for the non-relativistic system. This is because the volume dependence of any classical ideal gas, relativistic or not, is just the V^N term such that

$$F = -N k_{\mathrm{B}} T \log V + \overline{f}\left(T, N\right) \tag{5.42}$$

where $\overline{f}$ does not depend on V, i.e. does not affect the pressure. For the internal energy follows

$$U = F - F\frac{\partial F}{\partial T} = 3 N k_{\mathrm{B}} T \tag{5.43}$$

which is different from the non-relativistic limit.

Comments:

- In both ideal gases we had a free energy of the type

$$F = -Nk_{\mathrm{B}}Tf\left(\frac{(V/N)^{1/3}}{\lambda}\right) \qquad (5.44)$$

with known function $f(x)$. The main difference was the de Broglie wave length.

$$\lambda_{\mathrm{cl}} = \frac{h}{\sqrt{mk_{\mathrm{B}}T}}$$

$$\lambda_{\mathrm{rel}} = \frac{hc}{k_{\mathrm{B}}T} \qquad (5.45)$$

Assuming that there is a mean particle distance d, it holds $V \simeq d^3 N$ such that

$$F = -Nk_{\mathrm{B}}Tf\left(\frac{d}{\lambda}\right). \qquad (5.46)$$

Even if the particle-particle interaction is weak, we expect that quantum interference terms become important if the wavelength of the particles becomes comparable or larger than the inter-particle distance, i.e. quantum effects occur if

$$\lambda > d \qquad (5.47)$$

Considering now the ratio of λ for two systems (a classical and a relativistic one) at same temperature it holds

$$\left(\frac{\lambda_{\mathrm{rel}}}{\lambda_{\mathrm{cl}}}\right)^2 = \frac{mc^2}{k_{\mathrm{B}}T} \qquad (5.48)$$

Thus, if the thermal energy of the classical system is below mc^2 (which must be true by assumption that it is a non-relativistic object), it holds that $\frac{\lambda_{\mathrm{rel}}}{\lambda_{\mathrm{cl}}} \gg 1$. Thus, quantum effects should show up first in the relativistic system.

- One can also estimate at what temperature is the de Broglie wavelength comparable to the wavelength of 500nm, i.e. for photons in the visible part of the spectrum. This happens for

$$T \simeq \frac{hc}{k_{\mathrm{B}}500\mathrm{nm}} \simeq 28,000\mathrm{K} \qquad (5.49)$$

Photons which have a wave length equal to the de Broglie wavelength at room temperature are have a wavelength

$$\frac{hc}{k_{\mathrm{B}}300\mathrm{K}} \simeq 50\mu m \qquad (5.50)$$

- Considering the more general dispersion relation $\varepsilon(\mathbf{p}) = \sqrt{m^2 c^4 + c^2 p^2}$, the non-relativistic limit should be applicable if $k_{\mathrm{B}} T \ll mc^2$, whereas in the opposite limit $k_{\mathrm{B}} T \gg mc^2$ the ultra-relativistic calculation is the right one. This implies that one only needs to heat up a system above the temperature mc^2/k_{B} and the specific heat should increase from $\frac{3}{2} N k_{\mathrm{B}}$ to twice this value. Of course this consideration ignores the effects of particle antiparticle excitations which come into play in this regime as well.

5.1.4   Equipartition theorem

In classical mechanics one can make very general statements about the internal energy of systems with a Hamiltonian of the type

$$H = \sum_{i=1}^{N} \sum_{\alpha=1}^{l} \left(A_\alpha p_{i,\alpha}^2 + B_\alpha q_{i,\alpha}^2 \right) \tag{5.51}$$

where i is the particle index and α the index of additional degrees of freedom, like components of the momentum or an angular momentum. Here $p_{i,\alpha}$ and $q_{i,\alpha}$ are the generalized momentum and coordinates of classical mechanics. The internal energy is then written as

$$
\begin{aligned}
\langle H \rangle &= \frac{\int \prod_{i=1}^{N} \prod_{\alpha=1}^{l} \frac{dp_{i,\alpha} dx_{i,\alpha}}{h} H \exp(-\beta H)}{\int \prod_{i=1}^{N} \prod_{\alpha=1}^{l} \frac{dp_{i,\alpha} dx_{i,\alpha}}{h} \exp(-\beta H)} \\
&= N \sum_{\alpha=1}^{l} \left(\frac{\int dp_\alpha A_\alpha p_\alpha^2 e^{-\beta A_\alpha p_\alpha^2}}{\int dp_\alpha e^{-\beta A_\alpha p_\alpha^2}} + \frac{\int dq_\alpha B_\alpha q_\alpha^2 e^{-\beta B_\alpha q_\alpha^2}}{\int dq_\alpha e^{-\beta B_\alpha q_\alpha^2}} \right) \\
&= N \sum_{\alpha=1}^{l} \left(\frac{k_B T}{2} + \frac{k_B T}{2} \right) = N l k_B T. \tag{5.52}
\end{aligned}
$$

Thus, every quadratic degree of freedom contributes by a factor $\frac{k_{\mathrm{B}} T}{2}$ to the internal energy. In particular this gives for the non-relativistic ideal gas with $l = 3$, $A_\alpha = \frac{1}{2m}$ and $B_a = 0$ that $\langle H \rangle = \frac{3}{2} N k_{\mathrm{B}} T$ as expected. Additional rotational degrees of freedoms in more complex molecules will then increase this number.

5.2   Ideal quantum gases

5.2.1   Occupation number representation

So far we have considered only classical systems or (in case of the Ising model or the system of non-interacting spins) models of distinguishable quantum spins. If we want to consider quantum systems with truly distinguishable particles, one

has to take into account that the wave functions of fermions or bosons behave differently and that states have to be symmetrized or antisymmetric. I.e. in case of the partition sum

$$Z = \sum_n \left\langle \Psi_n \left| e^{-\beta H} \right| \Psi_n \right\rangle \tag{5.53}$$

we have to construct many particle states which have the proper symmetry under exchange of particles. This is a very cumbersome operation and turns out to be highly impractical for large systems.

The way out of this situation is the so called second quantization, which simply respects the fact that labeling particles was a stupid thing to begin with and that one should characterize a quantum many particle system differently. If the label of a particle has no meaning, a quantum state is completely determined if one knows which states of the system are occupied by particles and which not. The states of an ideal quantum gas are obviously the momenta since the momentum operator

$$\widehat{\mathbf{p}}_l = \frac{\hbar}{i} \nabla_l \tag{5.54}$$

commutes with the Hamiltonian of an ideal quantum system

$$H = \sum_{l=1}^{N} \frac{\widehat{\mathbf{p}}_l^2}{2m}. \tag{5.55}$$

In case of interacting systems the set of allowed momenta do not form the eigenstates of the system, but at least a complete basis the eigenstates can be expressed in. Thus, we characterize a quantum state by the set of numbers

$$n_1, n_2, ...n_M \tag{5.56}$$

which determine how many particles occupy a given quantum state with momentum p_1, p_2, ...p_M. In a one dimensional system of size L those momentum states are

$$p_l = \frac{\hbar 2\pi l}{L} \tag{5.57}$$

which guarantee a periodic wave function. For a three dimensional system we have

$$\mathbf{p}_{l_x,l_y,l_z} = \frac{\hbar 2\pi \left(l_x \mathbf{e}_x + l_y \mathbf{e}_y + l_z \mathbf{e}_z \right)}{L}. \tag{5.58}$$

A convenient way to label the occupation numbers is therefore $n_{\mathbf{p}}$ which determined the occupation of particles with momentum eigenvalue $\mathbf{p}$. Obviously, the total number of particles is:

$$N = \sum_{\mathbf{p}} n_{\mathbf{p}} \tag{5.59}$$

whereas the energy of the system is

$$E = \sum_{\mathbf{p}} n_{\mathbf{p}} \varepsilon \left(\mathbf{p} \right) \tag{5.60}$$

If we now perform the summation over all states we can just write

$$Z = \sum_{\{n_\mathbf{p}\}} \exp\left(-\beta \sum_\mathbf{p} n_\mathbf{p}\varepsilon\left(\mathbf{p}\right)\right) \delta_{N,\sum_\mathbf{p} n_\mathbf{p}}. \tag{5.61}$$

where the Kronecker symbol $\delta_{N,\sum_\mathbf{p} n_\mathbf{p}}$ ensures that only configurations with correct particle number are taken into account.

5.2.2 Grand canonical ensemble

At this point it turns out to be much easier to not analyze the problem for fixed particle number, but solely for fixed averaged particle number $\langle N \rangle$. We already expect that this will lead us to the grand-canonical potential

$$\Omega = F - \mu N = U - TS - \mu N \tag{5.62}$$

with

$$d\Omega = -SdT - pdV - Nd\mu \tag{5.63}$$

such that

$$\frac{\partial \Omega}{\partial \mu} = -N. \tag{5.64}$$

In order to demonstrate this we generalize our derivation of the canonical ensemble starting from the principle of maximum entropy. We have to maximize

$$S = -k_\mathrm{B} \sum_{i=1}^{\mathcal{N}} p_i \log p_i. \tag{5.65}$$

with $\mathcal{N}$ the total number of macroscopic states, under the conditions

$$1 \;\; = \;\; \sum_{i=1}^{\mathcal{N}} p_i \tag{5.66}$$

$$\langle E \rangle \;\; = \;\; \sum_{i=1}^{\mathcal{N}} p_i E_i \tag{5.67}$$

$$\langle N \rangle \;\; = \;\; \sum_{i=1}^{\mathcal{N}} p_i N_i \tag{5.68}$$

where N_i is the number of particles in the state with energy E_i. Obviously we are summing over all many body states of all possible particle numbers of the system. We have to minimize

$$I = S + \lambda \left(\sum_{i=1}^{\mathcal{N}} p_i - 1\right) - k_\mathrm{B}\beta \left(\sum_{i=1}^{\mathcal{N}} p_i E_i - \langle E \rangle\right) + k_\mathrm{B}\nu \left(\sum_{i=1}^{\mathcal{N}} p_i N_i - \langle N \rangle\right) \tag{5.69}$$

We set the derivative of I w.r.t p_i equal zero:

$$\frac{\partial I}{\partial p_j} = -k_{\mathrm{B}} \left(\log p_i + 1 \right) + \lambda - k_{\mathrm{B}} \beta E_i + k_{\mathrm{B}} \nu N_i = 0 \tag{5.70}$$

which gives

$$p_i = \exp \left(\frac{\lambda}{k_{\mathrm{B}}} - 1 - \beta E_i + \nu N_i \right) = \frac{1}{Z_g} \exp \left(-\beta E_i + \nu N_i \right) \tag{5.71}$$

where the constant Z_g (and equivalently the Lagrange multiplier λ) are determined by

$$Z_g = \sum_i \exp \left(-\beta E_i + \nu N_i \right) \tag{5.72}$$

which guarantees normalization of the probabilities.

The Lagrange multiplier β is now determined via

$$\langle E \rangle = \frac{1}{Z_g} \sum_i E_i \exp \left(-\beta E_i + \nu N_i \right) = -\frac{\partial}{\partial \beta} \log Z_g. \tag{5.73}$$

whereas

$$\langle N \rangle = \frac{1}{Z_g} \sum_i N_i \exp \left(-\beta E_i + \nu N_i \right) = \frac{\partial}{\partial \nu} \log Z_g \tag{5.74}$$

For the entropy $S = -\frac{\Omega - U + \mu N}{T}$ follows $(\log p_i = -\beta E_i + \nu N_i - \log Z_g)$

$$\begin{aligned}
S &= -k_{\mathrm{B}} \sum_{i=1}^{\mathcal{N}} p_i \log p_i = k_{\mathrm{B}} \sum_{i=1}^{\mathcal{N}} p_i \left(\beta E_i - \nu N_i + \log Z \right) \\
&= k_{\mathrm{B}} \beta \langle E \rangle - k_{\mathrm{B}} \nu \langle N \rangle + k_{\mathrm{B}} \log Z_g = \\
&= \langle E \rangle - \frac{\nu}{\beta} \langle N \rangle + k_{\mathrm{B}} T \log Z_g \\
&\quad -k_{\mathrm{B}} \beta \frac{\partial}{\partial \beta} \log Z + k_{\mathrm{B}} \log Z
\end{aligned} \tag{5.75}$$

which implies

$$\Omega = -k_{\mathrm{B}} T \log Z_g. \tag{5.76}$$

We assumed again that $\beta = \frac{1}{k_{\mathrm{B}} T}$ which can be verified since indeed $S = -\frac{\partial \Omega}{\partial T}$ is fulfilled. Thus we can identify the chemical potential

$$\mu = \frac{\nu}{\beta} \tag{5.77}$$

which indeed reproduces that

$$\langle N \rangle = \frac{\partial}{\partial \nu} \log Z_g = \beta \frac{\partial}{\partial \mu} \log Z_g = -\frac{\partial \Omega}{\partial \mu}. \tag{5.78}$$

Thus, we can obtain all thermodynamic variables by working in the grand canonical ensemble instead.

5.2.3   Partition function of ideal quantum gases

Returning to our earlier problem of non-interacting quantum gases we therefore find

$$Z_g = \sum_{\{n_{\mathbf{p}}\}} \exp\left(-\beta \sum_{\mathbf{p}} n_{\mathbf{p}}\left(\varepsilon\left(\mathbf{p}\right) - \mu\right)\right) \tag{5.79}$$

for the grand partition function. This can be rewritten as

$$Z_g = \sum_{n_{\mathbf{p}_1}} \sum_{n_{\mathbf{p}_2}} \cdots \prod_{\mathbf{p}} e^{-\beta n_{\mathbf{p}}(\varepsilon(\mathbf{p}) - \mu)} = \prod_{\mathbf{p}} \sum_{n_{\mathbf{p}}} e^{-\beta n_{\mathbf{p}}(\varepsilon(\mathbf{p}) - \mu)} \tag{5.80}$$

Fermions: In case of fermions $n_{\mathbf{p}} = 0, 1$ such that

$$Z_{g\text{FD}} = \prod_{\mathbf{p}} \left(1 + e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}\right) \tag{5.81}$$

which gives (FD stands for Fermi-Dirac)

$$\Omega_{\text{FD}} = -k_{\text{B}}T \sum_{\mathbf{p}} \log\left(1 + e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}\right) \tag{5.82}$$

Bosons: In case of bosons $n_{\mathbf{p}}$ can take any value from zero to infinity and we obtain

$$\sum_{n_{\mathbf{p}}} e^{-\beta n_{\mathbf{p}}(\varepsilon(\mathbf{p}) - \mu)} = \sum_{n_{\mathbf{p}}} \left(e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}\right)^{n_{\mathbf{p}}} = \frac{1}{1 - e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}} \tag{5.83}$$

which gives (BE stands for Bose-Einstein)

$$Z_{g\text{BE}} = \prod_{\mathbf{p}} \left(1 - e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}\right)^{-1} \tag{5.84}$$

as well as

$$\Omega_{\text{BE}} = k_{\text{B}}T \sum_{\mathbf{p}} \log\left(1 - e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}\right). \tag{5.85}$$

5.2.4   Classical limit

Of course, both results should reproduce the classical limit. For large temperature follows via Taylor expansion:

$$\Omega_{\text{class}} = -k_{\text{B}}T \sum_{\mathbf{p}} e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}. \tag{5.86}$$

This can be motivated as follows: in the classical limit we expect the mean particle distance, d_0 ($\frac{\langle N \rangle}{V} \simeq d_0^{-d}$) to be large compared to the de Broglie wave length λ, i.e. classically

$$d \gg \lambda. \tag{5.87}$$

The condition which leads to Eq.5.86 is $e^{-\beta(\varepsilon(\mathbf{p})-\mu)} \ll 1$. Since $\varepsilon(\mathbf{p}) > 0$ this is certainly fulfilled if $e^{\beta\mu} \ll 1$. Under this condition holds for the particle density

$$
\begin{aligned}
\frac{\langle N \rangle}{V} &= \frac{1}{V} \sum_{\mathbf{p}} e^{-\beta(\varepsilon(\mathbf{p})-\mu)} = e^{\beta\mu} \int \frac{d^d p}{h^d} \exp\left(-\beta\varepsilon(\mathbf{p})\right) \\
&= \frac{e^{\beta\mu}}{\lambda^d}
\end{aligned}
\tag{5.88}
$$

where we used that the momentum integral always yields the inverse de Broglie length^d. Thus, indeed if $e^{-\beta(\varepsilon(\mathbf{p})-\mu)} \ll 1$ it follows that we are in the classical limit.

Analyzing further, Eq.5.86, we can write

$$
\sum_{\mathbf{p}} f(\varepsilon(\mathbf{p})) = \frac{\Delta\mathbf{p}}{\Delta\mathbf{p}} \sum_{\mathbf{p}} f(\varepsilon(\mathbf{p})) = \Delta\mathbf{p}^{-1} \int d^3 p f(\varepsilon(\mathbf{p}))
\tag{5.89}
$$

with

$$
\Delta\mathbf{p} = \left(\frac{h}{L}\right)^3
\tag{5.90}
$$

due to

$$
\mathbf{p}_{l_x, l_y, l_z} = \frac{\hbar 2\pi (l_x \mathbf{e}_x + l_y \mathbf{e}_y + l_z \mathbf{e}_z)}{L}
\tag{5.91}
$$

such that

$$
\Omega_{\text{class}} = -k_{\text{B}} T V \int \frac{d^3 p}{h^3} e^{-\beta(\varepsilon(\mathbf{p})-\mu)}
\tag{5.92}
$$

In case of the direct calculation we can use that the grand canonical partition function can be obtained from the canonical partition function as follows. Assume we know the canonical partition function

$$
Z(N) = \sum_{i \text{ for fixed } N} e^{-\beta E_i(N)}
\tag{5.93}
$$

then the grand canonical sum is just

$$
Z_g(\mu) = \sum_{N=0}^{\infty} \sum_{i \text{ for fixed } N} e^{-\beta(E_i(N)-\mu N)} = \sum_{N=0}^{\infty} e^{\beta\mu N} Z(N)
\tag{5.94}
$$

Applying this to the result we obtained for the canonical partition function

$$
Z(N) = \frac{1}{N!} \left(\frac{V}{\lambda^3}\right)^N = \frac{1}{N!} \left(V \int \frac{d^3 p}{h^3} e^{-\beta\varepsilon(\mathbf{p})}\right)^N
\tag{5.95}
$$

Thus

$$
\begin{aligned}
Z_g(\mu) &= \sum_{N=0}^{\infty} \frac{1}{N!} \left(V \int \frac{d^3 p}{h^3} e^{-\beta(\varepsilon(\mathbf{p})-\mu)}\right)^N \\
&= \exp\left(V \int \frac{d^3 p}{h^3} e^{-\beta(\varepsilon(\mathbf{p})-\mu)}\right)
\end{aligned}
\tag{5.96}
$$

and it follows the expected result

$$\Omega_{\text{class}} = -k_{\text{B}}TV \int \frac{d^3p}{h^3} e^{-\beta(\varepsilon(\mathbf{p})-\mu)} \tag{5.97}$$

From the Gibbs Duhem relation

$$U = TS - pV + \mu N. \tag{5.98}$$

we found earlier

$$\Omega = -pV \tag{5.99}$$

for the grand canonical ensemble. Since

$$\langle N \rangle = -\frac{\partial \Omega_{\text{class}}}{\partial \mu} = -\beta \Omega_{\text{class}} \tag{5.100}$$

follows

$$pV = Nk_{\text{B}}T \tag{5.101}$$

which is the expected equation of state of the grand-canonical ensemble. Note, we obtain the result, Eq.5.86 or Eq.5.92 purely as the high temperature limit and observe that the indistinguishability, which is natural in the quantum limit, survives the classical limit since our result agrees with the one obtained from the canonical formalism with Gibbs correction factor. Also, the factor $\frac{1}{h^3}$ in the measure follows naturally.

5.2.5   Analysis of the ideal fermi gas

We start from

$$\Omega = -k_{\text{B}}T \sum_{\mathbf{p}} \log \left(1 + e^{-\beta(\varepsilon(\mathbf{p})-\mu)} \right) \tag{5.102}$$

which gives

$$\langle N \rangle = -\frac{\partial \Omega}{\partial \mu} = \sum_{\mathbf{p}} \frac{e^{-\beta(\varepsilon(\mathbf{p})-\mu)}}{1 + e^{-\beta(\varepsilon(\mathbf{p})-\mu)}} = \sum_{\mathbf{p}} \frac{1}{e^{\beta(\varepsilon(\mathbf{p})-\mu)} + 1} = \sum_{\mathbf{p}} \langle n_{\mathbf{p}} \rangle \quad (5.103)$$

i.e. we obtain the averaged occupation number of a given quantum state

$$\langle n_{\mathbf{p}} \rangle = \frac{1}{e^{\beta(\varepsilon(\mathbf{p})-\mu)} + 1} \tag{5.104}$$

Often one uses the symbol $f(\varepsilon(\mathbf{p}) - \mu) = \langle n_{\mathbf{p}} \rangle$. The function

$$f(\omega) = \frac{1}{e^{\beta\omega} + 1} \tag{5.105}$$

is called Fermi distribution function. For $T = 0$ this simplifies to

$$\langle n_{\mathbf{p}} \rangle = \begin{cases} 1 & \varepsilon(\mathbf{p}) < \mu \\ 0 & \varepsilon(\mathbf{p}) > \mu \end{cases} \tag{5.106}$$

States below the energy μ are singly occupied (due to Pauli principle) and states above μ are empty. $\mu(T=0) = E_F$ is also called the Fermi energy.

In many cases will we have to do sums of the type

$$I = \frac{1}{V} \sum_{\mathbf{p}} f(\varepsilon(\mathbf{p})) = \frac{1}{h^3} \int d^3 p f(\varepsilon(\mathbf{p})) \tag{5.107}$$

these three dimensional integrals can be simplified by introducing the density of states

$$\rho(\omega) = \int \frac{d^3 p}{h^3} \delta(\omega - \varepsilon(\mathbf{p})) \tag{5.108}$$

such that

$$I = \int d\omega \rho(\omega) f(\omega) \tag{5.109}$$

We can determine $\rho(\omega)$ by simply performing a substitution of variables $\omega = \varepsilon(p)$ if $\varepsilon(\mathbf{p}) = \varepsilon(p)$ only depends on the magnitude $|\mathbf{p}| = p$ of the momentum

$$I = \frac{4\pi}{h^3} \int p^2 dp f(\varepsilon(p)) = \frac{4\pi}{h^3} \int d\omega \frac{dp}{d\omega} p^2(\omega) f(\omega) \tag{5.110}$$

such that

$$\rho(\omega) = \frac{4\pi m}{h^3} \sqrt{2m\omega} = V A_0 \sqrt{\omega} \tag{5.111}$$

with $A_0 = \frac{4\pi}{h^3} \sqrt{2} m^{3/2}$. Often it is more useful to work with the density of states per particle

$$\rho_0(\omega) = \frac{V \rho(\omega)}{\langle N \rangle} = \frac{V}{\langle N \rangle} A_0 \sqrt{\omega}. \tag{5.112}$$

We determine the chemical potential as function of $\langle N \rangle$ for $T = 0$.

$$\begin{aligned} \langle N \rangle &= \sum_{\mathbf{p}} f(\varepsilon(\mathbf{p})) = V \int \rho(\omega) n(\omega) d\omega \\ &= V A_0 \int_0^{E_F} \omega^{1/2} d\omega = V \frac{2}{3} A_0 E_F^{3/2} \end{aligned} \tag{5.113}$$

which gives

$$E_F = \frac{\hbar^2}{2m} \left(\frac{6\pi^2 \langle N \rangle}{V} \right)^{2/3} \tag{5.114}$$

If $V = d^3 N$ it holds that $E_F \sim \frac{\hbar^2}{2m} d^{-2}$. Furthermore it holds that

$$\rho_0(E_F) = \frac{V}{\langle N \rangle} \frac{2m}{4\pi^2 \hbar^2} \left(\frac{6\pi^2 \langle N \rangle}{V} \right)^{1/3} = \frac{3}{2} \frac{1}{E_F} \tag{5.115}$$

Equally we can analyze the internal energy

$$\begin{aligned} U &= -\frac{\partial}{\partial \beta} \log Z_g = -\frac{\partial}{\partial \beta} \sum_{\mathbf{p}} \log\left(1 + e^{-\beta(\varepsilon(\mathbf{p}) - \mu)}\right) \\ &= \sum_{\mathbf{p}} \frac{\varepsilon(\mathbf{p})}{e^{\beta(\varepsilon(\mathbf{p}) - \mu)} + 1} \end{aligned} \tag{5.116}$$

such that

$$U = \sum_{\mathbf{p}} \varepsilon\left(\mathbf{p}\right)\left\langle n_{\mathbf{p}}\right\rangle = \int \rho\left(\omega\right)\omega n\left(\omega - \mu\right)d\omega = \frac{3}{5}\left\langle N\right\rangle E_{\mathrm{F}} \qquad (5.117)$$

At finite temperatures, the evaluation of the integrals is a bit more subtle. The details, which are only technical, will be discussed in a separate handout. Here we will concentrate on qualitative results. At finite but small temperatures the Fermi function only changes in a regime $\pm k_{\mathrm{B}}T$ around the Fermi energy. In case of metals for example the Fermi energy with $d \simeq 1 - 10\mathring{A}$ leads to $E_{\mathrm{F}} \simeq$ 1...10eV i.e. $E_{F}/k_{\mathrm{B}} \simeq 10^{4}...10^{5}\mathrm{K}$ which is huge compared to room temperature. Thus, metals are essentially always in the quantum regime whereas low density systems like doped semiconductors behave more classically.

If we want to estimate the change in internal energy at a small but finite temperature one can argue that there will only be changes of electrons close to the Fermi level. Their excitation energy is $\sim k_{\mathrm{B}}T$ whereas the relative number of excited states is only $\rho_{0}\left(E_{F}\right)k_{\mathrm{B}}T$. Due to $\rho_{0}\left(E_{F}\right) \sim \frac{1}{E_{\mathrm{F}}}$ it follows in metals $\rho_{0}\left(E_{F}\right)k_{\mathrm{B}}T \ll 1$. We therefore estimate

$$U \simeq \frac{3}{5}\left\langle N\right\rangle E_{\mathrm{F}} + \left\langle N\right\rangle \rho_{0}\left(E_{F}\right)\left(k_{\mathrm{B}}T\right)^{2} + ... \qquad (5.118)$$

at lowest temperature. This leads then to a specific heat at constant volume and particle number

$$\begin{aligned} C\left(T,V,N\right) &= \frac{\partial U\left(T,V,N\right)}{\partial T} \sim 2k_{\mathrm{B}}^{2}\left\langle N\right\rangle \rho_{0}\left(E_{F}\right)T \\ &= 2k_{\mathrm{B}}^{2}V\rho\left(E_{F}\right)T = \gamma T \end{aligned} \qquad (5.119)$$

which is linear, with a coefficient determined by the density of states at the Fermi level. The correct result (see below) is

$$\gamma = \frac{\pi^{2}}{3}k_{\mathrm{B}}^{2}\rho_{0}\left(E_{F}\right) \qquad (5.120)$$

which is almost identical to the one we estimated here. Note, this result does not depend on the specific form of the density of states as long as $\rho_{0}\left(E_{F}\right)$ is a finite constant. It is much more general than the free electron case with a square root density of states.

Let's be a bit more careful with the analysis of the heat capacity. Suppose we want to analyze

$$C\left(T,V,\mu\right) = T\frac{\partial S\left(T,V,\mu\right)}{\partial T}. \qquad (5.121)$$

We then consider the internal energy

$$U\left(T,V,\mu\right) = U\left(S\left(T,V,\mu\right),V,N\left(T,V,\mu\right)\right) \qquad (5.122)$$

such that

$$\frac{\partial U\left(T,V,\mu\right)}{\partial T} = \frac{\partial U}{\partial S}\frac{\partial S\left(T,V,\mu\right)}{\partial T} + \frac{\partial U}{\partial N}\frac{\partial N\left(T,V,\mu\right)}{\partial T} \qquad (5.123)$$

from

$$dU = TdS - pdV + \mu dN \tag{5.124}$$

follows

$$\frac{\partial U(T,V,\mu)}{\partial T} = T\frac{\partial S(T,V,\mu)}{\partial T} + \mu\frac{\partial N(T,V,\mu)}{\partial T} \tag{5.125}$$

such that

$$C(T,V,\mu) = \frac{\partial}{\partial T}(U(T,V,\mu) - \mu N(T,V,\mu)) \tag{5.126}$$

It holds

$$U - \mu N = V\int \rho(\omega)\frac{\omega - \mu}{e^{\beta(\varepsilon-\mu)} + 1}d\omega$$

such that

$$C(T,V,\mu) = V\int \rho(\omega)\frac{(\omega - \mu)^2}{k_B T^2}\frac{e^{\beta(\varepsilon-\mu)}}{\left(e^{\beta(\varepsilon-\mu)} + 1\right)^2}d\omega.$$

The integral is confined to energies $\varepsilon \simeq \mu \pm k_B T$. If the density of states is indeed constant in this regime, we can approximate it by $\rho(\mu)$ and obtain

$$\begin{aligned}
C(T,V,\mu) &= V\rho(\mu)\int \frac{(\omega - \mu)^2}{k_B T^2}\frac{e^{\beta(\varepsilon-\mu)}}{\left(e^{\beta(\varepsilon-\mu)} + 1\right)^2}d\omega \\
&= V\rho(\mu)k_B^2 T\int_{-\infty}^{\infty}\frac{x^2 e^x}{(e^x + 1)^2}dx
\end{aligned}$$

It holds indeed $\int_{-\infty}^{\infty}\frac{x^2 e^x}{(e^x+1)^2}dx = \frac{\pi^2}{3}$ and we obtain

$$C(T,V,\mu) = V\frac{\pi^2}{3}\rho(\mu)k_B^2 T.$$

To compare this result with the heat capacity at constant particle number

$$\begin{aligned}
C(T,V,N) &= T\frac{\partial}{\partial T}S(T,V,\mu(T,N,V)) \\
&= C(T,V,\mu) + T\frac{\partial S(T,V,\mu)}{\partial\mu}\frac{\partial\mu(T,N,V)}{\partial T} \\
&= C(T,V,\mu) + T\frac{\partial N(T,V,\mu)}{\partial T}\frac{\partial\mu(T,N,V)}{\partial T}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial N(T,V,\mu)}{\partial T} &= V\frac{\partial}{\partial T}\int \rho(\omega)\frac{1}{e^{\beta(\varepsilon-\mu)} + 1}d\omega \\
&= V\int \rho(\omega)\frac{(\omega - \mu)}{k_B T^2}\frac{e^{\beta(\varepsilon-\mu)}}{\left(e^{\beta(\varepsilon-\mu)} + 1\right)^2}d\omega.
\end{aligned}$$

If one now assumes a constant density of states, the integral vanishes. Thus we assume

$$\rho(\omega) = \rho(\mu) + \rho'(\mu)(\varepsilon - \mu)$$

and it follows

$$\frac{\partial N\left(T,V,\mu\right)}{\partial T} = V\rho'\left(\mu\right)\frac{\pi^2}{3}k_B^2 T$$

Thus, even if $\frac{\partial\mu(T,N,V)}{\partial T}$ stays constant, we can at small temperatures neglect the difference between $C\left(T,V,\mu\right)$ and $C\left(T,V,N\right)$.

Similarly one can analyze the magnetic susceptibility of a metal. Here the energy of the up ad down spins is different once a magnetic field is applied, such that a magnetization

$$
\begin{aligned}
M &= \mu_B\left(\langle N_\uparrow\rangle - \langle N_\downarrow\rangle\right) \\
&= \mu_B\left\langle N\right\rangle\left(\int_0^{E_{\mathrm{F}}}\rho_0\left(\omega+\mu_B B\right)-\rho_0\left(\omega-\mu_B B\right)\right)d\omega \quad (5.127)
\end{aligned}
$$

For small field we can expand $\rho_0\left(\omega+\mu_B B\right)\simeq\rho_0\left(\omega\right)+\frac{\partial\rho_0(\omega)}{\partial\omega}\mu_B B$ which gives

$$
\begin{aligned}
M &= 2\mu_B^2\left\langle N\right\rangle B\int_0^{E_{\mathrm{F}}}\frac{\partial\rho_0\left(\omega\right)}{\partial\omega}d\omega \\
&= 2\mu_B^2\left\langle N\right\rangle B\rho_0\left(E_F\right) \quad\quad\quad\quad (5.128)
\end{aligned}
$$

This gives for the susceptibility

$$\chi = \frac{\partial M}{\partial B} = 2\mu_B^2\left\langle N\right\rangle\rho_0\left(E_F\right). \quad\quad\quad (5.129)$$

Thus, one can test the assumption to describe electrons in metals by considering the ratio of χ and C_V which are both proportional to the density of states at the Fermi level.

5.2.6   The ideal Bose gas

Even without calculation is it obvious that ideal Bose gases behave very differently at low temperatures. In case of Fermions, the Pauli principle enforced the occupation of all states up to the Fermi energy. Thus, even at $T = 0$ are states with rather high energy involved. The ground state of a Bose gas is clearly different. At $T = 0$ all bosons occupy the state with lowest energy, which is in our case $\mathbf{p} = \mathbf{0}$. An interesting question is then whether this macroscopic occupation of one single state remains at small but finite temperatures. Here, a macroscopic occupation of a single state implies

$$\lim_{\langle N\rangle\to\infty}\frac{\langle n_{\mathbf{p}}\rangle}{\langle N\rangle} > 0. \quad\quad\quad\quad (5.130)$$

We start from the partition function

$$\Omega_{\mathrm{BE}} = k_{\mathrm{B}}T\sum_{\mathbf{p}}\log\left(1-e^{-\beta(\varepsilon(\mathbf{p})-\mu)}\right) \quad\quad\quad (5.131)$$

which gives for the particle number

$$\langle N \rangle = -\frac{\partial \Omega}{\partial \mu} = \sum_{\mathbf{p}} \frac{1}{e^{\beta(\varepsilon(\mathbf{p}) - \mu)} - 1}. \tag{5.132}$$

Thus, we obtain the averaged occupation of a given state

$$\langle n_{\mathbf{p}} \rangle = \frac{1}{e^{\beta(\varepsilon(\mathbf{p}) - \mu)} - 1}. \tag{5.133}$$

Remember that Eq.5.132 is an implicit equation to determine $\mu(\langle N \rangle)$. We rewrite this as

$$\langle N \rangle = V \int d\omega \rho(\omega) \frac{1}{e^{\beta(\omega - \mu)} - 1}. \tag{5.134}$$

The integral diverges if $\mu > 0$ since then for $\omega \simeq \mu$

$$\langle N \rangle \simeq V \int d\omega \frac{\rho(\omega)}{\beta(\omega - \mu)} \to \infty \tag{5.135}$$

if $\rho(\mu) \neq 0$. Since $\rho(\omega) = 0$ if $\omega < 0$ it follows

$$\mu \leq 0. \tag{5.136}$$

The case $\mu = 0$ need special consideration. At least for $\rho(\omega) \sim \omega^{1/2}$, the above integral is convergent and we should not exclude $\mu = 0$.

Lets proceed by using

$$\rho(\omega) = A_0 \sqrt{\omega} \tag{5.137}$$

with $A_0 = \frac{4\pi}{h^3} \sqrt{2} m^{3/2}$. Then follows

$$\begin{aligned}
\frac{\langle N \rangle}{V} &= A_0 \int_0^\infty d\omega \frac{\sqrt{\omega}}{e^{\beta(\omega - \mu)} - 1} \\
&< A_0 \int_0^\infty d\omega \frac{\sqrt{\omega}}{e^{\beta\omega} - 1} \\
&= A_0 (k_B T)^{3/2} \int_0^\infty dx \frac{x^{1/2}}{e^x - 1}
\end{aligned} \tag{5.138}$$

It holds

$$\int_0^\infty dx \frac{x^{1/2}}{e^x - 1} = \frac{\sqrt{\pi}}{2} \varsigma\left(\frac{3}{2}\right) \simeq 2.32 \tag{5.139}$$

We introduce

$$k_B T_0 = a_0 \frac{\hbar^2}{m} \left(\frac{\langle N \rangle}{V}\right)^{2/3} \tag{5.140}$$

with

$$a_0 = \frac{2\pi}{\varsigma\left(\frac{3}{2}\right)^{2/3}} \simeq 3.31. \tag{5.141}$$

The above inequality is then simply:

$$T_0 < T. \tag{5.142}$$

Our approach clearly is inconsistent for temperatures below T_0 (Note, except for pre-factors, $k_B T_0$ is a similar energy scale than the Fermi energy in ideal fermi systems). Another way to write this is that

$$\langle N \rangle < \langle N \rangle \left(\frac{T}{T_0} \right)^{3/2}. \tag{5.143}$$

Note, the right hand side of this equation does not depend on $\langle N \rangle$. It reflects that we could not obtain all particle states below T_0.

The origin of this failure is just the macroscopic occupation of the state with $\mathbf{p} = \mathbf{0}$. It has zero energy but has been ignored in the density of states since $\rho(\omega = 0) = 0$. By introducing the density of states we assumed that no single state is relevant (continuum limit). This is obviously incorrect for $\mathbf{p} = \mathbf{0}$. We can easily repair this if we take the state $\mathbf{p} = \mathbf{0}$ explicitly into account.

$$\langle N \rangle = \sum_{\mathbf{p} > \mathbf{0}} \frac{1}{e^{\beta(\varepsilon(\mathbf{p}) - \mu)} - 1} + \frac{1}{e^{-\beta\mu} - 1} \tag{5.144}$$

for all finite momenta we can again introduce the density of states and it follows

$$\langle N \rangle = V \int d\omega \rho(\omega) \frac{1}{e^{\beta(\omega - \mu)} - 1} + \frac{1}{e^{-\beta\mu} - 1} \tag{5.145}$$

The contribution of the last term

$$N_0 = \frac{1}{e^{-\beta\mu} - 1} \tag{5.146}$$

is only relevant if

$$\lim_{\langle N \rangle \to \infty} \frac{N_0}{\langle N \rangle} > 0. \tag{5.147}$$

If $\mu < 0$, N_0 is finite and $\lim_{\langle N \rangle \to \infty} \frac{N_0}{\langle N \rangle} = 0$. Thus, below the temperature $T = T_0$ the chemical potential must vanish in order to avoid the above inconsistency. For $T < T_0$ follows therefore

$$\langle N \rangle = \langle N \rangle \left(\frac{T}{T_0} \right)^{3/2} + N_0 \tag{5.148}$$

which gives us the temperature dependence of the occupation of the $\mathbf{p} = \mathbf{0}$ state: If $T < T_0$

$$N_0 = \langle N \rangle \left(1 - \left(\frac{T}{T_0} \right)^{3/2} \right). \tag{5.149}$$

and $N_0 = 0$ for $T > T_0$. Then $\mu < 0$.

For the internal energy follows

$$U = V \int d\omega \rho(\omega) \, \omega \frac{1}{e^{-\beta(\omega-\mu)} - 1} \tag{5.150}$$

which has no contribution from the condensate which has $\omega = 0$. The way the existence of the condensate is visible in the energy is via $\mu(T < T_0) = 0$ such that for $T < T_0$

$$U = V A_0 \int d\omega \frac{\omega^{3/2}}{e^{-\beta\omega} - 1} = V A_0 (k_B T)^{5/2} \int_0^\infty dx \frac{x^{3/2}}{e^{-x} - 1} \tag{5.151}$$

It holds again $\int_0^\infty dx \frac{x^{3/2}}{e^{-x}-1} = \frac{3}{4}\sqrt{\pi}\varsigma(5/2) \simeq 1.78$. This gives

$$U = 0.77 \langle N \rangle k_B T \left(\frac{T}{T_0}\right)^{3/2} \tag{5.152}$$

leading to a specific heat (use $U = \alpha T^{5/2}$)

$$C = \frac{\partial U}{\partial T} = \frac{5}{2}\alpha T^{3/2} = \frac{5}{2}\frac{U}{T} \sim T^{3/2}. \tag{5.153}$$

This gives

$$S = \int_0^T \frac{c(T')}{T'} dT' = \frac{5}{2}\alpha \int_0^T T'^{1/2} dT' = \frac{5}{3}\alpha \, T^{3/2} = \frac{5}{3}\frac{U}{T} \tag{5.154}$$

which leads to

$$\Omega = U - TS - \mu N = -\frac{2}{3}U \tag{5.155}$$

The pressure below T_0 is

$$p = -\frac{\partial \Omega}{\partial V} = \frac{5}{3}\frac{\partial U}{\partial V} = 0.08\frac{m^{3/2}}{\hbar^3}(k_B T)^{5/2} \tag{5.156}$$

which is independent of V. This determines the phase boundary

$$p_c = p_c(v_c) \tag{5.157}$$

with specific volume $v = \frac{V}{\langle N \rangle}$ at the transition:

$$p_c = 1.59\frac{\hbar^2}{m}v^{-5/3}. \tag{5.158}$$

5.2.7   Photons in equilibrium

A peculiarity of photons is that they do not interact with each other, but only with charged matter. Because of this, photons need some amount of matter to equilibrate. This interaction is then via absorption and emission of photons, i.e. via a mechanism which changes the number of photons in the system.

Another way to look at this is that in relativistic quantum mechanics, particles can be created by paying an energy which is at least mc^2. Since photons are massless (and are identical to their antiparticles) it is possible to create, without any energy an arbitrary number of photons in the state with $\varepsilon(\mathbf{p}) = 0$. Thus, it doesn't make any sense to fix the number of photons. Since adding a photon with energy zero to the equilibrium is possible, the chemical potential takes the value $\mu = 0$. It is often argued that the number of particles is adjusted such that the free energy is minimal: $\frac{\partial F}{\partial N} = 0$, which of course leads with $\mu = \frac{\partial F}{\partial N}$ to the same conclusion that μ vanishes.

Photons are not the only systems which behave like this. Phonons, the excitations which characterize lattice vibrations of atoms also adjust their particle number to minimize the free energy. This is most easily seen by calculating the canonical partition sum of a system of N_{at} atoms vibrating in a harmonic potential. As usual we find for non-interacting oscillators that $Z(N_{\text{at}}) = Z(1)^{N_{\text{at}}}$ with

$$Z(1) = \sum_{n=0}^{\infty} e^{-\beta\hbar\omega_0\left(n+\frac{1}{2}\right)} = e^{-\frac{\beta\hbar\omega_0}{2}} \frac{1}{1 - e^{-\beta\hbar\omega_0}} \tag{5.159}$$

Thus, the free energy

$$F = N_{\text{at}} \frac{\hbar\omega_0}{2} + k_{\text{B}} T N_{\text{at}} \log\left(1 - e^{-\beta\hbar\omega_0}\right) \tag{5.160}$$

is (ignoring the zero point energy) just the grand canonical potential of bosons with energy $\hbar\omega_0$ and zero chemical potential. The density of states is

$$\rho(\varepsilon) = N_{\text{at}} \delta(\varepsilon - \hbar\omega_0). \tag{5.161}$$

This theory can be more developed and one can consider coupled vibrations between different atoms. Since any system of coupled harmonic oscillators can be mapped onto a system of uncoupled oscillators with modified frequency modes we again obtain an ideal gas of bosons (phonons). The easiest way to determine these modified frequency for low energies is to start from the wave equation for sound

$$\frac{1}{c_s^2} \frac{\partial^2 u}{\partial t^2} = \nabla^2 u \tag{5.162}$$

which gives with the ansatz $u(\mathbf{r}, t) = u_0 \exp(i(\omega t - \mathbf{q} \cdot \mathbf{r}))$ leading to $\omega(\mathbf{q}) = c_s |\mathbf{q}|$, with sound velocity c_s. Thus, we rather have to analyze

$$F = \sum_{\mathbf{q}} \frac{\hbar\omega(\mathbf{q})}{2} + k_{\text{B}} T \sum_{\mathbf{q}} \log\left(1 - e^{-\beta\hbar\omega(\mathbf{q})}\right) \tag{5.163}$$

which is indeed the canonical (or grand canonical since $\mu = 0$) partition sum of ideal bosons. Thus, if we do a calculation for photons we can very easily apply the same results to lattice vibrations at low temperatures. The energy momentum dispersion relation for photons is

$$\varepsilon\left(\mathbf{p}\right) = c\left|\mathbf{p}\right| \tag{5.164}$$

with velocity of light c. This gives

$$U = \sum_{\mathbf{p}} \varepsilon\left(\mathbf{p}\right) \frac{1}{e^{\beta\varepsilon(\mathbf{p})} - 1} = V \int d\varepsilon \frac{\rho\left(\varepsilon\right)\varepsilon}{e^{\beta\varepsilon} - 1}. \tag{5.165}$$

The density of states follows as:

$$
\begin{aligned}
I &= \frac{1}{V} \sum_{\mathbf{p}} F\left(\varepsilon\left(\mathbf{p}\right)\right) = \frac{1}{h^3} \int d^3 p\, F\left(cp\right) \\
&= \frac{1}{h^3} 4\pi \int p^2 dp\, F\left(cp\right) = \frac{4\pi}{c^3 h^3} \int d\varepsilon\, \varepsilon^2 F\left(\varepsilon\right).
\end{aligned} \tag{5.166}
$$

which gives for the density of states

$$\rho\left(\varepsilon\right) = g \frac{4\pi}{c^3 h^3} \varepsilon^2. \tag{5.167}$$

Here $g = 2$ determines the number of photons per energy. This gives the radiation energy as function of frequency $\varepsilon = \hbar\omega$:

$$U = \frac{gV\hbar}{c^3 2\pi^2} \int d\omega \frac{\omega^3}{e^{\beta\hbar\omega} - 1} \tag{5.168}$$

The energy per frequency interval is

$$\frac{dU}{d\omega} = \frac{gV\hbar}{c^3 2\pi^2} \frac{\omega^3}{e^{\beta\hbar\omega} - 1}. \tag{5.169}$$

This gives the famous Planck formula which was actually derived using thermodynamic arguments and trying to combine the low frequency behavior

$$\frac{dU}{d\omega} = \frac{gV k_{\mathrm{B}} T}{c^3 2\pi^2} \omega^2 \tag{5.170}$$

which is the Rayleigh-Jeans law and the high frequency behavior

$$\frac{dU}{d\omega} = \frac{gV\hbar}{c^3 2\pi^2} \omega^3 e^{\beta\hbar\omega} \tag{5.171}$$

which is Wien's formula.

In addition we find from the internal energy that $x = \beta\hbar\omega$

$$
\begin{aligned}
U &= \frac{gV}{\hbar^3 c^3 2\pi^2} \left(k_{\mathrm{B}} T\right)^4 \int dx \frac{x^3}{e^x - 1} \tag{5.172} \\
&= \frac{gV\pi^2}{\hbar^3 c^3 30} \left(k_{\mathrm{B}} T\right)^4 \tag{5.173}
\end{aligned}
$$

where we used $\int dx \frac{x^3}{e^x-1} = \frac{\pi^4}{15}$. U can then be used to determine all other thermodynamic properties of the system.

Finally we comment on the applicability of this theory to lattice vibrations. As discussed above, one important quantitative distinction is of course the value of the velocity. The sound velocity, c_s, is about 10^{-6} times the value of the speed of light $c = 2.99 \times 10^8 \mathrm{ms}^{-1}$. In addition, the specific symmetry of a crystal matters and might cause different velocities for different directions of the sound propagation. Considering only cubic crystals avoids this complication. The option of transverse and longitudinal vibrational modes also changes the degeneracy factor to $g = 3$ in case of lattice vibrations. More fundamental than all these distinctions is however that sound propagation implies that the vibrating atom are embedded in a medium. The interatomic distance, a, will then lead to an lower limit for the wave length of sound ($\lambda > 2\alpha$) and thus to an upper limit $\sim h/(2a) = \hbar\frac{\pi}{a}$. This implies that the density of states will be cut off at high energies.

$$U_{\mathrm{phonons}} = g\frac{4\pi V}{c_s^3 h^3} \int_0^{k_{\mathrm{B}}\theta_D} d\varepsilon \frac{\varepsilon^3}{e^{\beta\varepsilon}-1}. \tag{5.174}$$

The cut off is expressed in terms of the Debye temperature θ_D. The most natural way to determine this scale is by requiring that the number of atoms equals the total integral over the density of states

$$N_{\mathrm{at}} = V \int_0^{k_{\mathrm{B}}\theta_D} \rho\left(\varepsilon\right) d\varepsilon = g\frac{4\pi V}{c^3 h^3} \int_0^{k_{\mathrm{B}}\theta_D} \varepsilon^2 d\varepsilon = g\frac{4\pi V}{3c^3 h^3} \left(k_{\mathrm{B}}\theta_D\right)^3 \tag{5.175}$$

This implies that the typical wave length at the cut off, λ_D, determined by $hc\lambda_D^{-1} = k_{\mathrm{B}}\theta_D$ is

$$\frac{N_{\mathrm{at}}}{V} = g\frac{4\pi}{3}\lambda_D^{-3} \tag{5.176}$$

If one argues that the number of atoms per volume determines the interatomic spacing as $V = \frac{4\pi}{3}a^3 N_{\mathrm{at}}$ leads finally to $\lambda_D = 3.7a$ as expected. Thus, at low temperatures $T \ll \theta_D$ the existence of the upper cut off is irrelevant and

$$U_{\mathrm{phonons}} = \frac{\pi^2 V}{c_s^3 \hbar^3} \frac{g}{30} \left(k_{\mathrm{B}}T\right)^4 \tag{5.177}$$

leading to a low temperature specific heat $C \sim T^3$, whereas for high temperatures $T \gg \theta_D$

$$U_{\mathrm{phonons}} = \frac{\pi^2 V}{3c_s^3 \hbar^3} \frac{g}{30} k_{\mathrm{B}}T \left(k_{\mathrm{B}}\theta_D\right)^3 = N_{\mathrm{at}}k_{\mathrm{B}}T \tag{5.178}$$

which is the expected behavior of a classical system. This makes us realize that photons will never recover this classical limit since they do not have an equivalent to the upper cut off θ_D.

5.2.8 MIT-bag model for hadrons and the quark-gluon plasma

Currently, the most fundamental building blocks of nature are believed to be families of quarks and leptons. The various interactions between these particles are mediated by so called intermediate bosons. In case of electromagnetism the intermediate bosons are photons. The weak interaction is mediated by another set of bosons, called W and Z. In distinction to photons these bosons turn out to be massive and interact among each other (remember photons only interact with electrically charged matter, not with each other). Finally, the strong interaction, which is responsible for the formation of protons, neutrons and other hadrons, is mediated by a set of bosons which are called gluons. Gluons are also self interacting. The similarity between these forces, all being mediated by bosons, allowed to unify their description in terms of what is called the standard model.

A particular challenge in the theory of the strong interaction is the formation of bound states like protons etc. which can not be understood by using perturbation theory. This is not too surprising. Other bound states like Cooper pairs in the theory of superconductivity or the formation of the Hydrogen atom, where a proton and an electron form a localized bound state, are not accessible using perturbation theory either. There is however something special in the strong interaction which goes under the name asymptotic freedom. The interaction between quarks increases (!) with the distance between them. While at long distance, perturbation theory fails, it should be possible to make some progress on short distances. This important insight by Wilczek, Gross and Politzer led to the 2004 Nobel price in physics. Until today hadronization (i.e. formation of hadrons) is at best partly understood and qualitative insight was obtained mostly using rather complex (and still approximate) numerical simulations.

In this context one should also keep in mind that the mass of the quarks ($m_u \simeq 5\text{MeV}$, $m_d \simeq 10\text{MeV}$) is much smaller than the mass of the proton $m_p \simeq 1\text{GeV}$. Here we use units with $c = 1$ such that masses are measured in energy units. Thus, the largest part of the proton mass stems from the kinetic energy of the quarks in the proton.

A very successful phenomenological theory with considerable predictive power are the MIT and SLAC bag models. The idea is that the confinement of quarks can be described by assuming that the vacuum is dia-electric with respect to the color-electric field. One assumes a spherical hadron with distance R. The hadron constantly feels an external pressure from the outside vacuum. This is described by an energy

$$U_B = \frac{4\pi}{3} B R^3 \qquad (5.179)$$

where the so called bag constant B is an unknown constant. Since $U_B \simeq RF = RAp$ with pressure p and bag area A it holds that B is an external pressure acting on the bag. To determine B requires to solve the full underlying quantum chromodynamics of the problem. Within the bag, particles are weakly interacting and for our purposes we assume that they are non-interacting, i.e. quarks and gluons are free fermions and bosons respectively. Since these particles are

confined in a finite region their typical energy is

$$\varepsilon\left(\mathbf{p}\right) \simeq cp \simeq c\frac{h}{R} \tag{5.180}$$

and the total energy is of order

$$U = c\frac{h}{R} + \frac{4\pi}{3}BR^3 \tag{5.181}$$

Minimizing this w.r.t. R yields

$$R_0 = \left(\frac{ch}{4\pi}\right)^{1/4} B^{-1/4} \tag{5.182}$$

using a the known size of a proton $R_0 \simeq 1\text{fm} = 10^{-13}\text{cm}$ gives $B \simeq 60\text{MeV/fm}^3$.
In units where energy, mass, frequency and momentum are measured in electron
volts and length in inverse electron volts ($c = \frac{h}{2\pi} = 1$) this yields $B \simeq 160\text{MeV}$.

Using this simple picture we can now estimate the temperature needed to
melt the bag. If this happens the proton should seize to be a stable hadron and
a new state of matter, called the quark gluon plasma, is expected to form. This
should be the case when the thermal pressure of the gluons and quarks becomes
larger than the bag pressure B

$$p_Q + p_G = B \tag{5.183}$$

Gluons and quarks are for simplicity assumed to be massless. In case of gluons
it follows, just like for photons, that ($k_B = 1$)

$$p_G = g_G\frac{\pi^2}{90}T^4 \tag{5.184}$$

where $g_G = 16$ is the degeneracy factor of the gluon. The calculation for quarks
is more subtle since we need to worry about the chemical potential of these
fermions. In addition, we need to take into account that we can always thermally
excite antiparticles. Thus we discuss the ultra-relativistic Fermi gas in the next
paragraph in more detail.

5.2.9   Ultra-relativistic fermi gas

In the ultra-relativistic limit $k_B T$ can be as large as mc^2 and we need to take into

account that fermions can generate their antiparticles (e.g. positrons in addition
to electrons are excited). Electrons and positrons (quarks and antiquarks) are
always created and annihilated in pairs.

The number of observable electrons is

$$N_e = \sum_{\varepsilon > 0} \frac{1}{e^{\beta(\varepsilon - \mu)} - 1} \tag{5.185}$$

Since positrons are just 1−not observable electrons at negative energy, it follows

$$N_p = \sum_{\varepsilon<0} \left(1 - \frac{1}{e^{\beta(\varepsilon-\mu)}-1}\right) = \sum_{\varepsilon>0} \frac{1}{e^{\beta(\varepsilon+\mu)}-1} \qquad (5.186)$$

The particle excess is then the one unaffected by creation and annihilation of pairs

$$N = N_+ - N_- \qquad (5.187)$$

We conclude that electrons and positrons (quarks and antiquarks) can be considered as two independent ideal fermi systems with positive energy but chemical potential of opposite sign

$$\mu_e = -\mu_p \equiv \mu. \qquad (5.188)$$

It follows with $\varepsilon = cp$

$$
\begin{aligned}
\log Z_g &= g \sum_{\mathbf{P}} \left(\log\left(1 + e^{-\beta(\varepsilon(\mathbf{P})-\mu)}\right) + \log\left(1 + e^{-\beta(\varepsilon(\mathbf{P})+\mu)}\right) \right) \\
&= g \frac{4\pi V}{h^3 c^3} \int \omega^2 d\omega \log\left(1 + e^{-\beta(\omega-\mu)}\right) + \log\left(1 + e^{-\beta(\omega+\mu)}\right)
\end{aligned} \qquad (5.189)
$$

Performing a partial integration gives

$$\log Z_g = g \frac{4\pi V}{h^3 c^3} \frac{\beta}{3} \int \omega^3 d\omega \left(\frac{1}{e^{\beta(\omega-\mu)}+1} + \frac{1}{e^{\beta(\omega+\mu)}+1} \right) \qquad (5.190)$$

substitute $x = \beta(\omega-\mu)$ and $y = \beta(\omega+\mu)$

$$
\begin{aligned}
\log Z_g &= g \frac{4\pi V}{h^3 c^3} \frac{1}{3} \left[\int_{-\beta\mu}^{\infty} dx \frac{\left(\frac{x}{\beta}+\mu\right)^3}{e^{\beta x}+1} + \int_{\beta\mu}^{\infty} dy \frac{\left(\frac{y}{\beta}-\mu\right)^3}{e^{\beta x}+1} \right] \\
&= g \frac{4\pi V}{h^3 c^3} \frac{\beta^{-3}}{3} \left[\int_{0}^{\infty} dx \frac{(x+\beta\mu)^3}{e^{\beta x}+1} + \int_{0}^{\infty} dy \frac{(y-\beta\mu)^3}{e^{\beta x}+1} \right. \\
&\qquad \left. + \int_{-\beta\mu}^{0} dx \frac{(x+\beta\mu)^3}{e^{\beta x}+1} - \int_{0}^{\beta\mu} dy \frac{(y-\beta\mu)^3}{e^{\beta x}+1} \right]
\end{aligned} \qquad (5.191)
$$

The first two integrals can be directly combined, the last two after substitution $y = -x$

$$\log Z_g = g \frac{4\pi V}{h^3 c^3} \frac{\beta^{-3}}{3} \left[\int_{0}^{\infty} dx \frac{2x^3 + 6x(\beta\mu)^2}{e^{\beta x}+1} + \int_{0}^{\beta\mu} dz z^3 \right] \qquad (5.192)$$

with $z = x + \beta\mu$. Using

$$
\begin{aligned}
\int dx \frac{x^3}{e^x - 1} &= \frac{7\pi^4}{120} \\
\int dx \frac{x}{e^x - 1} &= \frac{\pi^2}{12}
\end{aligned} \qquad (5.193)
$$

follows finally

$$\log Z_g = \frac{gV}{h^3c^3}\frac{4\pi}{3}(kT)^3\left(\frac{7\pi^4}{120}+\frac{\pi^2}{2}\left(\frac{\mu}{kT}\right)^2+\frac{1}{4}\left(\frac{\mu}{kT}\right)^4\right) \qquad (5.194)$$

Similarly it follows

$$N = \frac{g4\pi V}{h^3c^3}(kT)^3\left(\frac{\pi^2}{3}\frac{\mu}{kT}+\frac{1}{3}\left(\frac{\mu}{kT}\right)^4\right) \qquad (5.195)$$

It follows for the pressure immediately

$$p = \frac{g}{h^3c^3}\frac{4\pi}{3}(kT)^4\left(\frac{7\pi^4}{120}+\frac{\pi^2}{2}\left(\frac{\mu}{kT}\right)^2+\frac{1}{4}\left(\frac{\mu}{kT}\right)^4\right) \qquad (5.196)$$

Using these results we can now proceed and determine, for a given density of nucleons (or quarks) the chemical potential at a given temperature. For example, in order to obtain about five times nuclear density

$$n_Q = 2.55\frac{1}{\text{fm}^3} \qquad (5.197)$$

at a temperature $T \simeq 150\text{MeV}$ one has a value $\mu \simeq 2.05T$.

Using the above value for the bag constant we are then in a position to analyze our estimate for the transition temperature of the quark gluon plasma

$$p_Q + p_G = B \qquad (5.198)$$

which leads to

$$B = T_c^4\left(\frac{37\pi^2}{90}+\left(\frac{\mu_c}{T}\right)^2+\frac{1}{2\pi^2}\left(\frac{\mu_c}{T_c}\right)^4\right) \qquad (5.199)$$

For example at $\mu_c = 0$ it follows $T_c \simeq 0.7B^{1/4} \simeq 112\text{MeV}$ and for $T_c = 0$ it holds $\mu_c = 2.1B^{1/4} \simeq 336\text{MeV}$.

To relate density and chemical potential one only has to analyze

$$y = x + \frac{x^3}{\pi^2} \qquad (5.200)$$

with $x = \mu/T$ and $y = \frac{54}{g_Q}\frac{n_Q}{T^3}$ with $g_Q = 12$.

Chapter 6

Interacting systems and phase transitions

6.1 The classical real gas

In our investigation of the classical ideal gas we ignored completely the fact that the particles interact with each other. If we continue to use a classical theory, but allow for particle-particle pair interactions with potential $U\left(\mathbf{r}_i - \mathbf{r}_j\right)$, we obtain for the partition function

$$Z = \int \frac{d^{3N}p d^{3N}r}{h^{3N} N!} e^{-\beta\left(\sum_i \frac{\mathbf{p}_i^2}{2m} + \sum_{i<j} U\left(\mathbf{r}_i - \mathbf{r}_j\right)\right)}$$

with $d^{3N}r = \prod_i d^3r$ and similar for $d^{3N}p$. The integration over momentum can be performed in full analogy to the ideal gas and we obtain:

$$Z = \frac{Q_N\left(V, T\right)}{N! \lambda^{dN}} \tag{6.1}$$

where we have:

$$
\begin{aligned}
Q_N\left(V, T\right) &= \int d^{3N}r \exp\left(-\beta \sum_{i<j} U\left(\mathbf{r}_i - \mathbf{r}_j\right)\right) \\
&= \int d^{3N}r \prod_{i<j} e^{-\beta U_{ij}} \tag{6.2}
\end{aligned}
$$

If $\beta U_{ij} \ll 1$ it still holds that $e^{-\beta U_{ij}} \simeq 1$ and an expansion is nontrivial, however one can consider instead

$$f_{ij} = e^{-\beta U_{ij}} - 1 \tag{6.3}$$

which is also well defined in case of large U_{ij}. Then

$$\prod_{i<j}(1+f_{ij}) = 1 + \sum_{i<j}f_{ij} + \sum_{i<k;l<m}f_{ik}f_{lm} + \dots \tag{6.4}$$

And it follows

$$Q_N(V,T) = V^N + V^{N-1}\frac{N(N-1)}{2}\int d^3r\left(e^{-\beta U(r)}-1\right) \tag{6.5}$$

where we took into account that there are $\frac{N(N-1)}{2}$ pairs $i<j$. If we set

$$a(T) = \int d^3r\left(e^{-\beta U(r)}-1\right) \tag{6.6}$$

follows

$$Z = \frac{V^N}{N!\lambda^{dN}}\left(1+\frac{N^2}{2V}a(T)\right) \tag{6.7}$$

It follows for the equation of state that

$$
\begin{aligned}
p &= -\frac{\partial F}{\partial V} = k_{\mathrm{B}}T\frac{\partial\log Z}{\partial V} = \frac{Nk_{\mathrm{B}}T}{V} - k_{\mathrm{B}}T\frac{\frac{a}{2}\left(\frac{N}{V}\right)^2}{1+\frac{a}{2}\left(\frac{N}{V}\right)^2}\\
&\simeq \frac{Nk_{\mathrm{B}}T}{V}\left(1-\frac{a}{2}\frac{N}{V}\right)
\end{aligned} \tag{6.8}
$$

An often used potential in this context is the Lennard-Jones potential

$$U(r) = U_0\left(\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^6\right) \tag{6.9}$$

which has a minimum at $r=r_0$ and consists of a short range repulsive and a longer range attractive part. For simplicity we approximate this by

$$U(r) = \begin{cases} \infty & r<r_0 \\ -U_0\left(\frac{r_0}{r}\right)^6 & r \geq r_0 \end{cases} \tag{6.10}$$

called Sutherland potential. Then

$$a(T) = -4\pi\int_0^{r_0}r^2dr + 4\pi\int_{r_0}^{\infty}r^2\left(e^{\beta U_0\left(\frac{r_0}{r}\right)^6}-1\right)dr \tag{6.11}$$

expanding for small potential gives with $4\pi\beta\int_{r_0}^{\infty}r^2U_0\left(\frac{r_0}{r}\right)^6dr = \frac{4\pi}{3}r_0^3\beta U_0$ such that

$$a(T) = -\frac{4\pi}{3}r_0^3\left(1-\beta U_0\right). \tag{6.12}$$

This gives with $v = \frac{V}{N}$

$$p = \frac{Nk_{\mathrm{B}}T}{V}\left(1+\frac{2\pi}{3v}r_0^3\left(1-\beta U_0\right)\right) \tag{6.13}$$

or

$$p + \frac{2\pi}{3v^2}r_0^3 U_0 = \frac{k_B T}{v}\left(1 + \frac{2\pi r_0^3}{3v}\right) = k_B T \left(v - \frac{2\pi r_0^3}{3}\right)^{-1} \tag{6.14}$$

which gives

$$\left(p + \frac{a}{v^2}\right)(v - b) = k_B T \tag{6.15}$$

which is the van der Waals equation of state with

$$a = \frac{2\pi}{3}r_0^3 U_0$$
$$b = \frac{2\pi r_0^3}{3} \tag{6.16}$$

The analysis of this equation of state yields that for temperatures below

$$T_{cr} = \frac{a}{27bk_B} = \frac{U_0}{27k_B} \tag{6.17}$$

there are three solutions V_i of Eq.**??** for a given pressure. One of these solutions (the one with intermediate volume) can immediately be discarded since here $\frac{dp}{dV} > 0$, i.e. the system has a negative compressibility (corresponding to a local maximum of the free energy). The other two solutions can be interpreted as coexistent high density (liquid) and low density (gas) fluid.

6.2 Classification of Phase Transitions

Phase transitions are transitions between qualitatively distinct equilibrium states of matter such as solid to liquid, ferromagnet to paramagnet, superconductor to normal conductor etc. The first classification of phase transitions goes back to 1932 when Paul Ehrenfest argued that a phase transition of nth-order occurs when there is a discontinuity in the nth-derivative of a thermodynamic potential. Thus, at a 1st-order phase transition the free energy is assumed to be continuous but has a discontinuous change in slope at the phase transition temperature T_c such that the entropy (which is a first derivative) jumps from a larger value in the high temperature state $S(T_c + \varepsilon)$ to a smaller value $S(T_c - \varepsilon)$ in the low temperature state, where ε is infinitesimally small. Thus a latent heat

$$\Delta Q = T_c \Delta S = T_c\left(S(T_c + \varepsilon) - S(T_c - \varepsilon)\right) \tag{6.18}$$

is needed to go from the low to the high temperature state. Upon cooling, the system jumps into a new state with very different internal energy (due to $F = U - TS$ must U be discontinuous if S is discontinuous).

Following Ehrenfest's classification, a second order phase transition should have a discontinuity in the second derivative of the free energy. For example the entropy should then be continuous but have a change in slope leading to a jump in the value of the specific heat. This is indeed what one finds in approximate, so called mean field theories. However a more careful analysis of

the role of spatial fluctuations (see below) yields that the specific heat rather behaves according to a power-law with $C \sim (T - T_c)^{-\alpha}$ or (like in the two dimensional Ising model) diverges logarithmically. In some cases $\alpha < 0$ occurs making it hard to identify the effects of these fluctuations. In other cases, like conventional superconductors, the quantitative effect of fluctuations is so small that the experiment is virtually indistinguishable from the mean field expectation and Ehrenfest's picture is a very sensible one.

More generally one might classify phase transitions in a way that at a nth-order transition a singularity of some kind occurs in the nth-derivative of the free energy, whereas all lower derivatives are continuous. The existence of a latent heat in first order transitions and its absence in a second order transition is then valid even if one takes fluctuations into account.

Finally one needs to realize that strictly no phase transition exists in a finite system. For a finite system the partition sum is always finite. A finite sum of analytic functions $\sim e^{-\beta E_i}$ is analytic itself and does not allow for similarities in its derivatives. Thus, the above classification is valid only in the thermodynamic limit of infinite systems.

6.3 Gibbs phase rule and first order transitions

We next discuss the issue of how many state variables are necessary to uniquely determine the state of a system. To this end, we start from an isolated system which contains K different particle species (chemical components) and P different phases (solid, liquid, gaseous,...) that coexist. Each phase can be understood as a partial system of the total system and one can formulate the first law for each phase, where we denote quantities of the ith phase by superscript $i = 1, ..., P$. We have

$$dU^{(i)} = T^{(i)}dS^{(i)} - p^{(i)}dV^{(i)} + \sum_{l=1}^{P} \mu_l^{(i)}dN_l^{(i)} \qquad (6.19)$$

Other terms also may appear, if electric or magnetic effects play a role. In this formulation of the first law, $U^{(i)}$ of phase i is a function of the extensive state variables $S^{(i)}$, $V^{(i)}$, $N_l^{(i)}$, i.e., it depends on $K + 2$ variables. Altogether we therefore have $P(K + 2)$ extensive state variables. If the total system is in thermodynamic equilibrium, we have $T^{(i)} = T$, $p^{(i)} = p$ and $\mu_l^{(i)} = \mu_l$. Each condition contains $P-1$ equations, so that we obtain a system of $(P - 1)(K + 2)$ equations. Since $T^{(i)}$, $p^{(i)}$, and $\mu_l^{(i)}$ are functions of $S^{(i)}$, $V^{(i)}$, $N_l^{(i)}$ we can eliminate one variable with each equation. Thus, we only require

$$(K + 2)P - (K + 2)(P - 1) = K + 2 \qquad (6.20)$$

extensive variables to determine the equilibrium state of the total system. As we see, this number is independent of the number of phases. If we now consider

that exactly P extensive variables (e.g., $V^{(i)}$ with $i = 1, ..., P$) determine the size of the phases (i.e., the volumes occupied by each), one needs

$$F = K + 2 - P$$

intensive variables. This condition is named after J.W. Gibbs and is called Gibbs' phase rule. It is readily understood with the help of concrete examples. Let us for instance think of a closed pot containing a vapor. With $K = 1$ we need $3 = K+2$ extensive variables for a complete description of the system, e.g., S, V, and N. One of these (e.g., V), however, determines only the size of the system. The intensive properties are completely described by $F = 1 + 2 - 1 = 2$ intensive variables, for instance by the pressure and the temperature. Then also U/V, S/V, N/V, etc. are fixed and by additionally specifying V one can also obtain all extensive quantities.

If both vapor and liquid are in the pot and if they are in equilibrium, we can only specify one intensive variable, $F - 1 + 2 - 2 = 1$, e.g., the temperature. The vapor pressure assumes automatically its equilibrium value. All other intensive properties of the phases are also determined. If one wants in addition to describe the extensive properties, one has to specify for instance V^{liq} and V^{vap}, i.e., one extensive variable for each phase, which determines the size of the phase (of course, one can also take N^{liq} and N^{vap}, etc.). Finally, if there are vapor, liquid, and ice in equilibrium in the pot, we have $F = 1 + 2 - 3 = 0$. This means that all intensive properties are fixed: pressure and temperature have definite values. Only the size of the phases can be varied by specifying V^{liq}, V^{sol}, and V^{vap}. This point is also called triple point of the system.

6.4 The Ising model

Interacting spins which are allowed to take only the two values $S_i = \pm 1$ are often modeled in terms of the Ising model

$$H = -\sum_{i,j} J_{ij} S_i S_{i+1} - \mu B \sum_i S_i \qquad (6.21)$$

where J_{ij} is the interaction between spins at sites i and j. The microscopic origin of the J_{ij} can be the dipol-dipol interaction between spins or exchange interaction which has its origin in a proper quantum mechanical treatment of the Coulomb interaction. The latter is dominant in many of the known $3d$, $4f$ and $5f$ magnets. Often the Ising model is used in a context unrelated to magnetism, like the theory of binary alloys where $S_i = \pm 1$ corresponds to the two atoms of the alloy and J_{ij} characterizes the energy difference between two like and unlike atoms on sites i and j. This and many other applications of this model make the Ising model one of the most widely used concepts and models in statistical mechanics. The model has been solved in one and two spatial dimensions and for situations where every spin interacts with every other spin with an interaction J_{ij}/N. No analytic solution exists for three dimensions even though computer

simulations for a model with nearest neighbor coupling J demonstrate that the model is ordered, with $\lim_{B \to 0} \langle S_i \rangle \neq 0$, below a temperature $T_c \simeq 4.512J$. Similarly the solution for the square lattice in $d = 2$ yields an ordered state below $T_c = 2J/\mathrm{arc\,coth}\,\sqrt{2} \simeq 2.269J$, while the Ising model in one spatial dimension has $T_c = 0$, i.e. the ground state is ordered while no zero field magnetization exists at a finite temperature. The latter is caused by the fact that any domain wall in a d-dimensional model (with short range interactions) costs an energy $E_d = JN_d^{d-1}$, where N_d is the number of spins in the domain. While this is a large excitation energy for $d > 1$ it is only of order J in $d = 1$ and domains with opposite magnetization of arbitrary size can easily be excited at finite T. This leads to a breakdown of long range order in $d = 1$.

6.4.1   Exact solution of the one dimensional model

A first nontrivial model of interacting particles is the one dimensional Ising model in an external field, with Hamiltonian

$$
\begin{aligned}
H &= -J \sum_i S_i S_{i+1} - \mu B \sum_i S_i \\
&= -J \sum_i S_i S_{i+1} - \frac{\mu B}{2} \sum_i (S_i + S_{i+1})
\end{aligned}
\tag{6.22}
$$

The partition function is

$$
\begin{aligned}
Z &= \sum_{\{S_i\}} e^{-\beta H} \\
&= \sum_{S_1 = \pm 1} \cdots \sum_{S_N = \pm 1} e^{\beta \sum_i \left[J S_i S_{i+1} + \frac{\mu B}{2} S_i + S_{i+1} \right]}
\end{aligned}
\tag{6.23}
$$

We use the method of transfer matrices and define the operator T defined via its matrix elements:

$$
\langle S_i \left| T \right| S_{i+1} \rangle = e^{\beta \sum_i \left[J S_i S_{i+1} + \frac{\mu B}{2} (S_i + S_{i+1}) \right]}
\tag{6.24}
$$

The operator can be represented as 2×2 matrix

$$
T = \begin{pmatrix} e^{\beta(J + \mu_B B)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J - \mu_B B)} \end{pmatrix}
\tag{6.25}
$$

It holds then that

$$
\begin{aligned}
Z &= \sum_{S_1 = \pm 1} \cdots \sum_{S_N = \pm 1} \langle S_1 \left| T \right| S_2 \rangle \langle S_2 \left| T \right| S_3 \rangle \ldots \langle S_N \left| T \right| S_1 \rangle \\
&= \sum_{S_1 = \pm 1} \langle S_1 \left| T^N \right| S_1 \rangle = \mathrm{tr} T^N.
\end{aligned}
\tag{6.26}
$$

This can be expressed in terms of the eigenvalues of the matrix T

$$
\lambda_{\pm} = e^y \cosh x \pm \left[e^{-2y} + e^{2y} \sinh^2 x \right]^{1/2}
\tag{6.27}
$$

with

$$
\begin{aligned}
x &= \beta\mu_{\mathrm{B}}B \\
y &= \beta J.
\end{aligned}
\tag{6.28}
$$

It follows

$$
Z = \lambda_+^N + \lambda_-^N
\tag{6.29}
$$

yielding

$$
\begin{aligned}
F &= -k_{\mathrm{B}}T\left(N\log\lambda_+ + \log\left(1 + \left(\frac{\lambda_-}{\lambda_+}\right)^N\right)\right) \\
&= -k_{\mathrm{B}}TN\log\lambda_+
\end{aligned}
\tag{6.30}
$$

where we used in the last step that $\lambda_- < \lambda_+$ and took the limit $N \to \infty$.

For the non-interacting system ($y = 0$) we obtain immediately the result of non-interacting spins

$$
F = -Nk_{\mathrm{B}}T\log\left(2\cosh x\right).
\tag{6.31}
$$

For the magnetization

$$
M = -\frac{\partial F}{\partial B} = N\mu_B\frac{\sinh x}{\left(e^{-4y} + \sinh^2 x\right)^{1/2}}.
\tag{6.32}
$$

For $T = 0$ follows $M = N\mu_B$, i.e. the system is fully polarized. In particular $M(B \to 0, T = 0) \neq 0$. On the other hand it holds for any finite temperature that

$$
M(T, B \to 0) \to 0.
\tag{6.33}
$$

Thus, there is no ordered state with finite zero-field magnetization at finite temperature. In other words, the one dimensional Ising model orders only at zero temperature.

6.4.2 Mean field approximation

In this section we discuss an approximate approach to solve the Ising model which is based on the assumption that correlations, i.e. simultaneous deviations from mean values like

$$
(S_i - \langle S_i\rangle)(S_j - \langle S_j\rangle)
\tag{6.34}
$$

are small and can be neglected. Using the identity

$$
S_iS_j = (S_i - \langle S_i\rangle)(S_j - \langle S_j\rangle) + S_i\langle S_j\rangle + \langle S_i\rangle S_j - \langle S_i\rangle\langle S_j\rangle
\tag{6.35}
$$

we can ignore the first term and write in the Hamiltonian of the Ising model, assuming $\langle S_i\rangle = \langle S\rangle$ independent of i:

$$
H = -\sum_i \left(zJ\langle S\rangle + \mu B\right)S_i - \frac{z}{2}N\langle S\rangle^2
\tag{6.36}
$$

This model is, except for the constant $\frac{-z}{2} N \langle S \rangle^2$ equal to the energy of non-interacting spins in an effective magnetic field

$$B_{eff} = B + \frac{zJ}{\mu} \langle S \rangle . \tag{6.37}$$

Thus, we can use our earlier result for this model to find the expectation value $\langle S \rangle$ in terms of the field

$$\langle S \rangle = \tanh \left(\frac{\mu B_{eff}}{k_{\mathrm{B}} T} \right) \tag{6.38}$$

Setting now the external field $B = 0$, we obtain

$$\langle S \rangle = \tanh \left(\frac{zJ}{k_{\mathrm{B}} T} \langle S \rangle \right) \tag{6.39}$$

If

$$T > T_c = \frac{zJ}{k_{\mathrm{B}}} \tag{6.40}$$

this nonlinear equation has only the solution $\langle S \rangle = 0$. However, for T below T_c another solution with finite $\langle S \rangle$ emerges continuously. This can be seen more directly by expanding the above tanh for small argument, and it follows

$$\langle S \rangle = \tanh \left(\frac{T_c}{T} \langle S \rangle \right) \simeq \frac{T_c}{T} \langle S \rangle - \frac{1}{3} \left(\frac{T_c}{T} \right)^2 \langle S \rangle^3 \tag{6.41}$$

which yields

$$\langle S \rangle \propto (T_c - T)^{1/2} \tag{6.42}$$

.i.e. the magnetization vanishes continuously. Right at T_c a small external field B causes a finite magnetization which is determines by

$$\langle S \rangle = \frac{\mu B + zJ \langle S \rangle}{k_{\mathrm{B}} T_c} - \frac{1}{3} \left(\frac{\mu B + zJ \langle S \rangle}{k_{\mathrm{B}} T_c} \right)^3 \tag{6.43}$$

which yields

$$\langle S \rangle = \left(3 \frac{\mu B}{k_{\mathrm{B}} T_c} \right)^{1/3} . \tag{6.44}$$

Finally we can analyze the magnetic susceptibility $\chi = \frac{\partial M}{\partial B}$ with $M = \mu N \langle S \rangle$. We first determine $\chi_S = \frac{\partial \langle S \rangle}{\partial B}$ above T_c. It follows for small field

$$\langle S \rangle \simeq \frac{\mu B + k_{\mathrm{B}} T_c \langle S \rangle}{k_{\mathrm{B}} T} \tag{6.45}$$

such that

$$\chi_S = \frac{\mu + k_{\mathrm{B}} T_c \chi_S}{k_{\mathrm{B}} T} \tag{6.46}$$

yielding

$$\chi_S = \frac{\mu}{k_B\,(T - T_c)} \tag{6.47}$$

and we obtain

$$\chi = \frac{C}{T - T_c} \tag{6.48}$$

with Curie constant $C = \mu^2 N/k_B$. This is the famous Curie-Weiss law which demonstrates that the uniform susceptibility diverges at an antiferromagnetic phase transition.

6.5 Landau theory of phase transitions

Landau proposed that one should introduce an order parameter to describe the properties close to a phase transition. This order parameter should vanish in the high temperature phase and be finite in the ordered low temperature phase. The mathematical structure of the order parameter depends strongly on the system under consideration. In case of an Ising model the order parameter is a scalar, in case of the Heisenberg model it is a vector. For example, in case of a superconductor or the normal fluid - superfluid transition of ^{4}He it is a complex scalar, characterizing the wave function of the coherent low temperature state. Another example are liquid crystals where the order parameter is a second rank tensor.

In what follows we will first develop a Landau theory for a scalar, Ising type order. Landau argued that one can expand the free energy density in a Taylor series with respect to the order parameter ϕ. This should be true close to a second order transition where ϕ vanishes continuously:

$$f\,(\phi) = f_0 - h\phi + \frac{a}{2}\phi^2 + \frac{b}{3}\phi^3 + \frac{c}{4}\phi^4 + ... \tag{6.49}$$

The physical order parameter is the determined as the one which minimizes f

$$\left.\frac{\partial f}{\partial \phi}\right|_{\phi=\phi_0} = 0. \tag{6.50}$$

If $c < 0$ this minimum will be at $\pm\infty$ which is unphysical. If indeed $c < 0$ one needs to take a term $\sim \phi^6$ into account and see what happens. In what follows we will always assume $c > 0$. In the absence of an external field should hold that $f\,(\phi) = f\,(-\phi)$, implying $h = b = 0$. Whether or not there is a minimum for $\phi \neq 0$ depends now on the sign of a. If $a > 0$ the only minimum of

$$f\,(\phi) = f_0 + \frac{a}{2}\phi + \frac{c}{4}\phi^4 \tag{6.51}$$

is at $\phi = 0$. However, for $a < 0$ there are two a new solutions $\phi = \pm\sqrt{\frac{-a}{c}}$. Since ϕ is expected to vanish at $T = T_c$ we conclude that $a\,(T)$ changes sign at T_c suggesting the simple ansatz

$$a\,(T) = a_0\,(T - T_c) \tag{6.52}$$

with $a_0 > 0$ being at most weakly temperature dependent. This leads to a temperature dependence of the order parameter $\sqrt{\frac{a_0(T_c - T)}{c}}$

$$\phi_0 = \begin{cases} \sqrt{\frac{a_0(T_c - T)}{c}} & T < T_c \\ 0 & T > T_c \end{cases} \tag{6.53}$$

It will turn out that a power-law relation like

$$\phi \sim (T_c - T)^\beta \tag{6.54}$$

is valid in a much more general context. The main change is the value of β. The prediction of the Landau theory is $\beta = \frac{1}{2}$.

Next we want to study the effect of an external field (= magnetic field in case ϕ characterizes the magnetization of an Ising ferromagnet). This is done by keeping the term $h\phi$ in the expansion for f. The actual external field will be proportional to h. Then we find that f is minimized by

$$a\phi_0 + c\phi_0^3 = h \tag{6.55}$$

Right at the transition temperature where $a = 0$ this gives

$$\phi_0^3 \sim h^{1/\delta} \tag{6.56}$$

where the Landau theory predicts $\delta = 3$. Finally we can analyze the change of the order parameter with respect to an external field. We introduce the susceptibility

$$\chi = \left. \frac{\partial \phi_0}{\partial h} \right|_{h \to 0} \tag{6.57}$$

and find from Eq.6.55

$$a\chi + 3c\phi_0^2(h = 0)\chi = 1 \tag{6.58}$$

using the above result for $\phi_0^2(h = 0) = \frac{a}{c}$ if $T < T_c$ and $\phi_0^2(h = 0) = 0$ above T_c gives

$$\chi = \begin{cases} \frac{1}{4a_0}(T_c - T)^{-\gamma} & T < T_c \\ \frac{1}{a_0}(T - T_c)^{-\gamma} & T > T_c \end{cases} \tag{6.59}$$

with exponent $\gamma = 1$.

Next we consider the specific heat where we insert our solution for ϕ_0 into the free energy density.

$$f = \frac{a(T)}{2}\phi_0^2 + \frac{c}{4}\phi_0^4 = \begin{cases} -\frac{a_0^2}{4c}(T - T_c)^2 & T < T_c \\ 0 & T > T_c \end{cases} \tag{6.60}$$

This yields for the specific heat per volume

$$c = -T\frac{\partial^2 f}{\partial T} = \begin{cases} \frac{a_0^2}{4c}T & T < T_c \\ 0 & T > T_c \end{cases} . \tag{6.61}$$

The specific heat is discontinuous. As we will see later, the general form of the specific heat close to a second order phase transition is

$$c(T) \sim (T - T_c)^{-\alpha} + const \tag{6.62}$$

where the result of the Landau theory is

$$\alpha = 0. \tag{6.63}$$

So far we have considered solely spatially constant solutions of the order parameter. It is certainly possible to consider the more general case of spatially varying order parameters, where the free energy

$$F = \int d^d r f [\phi(r)] \tag{6.64}$$

is given as

$$f[\phi(r)] = \frac{a}{2} \phi(r)^2 + \frac{c}{4} \phi(r)^4 - h(r) \phi(r) + \frac{b}{2} (\nabla \phi(r))^2 \tag{6.65}$$

where we assumed that it costs energy to induce an inhomogeneity of the order parameter ($b > 0$). The minimum of F is now determined by the Euler-Lagrange equation

$$\frac{\partial f}{\partial \phi} - \nabla \frac{\partial f}{\partial \nabla \phi} = 0 \tag{6.66}$$

which leads to the nonlinear partial differential equation

$$a\phi(r) + c\phi(r)^3 = h(r) + b\nabla^2 \phi(r) \tag{6.67}$$

Above the transition temperature we neglect again the non-linear term   and have to solve

$$a\phi(r) - b\nabla^2 \phi(r) = h(r) \tag{6.68}$$

It is useful to consider the generalized susceptibility

$$\delta\phi(r) = \int d^d r' \chi(r - r') \delta h(r') \tag{6.69}$$

which determines how much a local change in the order parameter is affected by a local change of an external field at a distance $r - r'$. This is often written as

$$\chi(r - r') = \frac{\delta\phi(r)}{\delta h(r')}. \tag{6.70}$$

We determine $\chi(r - r')$ by Fourier transforming the above differential equation with

$$\phi(r) = \int d^d k \, e^{ikr} \phi(k) \tag{6.71}$$

which gives

$$a\phi(k) + bk^2 \phi(k) = h(k) \tag{6.72}$$

In addition it holds for $\chi(k)$:

$$\delta\phi(k) = \chi(k)\,\delta h(k).\tag{6.73}$$

This leads to

$$\chi(k) = \frac{b^{-1}}{\xi^{-2} + k^2}\tag{6.74}$$

where we introduced the length scale

$$\xi = \sqrt{\frac{b}{a}} = \sqrt{\frac{b}{a_0}}\,(T - T_c)^{-1/2}\tag{6.75}$$

This result can now be back-transformed yielding

$$\chi(r - r') = \left(\frac{\xi}{r - r'}\right)^{\frac{d-1}{2}} \exp\left(-\frac{|r - r'|}{\xi}\right)\tag{6.76}$$

Thus, spins are not correlated anymore beyond the correlation length ξ. In general the behavior of ξ close to T_c can be written as

$$\xi \sim (T - T_c)^{-\nu}\tag{6.77}$$

with $\nu = \frac{1}{2}$.

A similar analysis can be performed in the ordered state. Starting again at

$$a\phi(r) + c\phi(r)^3 = h(r) + b\nabla^2\phi(r)\tag{6.78}$$

and assuming $\phi(r) = \phi_0 + \psi(r)$ where ϕ_0 is the homogeneous, $h = 0$, solution, it follows for small $\psi(r)$:

$$\left(a + 3c\phi_0^2\right)\psi(r) = h(r) + b\nabla^2\psi(r)\tag{6.79}$$

and it holds $a + 3c\phi_0^2 = -2a > 0$. Thus in momentum space

$$\chi(k) = \frac{d\psi(k)}{dh(k)} = \frac{b^{-1}}{\xi_<^{-2} + k^2}\tag{6.80}$$

with

$$\xi = \sqrt{\frac{b}{-2a}} = \sqrt{\frac{b}{2a_0}}\,(T_c - T)^{-1/2}\tag{6.81}$$

We can now estimate the role of fluctuations beyond the linearized form used. This can be done by estimating the size of the fluctuations of the order parameter compared to its mean value ϕ_0. First we note that

$$\chi(r - r') = \langle(\phi(r) - \phi_0)(\phi(r') - \phi_0)\rangle\tag{6.82}$$

Thus the fluctuations of $\phi(r)$ in the volume ξ^d is

$$\langle\delta\phi^2\rangle = \frac{1}{\xi^d}\int_{r<\xi} d^d r\,\chi(r)\tag{6.83}$$

$$\chi(r) = \int \frac{d^d k}{(2\pi)^d} \frac{b^{-1}}{k^2 + \xi^{-2}} e^{-ikr}$$

$$\int_{r<\xi} d^d r \chi(r) = \int \frac{d^d k}{(2\pi)^d} \frac{b^{-1}}{k^2 + \xi^{-2}} \int_{r<\xi} d^d r e^{-ikr}$$

$$= \int \frac{d^d k}{(2\pi)^d} \frac{b^{-1}}{k^2 + \xi^{-2}} \prod_{\alpha=1}^{d} \frac{\sin k_\alpha \xi}{k_\alpha} \tag{6.84}$$

where we used

$$\frac{1}{2} \int_{-\xi}^{\xi} dx e^{-ik_x x} \simeq \frac{\sin k_x \xi}{k_x} \tag{6.85}$$

The last integral can be evaluated by substituting $z_a = k_\alpha \xi$ leading to

$$\int_{r<\xi} d^d r \chi(r) = \xi^2 \int \frac{d^d z}{(2\pi)^d} \frac{b^{-1}}{z^2 + 1} \prod_{\alpha=1}^{d} \frac{\sin z_\alpha}{z_\alpha} \propto d^{-1}\xi^2 \tag{6.86}$$

Thus, it follows

$$\langle \delta\phi^2 \rangle \propto b^{-1} \xi^{2-d} \tag{6.87}$$

This must be compared with the mean value of the order parameter

$$\phi_0^2 = \frac{a}{c} \propto \frac{b}{c}\xi^{-2} \tag{6.88}$$

and it holds

$$\frac{\langle \delta\phi^2 \rangle}{\phi_0^2} \simeq \frac{c}{b^2}\xi^{4-d} \tag{6.89}$$

Thus, for $d > 4$ fluctuations become small as $\xi \to \infty$, whereas they cannot be neglected for $d < 4$. In $d = 4$, a more careful analysis shows that $\frac{\langle \delta\phi^2 \rangle}{\phi_0^2} \propto \log \xi$.

Role of an additional term ϕ^6:

$$f = \frac{1}{2}a\varphi^2 + \frac{c}{4}\varphi^4 + \frac{w}{6}\varphi^6 \tag{6.90}$$

The minimum is at

$$\frac{\partial E}{\partial \varphi} = a\varphi + c\varphi^3 + w\varphi^5 = 0 \tag{6.91}$$

which gives either $\varphi = 0$ or $r + c\varphi^2 + w\varphi^4 = 0$ with the two solutions

$$\varphi_\pm^2 = \frac{-c \pm \sqrt{c^2 - 4aw}}{2w} \tag{6.92}$$

If $c > 0$ we can exclude the two solutions φ_- which are purely imaginary. If $a > 0$, then the φ_+ are imaginary as well and the only solution is $\varphi = 0$. Close to the transition temperature at $a = 0$ holds for $r \to 0^-$: $\varphi = \sqrt{\frac{-a}{c}}$, i.e. the behavior is not affected by w.

If $c < 0$ then the solutions φ_- might become relevant if $c^2 > 4aw$. For $a^* = \frac{c^2}{4w}$ new solutions at $\varphi^* = \pm\sqrt{\frac{-c}{2w}}$ occur for the first time. for $a < a^*$ these solutions are given by

$$\varphi = \pm\sqrt{\frac{\sqrt{c^2 - 4aw} - c}{2d}} \qquad (6.93)$$

At $a_c = \frac{3c^2}{16w}$ the energy $f = \frac{c^3 - 6acw - \left(c^2 - 4aw\right)^{3/2}}{24w^2}$ of this solution equals the one for $\varphi = 0$. The order parameter at this point is $\varphi_c = \pm\sqrt{\frac{-3c}{4w}}$. Finally, for $a^{**} = 0$ the solution at $\varphi = 0$ which was metastable for $a^{**} < a < a_c$ disappears completely.

If $c = 0$ the solution is for $a < 0$:

$$\varphi = \left(\frac{-a}{w}\right)^{1/4} \qquad (6.94)$$

whereas it vanishes for $a > 0$.

$$\begin{aligned}
\phi &= \mu^{1/\delta} \\
\frac{d\phi}{d\mu} &= \mu^{1/\delta - 1} = \phi^{1-\delta}
\end{aligned} \qquad (6.95)$$

6.5.0.1 Statistical mechanics motivation of the Landau theory:

We start from the Ising model

$$H\left[S_i\right] = -\sum_{ij} J_{ij} S_i S_j \qquad (6.96)$$

in an spatially varying magnetic field. The partition function is $Z = \sum_{\{S_i\}} e^{-\beta H[S_i]}$. In order to map this problem onto a continuum theory we use the identity

$$\int \prod_i^N dx_i \exp\left(-\frac{1}{4}\sum_{ij} x_i \left(V^{-1}\right)_{ij} x_j + s_i x_i\right) = \left(2\sqrt{\pi}\right)^N \sqrt{\det V} e^{\sum_{i,j}^N V_{ij} s_i s_j} \qquad (6.97)$$

which can be shown to be correct by rotating the variables x_i into a representation where V is diagonal and using

$$\int dx \exp\left(-\frac{x^2}{4v} + sx\right) = 2\sqrt{4\pi v} e^{V s^2} \qquad (6.98)$$

This identity can now be used to transform the Ising model (use $V_{ij} = \beta J_{ij}$) according to

$$Z = \frac{\sum_{\{S_i\}} \int \prod_i^N dx_i \exp\left(-\frac{1}{4}\sum_{ij} x_i V_{ij}^{-1} x_j + x_i S_i\right)}{(4\pi)^{\frac{N}{2}} \sqrt{\det V}} \tag{6.99}$$

$$= \frac{\int \prod_i^N dx_i \exp\left(-\frac{1}{4}\sum_{ij} x_i V_{ij}^{-1} x_j\right) \sum_{\{S_i\}} e^{x_i S_i}}{(4\pi)^{\frac{N}{2}} \sqrt{\det V}} \tag{6.100}$$

The last term is just the partition function of free spins in an external field $\sim x_i$ and it holds

$$\sum_{\{S_i\}} e^{x_i S_i} \sim \exp\left(\sum_i \log\left(\cosh x_i\right)\right) \tag{6.101}$$

Transforming $\phi_i = \frac{1}{\sqrt{2}}\sum_j V_{ij}^{-1} x_j$ gives

$$Z \sim \int D\phi \exp\left(-\frac{\beta}{2}\sum_{ij} \phi_i J_{ij}\phi_j + \sum_i \log\left(\cosh\left(\sqrt{2}\sum_j \beta J_{ij}\phi_j\right)\right)\right) \tag{6.102}$$

where $D\phi = \prod_i d\phi_i$. Using

$$\log\left(\cosh z\right) \simeq \frac{z^2}{2} - \frac{z^4}{12} \tag{6.103}$$

we obtain

$$Z \sim \int D\phi \exp\left(-\beta H_{\text{eff}}\left[\phi\right]\right) \tag{6.104}$$

with

$$H_{\text{eff}}\left[\phi\right] = \frac{1}{2}\sum_{ij} \phi_i \left(J_{ij} - \frac{\beta}{4}\sum_l J_{il}J_{lj}\right)\phi_j + \frac{1}{12}\sum_{i,j,k,l,m} J_{ij}J_{ik}J_{il}J_{im}\phi_j\phi_k\phi_l\phi_m \tag{6.105}$$

It is useful to go into a momentum representation

$$\phi_i = \phi\left(\mathbf{R}_i\right) = \int \frac{d^D k}{(2\pi)^D} \phi_{\mathbf{k}} e^{-i\mathbf{k}\cdot\mathbf{R}} \tag{6.106}$$

which gives

$$H_{\text{eff}}\left[\phi\right] = \frac{1}{2}\int d^d k \phi_k \left(J_k - \frac{\beta}{4}J_k^2\right)\phi_{-k}$$
$$+ \frac{1}{4}\int d^d k_1 d^d k_2 d^d k_3 u\left(k_1, k_2, k_3\right)\phi_{k_1}\phi_{k_2}\phi_{k_3}\phi_{-k_1-k_2-k_3} \tag{6.107}$$

with

$$u(k_1, k_2, k_3) = \frac{\beta^4}{3} J_{k_1} J_{k_2} J_{k_3} J_{-k_1-k_2-k_3} \tag{6.108}$$

Using $J_{ij} = J$ for nearest neighbors and zero otherwise gives for a cubic lattice

$$J_k = 2J \sum_{\alpha=x,y,\ldots} \cos(k_\alpha \bar{a}) \simeq 2J\left(d - \bar{a}^2 \mathbf{k}^2\right) = J_0 \left(1 - \frac{\bar{a}^2}{d} \mathbf{k}^2\right) \tag{6.109}$$

Here $\bar{a}$ is the lattice constant and we expanded J_k for small momenta (wave length large compared to $\bar{a}$)

$$
\begin{aligned}
J_k - \frac{\beta}{4} J_k^2 &= J_0 - J_0 \frac{\bar{a}^2}{d} \mathbf{k}^2 - \frac{\beta}{4} J_0^2 + 2\frac{\beta}{4} J_0^2 \frac{\bar{a}^2}{d} \mathbf{k}^2 \\
&= J_0 - \frac{\beta}{4} J_0^2 + \left(2\frac{\beta}{4} J_0^2 - J_0\right) \frac{\bar{a}^2}{d} \mathbf{k}^2 \\
&= \frac{J_0}{T}\left(T - \frac{J_0}{4k_B}\right) + J_0\left(\frac{\beta J_0}{2} - 1\right) \frac{\bar{a}^2}{d} \mathbf{k}^2 \tag{6.110}
\end{aligned}
$$

At the transition $\beta J_0 = 4$ such that

$$J_k - \frac{\beta}{4} J_k^2 \simeq a_0 (T - T_c) + b\mathbf{k}^2 \tag{6.111}$$

with $T_c = \frac{J_0}{4k_B}$, $a_0 \simeq 4k_B$, $b \simeq J_0 \frac{\bar{a}^2}{d}$. Using $u \simeq \frac{1}{3}(\beta J_0)^3$ gives finally

$$
\begin{aligned}
H_{\text{eff}}[\phi] &= \frac{1}{2} \int d^d k \phi_k \left(a_0 (T - T_c) + b\mathbf{k}^2\right) \phi_{-k} \\
&\quad + \frac{u}{4} \int d^d k_1 d^d k_2 d^d k \; \phi_{k_1} \phi_{k_2} \phi_{k_3} \phi_{-k_1-k_2-k_3} \tag{6.112}
\end{aligned}
$$

This is precisely the Landau form of an Ising model, which becomes obvious if one returns to real space

$$H_{\text{eff}}[\phi] = \frac{1}{2} \int d^d r \left(a_0 (T - T_c) \phi^2 + b (\nabla\phi)^2 + \frac{u}{2} \phi^4\right). \tag{6.113}$$

From these considerations we also observe that the partition function is given as

$$Z = \int D\phi \exp\left(-\beta H_{\text{eff}}[\phi]\right) \tag{6.114}$$

and it is, in general, not the minimum of $H_{\text{eff}}[\phi]$ w.r.t. ϕ which is physically realized, instead one has to integrate over all values of ϕ to obtain the free energy. Within Landau theory we approximate the integral by the dominating contribution of the integral, i.e. we write

$$\int D\phi \exp\left(-\beta H_{\text{eff}}[\phi]\right) \simeq \exp\left(-\beta H_{\text{eff}}[\phi_0]\right) \tag{6.115}$$

where $\left.\frac{\delta H}{\delta\phi}\right|_{\phi=\phi_0} = 0$.

6.5.0.2 Ginzburg criterion

One can now estimate the range of applicability of the Landau theory. This is best done by considering the next order corrections and analyze when they are small. If this is the case, one can be confident that the theory is controlled. Before we go into this we need to be able to perform some simple calculations with these multidimensional integrals.

First we consider for simplicity a case where $H_{\text{eff}}[\phi]$ has only quadratic contributions. It holds

$$
\begin{aligned}
Z & = \int D\phi \exp\left(-\frac{1}{2}\int d^d k \phi_k \left(a + b\mathbf{k}^2\right)\phi_{-k}\right) \\
& = \int \prod_k d\phi_k \exp\left(-\frac{\Delta k}{2}\phi_k \left(a + b\mathbf{k}^2\right)\phi_{-k}\right) \\
& = \prod_k \left(\frac{(2\pi)^d}{a + b\mathbf{k}^2}\right)^{1/2} \\
& \sim \exp\left(\frac{1}{2}\int d^d k \log \chi(k)\right)
\end{aligned}
\tag{6.116}
$$

with

$$
\chi(k) = \frac{1}{a + b\mathbf{k}^2}.
\tag{6.117}
$$

It follows for the free energy

$$
F = -\frac{k_B T}{2}\int d^d k \log \chi(k)
\tag{6.118}
$$

One can also add to the Hamiltonian an external field

$$
H_{\text{eff}}[\phi] \to H_{\text{eff}}[\phi] - \int d^d k\, h(k)\,\phi(k)
\tag{6.119}
$$

Then it is easy to determine the correlation function

$$
\chi(k) = \langle \phi_k \phi_{-k} \rangle - \langle \phi_k \rangle \langle \phi_{-k} \rangle
\tag{6.120}
$$

via

$$
\begin{aligned}
\left.\frac{\delta \log Z}{\delta h_k \delta h_{-k}}\right|_{h\to 0} & = \frac{\delta}{\delta h_k}\frac{1}{Z}\int D\phi\, \phi_k e^{-\beta H_{\text{eff}}[\phi]} \\
& = \frac{1}{Z}\int D\phi\, \phi_k \phi_{-k} e^{-\beta H_{\text{eff}}[\phi]} - \frac{\left(\int D\phi\, \phi_k e^{-\beta H_{\text{eff}}[\phi]}\right)^2}{Z^2} \\
& = \chi(k)
\end{aligned}
\tag{6.121}
$$

This can again be done explicitly for the case with $u = 0$:

$$
\begin{aligned}
Z[h] & = \int D\phi \exp\left(-\frac{1}{2}\int d^d k \phi_k \left(a + b\mathbf{k}^2\right)\phi_{-k} + \int d^d k\, h(k)\,\phi_k\right) \\
& = Z[0]\exp\left(\frac{1}{2}\int d^d k\, h_k \chi(k)\, h_{-k}\right)
\end{aligned}
\tag{6.122}
$$

Performing the second derivative of $\log Z$ gives indeed $\langle \phi_k \phi_{-k} \rangle = \frac{1}{a+b\mathbf{k}^2}$. Thus, we obtain as expected

$$\chi(k) = \frac{\delta \phi_k}{\delta h_{-k}}. \tag{6.123}$$

Let us analyze the specific heat related to the free energy

$$F = -\frac{k_B T}{2} \int d^d k \log \chi(k) \tag{6.124}$$

It holds for the singular part of the specific heat

$$c \sim -\frac{\partial^2 F}{\partial a^2} \sim \int d^d k \chi(k)^2 \sim \int \frac{k^{d-1} dk}{\left(\xi^{-2} + k^2\right)^2} \sim \xi^{4-d} \tag{6.125}$$

Thus, as $\xi \to \infty$ follows that there is no singular (divergent) contribution to the specific heat if $d > 4$ just as we found in the Landau theory. However, for $d < 4$ the specific heat diverges and we obtain a behavior different from what Landau theory predicted.

Another way to see this is to study the role of inhomogeneous fluctuations as caused by the

$$H_{inh} = \frac{d}{2} \int d^d r \left(\nabla \phi\right)^2 \tag{6.126}$$

Consider a typical variation on the scale $\nabla \phi \sim \sqrt{\frac{-a}{u}} \xi^{-1}$ and integrate those over a volume of size ξ^d gives

$$H_{inh} \sim b\xi^{d-2} \frac{a}{u} \sim \frac{b^2}{u} \xi^{d-4} \tag{6.127}$$

Those fluctuations should be small compared to temperature in order to keep mean field theory valid. If their energy is large compared to $k_B T$ they will be rare and mean field theory is valid. Thus we obtain again that mean field theory breaks down for $d < 4$. This is called the Ginzburg criterion. Explicitly this criterion is

$$\xi^{-1} > \left(\frac{u}{b^2} k_B T\right)^{\frac{1}{4-d}}. \tag{6.128}$$

Note, if b is large for some reason, fluctuation physics will enter only very close to the transition. This is indeed the case for many so called conventional superconductors.

6.6 Scaling laws

A crucial observation of our earlier results of second order phase transitions was the divergence of the correlation length

$$\xi(T \to T_c) \to \infty. \tag{6.129}$$

This divergency implies that at the critical point no characteristic length scale exists, which is in fact an important reason for the emergence of the various power laws. Using h as a dimensionless number proportional to an external field and

$$t = \frac{T - T_c}{T_c} \tag{6.130}$$

as dimensionless measure of the distance to the critical point the various critical exponents are:

$$
\begin{aligned}
\xi\,(t, h = 0) &\sim t^{-\nu} \\
\phi\,(t, h = 0) &\sim |t|^{\beta} \\
\phi\,(t = 0, h) &\sim h^{1/\delta} \\
\chi\,(t, h = 0) &\sim t^{-\gamma} \\
c\,(t, h = 0) &\sim t^{-\alpha} \\
\chi\,(x \to \infty, t = 0) &\sim x^{2-d-\eta}.
\end{aligned}
\tag{6.131}
$$

where D is the spatial dimensionality. The values of the critical exponents for a number of systems are given in the following table

exponent	mean field	$d = 2$, Ising	$d = 3$, Ising
α	0	0	0.12
β	$\frac{1}{2}$	$\frac{1}{8}$	0.31
γ	1	$\frac{7}{4}$	1.25
ν	$\frac{1}{2}$	1	0.64
δ	3	15	5.0
η	0	$\frac{1}{4}$	0.04

It turn out that a few very general assumptions about the *scaling* behavior of the correlation function $\chi\,(q)$ and the free energy are sufficient to derive very general relations between these various exponents. Those relations are called *scaling laws*. We will argue that the fact that there is no typical length scale characterizing the behavior close to a second order phase transition leads to a power-law behavior of the singular contributions to the free energy and correlation function. For example, consider the result obtained within Landau theory

$$\chi\,(q, t) = \frac{1}{t + q^2} \tag{6.132}$$

where we eliminated irrelevant pre-factors. Rescaling all length r of the system according to $x \to x/b$, where b is an arbitrary dimensionless number, leads to $k \to kb$. Obviously, the mean field correlation function obeys

$$\chi\,(q, t) = b^2 \chi\,\left(bq, tb^2\right). \tag{6.133}$$

Thus, upon rescaling ($k \to kb$), the system is characterized by a correlation function which is the same up to a pre-factor and a readjustment of the distance

from the critical point. In what follows we will generalize this expression and assume that even beyond the mean field theory of Landau a similar relationship holds

$$\chi(q, t) = b^{2-\eta} \chi(bq, tb^y). \tag{6.134}$$

The mean field theory is obviously recovered if $y = 2$ and $\eta = 0$. Since b is arbitrary, we can for example chose $tb^y = 1$ implying $b = t^{-\frac{1}{y}}$ and we obtain directly from our above ansatz

$$\chi(q, t) = t^{-\frac{2-\eta}{y}} \chi\left(qt^{-\frac{1}{y}}, 1\right). \tag{6.135}$$

By definition, the correlation length is the length scale which characterizes the momentum variation of $\chi(q, t)$ i.e. $\chi(q, t) \sim f(q\xi)$, which leads to $\xi \sim t^{-\frac{1}{y}}$ and we obtain

$$\nu = y^{-1}. \tag{6.136}$$

The exponent y of our above ansatz for $\chi(q, t)$ is therefore directly related to the correlation length exponent. This makes it obvious why it was necessary to generalize the mean field behavior. $y = 2$ yields the mean field value of ν. Next we consider $t = 0$ and chose $bq = 1$ such that

$$\chi(q, t = 0) = \frac{1}{q^{2-\eta}} \chi(1, 0) \tag{6.137}$$

which gives

$$\chi(x, t = 0) = \int \frac{d^d q}{(2\pi)^d} \chi(q, t = 0) e^{ikx} \sim \int dq e^{ikx} \frac{q^{d-1}}{q^{2-\eta}} \tag{6.138}$$

substituting $z = kx$ gives

$$\chi(r, t = 0) \sim x^{2-d-\eta}. \tag{6.139}$$

Thus, the exponent η of Eq.6.134 is indeed the same exponent as the one given above. This exponent is often called anomalous dimension and characterizes the change in the power-law decay of correlations at the critical point (and more generally for length scales smaller than ξ). Thus we can write

$$\chi(q, t) = b^{2-\eta} \chi\left(bq, tb^{\frac{1}{\nu}}\right). \tag{6.140}$$

Similar to the correlation function can we also make an assumption for the free energy

$$F(t, h) = b^{-D} F(tb^y, hb^{y_h}). \tag{6.141}$$

The pre-factor b^{-d} is a simple consequence of the fact that an extensive quantity changes upon rescaling of length with a corresponding volume factor. Using $y = \nu^{-1}$ we can again use $tb^y = 1$ and obtain

$$F(t, h) = t^{D\nu} F\left(1, ht^{-\nu y_h}\right). \tag{6.142}$$

This enables us to analyze the specific heat at $h = 0$ as

$$c \sim \frac{\partial^2 F(t,0)}{\partial t^2} \sim t^{d\nu - 2} \tag{6.143}$$

which leads to

$$\alpha = 2 - d\nu. \tag{6.144}$$

This is a highly nontrivial relationship between the spatial dimensionality, the correlation length exponent and the specific heat exponent. It is our first scaling law. Interestingly, it is fulfilled in mean field (with $\alpha = 0$ and $\nu = \frac{1}{2}$) only for $d = 4$.

The temperature variation of the order parameter is given as

$$\phi(t) \sim \left. \frac{\partial F(t,h)}{\partial h} \right|_{h \to 0} \sim t^{\nu(d - y_h)} \tag{6.145}$$

which gives

$$\beta = \nu(d - y_h) = 2 - \alpha - \nu y_h \tag{6.146}$$

This relationship makes a relation between y_h and the critical exponents just like y was related to the exponent ν. Within mean field

$$y_h = 3 \tag{6.147}$$

Alternatively we can chose $hb^{y_h} = 1$ and obtain

$$F(t,h) = h^{\frac{d}{y_h}} F\left(th^{-\frac{1}{\nu y_h}}, 0\right) \tag{6.148}$$

This gives for the order parameter at the critical point

$$\phi(t = 0, h) \sim \frac{\partial F(t = 0, h)}{\partial h} \sim h^{\frac{d}{y_h} - 1} \tag{6.149}$$

and gives $\frac{1}{\delta} = \frac{d}{y_h} - 1$. One can simplify this to

$$\delta = \frac{y_h}{d - y_h} = \frac{2 - \alpha - \beta}{\beta} \tag{6.150}$$

and yields

$$\beta(1 + \delta) = 2 - \alpha \tag{6.151}$$

Note, the mean field theory obeys $\delta = \frac{y_h}{y_h - d}$ only for $d = 4$. whereas $\delta = \frac{2 - \alpha - \beta}{\beta}$ is obeyed by the mean field exponents for all dimensions. This is valid quite generally, scaling laws where the dimension, d, occurs explicitly are fulfilled within mean field only for $d = 4$ whereas scaling laws where the dimensionality does not occur are valid more generally.

The last result allows us to rewrite our original ansatz for the free energy

$$F(t,h) = b^{(2-\alpha)\nu^{-1}} F\left(tb^{\frac{1}{\nu}}, hb^{\frac{\beta\delta}{\nu}}\right). \tag{6.152}$$

such that $tb^{\frac{1}{\nu}} = 1$ leads to

$$F(t, h) = t^{2-\alpha} F\left(1, ht^{-\beta\delta}\right) \tag{6.153}$$

We next analyze how the susceptibility diverges at the critical point. It holds

$$\chi \sim \left.\frac{\partial^2 F(t, h)}{\partial h^2}\right|_{h \to 0} \sim t^{2-\alpha-2\beta\delta} \tag{6.154}$$

which leads to

$$\gamma = \alpha - 2 + 2\beta\delta \tag{6.155}$$

which is yet another scaling relation.

The last scaling law follows from the fact that the correlation function $\chi(q, t)$ taken at $q = 0$ equals the susceptibility χ just analyzed. This gives

$$\chi(t) = b^{2-\eta} \chi(tb^y) \tag{6.156}$$

and choosing again $tb^y = 1$ yields

$$\chi(t) = t^{-\nu(2-\eta)} \tag{6.157}$$

such that

$$\gamma = \nu(2 - \eta). \tag{6.158}$$

To summarize, we have identified all the exponents in the assumed scaling relations of $F(t, h)$ and $\chi(q, t)$ with critical exponents (see Eqn.6.140 and6.152). In addition we have four relationships the six exponents have to fulfill at the same time which are collected here:

$$\alpha + 2\beta + \gamma = 2 \tag{6.159}$$

$$\begin{aligned} \alpha &= 2 - d\nu. \\ \beta(1 + \delta) &= 2 - \alpha \\ 2\beta\delta - \gamma &= 2 - \alpha \\ \gamma &= \nu(2 - \eta) \end{aligned} \tag{6.160}$$

One can easily check that the exponents of the two and three dimensional Ising model given above indeed fulfill all these scaling laws. If one wants to calculate these exponents, it turns out that one only needs to determine two of them, all others follow from scaling laws.

6.7 Renormalization group

6.7.1 Perturbation theory

A first attempt to make progress in a theory beyond the mean field limit would be to consider the order parameter

$$\phi(x) = \phi_0 + \psi(x) \tag{6.161}$$

and assume that ϕ_0 is the result of the Landau theory and then consider $\psi(x)$ as a small perturbation. We start from

$$H[\phi] = \frac{1}{2}\int d^d x \phi(x)\left(r - \nabla^2\right)\phi(x) + \frac{u}{4}\int d^d x \phi(x)^4. \qquad (6.162)$$

where we have introduced

$$\begin{aligned} r &= \frac{a}{b}, \\ u &= \frac{c}{Tb^2}, \end{aligned} \qquad (6.163)$$

and a new field variable $\phi_{\text{new}} = \sqrt{Td}\phi$, where the suffix new is skipped in what follows. We also call H_{eff} simply H.

We expand up to second order in $\psi(x)$:

$$H[\phi] = V\left(\frac{r}{2}\phi_0^2 + \frac{u}{4}\phi_0^4\right) + \frac{1}{2}\int d^d x \psi(x)\left(r - \nabla^2 + 3u\phi_0^2\right)\psi(x) \qquad (6.164)$$

The fluctuation term is therefore of just the type discussed in the context of the $u = 0$ Gaussian theory, only with a changes value of $r \to r + 3u\phi_0^2$. Thus, we can use our earlier result for the free energy of the Gaussian model $F_{\text{Gauss}} = -\frac{1}{2}\int d^d k \log \chi(k)$ and obtain in the present case

$$\frac{F}{V} = \frac{r}{2}\phi_0^2 + \frac{u}{4}\phi_0^4 + \frac{1}{2V}\int d^d k \log\left(r + k^2 + 3u\phi_0^2\right) \qquad (6.165)$$

We can now use this expression to expand this free energy again up to fourth order in ϕ_0 using:

$$\log\left(r + k^2 + 3u\phi_0^2\right) \simeq \log\left(r + k^2\right) + \frac{3u\phi_0^2}{r + k^2} - \frac{1}{2}\frac{9u^2\phi_0^4}{\left(r + k^2\right)^2} \qquad (6.166)$$

it follows

$$\frac{F}{V} = \frac{F_0}{V} + \frac{r'}{2}\phi_0^2 + \frac{u'}{4}\phi_0^4 \qquad (6.167)$$

with

$$\begin{aligned} r' &= r + 3u\int d^d k \frac{1}{r + k^2}, \\ u' &= u - 9u^2\int d^d k \frac{1}{\left(r + k^2\right)^2}. \end{aligned} \qquad (6.168)$$

These are the fluctuation corrected values of the Landau expansion. If these corrections were finite, Landau theory is consistent, if not, one has to use a qualitatively new approach to describe the fluctuations. At the critical point $r = 0$ and we find that

$$\int d^d k \frac{1}{k^2} \propto \int^\Lambda \frac{k^{d-1}}{k^2} dk \propto \begin{cases} \Lambda^{d-2} & d > 2 \\ \infty & d \leq 2 \end{cases}$$

where Λ is some upper limit of the momentum integration which takes into account that the system under consideration is always embedded in a lattice of atoms with interatomic spacing $\bar{a}$, such that $\Lambda \sim \bar{a}^{-1}$. It holds that r' is finite for $D > 2$ which we assume to be the case in what follows. The finite correction to r' only shifts the transition temperature by a finite amount. Since the value of T_c was not so much our concern, this is a result which is acceptable. However, the correction, u', to u diverges for $d \leq 4$:

$$\int d^d k \frac{1}{k^4} \propto \int^\Lambda \frac{k^{d-1}}{k^4} dk \propto \left\{ \begin{array}{cc} \Lambda^{d-4} & d > 4 \\ \infty & d \leq 4 \end{array} \right.$$

and the nonlinearity (interactions) of the theory become increasingly stronger. This is valid for arbitrarily small values of u itself, i.e. a perturbation theory in u will fail for $d \leq 4$. The dimension below which such strong fluctuations set in is called *upper critical dimension*.

The strategy of the above scaling laws (i.e. the attempt to see what happens as one rescales the characteristic length scales) will be the heart of the renormalization group theory which we employ to solve the dilemma below four dimension.

6.7.2 Fast and slow variables

The divergency which caused the break down of Landau theory was caused by long wave length, i.e. the $k \to 0$ behavior of the integral which renormalized $u \to u'$. One suspicion could be that only long wave length are important for an understanding of this problem. However, this is not consistent with the scaling concept, where the rescaling parameter was always assumed to be arbitrary. In fact it fluctuations on *all* length scales are crucial close to a critical point. This is on the one hand a complication, on the other hand one can take advantage of this beautiful property. Consider for example the scaling properties of the correlation function

$$\chi(q, t) = b^{2-\eta} \chi \left(bq, tb^{\frac{1}{\nu}} \right). \tag{6.169}$$

Repeatedly we chose $tb^{\frac{1}{\nu}} = 1$ such that $b = t^{-\nu} \to \infty$ as one approaches the critical point. However, if this scaling property (and the corresponding scaling relation for the free energy) are correct for generic b (of course only if the system is close to T_c) one might analyze a rescaling for b very close to 1 and infer the exponents form this more innocent regime. If we obtain a scaling property of $\chi(q, t)$ it simply doesn't matter how we determined the various exponents like ν, η etc.

This, there are two key ingredients of the renormalization group. The first is the assumption that scaling is a sensible approach, the second is a decimation procedure which makes the scaling transformation $x \to x/b$ explicit for $b \simeq 1$. A convenient way to do this is by considering $b = e^l$ for small l. Lets consider a field variable

$$\phi(\mathbf{k}) = \int d^d x \exp(i\mathbf{k} \cdot \mathbf{x}) \phi(\mathbf{x}) \tag{6.170}$$

Since there is an underlying smallest length-scale $\bar{a}$ ($\simeq$interatomic spacing), no waves with wave number larger than a given upper cut off $\Lambda \simeq \bar{a}^{-1}$ should occur. For our current analysis the precise value of Λ will be irrelevant, what matters is that such a cut off exists. Thus, be observe that $\phi(k) = 0$ if $k > \Lambda$.

We need to develop a scheme which allows us to explicitly rescale length or momentum variables. How to do this goes back to the work of Leo Kadanoff and Kenneth G. Wilson in the early 70th of the last century. The idea is to divide the typical length variations of $\phi(k)$ into short and long wave length components

$$\phi(k) = \left\{ \begin{array}{ll} \phi^<(k) & 0 < k \leq \Lambda/b \\ \phi^>(k) & \Lambda/b < k \leq \Lambda \end{array} \right. . \tag{6.171}$$

If one now eliminates the degrees of freedoms $\phi^>$ one obtains a theory for $\phi^<$ only

$$\exp\left(-\widetilde{H}\left[\phi^<\right]\right) = \int D\phi^> \exp\left(-H\left[\phi^<, \phi^>\right]\right). \tag{6.172}$$

The momenta in $\widetilde{H}[\phi^<]$ are confined to the smaller region $0 < k \leq \Lambda/b$. We can now rescale simply according to

$$k' = bk \tag{6.173}$$

such that the new variable k' is restricted to the original scales $0 < k' \leq \Lambda$. The field variable is then $\phi^<\left(\frac{k'}{b}\right)$ and will conveniently be called

$$\phi'(k') = b^{-\rho}\phi^<\left(\frac{k'}{b}\right) \tag{6.174}$$

where the pre-factor $b^{-\rho}$ is only introduced for later convenience to be able to keep the pre-factor of the k^2 term in the Hamiltonian the same. The *renormalized* Hamiltonian is then determined by

$$H'[\phi'] = \widetilde{H}\left[b^{-\rho}\phi'\right]. \tag{6.175}$$

In practice this is then a theory of the type where the initial Hamiltonian

$$\begin{aligned} H[\phi] &= \frac{1}{2}\int d^d k\,(r + k^2)\,|\phi(k)| + \\ &\quad \frac{u}{4}\int d^d k_1 d^d k_2 d^d k_3 \phi(k_1)\,\phi(k_2)\,\phi(k_3)\,\phi(-k_1 - k_2 - k_3) \end{aligned} \tag{6.176}$$

leads to a renormalized Hamiltonian

$$\begin{aligned} H'[\phi'] &= \frac{1}{2}\int d^d k'\,\left(r(l) + k'^2\right)\,|\phi'(k')| + \\ &\quad \frac{u(l)}{4}\int d^d k'_1 d^d k'_2 d^d k'_3 \phi'(k'_1)\,\phi'(k'_2)\,\phi'(k'_3)\,\phi'(-k'_1 - k'_2 - k'_3) \end{aligned} \tag{6.177}$$

If this is the case one may as well talk about a mapping

$$H(r, u) \to H'(r(l), u(l)) \tag{6.178}$$

and one only needs to analyze where this mapping takes one.

If one now analyzes the so called flow equation of the parameters $r(l)$, $u(l)$ etc. there are a number of distinct cases. The most distinct case is the one where a fixed point is approaches where $r(l \to \infty) = r^*$, $u(l \to \infty) = u^*$ etc. If this is the case the low energy behavior of the system is identical for all initial values which reach the fixed point. Before we go into this we need to make sure that the current procedure makes any sense and reproduces the idea of scaling.

6.7.3   Scaling behavior of the correlation function:

We start from $H[\phi]$ characterized by a cut off Λ. The new Hamiltonian with cut off Λ/b, which results from the shell integration, is then determined by

$$e^{-\widetilde{H}[\phi^<]} = \int D\phi^> e^{-H[\phi^<,\phi^>]}, \qquad (6.179)$$

which is supplemented by the rescaling

$$\phi^<(k) = b^\rho \phi'(bk)$$

which yields the new Hamiltonian $H'[\phi']$ which is also characterized by the cut off Λ. If one considers states with momenta with $k < \Lambda/b$, it is possible to determine the corresponding correlation function either from $H[\phi]$ or from $H'[\phi']$. Thus, we can either start from the original action:

$$\langle \phi(k_1)\phi(k_2)\rangle = \int \frac{D\phi e^{-H[\phi]}}{Z}\phi(k_1)\phi(k_2) = \chi(k_1)\delta(k_1 + k_2) \qquad (6.180)$$

or, alternatively, use the renormalized action:

$$\begin{aligned} \langle \phi(k_1)\phi(k_2)\rangle &= \int \frac{D\phi' e^{-H'[\phi']}}{Z'} b^{2\rho}\phi'(bk_1)\phi'(bk_1) \\ &= b^{2\rho}\chi'(kb_1)\delta(bk_1 + bk_2) \\ &= b^{2\rho-d}\chi'(bk_1)\delta(k_1 + k_2) \qquad (6.181) \end{aligned}$$

where $\chi'(bk) = \chi(bk, r(l), u(l))$ is the correlation function evaluated for H' i.e. with parameters $r(l)$ and $u(l)$ instead of the bare ones r and u, respectively. It follows

$$\chi(k, r, u) = b^{2\rho-d}\chi(k, r(l), u(l)) \qquad (6.182)$$

This is close to an actual derivation of the above scaling assumption and suggests to identify

$$2\rho - d = 2 - \eta. \qquad (6.183)$$

What is missing is to demonstrate that $r(l)$ and $u(l)$ give rise to a behavior $te^{yl} = tb^y$ of some quantity t which vanishes at the phase transition. To see this is easier if one performs the calculation explicitly.

6.7.4 ε-expansion of the ϕ^4-theory

We will now follow the recipe outlined in the previous paragraphs and explicitly calculate the functions $r(l)$ and $u(l)$. It turns out that this can be done in a controlled fashion for spatial dimensions close to $d = 4$ and we therefore perform an expansion in $\varepsilon = 4 - d$. In addition we will always assume that the initial coupling constant u is small. We start from the Hamiltonian

$$
\begin{aligned}
H[\phi] \;=\; & \frac{1}{2} \int d^d k \left(r + k^2 \right) |\phi(k)|^2 \\
& + \frac{u}{4} \int d^d k_1 d^d k_2 d^d k_3 \phi(k_1) \phi(k_2) \phi(k_3) \phi(-k_1 - k_2 - k_3)
\end{aligned} \tag{6.184}
$$

Concentrating first on the quadratic term it follows

$$
\begin{aligned}
H_0\left[\phi^>, \phi^<\right] \;=\; & \frac{1}{2} \int_{\Lambda/b < k < \Lambda} d^d k \left(r + k^2 \right) \left|\phi^>(k)\right|^2 \\
& + \frac{1}{2} \int_{k < \Lambda/b} d^d k \left(r + k^2 \right) \left|\phi^<(k)\right|^2
\end{aligned} \tag{6.185}
$$

There is no coupling between the $\phi^>$ and $\phi^<$ and therefore (ignoring constants)

$$
\widetilde{H}_0\left[\phi^<\right] = \frac{1}{2} \int_{k < \Lambda/b} d^d k \left(r + k^2 \right) \left|\phi^<(k)\right| \tag{6.186}
$$

Now we can perform the rescaling $\phi^<(k) = b^\rho \phi'(bk)$ and obtain with $k' = bk$

$$
\begin{aligned}
H_0' \;=\; & \frac{b^{2\rho - d}}{2} \int_{k' < \Lambda} d^d k' \left(r + b^{-2} k^2 \right) |\phi'(k')| \\
\;=\; & \frac{b^{2\rho - d - 2}}{2} \int_{k' < \Lambda} d^d k' \left(b^2 r + k^2 \right) |\phi'(k')|
\end{aligned} \tag{6.187}
$$

This suggests $\rho = \frac{d+2}{2}$ and gives $r(l) = e^{2l} r$.

Next we consider the quartic term

$$
H_{\text{int}} = \frac{u}{4} \int d^d k_1 d^d k_2 d^d k_3 \phi(k_1) \phi(k_2) \phi(k_3) \phi(-k_1 - k_2 - k_3) \tag{6.188}
$$

which does couple $\phi^>$ and $\phi^<$. If all three momenta are inside the inner shell, we can easily perform the rescaling and find

$$
H_{\text{int}}' = \frac{u b^{4\rho - 3d}}{4} \int d^D k_1' d^D k_2' d^D k_3' \phi(k_1') \phi(k_2') \phi(k_3') \phi(-k_1' - k_2' - k_3') \tag{6.189}
$$

which gives with the above result for ρ:

$$
4\rho - 3D = 4 - d \tag{6.190}
$$

yielding

$$
u(l) = u e^{\varepsilon l}. \tag{6.191}
$$

The leading term for small u gives therefore the expected behavior that $u\,(l \to \infty) \to 0$ if $d > 4$ and that u grows if $d < 4$. If d grows we cannot trust the leading behavior anymore and need to go to the next order perturbation theory. Technically this is done using techniques based on Feynman diagrams. The leading order terms can however be obtained quite easily in other ways and we don't need to spend our time on introducing technical tools. It turns out that the next order corrections are identical to the direct perturbation theory,

$$
\begin{aligned}
r' &= e^{2l}r + 3u \int_{\Lambda/b < k < \Lambda} d^d k \frac{1}{r + k^2} \\
u' &= e^{\varepsilon l}u - 9u^2 \int_{\Lambda/b < k < \Lambda} d^d k \frac{1}{\left(r + k^2\right)^2}\,.
\end{aligned}
\tag{6.192}
$$

with the important difference that the momentum integration is restricted to the shell with radius between Λ/b and Λ. This avoids all the complications of our earlier direct perturbation theory where a divergency in u' resulted from the lower limit of the integration (long wave lengths). Integrals of the type

$$
I = \int_{\Lambda/b < k < \Lambda} d^d k f\,(k)
\tag{6.193}
$$

can be easily performed for small l:

$$
\begin{aligned}
I &= K_d \int_{\Lambda e^{-l}}^{\Lambda} k^{d-1} f\,(k)\,dk \simeq K_d \Lambda^{d-1} f\,(\Lambda)\left(\Lambda - \Lambda e^{-l}\right) \\
&\simeq K_d \Lambda^d f\,(\Lambda)\,l
\end{aligned}
\tag{6.194}
$$

It follows therefore

$$
\begin{aligned}
r' &= (1 + 2l)\,r + \frac{3K_d \Lambda^d}{r + \Lambda^2} ul \\
u' &= (1 + \varepsilon l)\,u - \frac{9K_d \Lambda^d}{\left(r + \Lambda^2\right)^2} u^2 l\,.
\end{aligned}
\tag{6.195}
$$

which is due to the small l limit conveniently written as a differential equation

$$
\begin{aligned}
\frac{dr}{dl} &= 2r + \frac{3K_d \Lambda^d}{r + \Lambda^2} u \\
\frac{du}{dl} &= \varepsilon u - \frac{9K_d \Lambda^d}{\left(r + \Lambda^2\right)^2} u^2.
\end{aligned}
\tag{6.196}
$$

Before we proceed we introduce more convenient variables

$$
\begin{aligned}
r &\to \frac{r}{\Lambda^2} \\
u &\to K_d \Lambda^{d-4} u
\end{aligned}
\tag{6.197}
$$

which are dimensionless and obtain the differential equations

$$
\begin{aligned}
\frac{dr}{dl} &= 2r + \frac{3u}{1+r} \\
\frac{du}{dl} &= \varepsilon u - \frac{9u^2}{(1+r)^2}.
\end{aligned}
$$
(6.198)

The system has indeed a fixed point (where $\frac{dr}{dl} = \frac{du}{dl} = 0$) determined by

$$
\begin{aligned}
\varepsilon &= \frac{9u^*}{(1+r^*)^2} \\
2r^* &= -\frac{3u^*}{1+r^*}
\end{aligned}
$$
(6.199)

This simplifies at leading order in ε to

$$
\begin{aligned}
u^* &= \frac{\varepsilon}{9} \text{ or } 0 \\
r^* &= -\frac{3}{2}u^*
\end{aligned}
$$
(6.200)

If the system reaches this fixed point it will be governed by the behavior it its immediate vicinity, allowing us to linearize the flow equation in the vicinity of the fixed point, i.e. for small

$$
\begin{aligned}
\delta r &= r - r^* \\
\delta u &= u - u^*
\end{aligned}
$$
(6.201)

Consider first the fixed point with $u^* = r^* = 0$ gives

$$
\frac{d}{dl}\begin{pmatrix} \delta r \\ \delta u \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & \varepsilon \end{pmatrix}\begin{pmatrix} \delta r \\ \delta u \end{pmatrix}
$$
(6.202)

with eigenvalues $\lambda_1 = 2$ and $\lambda_2 = \varepsilon$. Both eigenvalues are positive for $\varepsilon > 0$ $(D < 4)$ such that there is no scenario under which this fixed point is ever governing the low energy physics of the problem.

Next we consider $u^* = \frac{\varepsilon}{9}$ and $r^* = -\frac{\varepsilon}{6}$. It follows

$$
\frac{d}{dl}\begin{pmatrix} \delta r \\ \delta u \end{pmatrix} = \begin{pmatrix} 2 - \frac{\varepsilon}{3} & 3 + \frac{\varepsilon}{2} \\ 0 & -\varepsilon \end{pmatrix}\begin{pmatrix} \delta r \\ \delta u \end{pmatrix}
$$
(6.203)

with eigenvalues

$$
\begin{aligned}
y &= 2 - \frac{\varepsilon}{2} \\
y' &= -\varepsilon
\end{aligned}
$$
(6.204)

the corresponding eigenvectors are

$$
\begin{aligned}
e &= (1,0) \\
e' &= \left(-\frac{3}{2} + \frac{\varepsilon}{8}, 1\right)
\end{aligned}
$$
(6.205)

Thus, a variation along the $e$direction (which is varying r) causes the system to leave the fixed point (positive eigenvalue), whereas it will approach the fixed point if

$$(r, u) \sim e'$$ (6.206)

this gives

$$r = u \left(-\frac{3}{2} + \frac{\varepsilon}{8} \right)$$ (6.207)

which defines the *critical surface* in parameter space. If a system is on this surface it approaches the fixed point. If it is slightly away, the quantity

$$t = r - u \left(-\frac{3}{2} + \frac{\varepsilon}{8} \right)$$ (6.208)

is non-zero and behaves as

$$t(l) = te^{yl} = tb^y.$$ (6.209)

The flow behavior for large l is only determined by the value of t which is the only scaling variable, which vanishes at the critical point. Returning now to the initial scaling behavior of the correlation function we can write explicitly

$$\chi(k, t) = b^2 \chi(k, tb^y)$$ (6.210)

comparing this with $\chi(q, t) = b^{2-\eta} \chi\left(bq, tb^{\frac{1}{\nu}}\right)$ gives immediately the two critical exponents

$$\begin{aligned} \eta &= \mathcal{O}\left(\varepsilon^2\right) \\ \nu &\simeq \frac{1}{2} + \frac{\varepsilon}{8}. \end{aligned}$$ (6.211)

Extrapolating this to the $\varepsilon = 1$ case gives numerical results for the critical exponents which are much closer to the exact ones (obtained via numerical simulations)

exponent	$\varepsilon -$ expansion	$d = 3$, Ising
α	0.125	0.12
β	0.3125	0.31
γ	1.25	1.25
ν	0.62	0.64
δ	5	5.0
η	0	0.04

A systematic improvement of these results occurs if one includes higher order terms of the εexpansion. Thus, the renormalization group approach is a very powerful tool to analyze the highly singular perturbation expansion of the ϕ^4-theory below its upper critical dimension. How is it possible that one can obtain so much information by essentially performing a low order expansion in u for

a small set of high energy degrees of freedom? The answer is in the power of the scaling concept. We have assumed that the form $\chi(q,t) = b^{2-\eta}\chi\left(bq, tb^{\frac{1}{\nu}}\right)$ which we obtained for very small deviations of b from unity is valid for all b. If for example the value of ν and η would change with l there would be no way that we could determine the critical exponents from such a procedure. If scaling does not apply, no critical exponent can be deduced from the renormalization group.

6.7.5 Irrelevant interactions

Finally we should ask why we restricted ourself to the quartic interaction only. For example, one might have included a term of the type

$$H_{(6)} = \frac{v}{6}\int d^d x \phi(x)^6 \tag{6.212}$$

which gives in momentum space

$$
\begin{aligned}
H_{(6)} &= \frac{v}{6}\int d^d k_1 ... d^d k_5 \phi(k_1)\phi(k_2)\phi(k_3)(k_4)\phi(k_5)\\
&\times \phi(-k_1 - k_2 - k_3 - k_4 - k_5)
\end{aligned} \tag{6.213}
$$

The leading term of the renormalization group is the one where all three momenta are inside the inner shell, and we can perform the rescaling immediately:

$$
\begin{aligned}
H'_{\text{int}} &= \frac{vb^{6\rho-5d}}{5}\int d^d k'_1 ... d^d k'_5 \phi(k'_1)\phi(k'_2)\phi(k'_3)\phi(k'_4)\phi(k'_5)\\
&\times \phi(-k'_1 - k'_2 - k'_3 - k'_4 - k'_5)
\end{aligned} \tag{6.214}
$$

and the v dependence is with $\rho = \frac{2+d}{2}$ and $6\rho - 5d = 2(3-d) = -2(1-\varepsilon)$

$$v(l) = ve^{-2(1-\varepsilon)l} \tag{6.215}$$

Thus, in the strict sense of the ε expansion such a term will never play a role. Only if $u \ll \varepsilon$ initially is it important to keep these effects into account. This happens in the vicinity of a so called tricritical point.

Chapter 7

Density matrix and fluctuation dissipation theorem

One can make a number of fairly general statements about quantum statistical systems using the concept of the density matrix. In equilibrium we found that the expectation value of a physical observable is given by

$$\langle O \rangle_{\text{eq}} = \text{tr}\left(\rho_{\text{eq}} O\right) \tag{7.1}$$

with density operator

$$\rho_{\text{eq}} = \frac{1}{Z} e^{-\beta H}. \tag{7.2}$$

The generalization to the grand canonical ensemble is straight forward. The density operator (or often called density matrix) is now given as $\rho_{\text{eq}} = \frac{1}{Z_g} e^{-\beta(H - \mu N)}$, where N is the particle number operator.

Considering now a system in a quantum state $|\psi_i\rangle$ with energy E_i, the expectation value of a physical observable is in that state is

$$O_i = \langle \psi_i | O | \psi_i \rangle. \tag{7.3}$$

If the system is characterized by a distribution function where a state $|\psi_i\rangle$ occurs with probability p_i it follows that the actual expectation value of O is

$$\langle O \rangle = \sum_i p_i O_i. \tag{7.4}$$

This can be written in a formal sense as

$$\langle O \rangle = \text{tr}\left(\rho O\right) \tag{7.5}$$

with the density operator

$$\rho = \sum_i |\psi_i\rangle \, p_i \, \langle\psi_i| \, . \tag{7.6}$$

Inserting this expression into Eq.7.5 gives the above result of one uses that the $|\psi_i\rangle$ are orthonormal.

A state is called pure if $p_i = 0$ for all $i \neq 0$ and $p_0 = 1$. Then

$$\rho_{\text{pure}} = |\psi_0\rangle \, \langle\psi_0| \tag{7.7}$$

which implies $\rho_{\text{pure}}^2 = \rho_{\text{pure}}$. States which are not pure are called mixed. Indeed, it holds in general

$$\text{tr}\,(\rho) = \sum_i p_i = 1 \tag{7.8}$$

which gives

$$\text{tr}\,(\rho^2) = \sum_i p_i^2 \leq 1. \tag{7.9}$$

The equal sign only holds for pure states making $\text{tr}\,(\rho^2)$ a general criterion for a state to be mixed.

Next we determine the equation of motion of the density matrix under the assumption that the probabilities are fixed: $\frac{dp_i}{dt} = 0$. We use

$$i\hbar \frac{\partial}{\partial t} |\psi_i\rangle = H |\psi_i\rangle \tag{7.10}$$

and

$$-i\hbar \frac{\partial}{\partial t} \langle\psi_i| = \langle\psi_i| H \tag{7.11}$$

which gives

$$
\begin{aligned}
i\hbar \frac{\partial}{\partial t} \rho &= \sum_i H |\psi_i\rangle \, p_i \, \langle\psi_i| - |\psi_i\rangle \, p_i \, \langle\psi_i| H = H\rho - \rho H \\
&= [H, \rho] \tag{7.12}
\end{aligned}
$$

which is called von Neuman equation.

The von Neuman equation should not be confused with the Heisenberg equation of operators in Heisenberg picture. The latter is given by

$$i\hbar \frac{d}{dt} A_t\,(t) = i\hbar \frac{\partial}{\partial t} A_t\,(t) + [A_t\,(t)\,, H] \tag{7.13}$$

The first term is the explicit variation of the operator A_t which might be there even in Schrödinger picture. The second term results from the unitary transformation $A_t\,(t) = e^{-iHt/\hbar} A_t e^{iHt/\hbar}$.

Once we know ρ, we can analyze for example the entropy

$$S = -k_{\text{B}} \text{tr} \rho \log \rho. \tag{7.14}$$

Using the von Neuman equation gives for the time variation of the entropy

$$
\begin{aligned}
\frac{\partial S}{\partial t} &= -k_{\text{B}}\text{tr}\left[\frac{\partial \rho}{\partial t}\left(\log\rho + 1\right)\right] \\
&= i\frac{k_{\text{B}}}{\hbar}\text{tr}\left[[\text{H},\rho]\left(\log\rho + 1\right)\right] = 0.
\end{aligned} \tag{7.15}
$$

Obviously this is a consequence of the initial assumption $\frac{dp_i}{dt} = 0$. We conclude that the von Neuman equation will not allow us to draw conclusions about the change of entropy as function of time.

7.1 Density matrix of subsystems

Conceptually very important information can be obtained if one considers the behavior of the density matrix of a subsystem of a bigger system. The bigger system is then assumed to be in a pure quantum state. We denote the variables of our subsystem with x and the variables of the bigger system which don't belong to our subsystem with Y. The wave function of the pure quantum mechanical state of the entire system is

$$
\Psi\left(Y, x, t\right) \tag{7.16}
$$

and we expand it's x-dependence in terms of a complete set of functions acting on the subsystem. Without loss of generality we can say

$$
\Psi\left(Y, x, t\right) = \sum_{\alpha}\Phi_{\alpha}\left(Y, t\right)\varphi_{\alpha}\left(x\right). \tag{7.17}
$$

Let $O\left(x\right)$ be some observable of the subsystem, i.e. the operator O does not act on the coordinates Y. If follows

$$
\langle O\rangle = \langle\Psi\left|O\right|\Psi\rangle = \sum_{\alpha,\alpha'}\langle\Phi_a\left(t\right)|\Phi_{\alpha'}\left(t\right)\rangle\langle\varphi_{\alpha}|O|\varphi_{\alpha'}\rangle \tag{7.18}
$$

This suggests to introduce the density operator

$$
\rho_{a'\alpha}\left(t\right) = \langle\Phi_a\left(t\right)|\Phi_{\alpha'}\left(t\right)\rangle \tag{7.19}
$$

such that

$$
\langle O\rangle = \text{tr}\rho\text{O}, \tag{7.20}
$$

where the trace is only with respect to the quantum numbers α of the subsystem. Thus, if one assumes that one can characterize the expectation value exclusively within the quantum numbers and coordinates of a subsystem, one is forced to introduce mixed quantum states and a density matrix.

Lets analyze the equation of motion of the density matrix.

$$
i\hbar\frac{\partial}{\partial t}\rho_{a'\alpha}\left(t\right) = i\hbar\int dY\left(\frac{\partial\Phi_a^*\left(Y, t\right)}{\partial t}\Phi_{\alpha'}\left(Y, t\right) + \Phi_a^*\left(Y, t\right)\frac{\partial\Phi_{\alpha'}\left(Y, t\right)}{\partial t}\right) \tag{7.21}
$$

where $\int dY...$ is the matrix element with respect to the variables Y. Since

$\Psi(Y, x, t)$ obeys the Schrödinger equation it follows

$$i\hbar \sum_\alpha \frac{\partial \Phi_\alpha(Y,t)}{\partial t} \varphi_\alpha(x) = H(Y,x) \sum_\alpha \Phi_\alpha(Y,t) \varphi_\alpha(x) \qquad (7.22)$$

Multiplying this by $\varphi_{\alpha'}^*(x)$ and integrating over x gives

$$i\hbar \frac{\partial \Phi_{a'}(Y,t)}{\partial t} = \sum_\beta H_{\alpha'\beta}(Y) \Phi_\beta(Y,t) \qquad (7.23)$$

$$i\hbar \frac{\partial \Phi_a^*(Y,t)}{\partial t} = -\sum_\beta \Phi_\beta^*(Y,t) H_{\beta\alpha}(Y) \qquad (7.24)$$

where

$$H_{\beta\alpha}(Y) = \langle \varphi_\beta | H(Y,x) | \varphi_\alpha \rangle. \qquad (7.25)$$

It follows

$$i\hbar \frac{\partial}{\partial t} \rho_{a'\alpha}(t) = \int dY \sum_\beta \left(\Phi_a^*(Y,t) H_{\alpha'\beta}(Y) \Phi_\beta(Y,t) - \Phi_\beta^*(Y,t) H_{\beta\alpha}(Y) \Phi_{\alpha'}(Y,t) \right)$$
$$(7.26)$$

Lets assume that

$$H(Y,x) = H_0(x) + W(Y,x) \qquad (7.27)$$

where H_0 does not depend on the environment coordinates. It follows

$$i\hbar \frac{\partial}{\partial t} \rho_{a'\alpha}(t) = \sum_\beta (H_{0,\alpha'\beta}\rho_{\beta\alpha} - \rho_{\alpha'\beta}H_{0,\beta\alpha}) - i\gamma_{\alpha,\alpha'}(t)\rho_{a'\alpha}(t) \qquad (7.28)$$

with

$$\gamma_{\alpha,\alpha'} = i\frac{\int dY \sum_\beta \left(\Phi_a^*(Y,t) W_{\alpha'\beta}(Y) \Phi_\beta(Y,t) - \Phi_\beta^*(Y,t) W_{\beta\alpha}(Y) \Phi_{\alpha'}(Y,t) \right)}{\int dY \Phi_{a'}^*(Y,t) \Phi_\alpha(Y,t)}$$
$$(7.29)$$

which obeys $\gamma_{\alpha'\alpha} = \gamma_{\alpha,\alpha'}^*$.

$$i\hbar \frac{\partial}{\partial t} \rho(t) = [H, \rho] - i\Gamma\rho. \qquad (7.30)$$

Thus, only if the subsystem is completely decoupled from the environment do we recover the von Neuman equation. In case there is a coupling between subsystem and environment the equation of motion of the subsystem is more complex, implying $\frac{dp_i}{dt} \neq 0$.

7.2 Linear response and fluctuation dissipation theorem

Lets consider a system coupled to an external field

$$W_t = \int \frac{d\omega}{2\pi} W(\omega) e^{-i(\omega+i\delta)t} \tag{7.31}$$

such that $W_{t\to-\infty} \to 0$. Next we consider the time evolution of a physical quantity A:

$$\langle A \rangle_t = \text{tr}(\rho_t A) \tag{7.32}$$

where

$$i\hbar \frac{\partial}{\partial t} \rho_t = [H + W_t, \rho_t] \tag{7.33}$$

We assume the system is in equilibrium at $t \to -\infty$:

$$\rho_{t\to-\infty} = \rho = \frac{1}{Z} e^{-\beta H}. \tag{7.34}$$

Lets go to the interaction representation

$$\rho_t = e^{-iHt/\hbar} \rho_t(t) e^{iHt/\hbar} \tag{7.35}$$

gives

$$i\hbar \frac{\partial \rho_t}{\partial t} = [H, \rho_t] + e^{-iHt/\hbar} i\hbar \frac{\partial \rho_t(t)}{\partial t} e^{iHt/\hbar} \tag{7.36}$$

which gives

$$i\hbar \frac{\partial \rho_t(t)}{\partial t} = [W_t(t), \rho_t(t)] \tag{7.37}$$

which is solved by

$$\begin{aligned} \rho_t(t) &= \rho - i\hbar \int_{-\infty}^t dt' [W_{t'}(t'), \rho_{t'}(t')] \\ \rho_t &= \rho - i\hbar \int_{-\infty}^t dt' e^{-iH(t-t')/\hbar} [W_{t'}, \rho_{t'}] e^{iH(t-t')/\hbar}. \end{aligned} \tag{7.38}$$

Up to leading order this gives

$$\rho_t = \rho - i\hbar \int_{-\infty}^t dt' e^{-iH(t-t')/\hbar} [W_{t'}, \rho] e^{iH(t-t')/\hbar}. \tag{7.39}$$

We can now determine the expectation value of A:

$$\langle A \rangle_t = \langle A \rangle - i\hbar \int_{-\infty}^t dt' \text{tr}([W_{t'}(t'), \rho] A(t)) \tag{7.40}$$

one can cyclically change order under the trace operation

$$\text{tr}(W\rho - \rho W) A = \text{tr}(AW - WA) \rho \tag{7.41}$$

which gives

$$\langle A \rangle_t = \langle A \rangle - i\hbar \int_{-\infty}^{t} dt' \langle [A(t), W_{t'}(t')] \rangle \tag{7.42}$$

It is useful to introduce (the retarded Green's function)

$$\langle\langle A(t); W_{t'}(t') \rangle\rangle = -i\hbar\theta(t-t') \langle [A(t), W_{t'}(t')] \rangle \tag{7.43}$$

such that

$$\langle A \rangle_t = \langle A \rangle + \int_{-\infty}^{\infty} dt' \langle\langle A(t); W_{t'}(t') \rangle\rangle. \tag{7.44}$$

The interesting result is that we can characterize the deviation from equilibrium (dissipation) in terms of fluctuations of the equilibrium (equilibrium correlation function).

Example, conductivity:

Here we have an interaction between the electrical field and the electrical polarization:

$$W_t = -\mathbf{p} \cdot \mathbf{E}_t \tag{7.45}$$

with

$$\mathbf{E}_t = \mathbf{E}_0 \exp\left(-i\left(\omega + i\delta\right)t\right) \tag{7.46}$$

If we are interested in the electrical current it follows

$$\langle j^\alpha \rangle_t = -\int_{-\infty}^{\infty} dt' \langle\langle j^\alpha(t); p^\beta(t') \rangle\rangle E^\beta e^{-i(\omega+i\delta)t'} \tag{7.47}$$

which gives in Fourier space

$$\langle j^\alpha \rangle_\omega = -\langle\langle j^\alpha; p^\beta \rangle\rangle_\omega E_0 \tag{7.48}$$

which gives for the conductivity

$$\sigma(\omega) = -\langle\langle j^\alpha; p^\beta \rangle\rangle_\omega. \tag{7.49}$$

Obviously, a conductivity is related to dissipation whereas the correlation function $\langle [j^\alpha(t), p^\beta(t')] \rangle$ is just an equilibrium fluctuation.

Chapter 8

Brownian motion and stochastic dynamics

We start our considerations by considering the diffusion of a particle with density $\rho(\mathbf{x},t)$. Diffusion should take place if there is a finite gradient of the density $\nabla\rho$. To account for the proper bookkeeping of particles, one starts from the continuity equation

$$\frac{\partial\rho}{\partial t} = \nabla \cdot \mathbf{j} \tag{8.1}$$

with a current given by $\mathbf{j} \simeq D\nabla\rho$, which is called Fick's law. The pre-factor D is the diffusion constant and we obtain $\frac{\partial\rho}{\partial t} = D\nabla^2\rho$. This is the diffusion equation, which is conveniently solved by going into Fourier representation

$$\rho(\mathbf{x},t) = \int \frac{d^d k}{(2\pi)^d} \rho(\mathbf{k},t) e^{-i\mathbf{k}\cdot\mathbf{x}}, \tag{8.2}$$

yielding an ordinary differential equation with respect to time:

$$\frac{\partial\rho(\mathbf{k},t)}{\partial t} = -Dk^2\rho(\mathbf{k},t), \tag{8.3}$$

with solution

$$\rho(\mathbf{k},t) = \rho_0(\mathbf{k}) e^{-Dk^2 t}. \tag{8.4}$$

where $\rho_0(\mathbf{k}) = \rho(\mathbf{k},t=0)$. Assuming that $\rho(\mathbf{x},t=0) = \delta(\mathbf{x})$, i.e. a particle at the origin, it holds $\rho_0(\mathbf{k}) = 1$ and we obtain

$$\rho(\mathbf{x},t) = \int \frac{d^d k}{(2\pi)^d} e^{-Dk^2 t} e^{-i\mathbf{k}\cdot\mathbf{x}}. \tag{8.5}$$

The Fourier transformation is readily done and it follows

$$\rho(x,t) = \frac{1}{(4\pi Dt)^{d/2}} e^{-\mathbf{x}^2/(4Dt)}. \tag{8.6}$$

In particular, it follows that $\langle x^2(t)\rangle = 2Dt$ grows only linearly in time, as opposed to the ballistic motion of a particle where $\langle x(t)\rangle = vt$.

8.1 Langevin equation

A more detailed approach to diffusion and Brownian motion is given by using the concept of a stochastic dynamics. We first consider one particle embedded in a fluid undergoing Brownian motion. Later we will average over all particles of this type. The fluid is modeled to cause friction and to randomly push the particle. This leads to the equation of motion for the velocity $v = \frac{dx}{dt}$:

$$\frac{dv(t)}{dt} = -\frac{\gamma}{m} v(t) + \frac{1}{m} \xi(t). \qquad (8.7)$$

Here γ is a friction coefficient proportional to the viscosity of the host fluid. If we consider large Brownian particles, the friction term can be expressed in terms of the shear viscosity, η, of the fluid and the radius, R, of the particle: $\gamma = 6\pi\eta R$. $\xi(t)$ is a random force, simulating the scattering of the particle with the fluid and is characterized by the correlation functions

$$
\begin{aligned}
\langle \xi(t) \rangle_\xi &= 0 \\
\langle \xi(t) \xi(t') \rangle_\xi &= g\delta(t - t').
\end{aligned}
\qquad (8.8)
$$

The pre-factor g is the strength of this noise. As the noise is uncorrelated in time, it is also referred to as white noise (all spectral components are equally present in the Fourier transform of $\langle \xi(t) \xi(t') \rangle_\xi$, similar to white light).

The above stochastic differential equation can be solved analytically for arbitrary ξ yielding

$$v(t) = v_0 e^{-\gamma t/m} + \frac{1}{m} \int_0^t ds\, e^{-\gamma(t-s)/m} \xi(s) \qquad (8.9)$$

and

$$x(t) = x_0 + \frac{mv_0}{\gamma} \left(1 - e^{-\gamma t/m}\right) + \frac{1}{\gamma} \int_0^t ds \left(1 - e^{-\gamma(t-s)/m}\right) \xi(s). \qquad (8.10)$$

We can now directly perform the averages of this result. Due to $\langle \xi(t) \rangle_\xi = 0$ follows

$$
\begin{aligned}
\langle v(t) \rangle_\xi &= v_0 e^{-\gamma t/m} \\
\langle x(t) \rangle_\xi - x_0 &= \frac{mv_0}{\gamma} \left(1 - e^{-\gamma t/m}\right)
\end{aligned}
\qquad (8.11)
$$

which implies that a particle comes to rest at a time $\tau \simeq \frac{m}{\gamma}$, which can be long if the viscosity of the fluid is small (γ is small). More interesting are the correlations between the velocity and positions at distant times. Inserting the results for $v(t)$ and $x(t)$ and using $\langle \xi(t) \xi(t') \rangle_\xi = g\delta(t - t')$ gives

$$\langle v(t) v(t') \rangle_\xi = \left(v_0^2 - \frac{g}{2m\gamma}\right) e^{-\gamma(t+t')/m} + \frac{g}{2m\gamma} e^{-\gamma(t-t')/m} \qquad (8.12)$$

and similarly

$$\left\langle (x(t) - x_0)^2 \right\rangle = \frac{m^2}{\gamma^2} \left(v_0^2 - \frac{g}{2m\gamma} \right) \left(1 - e^{-\gamma t/m} \right)^2$$
$$+ \frac{g}{\gamma^2} \left(t - \frac{m}{\gamma} \left(1 - e^{-\gamma t/m} \right) \right). \tag{8.13}$$

If the Brownian particle is in equilibrium with the fluid, we can average over all the directions and magnitudes of the velocity $v_0^2 \to \left\langle v_0^2 \right\rangle_T = \frac{k_B T}{m}$. Where $\langle ... \rangle_T$ refers to the thermal average over all particles (so far we only considered one of them embedded in the fluid). Furthermore, in equilibrium the dynamics should be stationary, i.e.

$$\left\langle \langle v(t) v(t') \rangle_\xi \right\rangle_T = f(t - t') \tag{8.14}$$

should only depend on the relative time $t - t'$ and not on some absolute time point, like t, t' or $t + t'$. This is fulfilled if

$$\left\langle v_0^2 \right\rangle_T = \frac{g}{2m\gamma} \tag{8.15}$$

which enables us to express the noise strength at equilibrium in terms of the temperature and the friction coefficient

$$g = 2\gamma k_B T \tag{8.16}$$

which is one of the simplest realization of the fluctuation dissipation theorem. Here the friction is a dissipative effect whereas the noise a fluctuation effect. Both are closely related in equilibrium.

This allows us to analyze the mean square displacement in equilibrium

$$\left\langle (x(t) - x_0)^2 \right\rangle_T = \frac{2k_B T}{\gamma} \left(t - \frac{m}{\gamma} \left(1 - e^{-\gamma t/m} \right) \right). \tag{8.17}$$

In the limit of long times $t \gg \tau = \frac{m}{\gamma}$ holds

$$\left\langle (x(t) - x_0)^2 \right\rangle_T \simeq \frac{2k_B T}{\gamma} t \tag{8.18}$$

which the result obtained earlier from the diffusion equation if we identify $D = \frac{k_B T}{\gamma}$. Thus, even though the mean velocity of the particle vanishes for $t \gg \tau$ it does not mean that it comes to rest, the particle still increases its mean displacement, only much slower than via a ballistic motion. This demonstrates how important it is to consider, in addition to mean values like $\langle v(t) \rangle_\xi$, correlation functions of higher complexity.

8.2 Random electrical circuits

Due to the random motion and discrete nature of electrons, an LRC series circuit experiences a random potential $\xi(t)$. This in turn induces a randomly varying

charge $Q(t)$ on the capacitor plates and a random current

$$I(t) = \frac{dQ}{dt} \tag{8.19}$$

through the resistor and inductor. The random charge satisfies

$$L\frac{d^2Q(t)}{dt^2} + R\frac{dQ(t)}{dt} + \frac{Q(t)}{C} = \xi(t). \tag{8.20}$$

Lets assume that the random potential is correlated according to.

$$
\begin{aligned}
\langle \xi(t) \rangle_\xi &= 0 \\
\langle \xi(t)\xi(t') \rangle_\xi &= g\delta(t - t').
\end{aligned} \tag{8.21}
$$

In addition we use that the energy of the circuit

$$E(Q, I) = \frac{1}{2C}Q^2 + \frac{L}{2}I^2 \tag{8.22}$$

implies via equipartition theorem that

$$
\begin{aligned}
\langle Q_0^2 \rangle_T &= Ck_BT \\
\langle I_0^2 \rangle_T &= \frac{k_BT}{L}.
\end{aligned} \tag{8.23}
$$

The above equation is solved for the current as:

$$
\begin{aligned}
I(t) &= I_0 e^{-\Gamma t}C(t) - \frac{1}{CL\Delta}Q_0 e^{-\Gamma t}\sinh(\Delta t) \\
&\quad + \frac{1}{L}\int_0^t ds\,\xi(s)\,e^{-\Gamma(t-t')}C(t - t')
\end{aligned} \tag{8.24}
$$

with damping rate

$$\Gamma = \frac{R}{L} \tag{8.25}$$

and time constant

$$\Delta = \sqrt{\frac{R^2 - L/C}{L^2}}, \tag{8.26}$$

as well as

$$C(t) = \cosh(\Delta t) - \frac{\Gamma}{\Delta}\sinh(\Delta t). \tag{8.27}$$

Assuming $\langle Q_0 I_0 \rangle_T = 0$ gives (assume $t > t'$)

$$
\begin{aligned}
\left\langle \langle I(t) I(t') \rangle_\xi \right\rangle_T &= e^{-\Gamma(t+t')}C(t)C(t')\langle I_0^2 \rangle_T \\
&\quad + \left(\frac{1}{CL\Delta}\right)^2 \langle Q_0^2 \rangle_T\, e^{-\Gamma(t+t')}\sinh(\Delta t)\sinh(\Delta t') \\
&\quad + \frac{g}{L^2}\int_0^{t'} ds\,e^{-\Gamma(t+t'-2s)}C(t - s)C(t' - s)
\end{aligned} \tag{8.28}
$$

The last integral can be performed analytically, yielding

$$
\frac{e^{-\Gamma(t+t')}}{4\Delta^2\Gamma}\left(\Gamma^2 - \Delta^2\right)\cosh\left(\Delta\left(t'-t\right)\right)
$$
$$
+\frac{e^{\Gamma(t-t')}}{4\Delta}\left(\sinh\left(\Delta\left(t'-t\right)\right) + \frac{\Delta}{\Gamma}\cosh\left(\Delta\left(t'-t\right)\right)\right)
$$
$$
+\frac{e^{\Gamma(t+t')}}{4\Delta^2}\left(\Gamma\cosh\left(\Delta\left(t'+t\right)\right) + \Delta\sinh\left(\Delta\left(t'+t\right)\right)\right) \qquad (8.29)
$$

It is possible to obtain a fully stationary current-current correlation function if

$$
g = 4Rk_BT \qquad (8.30)
$$

which is again an example of the fluctuation dissipation theorem.

It follows $(t' > t)$

$$
\left\langle \langle I\left(t\right) I\left(t'\right)\rangle_\xi \right\rangle_T = \frac{k_BT}{L}e^{-\Gamma(t'-t)}\left(\cosh\left(\Delta\left(t'-t\right)\right) - \frac{\Gamma}{\Delta}\sinh\left(\Delta\left(t'-t\right)\right)\right) \qquad (8.31)
$$

If $C \to 0$, it holds $\Delta \to \Gamma$ and Γ^{-1} is the only time scale of the problem. Current-current correlations decay exponentially. If $0 < \Delta < \Gamma$, the correlation function changes sign for $t' - t \simeq \Delta^{-1}$. Finally, if $R^2 < L/C$, current correlations decay in an oscillatory way (use $\cosh\left(ix\right) = \cos x$ and $\sinh\left(ix\right) = i\sin\left(x\right)$). Then $\Delta = i\delta$ with $\delta = \sqrt{\frac{L/C-R^2}{L^2}}$ and

$$
\left\langle \langle I\left(t\right) I\left(t'\right)\rangle_\xi \right\rangle_T = \frac{k_BT}{L}e^{-\Gamma(t'-t)}\left(\cos\left(\delta\left(t'-t\right)\right) - \frac{\Gamma}{\delta}\sin\left(\delta\left(t'-t\right)\right)\right). \qquad (8.32)
$$

For the fluctuating charge follows

$$
Q\left(t\right) = Q_0 + \int_0^t I\left(s\right)ds. \qquad (8.33)
$$

Chapter 9

Boltzmann transport equation

9.1 Transport coefficients

In the phenomenological theory of transport coefficients one considers a relationship between generalized currents, J_i, and forces, X_i which, close to equilibrium and for small forces is assumed to be linear:

$$J_i = \sum_j L_{ij} X_j \qquad (9.1)$$

Here, the precise definition of the forces is such that in each case an entropy production of the kind

$$\frac{dS}{dt} = \sum_i X_i J_i \qquad (9.2)$$

occurs. If this is the case, the coefficients L_{ij} are symmetric:

$$L_{ij} = L_{ji}, \qquad (9.3)$$

a result originally obtained by Onsager. The origin of this symmetry is the time reversal invariance of the microscopic processes causing each of these transport coefficients. This implies that in the presence of an external magnetic field holds

$$L_{ij}\left(\mathbf{H}\right) = L_{ji}\left(-\mathbf{H}\right). \qquad (9.4)$$

For example in case of an electric current it holds that from Maxwell's equations follows

$$\frac{dS_{\mathrm{E}}}{dt} = J\frac{E}{T} \qquad (9.5)$$

which gives $X_{\mathrm{E}} = \frac{E}{T}$ for the force (note, not the electrical field E itself). Considering a heat flux J_Q gives entropy flux J_Q/T. Then there should be a continuity

equation

$$\frac{dS_{\mathrm{Q}}}{dt} = \nabla \left(J_Q/T \right) = J_Q \nabla \frac{1}{T} \tag{9.6}$$

which gives $X_{\mathrm{Q}} = -\frac{1}{T^2}\nabla T$ for the generalized forces and it holds

$$\begin{aligned} J_E &= L_{EE}X_E + L_{EQ}X_Q \\ &= \frac{L_{EE}}{T}E - \frac{L_{EQ}}{T^2}\nabla T \end{aligned} \tag{9.7}$$

$$\begin{aligned} J_Q &= L_{QE}X_E + L_{QQ}X_Q \\ &= \frac{L_{QE}}{T}E - \frac{L_{QQ}}{T^2}\nabla T. \end{aligned} \tag{9.8}$$

We can now consider a number of physical scenario. For example the electrical current in the absence of a temperature gradient is the determined by the conductivity, σ, via

$$J_E = \sigma E \tag{9.9}$$

which gives

$$\sigma = \frac{L_{EE}}{T}. \tag{9.10}$$

On the other hand, the thermal conductivity is defined as the relation between heat current and temperature gradient in the absence of an electrical current

$$J_Q = -\kappa \nabla T. \tag{9.11}$$

This implies $E = \frac{L_{EQ}}{L_{EE}T}\nabla T$ for the electrical field yielding

$$\kappa = \frac{1}{T^2}\left(L_{QQ} - \frac{L_{QE}^2}{L_{EE}} \right). \tag{9.12}$$

Finally we can consider the Thermopower, which is the above established relation between E and ∇T for $J_E = 0$

$$E = S\nabla T \tag{9.13}$$

with

$$S = \frac{L_{EQ}}{L_{EE}T}. \tag{9.14}$$

One approach to determine these coefficients from microscopic principles is based on the Boltzmann equation.

9.2 Boltzmann equation for weakly interacting fermions

For quasi-classical description of electrons we introduce the Boltzmann distribution function $f_n(\mathbf{k},\mathbf{r},t)$. This is the probability to find an electron in state

$n, \mathbf{k}$ at point $\mathbf{r}$ at time t. More precisely is f/V the probability density to find an electron in state $n, \mathbf{k}$ in point $\mathbf{r}$. This means the probability to find it in a volume element dV is given by $f dV/V$.

We consider both $\mathbf{k}$ and $\mathbf{r}$ defined. This means that we consider wave packets with both $\mathbf{k}$ and $\mathbf{r}$ (approximately) defined, however always such that the uncertainty relation $\Delta k \Delta r \sim 1$ holds.

The electron density and the current density are given by

$$n(\mathbf{r}, t) = \frac{1}{V} \sum_{n, \mathbf{k}, \sigma} f_n(\mathbf{k}, \mathbf{r}, t) \tag{9.15}$$

$$\mathbf{j}(\mathbf{r}, t) = -\frac{e}{V} \sum_{n, \mathbf{k}, \sigma} \mathbf{v_k} f_n(\mathbf{k}, \mathbf{r}, t) \tag{9.16}$$

The equations of motion of non-interacting electrons in a periodic solid and weak external fields are

$$\frac{d\mathbf{r}}{dt} = \mathbf{v_k} = \frac{1}{\hbar} \left(\frac{\partial \varepsilon_n(\mathbf{k})}{\partial \mathbf{k}} \right) , \tag{9.17}$$

and

$$\hbar \frac{d\mathbf{k}}{dt} = -e\mathbf{E} - \frac{e}{c} (\mathbf{v} \times \mathbf{B}) . \tag{9.18}$$

They determine the evolution of the individual $\mathbf{k}(t)$ and $\mathbf{r}(t)$ of each wave packet.

If the electron motion would be fully determined by the equations of motion, the distribution function would satisfy

$$f_n(\mathbf{k}(t), \mathbf{r}(t), t) = f_n(\mathbf{k}(0), \mathbf{r}(0), 0) \tag{9.19}$$

Thus, the full time derivative would vanish

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f + \frac{\partial \mathbf{r}}{\partial t} \cdot \nabla_{\mathbf{r}} f = 0 \tag{9.20}$$

However, there are processes which change the distribution function. These are collisions with impurities, phonons, other electrons. The new equation reads

$$\frac{df}{dt} = \frac{\partial f}{\partial t} + \frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f + \frac{\partial \mathbf{r}}{\partial t} \cdot \nabla_{\mathbf{r}} f = \left(\frac{\partial f}{\partial t} \right)_{\text{Coll}} , \tag{9.21}$$

where $\left(\frac{\partial f}{\partial t} \right)_{\text{Coll}} = I[f]$ is called the collision integral.

Using the equations of motion we obtain the celebrated Boltzmann equation

$$\frac{\partial f}{\partial t} - \frac{e}{\hbar} \left(\mathbf{E} + \frac{1}{c} (\mathbf{v} \times \mathbf{B}) \right) \cdot \nabla_{\mathbf{k}} f + \mathbf{v}_{\mathbf{k}, n} \cdot \nabla_{\mathbf{r}} f = I[f] . \tag{9.22}$$

The individual contributions to $\frac{df}{dt}$ can be considered as a consequence of spatial inhomogeneity effects, such as temperature or chemical potential gradients (carriers of a given state enter from adjacent regions enter into $\mathbf{r}$ whilst others leave):

$$\mathbf{v_k} \cdot \nabla_{\mathbf{r}} f \simeq \mathbf{v_k} \cdot \frac{\partial f}{\partial T} \nabla T \tag{9.23}$$

In addition, there are effects due to external fields (changes of the **k**-vector at the rate)

$$-\frac{e}{\hbar}\left(\mathbf{E}+\frac{1}{c}\left(\mathbf{v}\times\mathbf{B}\right)\right)\cdot\nabla_{\mathbf{k}}f \tag{9.24}$$

Finally there are scattering processes, characterized by $I[f]$ which are determined by the difference between the rate at which the state **k** is entered and the rate at which carriers are lost from it.

9.2.1   Collision integral for scattering on impurities

The collision integral describes processes that bring about change of the state of the electrons, i.e., transitions. There are several reasons for the transitions: phonons, electron-electron collisions, impurities. Here we consider only one: scattering off impurities.

Scattering in general causes transitions in which an electron which was in the state $n_1, \mathbf{k}_1$ is transferred to the state $n_2, \mathbf{k}_2$. We will suppress the band index as in most cases we consider scattering within a band. The collision integral has two contribution: in and out: $I = I_{\text{in}} + I_{\text{out}}$.

The in-part describes transitions from all the states to the state **k**:

$$I_{\text{in}}[f] = \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k}) f(\mathbf{k}_1, \mathbf{r})[1 - f(\mathbf{k}, \mathbf{r})] \, , \tag{9.25}$$

where $W(\mathbf{k}_1, \mathbf{k})$ is the transition probability per unit of time (rate) from state $\mathbf{k}_1$ to state **k** given the state $\mathbf{k}_1$ is initially occupied and the state **k** is initially empty. The factors $f(\mathbf{k}_1)$ and $1 - f(\mathbf{k})$ take care for the Pauli principle.

The out-part describes transitions from the state **k** to all other states:

$$I_{\text{out}}[f] = -\sum_{\mathbf{k}_1} W(\mathbf{k}, \mathbf{k}_1) f(\mathbf{k}, \mathbf{r})[1 - f(\mathbf{k}_1, \mathbf{r})] \, , \tag{9.26}$$

The collision integral should vanish for the equilibrium state in which

$$f(\mathbf{k}) = f_0(\mathbf{k}) = \frac{1}{\exp\left[\frac{\varepsilon(\mathbf{k}) - \mu}{k_{\text{B}}T}\right] + 1} \, . \tag{9.27}$$

This can be rewritten as

$$\exp\left[\frac{\varepsilon(\mathbf{k}) - \mu}{k_{\text{B}}T}\right] f_0 = 1 - f_0 \, . \tag{9.28}$$

The requirement $I_{\text{in}}[f_0] + I_{\text{out}}[f_0]$ is satisfied if

$$W(\mathbf{k}, \mathbf{k}_1) \exp\left[\frac{\varepsilon(\mathbf{k}_1)}{k_{\text{B}}T}\right] = W(\mathbf{k}_1, \mathbf{k}) \exp\left[\frac{\varepsilon(\mathbf{k})}{k_{\text{B}}T}\right] \, . \tag{9.29}$$

We only show here that this is sufficient but not necessary. The principle that it is always so is called *detailed balance principle*. In particular, for elastic processes, in which $\varepsilon(\mathbf{k}) = \varepsilon(\mathbf{k}_1)$, we have

$$W(\mathbf{k}, \mathbf{k}_1) = W(\mathbf{k}_1, \mathbf{k}) \; . \tag{9.30}$$

In this case (when only elastic processes are present we obtain)

$$
\begin{aligned}
I[f] &= \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k}) f(\mathbf{k}_1)[1 - f(\mathbf{k})] - \sum_{\mathbf{k}_1} W(\mathbf{k}, \mathbf{k}_1) f(\mathbf{k})[1 - f(\mathbf{k}_1)] \\
&= \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k}) \left(f(\mathbf{k}_1) - f(\mathbf{k}) \right) \; .
\end{aligned}
\tag{9.31}
$$

9.2.2 Relaxation time approximation

We introduce $f = f_0 + \delta f$. Since $I[f_0] = 0$ we obtain

$$I[f] = \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k}) \left(\delta f(\mathbf{k}_1) - \delta f(\mathbf{k}) \right).$$

Assume the rates W are all equal and $\sum_{\mathbf{k}_1} \delta f(\mathbf{k}_1) = 0$ (no change in total density), then $I[f] \sim -\delta f(\mathbf{k})$. We introduce the relaxation time τ such that

$$I[f] = -\frac{\delta f}{\tau} \; . \tag{9.32}$$

This form of the collision integral is more general. That is it can hold not only for the case assumed above. Even if this form does not hold exactly, it serves as a simple tool to make estimates.

More generally, one can assume τ is $\mathbf{k}$-dependent, $\tau_{\mathbf{k}}$. Then

$$I[f(\mathbf{k})] = -\frac{\delta f(\mathbf{k})}{\tau_{\mathbf{k}}} \; . \tag{9.33}$$

9.2.3 Conductivity

Within the τ-approximation we determine the electrical conductivity. Assume an oscillating electric field is applied, where $\mathbf{E}(t) = \mathbf{E} e^{-i\omega t}$. The Boltzmann equation reads

$$\frac{\partial f}{\partial t} - \frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} f + \mathbf{v}_{\mathbf{k}} \cdot \nabla_{\mathbf{r}} f = -\frac{f - f_0}{\tau_{\mathbf{k}}} \; . \tag{9.34}$$

Since the field is homogeneous we expect homogeneous response $\delta f(t) = \delta f e^{-i\omega t}$. This gives

$$-\frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} f = \left(i\omega - \frac{1}{\tau_{\mathbf{k}}} \right) \delta f \; . \tag{9.35}$$

If we are only interested in the linear response with respect to the electric field, we can replace f by f_0 in the l.h.s. This gives

$$-\frac{e}{\hbar} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \hbar \mathbf{v}_{\mathbf{k}} \cdot \mathbf{E} = \left(i\omega - \frac{1}{\tau_{\mathbf{k}}} \right) \delta f \; . \tag{9.36}$$

and we obtain

$$\delta f = \frac{e\tau_{\mathbf{k}}}{1 - i\omega\tau_{\mathbf{k}}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \mathbf{v}_{\mathbf{k}} \cdot \mathbf{E} \tag{9.37}$$

For the current density we obtain $\mathbf{j}(t) = \mathbf{j}e^{-i\omega t}$, where

$$
\begin{aligned}
\mathbf{j} &= -\frac{e}{V} \sum_{\mathbf{k},\sigma} \mathbf{v}_{\mathbf{k}} \delta f(\mathbf{k}) \\
&= -\frac{2e^2}{V} \sum_{\vec{k}} \frac{\tau_{\mathbf{k}}}{1 - i\omega\tau_{\mathbf{k}}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} (\mathbf{v}_{\mathbf{k}} \cdot \mathbf{E}) \, \mathbf{v}_{\mathbf{k}} \\
&= -2e^2 \int \frac{d^3k}{(2\pi)^3} \frac{\tau_{\mathbf{k}}}{1 - i\omega\tau_{\mathbf{k}}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} (\mathbf{v}_{\mathbf{k}} \cdot \mathbf{E}) \, \mathbf{v}_{\mathbf{k}}.
\end{aligned}
\tag{9.38}
$$

We define the conductivity tensor $\sigma_{\alpha\beta}$ via $j_\alpha = \sum_\alpha \sigma_{\alpha,\beta} E_\beta$. Thus

$$\sigma_{\alpha,\beta} = -2e^2 \int \frac{d^3k}{(2\pi)^3} \frac{\tau_{\mathbf{k}}}{1 - i\omega\tau_{\mathbf{k}}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} v_{\mathbf{k}\alpha} v_{\mathbf{k}\beta} .$$

At low enough temperatures, i.e., for $k_{\mathrm{B}}T \ll \mu$,

$$\frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \approx -\delta(\varepsilon_{\mathbf{k}} - \mu) - \frac{\pi^2}{6} (k_{\mathrm{B}}T)^2 \delta''(\varepsilon_{\mathbf{k}} - \mu) , \tag{9.39}$$

Assuming τ is constant and the band energy is isotropic (effective mass is simple) we obtain

$$
\begin{aligned}
\sigma_{\alpha,\beta} &= -\frac{e^2\tau}{1 - i\omega\tau} \int d\varepsilon \rho(\varepsilon) \frac{d\Omega}{4\pi} \frac{\partial f_0}{\partial \varepsilon} v_\alpha v_\beta \\
&= \frac{e^2\tau\rho_F}{1 - i\omega\tau} \int \frac{d\Omega}{4\pi} v_\alpha v_\beta = \frac{2e^2\tau\rho_F}{(1 - i\omega\tau)} \frac{v_F^2}{3} \delta_{\alpha,\beta} .
\end{aligned}
\tag{9.40}
$$

For the dc-conductivity, i.e., for $\omega = 0$ we obtain

$$\sigma_{\alpha,\beta} = \frac{e^2\tau\rho_F v_F^2}{3} \delta_{\alpha,\beta} \tag{9.41}$$

where ρ_F is the total density of states at the Fermi level.

9.2.4   Determining the transition rates

Impurities are described by an extra potential acting on electrons

$$U_{\mathrm{imp}}(\mathbf{r}) = \sum_j v(\mathbf{r} - \mathbf{c}_j) , \tag{9.42}$$

where $\mathbf{c}_j$ are locations of the impurities.

In the Born approximation (Golden Rule) the rates are given by

$$W(\mathbf{k}_1, \mathbf{k}) = \frac{2\pi}{\hbar} |U_{\mathrm{imp},\mathbf{k}_1,\mathbf{k}}|^2 \delta(\varepsilon(\mathbf{k}_1) - \epsilon(\mathbf{k})) , \tag{9.43}$$

where the delta function is meaningful since we use W in a sum over $\mathbf{k}_1$. $U_{\mathrm{imp},\mathbf{k}_1,\mathbf{k}}$ is the matrix element of the impurity potential w.r.t. the Bloch states

$$U_{\mathrm{imp},\mathbf{k}_1,\mathbf{k}} = \frac{1}{V}\sum_j \int dV\, v(\mathbf{r}-\mathbf{c}_j)u^*_{\mathbf{k}_1}(\mathbf{r})u_{\mathbf{k}}(\mathbf{r})e^{i(\mathbf{k}-\mathbf{k}_1)\cdot\mathbf{r}} \qquad (9.44)$$

We assume all impurities are equivalent. Moreover we assume that they all have the same position within the primitive cell. That is the only random aspect is in which cell there is an impurity. Then $\mathbf{c}_j = \mathbf{R}_j + \delta\mathbf{c}$. Shifting by $\mathbf{R}_j$ in each term of the sum and using the periodicity of the functions u we obtain

$$\begin{aligned}
U_{\mathrm{imp},\mathbf{k}_1,\mathbf{k}} &= \frac{1}{V}\sum_j e^{i(\mathbf{k}-\mathbf{k}_1)\cdot\mathbf{R}_j}\int dV\, v(\mathbf{r}-\delta\mathbf{c})u^*_{\mathbf{k}_1}(\mathbf{r})u_{\mathbf{k}}(\mathbf{r})e^{i(\mathbf{k}-\mathbf{k}_1)\cdot\mathbf{r}} \\
&= \frac{1}{V}v_{\mathbf{k}_1,\mathbf{k}}\sum_j e^{i(\mathbf{k}-\mathbf{k}_1)\cdot\mathbf{R}_j} \qquad (9.45)
\end{aligned}$$

where $v_{\mathbf{k}_1,\mathbf{k}}$ is the matrix element of a single impurity potential.

This gives

$$|U_{\mathrm{imp},\mathbf{k}_1,\mathbf{k}}|^2 = \frac{1}{V^2}|v_{\mathbf{k}_1,\mathbf{k}}|^2 \sum_{j,l} e^{i(\mathbf{k}-\mathbf{k}_1)\cdot(\mathbf{R}_j-\mathbf{R}_l)} \ . \qquad (9.46)$$

This result will be put into the sum over $\mathbf{k}_1$ in the expression for the collision integral I. The locations $\mathbf{R}_j$ are random. Thus the exponents will average out. What remains are only diagonal terms. Thus we replace

$$|U_{\mathrm{imp},\mathbf{k}_1,\mathbf{k}}|^2 \to \frac{1}{V^2}|v_{\mathbf{k}_1,\mathbf{k}}|^2 N_{\mathrm{imp}} \ , \qquad (9.47)$$

where N_{imp} is the total number of impurities.

This gives for the collision integral

$$\begin{aligned}
I[f] &= \sum_{\vec{k}_1} W(\mathbf{k}_1,\mathbf{k})\,(f(\mathbf{k}_1)-f(\mathbf{k})) \\
&= \frac{2\pi}{\hbar}\frac{N_{\mathrm{imp}}}{V^2}\sum_{\vec{k}_1}|v_{\mathbf{k}_1,\mathbf{k}}|^2\,\delta(\varepsilon(\mathbf{k}_1)-\varepsilon(\mathbf{k}))\,(f(\mathbf{k}_1)-f(\mathbf{k})) \\
&= \frac{2\pi}{\hbar}n_{\mathrm{imp}}\int\frac{d^3k_1}{(2\pi)^3}v_{\mathbf{k}_1,\mathbf{k}}|^2\,\delta(\varepsilon(\mathbf{k}_1)-\varepsilon(\mathbf{k}))\,(f(\mathbf{k}_1)-f(\mathbf{k})) \ , (9.48)
\end{aligned}$$

where $n_{\mathrm{imp}} \equiv N_{\mathrm{imp}}/V$ is the density of impurities.

9.2.5 Transport relaxation time

As we have seen the correction to the distribution function due to application of the electric field was of the form $\delta f \sim \mathbf{E}\cdot\mathbf{v_k}$. In a parabolic band (isotropic spectrum) this would be $\delta f \sim \mathbf{E}\cdot\mathbf{k}$. So we make an ansatz

$$\delta f = a\,(k)\,\mathbf{E}\cdot\mathbf{e_k} \ , \qquad (9.49)$$

where $\mathbf{e_k} \equiv \mathbf{k}/|\mathbf{k}|$. For isotropic spectrum conservation of energy means $|\mathbf{k}| = |\mathbf{k}_1|$, the matrix element $v_{\mathbf{k}_1,\mathbf{k}}$ depends on the angle between $\mathbf{k}_1$ and $\mathbf{k}$ only, the surface S is a sphere. Then we obtain

$$
\begin{aligned}
I[\delta f] &= \frac{2\pi}{\hbar}\, n_{\text{imp}}\, \rho_F \int \frac{d\Omega_1}{4\pi} |v_{\mathbf{k}_1,\mathbf{k}}|^2 \; (\delta f(\mathbf{k}_1) - \delta f(\mathbf{k})) \\
&= \frac{2\pi}{\hbar}\, n_{\text{imp}} \rho_F a\,(k)\, E \int \frac{d\Omega_1}{4\pi} |v(\theta_{\mathbf{k}_1,\mathbf{k}})|^2\,(\cos\theta_{\mathbf{k},\mathbf{E}} - \cos\theta_{\mathbf{k}_1\mathbf{E}}) \quad (9.50)
\end{aligned}
$$

We choose direction $\mathbf{k}$ as z. Then the vector $\mathbf{k}_1$ is described in spherical coordinates by $\theta_{\mathbf{k}_1} \equiv \theta_{\mathbf{k},\mathbf{k}_1}$ and $\varphi_{\mathbf{k}_1}$. Analogously the vector $\mathbf{E}$ is described by $\theta_{\mathbf{E}} = \theta_{\mathbf{k},\mathbf{E}}$ and $\varphi_{\mathbf{E}}$. Then $d\Omega_1 = \sin\theta_{\mathbf{k}_1} d\theta_{\mathbf{k}_1} d\varphi_{\mathbf{k}_1}$.

From simple vector analysis we obtain

$$
\cos\theta_{\mathbf{E},\mathbf{k}_1} = \cos\theta_{\mathbf{E}} \cos\theta_{\mathbf{k}_1} + \sin\theta_{\mathbf{E}} \sin\theta_{\mathbf{k}_1} \cos(\varphi_{\mathbf{E}} - \varphi_{\mathbf{k}_1})\ . \tag{9.51}
$$

The integration then gives

$$
\begin{aligned}
I[\delta f] &= \frac{n_{\text{imp}}\rho_F}{2\hbar}\, a\,(k)\, E \int \sin\theta_{\mathbf{k}_1} d\theta_{\mathbf{k}_1} d\varphi_{\mathbf{k}_1} |v(\theta_{\mathbf{k}_1})|^2 \times \\
&\quad \times (\cos\theta_{\mathbf{E}} - \cos\theta_{\mathbf{E}} \cos\theta_{\mathbf{k}_1} - \sin\theta_{\mathbf{E}} \sin\theta_{\mathbf{k}_1} \cos(\varphi_{\mathbf{E}} - \varphi_{\mathbf{k}_1})\,) \\
&= \frac{\pi n_{\text{imp}}\rho_F}{\hbar} a\,(k)\, E \cos\theta_{\mathbf{g}} \int \sin\theta_{\mathbf{k}_1} d\theta_{\mathbf{k}_1} |v(\theta_{\mathbf{k}_1})|^2 (1 - \cos\theta_{\mathbf{k}_1}) \quad (9.52)
\end{aligned}
$$

Noting that $a\,(k)\, E \cos\theta_{\mathbf{g}} = a\,(k)\, \mathbf{E} \cdot \mathbf{e_k} = -\delta f$ we obtain

$$
I[\delta f] = -\frac{\delta f}{\tau_{\text{tr}}}\ , \tag{9.53}
$$

where

$$
\frac{1}{\tau_{\text{tr}}} = \frac{\pi n_{\text{imp}}\nu}{\hbar} \int \sin\theta d\theta\, |v(\theta)|^2\,(1 - \cos\theta) \tag{9.54}
$$

Note that our previous relaxation time approximation was based on total omission of the in-term. That is in the τ-approximation we had

$$
I[f] = \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k})\,(\delta f(\mathbf{k}_1) - \delta f(\mathbf{k})) \approx -\delta f(\mathbf{k}) \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k})\ .
$$

Thus

$$
\frac{1}{\tau} = \sum_{\mathbf{k}_1} W(\mathbf{k}_1, \mathbf{k}) = \frac{\pi n_{\text{imp}}\nu}{\hbar} \int d\theta\, |v(\theta)|^2 \sin\theta\ .
$$

The difference between τ_{tr} (transport time) and τ (momentum relaxation time) is the factor $(1 - \cos\theta)$ which emphasizes backscattering. If $|v(\theta)|^2 = const.$ we obtain $\tau_{\text{tr}} = \tau$.

9.2.6 H-theorem

Further insight into the underlying non-equilibrium dynamics can be obtained from analyzing the entropy density

$$H = k_B \int \frac{d^d x d^d k}{(2\pi)^d} \left[f_{\mathbf{k}}(\mathbf{r}) \ln f_{\mathbf{k}}(\mathbf{r}) + (1 - f_{\mathbf{k}}(\mathbf{r})) \ln (1 - f_{\mathbf{k}}(\mathbf{r})) \right]$$

It follows for the time dependence of H that

$$
\begin{aligned}
\frac{\partial H}{\partial t} &= k_B \int \frac{d^d x d^d k}{(2\pi)^d} \frac{\partial f_{\mathbf{k}}(\mathbf{r})}{\partial t} \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})} \\
&= -k_B \int \frac{d^d x d^d k}{(2\pi)^d} \left(\frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f + \frac{\partial \mathbf{r}}{\partial t} \cdot \nabla_{\mathbf{r}} f \right) \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})} \\
&\quad + k_B \int \frac{d^d x d^d k}{(2\pi)^d} I_{\mathbf{k}}[f] \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})}
\end{aligned}
$$

where we used that $f_{\mathbf{k}}(\mathbf{r})$ is determined from the Boltzmann equation.

Next we use that

$$\nabla_{\mathbf{k}} f_{\mathbf{k}}(\mathbf{r}) \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})} = \nabla_{\mathbf{k}} y_{\mathbf{k}}(\mathbf{r})$$

with

$$y_{\mathbf{k}}(\mathbf{r}) = \log(1 - f_{\mathbf{k}}(\mathbf{r})) + f_{\mathbf{k}}(\mathbf{r}) \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})}$$

which allows us to write the term with $\frac{\partial \mathbf{k}}{\partial t} \cdot \nabla_{\mathbf{k}} f$ as a surface integral. The same can be done for the term with $\frac{\partial \mathbf{r}}{\partial t} \cdot \nabla_{\mathbf{r}} f$. Thus, it follows

$$
\begin{aligned}
\frac{\partial H}{\partial t} &= k_B \int \frac{d^d x d^d k}{(2\pi)^d} I_{\mathbf{k}}[f] \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})} \\
&= k_B \int \frac{d^d x d^d k}{(2\pi)^d} W(\mathbf{k}', \mathbf{k}) f_{\mathbf{k}'}(\mathbf{r})[1 - f_{\mathbf{k}}(\mathbf{r})] - W(\mathbf{k}, \mathbf{k}') f_{\mathbf{k}}(\mathbf{r})[1 - f_{\mathbf{k}'}(\mathbf{r})] \ln \frac{f_{\mathbf{k}}(\mathbf{r})}{1 - f_{\mathbf{k}}(\mathbf{r})} \\
&= \frac{k_B}{2} \int \frac{d^d x d^d k}{(2\pi)^d} W(\mathbf{k}', \mathbf{k}) \left(f_{\mathbf{k}'}(\mathbf{r})[1 - f_{\mathbf{k}}(\mathbf{r})] - f_{\mathbf{k}}(\mathbf{r})[1 - f_{\mathbf{k}'}(\mathbf{r})] \right) \ln \frac{f_{\mathbf{k}}(\mathbf{r})(1 - f_{\mathbf{k}'}(\mathbf{r}))}{f_{\mathbf{k}'}(\mathbf{r})(1 - f_{\mathbf{k}}(\mathbf{r}))} \\
&= \frac{k_B}{2} \int \frac{d^d x d^d k}{(2\pi)^d} W(\mathbf{k}', \mathbf{k}) \left(f_{\mathbf{k}'}(\mathbf{r})[1 - f_{\mathbf{k}}(\mathbf{r})] - f_{\mathbf{k}}(\mathbf{r})[1 - f_{\mathbf{k}'}(\mathbf{r})] \right) \ln \frac{f_{\mathbf{k}}(\mathbf{r})(1 - f_{\mathbf{k}'}(\mathbf{r}))}{f_{\mathbf{k}'}(\mathbf{r})(1 - f_{\mathbf{k}}(\mathbf{r}))}
\end{aligned}
$$

It holds

$$(1 - x) \log x \leq 0$$

where the equal sign is at zero. Thus, it follows

$$\frac{\partial H}{\partial t} < 0.$$

Only for

$$\frac{f_{\mathbf{k}}\left(\mathbf{r}\right)\left(1 - f_{\mathbf{k}'}\left(\mathbf{r}\right)\right)}{f_{\mathbf{k}'}\left(\mathbf{r}\right)\left(1 - f_{\mathbf{k}}\left(\mathbf{r}\right)\right)} = 1$$

is H a constant. Thus

$$\log \frac{f_{\mathbf{k}}\left(\mathbf{r}\right)}{\left(1 - f_{\mathbf{k}}\left(\mathbf{r}\right)\right)} = const.$$

which is the equilibrium distribution function.

9.2.7   Local equilibrium, Chapman-Enskog Expansion

Instead of global equilibrium with given temperature T and chemical potential μ in the whole sample, consider a distribution function $f(\mathbf{r}, \mathbf{k})$ corresponding to space dependent $T(\mathbf{r})$ and $\mu(\mathbf{r})$:

$$f_0 = \frac{1}{\exp\left[\frac{\epsilon_k - \mu(\mathbf{r})}{k_{\mathrm{B}} T(\mathbf{r})}\right] + 1} \ . \tag{9.55}$$

This sate is called local equilibrium because also for this distribution function the collision integral vanishes: $I[f_0] = 0$. However this state is not static. Due to the kinematic terms in the Boltzmann equation (in particular $\mathbf{v_k} \cdot \nabla_{\mathbf{r}} f$) the state will change. Thus we consider the state $f = f_0 + \delta f$ and substitute it into the Boltzmann equation. This gives (we drop the magnetic field)

$$\frac{\partial \delta f}{\partial t} - \frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}}(f_0 + \delta f) + \mathbf{v_k} \cdot \nabla_{\mathbf{r}}(f_0 + \delta f) = I[\delta f] \ . \tag{9.56}$$

We collect all the δf terms in the r.h.s.:

$$-\frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} f_0 + \mathbf{v_k} \cdot \nabla_{\mathbf{r}} f_0 = I[\delta f] + \frac{\partial \delta f}{\partial t} + \frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} \delta f - \mathbf{v_k} \cdot \nabla_{\mathbf{r}} \delta f \ . \tag{9.57}$$

We obtain

$$\nabla_{\mathbf{r}} f_0 = -\frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \left(\nabla_{\mathbf{r}} \mu + \frac{(\varepsilon_{\mathbf{k}} - \mu)}{T} \nabla_{\mathbf{r}} T \right) \tag{9.58}$$

and

$$-\frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \mathbf{v_k} \left((\nabla_{\mathbf{r}} \mu + e\mathbf{E}) + \frac{\varepsilon_{\mathbf{k}} - \mu}{T} \nabla_{\mathbf{r}} T \right) = I[\delta f] + \frac{\partial \delta f}{\partial t} + \frac{e}{\hbar} \mathbf{E} \cdot \nabla_{\mathbf{k}} \delta f - \mathbf{v_k} \cdot \nabla_{\mathbf{r}} \delta f \ . \tag{9.59}$$

In the stationary state, relaxation time approximation, and neglecting the last two terms (they are small at small fields) we obtain

$$-\frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \mathbf{v_k} \left((\nabla_{\mathbf{r}} \mu + e\mathbf{E}) + \frac{\varepsilon_{\mathbf{k}} - \mu}{T} \nabla_{\mathbf{r}} T \right) = -\frac{\delta f}{\tau_{\mathrm{tr}}} \ , \tag{9.60}$$

which yields:

$$\delta f = \tau_{\mathrm{tr}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} \mathbf{v_k} \left((\nabla_{\mathbf{r}} \mu + e\mathbf{E}) + \frac{\varepsilon_{\mathbf{k}} - \mu}{T} \nabla_{\mathbf{r}} T \right) \ . \tag{9.61}$$

Thus we see that there are two forces getting the system out of equilibrium: the electrochemical field: $\mathbf{E}_{\text{el.ch.}} \equiv \mathbf{E} + (1/e)\nabla\mu$ and the gradient of the temperature ∇T. More precisely one introduces the electrochemical potential $\phi_{\text{el.ch.}}$ such that $\mathbf{E}_{\text{el.ch.}} = \mathbf{E} + (1/e)\nabla\mu = -\nabla\phi_{\text{el.ch.}} = -\nabla\phi + (1/e)\nabla\mu$. Thus $\phi_{\text{el.ch.}} = \phi - (1/e)\mu$.

On top of the electric current

$$\mathbf{j}_E(\mathbf{r},t) = -\frac{e}{V}\sum_{\mathbf{k},\sigma}\mathbf{v_k}\delta f(\mathbf{k},\mathbf{r},t) \tag{9.62}$$

we define the heat current

$$\mathbf{j}_Q(\mathbf{r},t) = \frac{1}{V}\sum_{\mathbf{k},\sigma}(\varepsilon_\mathbf{k}-\mu)\mathbf{v_k}\delta f(\mathbf{k},\mathbf{r},t) \tag{9.63}$$

This expression for the heat current follows from the definition of heat $dQ = dU - \mu dN$.

This gives

$$\begin{pmatrix} \mathbf{j}_E \\ \mathbf{j}_Q \end{pmatrix} = \begin{pmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{pmatrix} \begin{pmatrix} \mathbf{E}_{\text{el.ch.}} \\ \nabla T/T \end{pmatrix} \tag{9.64}$$

Before we determine these coefficients, we give a brief interpretation of the various coefficients. In the absence of ∇T, holds

$$\begin{aligned} \mathbf{j}_E &= K_{11}\mathbf{E}_{\text{el.ch.}} \\ \mathbf{j}_Q &= K_{21}\mathbf{E}_{\text{el.ch.}}. \end{aligned}$$

The first term is the usual conductivity, i.e.

$$\sigma = K_{11}$$

, while the second term describes a heat current in case of an applied electric field. The heat current that results as consequence of an electric current is called Peltier effect

$$\mathbf{j}_Q = \beta_E\mathbf{j}_E.$$

where

$$\beta_E = K_{21}/K_{11}.$$

is the Peltier coefficient.

In the absence of $\mathbf{E}_{\text{el.ch.}}$ holds

$$\begin{aligned} \mathbf{j}_E &= \frac{K_{12}}{T}\nabla T \\ \mathbf{j}_Q &= \frac{K_{22}}{T}\nabla T \end{aligned}$$

Thus, with $\mathbf{j}_Q = -\kappa\nabla T$ follows

$$\kappa = -K_{22}/T$$

for the thermal conductivity, while K_{12} determines the electric current that is caused by a temperature gradient. Keep in mind that the relationship between the two currents is now $\mathbf{j}_Q = \beta_T \mathbf{j}_E$ with .

$$\beta_T = K_{22}/K_{12} = -T\kappa/K_{12}$$

Finally a frequent experiment is to apply a thermal gradient and allow for no current flow. Then, due to

$$0 = K_{11}\mathbf{E}_{\text{el.ch.}} + \frac{K_{12}}{T}\nabla T$$

follows that a voltage is being induced with associated electric field

$$\mathbf{E}_{\text{el.ch.}} = S\nabla T$$

where

$$S = \frac{K_{12}}{TK_{11}}$$

is the Seebeck coefficient (often refereed to as thermopower).

For the electrical current density we obtain

$$
\begin{aligned}
\mathbf{j}_E &= -\frac{e}{V}\sum_{\mathbf{k},\sigma}\mathbf{v_k}\delta f(\mathbf{k}) \\
&= -\frac{e}{V}\sum_{\mathbf{k},\sigma}\tau_{\text{tr}}\frac{\partial f_0}{\partial\varepsilon_\mathbf{k}}\left[\mathbf{v_k}\cdot\left(e\mathbf{E}_{\text{el.ch.}} + \frac{\varepsilon_\mathbf{k}-\mu}{T}\nabla T\right)\right]\mathbf{v_k} \ . \quad (9.65)
\end{aligned}
$$

Thus for K_{11} we obtain

$$K_{11\alpha\beta} = -\frac{e^2}{V}\sum_{\mathbf{k},\sigma}\tau_{\text{tr}}\frac{\partial f_0}{\partial\varepsilon_\mathbf{k}}v_{\mathbf{k},\alpha}v_{\mathbf{k},\beta} \ . \quad (9.66)$$

For K_{12} this gives

$$K_{12\alpha\beta} = -\frac{e}{V}\sum_{\mathbf{k},\sigma}\tau_{\text{tr}}\frac{\partial f_0}{\partial\varepsilon_\mathbf{k}}(\varepsilon_\mathbf{k}-\mu)v_{\mathbf{k},\alpha}v_{\mathbf{k},\beta} \ . \quad (9.67)$$

For the heat current density we obtain

$$
\begin{aligned}
\mathbf{j}_Q &= \frac{1}{V}\sum_{\vec{k},\sigma}(\varepsilon_\mathbf{k}-\mu)\mathbf{v_k}\delta f(\mathbf{k}) \\
&= \frac{1}{V}\sum_{\vec{k},\sigma}\tau_{\text{tr}}(\varepsilon_\mathbf{k}-\mu)\frac{\partial f_0}{\partial\varepsilon_\mathbf{k}}\left[\mathbf{v_k}\cdot\left(e\mathbf{E}_{\text{el.ch.}} + \frac{\varepsilon_\mathbf{k}-\mu}{T}\nabla T\right)\right]\mathbf{v_k} \ . \quad (9.68)
\end{aligned}
$$

Thus for K_{21} we obtain

$$K_{21\alpha\beta} = \frac{e}{V}\sum_{\mathbf{k},\sigma}\tau_{\text{tr}}\frac{\partial f_0}{\partial\varepsilon_\mathbf{k}}(\varepsilon_\mathbf{k}-\mu)v_{\mathbf{k},\alpha}v_{\mathbf{k},\beta} \ . \quad (9.69)$$

For K_{22} this gives

$$K_{22\alpha\beta} = \frac{1}{V} \sum_{\mathbf{k},\sigma} \tau_{\text{tr}} \frac{\partial f_0}{\partial \varepsilon_{\mathbf{k}}} (\varepsilon_{\mathbf{k}} - \mu)^2 v_{\mathbf{k},\alpha} v_{\mathbf{k},\beta} \ . \tag{9.70}$$

K_{11} is just the conductivity calculated earlier. $K_{12} = -K_{21}$. This is one of the consequences of Onsager relations. $K_{12} \neq 0$ only if the density of states is asymmetric around μ (no particle-hole symmetry). Finally, for K_{22} we use

$$\frac{\partial f_0}{\partial \epsilon} \approx -\delta(\epsilon - \mu) - \frac{\pi^2}{6} (k_{\text{B}}T)^2 \delta''(\epsilon - \mu) \ , \tag{9.71}$$

This gives

$$
\begin{aligned}
K_{22\alpha\beta} &= \frac{1}{V} \sum_{\vec{k},\sigma} \tau_{\text{tr}} \frac{\partial f_0}{\partial \epsilon} (\epsilon_k - \mu)^2 v_{k,\alpha} v_{k,\beta} \\
&= \tau_{\text{tr}} \int \rho(\varepsilon) \, d\varepsilon \frac{d\Omega}{4\pi} \frac{\partial f_0}{\partial \varepsilon} (\varepsilon - \mu)^2 v_\alpha v_\beta \\
&= -\tau_{\text{tr}} \frac{\pi^2}{3} (k_{\text{B}}T)^2 \rho_F \int \frac{d\Omega}{4\pi} v_\alpha v_\beta = -\frac{\pi^2}{9} (k_{\text{B}}T)^2 \rho_F v_F^2 \tau_{\text{tr}} \delta_{\alpha,\beta} \tag{9.72}
\end{aligned}
$$

Thus, for thermal conductivity κ defined via $\mathbf{j}_Q = -\kappa \nabla T$ we obtain

$$\kappa = -\frac{K_{22}}{k_{\text{B}}T} = \frac{\pi^2}{9} k_{\text{B}}^2 T \rho_F v_F^2 \tau_{\text{tr}} \tag{9.73}$$

Comparing with the electrical conductivity

$$\sigma = \frac{1}{3} \rho_F v_F^2 \tau_{\text{tr}} \tag{9.74}$$

We obtain the Wiedemann-Franz law:

$$\frac{\kappa}{\sigma} = \frac{k_{\text{B}}^2 T}{e^2} \frac{\pi^2}{3} \ . \tag{9.75}$$