Statistical Physics

Hendrik van Hees[1]
Goethe University Frankfurt

July 22, 2019

[1]e-mail: hees@itp.uni-frankfurt.de

Contents

Contents

Preface

In these notes I try to give a self contained treatment of statistical physics. The reader should be familiar with quantum theory and the formalism of the 2nd quantization. It was my goal to show that statistical physics is not a special topic of its own but a completion of the concept of physics at all.

The main idea of application of statistical concepts in physics in the 19th century was to give an explanation of the world we live in from properties of little particles. Although this idea is very old in history of science (reaching back to Democritus), the main problem for "the statistical physicists" was to prove the existence of the atoms at all! There was no direct proof for the existence of discrete constituents of the macroscopic world which looks so smooth.

Although the program of statistical physics was very successful in deriving the macroscopic properties, especially the explanation of phenomenological thermo dynamics, there were many physicists in doubt if the atoms were real. The solution of this problem was given by Einstein, using statistical methods to explain the Brownian motion which he treated as the fluctuation of "mesoscopic" particles due to collisions with the surrounding molecules of the fluid the particles are suspended in.

Of course in our days the meaning of statistical physics has changed completely because the existence of atoms and even smaller constituents of matter is not in doubt. The most important meaning of statistical physics is to explain the classical nature of the macroscopic world, we are so familiar with from every day life, on grounds of the underlying quantum world.

This point of view dictates the whole conception of these notes. Thus we start with some mathematical foundations of statistics which is not meant to be rigorous but to give some techniques for the practical calculations. It is important to see that probability theory, which can be summarized by Kolmogorov's axioms, gives no advice how to describe a certain situation with help of probability distributions but only the mathematical general properties of probabilities.

Looking at the problem from the physical side we find as the basic theory quantum theory for many particle problems which is most conveniently formulated as a *quantum field theory*. Quantum theory contains also elements of statistics because of the Born interpretation for the "wave function"[1]. But these statistics have nothing to do with the statistics used to describe the macroscopic world as constituted by very many little particles but because of the principal impossibility to give all measurable observables of a system a certain value (*indeterminism*). Maximally the properties described by a complete compatible set of observables can be used to describe the system, then known to be in a certain *pure state* given by a ray in Hilbert space.

Statistical physics is used to describe situations where it is not possible to determine the state ket because of the complexity of the system under consideration. This is especially the case for macroscopic systems which contain about 10^{24} particles. Then we shall use probability distributions for the state kets. It is not difficult to formulate the general properties of probability distributions in quantum physics. Especially quantum theory gives the dynamics of the probability distribution from first principles of quantum dynamics. This will be explained in the beginning of the 2nd chapter.

---

[1]We will use the operator formalism in the abstract Hilbert space, rather than the Schrödinger's wave function approach in order to pronounce the particle point of view which seems to be more adequate for the meaning of statistical physics

Again we have the problem to find this probability distribution. Now it is clear that there is no help from quantum physics to find this distribution because the only thing we know about our system are the initial values of some observable macroscopic quantities but of course not the sharp values of a complete set of observables.

The solution of the problem how to find the probability distribution must thus be a new principle not contained in probability theory or quantum physics. The idea of this principle is to use the probability distribution which describes nothing else than the knowledge we have about the system. The idea is thus to measure the *amount of information* which is hidden in the probability distribution. This is the aim of the *information theory* founded by Shannon for applying it in signal theory of electric enginering. Our feeling is that we know more about the system as the distribution becomes more concentrated around a certain value of the stochastic observable under consideration. It is this simple idea which is formulated a little bit more exactly which leads to the entropy well known in thermodynamics and statistics as the measure of the lack of information against "complete" information of the system. Now it is clear what to do to find the probability distribution which should be used: Given the information of the system at a initial time we have to use the probability distribution of all those respecting the given information which gives the *maximal entropy*. Since the entropy is the measure of lacking information we can be sure to describe the given (incomplete) information about the system by the very probability distribution which contains the least prejudice. The maximal entropy principle of phenomenological thermodynamics (valid only for the equilibrium case) is thus founded on purely statistical grounds, namely the Jaynes principle of least prejudice. The same time this leads to the more generalized meaning of the entropy principle *beyond equilibrium situations*.

These notes are organized as follows:

The first chapter gives a self-contained introduction in mathematical statistics and information theory. The second chapter starts with a short introduction to basic quantum physics and uses the statistical concepts of the first chapter to describe quantum statistics. The third chapter gives some general ideas about thermodynamic equilibrium and a statistical explanation for phenomenological thermodynamics, i.e., of course we shall apply this concepts at first to ideal gas problems.

Chapter 1

Mathematical Foundations of Statistics

1.1 Events and Probability

In this section we shall give an introduction to the mathematical formalism behind probability theory. The main goal of applied mathematical statistics is the analysis of experiments with unpredictable results. This is a rather vague idea and has to be founded axiomatically to build a nice mathematical theory. Nevertheless these axioms are oriented on the practical task.

An often used example is playing dice and we shall use this example to give an idea what is axiomatic statistics about. The first thing we have to do is to describe precisely the results of the stochastic experiment. All possible results of an experiment build *the set of results*. In our example of the dice a result of an experiment may be described by the number showing up. Thus the set of results is given by $\Omega = \{1, 2, 3, 4, 5, 6\}$.

But often there are situations where it is not important which number shows up after throwing the dice but other *events*, for instance if the number is 6 or not. An event thus might be described as partial sets of Ω with the meaning that a certain experiment belongs to the event under consideration if the outcome is contained in the set describing the event. For instance the event $E = \{6\} \subseteq \Omega$ is realized if the dice shows the number 6.

Another example might be $E = \{2, 4, 6\}$ or "even number". An experiment realizes this event if the number coming out is even. The set of all events is thus the potency set of Ω, i.e., the set of all subsets of Ω. *It is important to realize that the elements of this set are sets, namely the subsets of Ω.*

Now we can also investigate what is the meaning of elementary operations with the events. Given two events E_1 and E_2 we can ask for the event E_1 and E_2 being present simultaneously. But we have defined an event as a subset of Ω containing all outcomes of an experiment belonging to the event. Thus the event that both E_1 and E_2 happened is given by the cut of these events $E_{1 \text{ and } 2} = E_1 \cap E_2$. If E_1 and E_2 are contradicting (for instance in our example with the dice take $E_1 = \{1, 3, 5\}$ and $E_2 = \{2, 4, 6\}$), which means that there are no common results contained in both E_1 and E_2, we have $E_1 \cap E_2 = \emptyset$. This gives us the interpretation of the empty set contained in the set of events: it's the impossible event.

The same line of reasoning shows that the event "E_1 or E_2" happened is given by $E_1 \cup E_2$.

Given an event E it is also sensible to ask if "E did not happen", which is again an event, which is clearly given by the complement of E with respect to Ω: $\bar{E} = \Omega \setminus E$. Take again the empty set which had the meaning of an event which can never happen doing an experiment. Its negation is the event which is true for any experiment one can do. This is given by $\Omega = \Omega \setminus \emptyset$.

These considerations show that all elementary set operations which can be done on the events again give an event with a simple relation to the events involved. Now for the formal theory of statistics it is important that the three operations $\cup$, $\cap$ and $\Omega \setminus \cdot$ applied to the potency set of Ω are closed, i.e., all these operations are maps $\mathscr{P}\Omega \to \mathscr{P}\Omega$.

Now it is simple to realize that the structure $(\mathscr{P}\Omega, \cap, \cup, \bar{\cdot})$ builds an *Boolean algebra*:

$$A \cup B = B \cup A, \ A \cap B = B \cap A, \tag{1.1.1}$$

$$A \cap (B \cup C) = (A \cap B) \cup (B \cap C), \ A \cup (B \cap C) = (A \cup B) \cap (B \cup C), \tag{1.1.2}$$

$$A \cup \emptyset = A, \ A \cap \Omega = A, \tag{1.1.3}$$

$$A \cup \bar{A} = \Omega, \ A \cap \bar{A} = \emptyset \tag{1.1.4}$$

where $A, B, C \in \mathscr{P}(\Omega)$ and $\bar{A} = \Omega \setminus A$.

We shall not prove all the basic rules for calculations using this algebra.

1.2 Probabilities

Now we like to measure the *probability* of an event. In this section we shall give only the basic properties of such a measure, not the rules how to obtain it in practice. We also don't give a discussion of the principal problems of probability theory present in the philosophical literature. All these problems are solved by looking from the axiomatic point of view. Thus we give immediately *Kolmogorov's axioms*:

(P1) The probability is a function $P : \mathscr{P}\Omega \to \mathbb{R}_{\geq 0}$,

(P2) $P(\Omega) = 1$,

(P3) For any events E_1 and E_2 with $E_1 \cap E_2 = \emptyset$ we have $P(E_1 \cup E_2) = P(E_1) + P(E_2)$.

The idea behind this axioms is that the probability measures how often an experiment will give a result belonging to a given event.

The most simple idea of how to measure it comes again from playing with a perfect dice where all numbers are equally probable, i.e., $P(\{i\}) = 1/6$ for $i = 1 \ldots 6$. Now Kolmogorov's axioms allow to calculate the probability for any event from this simple ansatz.[1] Then P3 is for this case nothing else than the idea that the probability is given as the fraction of the number of elements contained in the event and the number of results which are possible at all.

As we shall show later the probabilities can be measured in practice by performing very much independent experiments. If $N(E)$ are the number the event E occurs and N is the number of experiments performed then the probability is given by the limit $N(E)/N$ for $N \to \infty$.

Now we prove some simple rules about probabilities

Lemma 1. *(1) For $E \in \mathscr{P}\Omega$ we have $P(\bar{E}) = 1 - P(E)$. For all $E \in \mathscr{P}\Omega$ we have $P(E) \leq 1$*

(2) $P(\emptyset) = 0$.

(3) For $A, B \in \mathscr{P}\Omega$: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

(4) Sylvester's formula: Let $A_i \in \mathscr{P}\Omega$ with $i = 1, \ldots, n$. Then we find

$$P\left(\bigcup_{i=1}^{n} A_i\right) = \sum_{i=1}^{n} P(A_i) - \sum_{i,j=1; i<j}^{n} P(A_i \cap A_j) + \cdots + (-1)^{n-1} P(A_1 \cap A_2 \cap \ldots \cap A_n). \tag{1.2.1}$$

---

[1]Keep in mind that the choice of the probabilities for the elementary events (which are defined as the subsets of Ω with 1 element) are only a guess. There is no other idea behind this than what is meant by a "ideal dice". If the material the dice is made of is not homogeneous this guess might be wrong.

Proof. (1) is simply shown by using the fact that $E \cup \bar{E} = \Omega$ and $E \cap \bar{E} = \emptyset$. Using P2 and P3 we find

$$1 = P(\Omega) = P(E \cup \bar{E}) = P(E) + P(\bar{E}). \tag{1.2.2}$$

Since $0 \le P(\bar{E}) = 1 - P(E)$ we conclude $P(E) \le 1$.

(2) From (1) we know

$$P(\emptyset) = 1 - P(\bar{\emptyset}) = 1 - P(\Omega) = 0 \tag{1.2.3}$$

where we have used P2 in the last step.

(3) We have $A \cup B = A \cup (B \cap \bar{A})$ and $A \cap (B \cap \bar{A}) = \emptyset$. Thus we can use P3:

$$P(A \cup B) = P(A) + P(B \cap \bar{A}). \tag{1.2.4}$$

On the other hand we have $B = (A \cap B) \cup (\bar{A} \cap B)$ and $(A \cap B) \cap (\bar{A} \cap B) = \emptyset$ so that P3 again applies:

$$P(B) = P(A \cap B) + P(\bar{A} \cap B). \tag{1.2.5}$$

With help of this we can eliminate $P(\bar{A} \cap B)$ from (1.2.4) which gives the claim.

(4) is a simple proof by induction making use of

$$\bigcup_{i=1}^{n} A_i = \bigcup_{i=1}^{n-1} A_i \cup A_n \tag{1.2.6}$$

The rest is done by straightforwardly writing out all the sums which is a nice exercise for the reader!   Q.E.D.

Now we define conditional probabilities as follows: The probability for the event $B \in \mathscr{P}\Omega$ under the condition that $A \in \mathscr{P}\Omega$ is true is defined for $P(A) \ne 0$ as

$$P_A(B) = \frac{P(A \cap B)}{P(A)}. \tag{1.2.7}$$

Lemma 2. *The $P_A : \mathscr{P}\Omega \to \mathbb{R}_{\ge 0}$ fulfills Kolmogorov's axioms (P1)-(P3).*

Proof. (P1) We have to show that the range of P_A is really $[0,1]$. Clearly we have $P_A(E) \ge 0$ because $P(A \cap B) \ge 0$ and $P(A) > 0$. That it is ≤ 1 follows from (P2) and (P3) because each event can be built as disjunct union of elementary events.

(P2) $P_A(\Omega) = P(A \cap \Omega)/P(A) = P(A)/P(A) = 1$.

(P3) Let $B \cap C = \emptyset$. Then

$$P_A(B \cup C) = \frac{P[A \cap (B \cup C)]}{P(A)} = \frac{P[(A \cap B) \cup (A \cap C)]}{P(A)} \tag{1.2.8}$$

Now $(A \cap B) \cap (A \cap C) = A \cap (B \cap C) = A \cap \emptyset = \emptyset$ and using (P3) for the probabilities P we get the claim.   Q.E.D.

Definition 1. *Two events $A, B \in \mathscr{P}\Omega$ are called stochastically independent if and only if*

$$P(A \cap B) = P(A)P(B) \tag{1.2.9}$$

It should be kept in mind that stochastic independence is defined on grounds of probabilities and are not a property of the events as sets.

1.3 Probability variables

Let $\xi : \mathscr{P}\Omega \to \mathbb{R}$ be a function. Then ξ is called a *probability variable*. A simple example in the case of playing is the number showing up.

Having defined probabilities P on $\mathscr{P}\Omega$ we define probabilities for the probability variable as follows. We define the *probability function* for the probability variable ξ as

$$P_\xi : D_\xi \to \mathbb{R} : P_\xi(x) = P(\{r_i | \xi(r_i) \le x\}). \tag{1.3.1}$$

Further we define the probability distribution as the weak derivative of P_ξ with respect to x. The weak derivative is meant to be taken in the sense of the theory of distributions.

Now an example is at place. In the case of the example with the dice we have defined $\xi(\{k\}) = k$ for $k = 1 \to k$. Taking the probabilities $P(\{k\}) = 1/6$ we have

$$P_\xi(x) = \frac{1}{6} \sum_{i=1}^{6} \Theta(x - i) \tag{1.3.2}$$

where we have defined Heaviside's unit step function as

$$\Theta : \mathbb{R} \to \mathbb{R} : \Theta(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } x \ge 0. \end{cases} \tag{1.3.3}$$

Now we show that the weak derivative of the Θ-function (which has to be interpreted as distribution defined on the $C_0^\infty(\mathbb{R})$-functions[2] now) is Dirac's δ distribution. To this end remember the definition of the weak derivative Θ' as the distribution fulfilling:

$$\forall f \in C_0^\infty(\mathbb{R}) : \int dx \Theta'(x - y) f(x) = -\int dx \Theta(x - y) f'(x). \tag{1.3.4}$$

Now clearly we have

$$-\int dx \Theta(x) f'(x) = -\int_y^\infty dx f(x) = f(y) = \int dx \delta(x - y) f(x). \tag{1.3.5}$$

Thus for our example we find

$$W_\xi(x) := P_\xi'(x) = \frac{1}{6} \sum_{i=1}^{6} \delta(x - i). \tag{1.3.6}$$

Using the theory of distributions has the advantage that we can summarize the case of continues range of probability variables and discrete ones.

Now we can give some important definitions

Definition 2. *Given a function $f : \mathbb{R} \to \mathbb{R}$, its expectation value with respect to the probability variable ξ is given by*

$$\langle f(\xi) \rangle = \int dx f(x) W_\xi(x) \tag{1.3.7}$$

provided the defining integral exists.

---

[2] Remember that $C_0^\infty(\mathbb{R})$ is the space of all functions $f : \mathbb{R} \to \mathbb{R}$ with compact support, where the support is defined as $\{x | f(x) \ne 0\}$.

Especially we can write

$$P_\xi(x) = \langle \Theta(x - \xi) \rangle \text{ and } W_\xi(x) = \langle \delta(x - \xi) \rangle. \tag{1.3.8}$$

We have also

$$\langle 1 \rangle = \int dx \, W_\xi(x) = \lim_{x \to \infty} P_\xi(x) - \lim_{x \to -\infty} P_\xi(x) = 1. \tag{1.3.9}$$

Now let $g : M \to \mathbb{R}$, where M is the range of the probability variable ξ. g might be differentiable in the sense of distributions. For any $y \in \mathbb{R}$ the function $x \mapsto g(x) - y$ might have only a finite number of zero points, which are all simple.

Then we want to find the probability distribution for $\eta = g \circ \xi$.

Using (1.3.8) we write

$$W_\eta(y) = \langle \delta[y - g(\xi)] \rangle = \int dx \, \delta[y - g(x)]. \tag{1.3.10}$$

With the definition $\{g_n^{-1}(y)\}_{n \in I(y)}$ for the set of all points x with $g(x) = y$ we obtain immediately

$$W_\eta(y) = \sum_{n \in I(y)} \frac{W_\xi[g_n^{-1}(y)]}{|g'[g_n^{-1}(y)]|}. \tag{1.3.11}$$

Definition 3. *The characteristic function of a probability distribution W_ξ is defined by its Fourier transform:*

$$C_\xi(u) = \int dx \, \exp(iux) W_\xi(x) = \langle \exp(iu\xi) \rangle. \tag{1.3.12}$$

From this definition we read off that

$$M_n := \langle \xi^n \rangle = \frac{1}{i^n} \left(\frac{d}{du} \right)^n C_\xi(u)|_{u=0}, \tag{1.3.13}$$

i.e. we can calculate the expectation values of all potencies of ξ by derivation with respect to u. M_n is also called the nth momentum of the probability distribution W_ξ. With Taylor's theorem we can express the characteristic functions in terms of the momenta:

$$C_\xi(u) = 1 + \sum_{n=1}^{\infty} (iu)^n \frac{M_n}{n!}. \tag{1.3.14}$$

Thus if we know all momenta of the distribution we can reconstruct the characteristic function and so the distribution.

Definition 4. *The cumulative distribution function is defined as the logarithm of the characteristic function:*

$$K_\xi(u) = \ln[C_\xi(u)] = \ln\left[1 + \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} M_n \right] = \sum_{n=1}^{\infty} \frac{(iu)^n}{n!} K_n. \tag{1.3.15}$$

The coefficients K_n are called the cumulant of the distribution.

Taylor expansion of the logarithm shows that one may express the cumulants with help of the momenta. Thus knowledge of the cumulants also allows the reconstruction of the distribution.

Here we are content with the first two cumulants:

$$K_1 = M_1, \, K_2 = M_2 - M_1^2. \tag{1.3.16}$$

Thus the first cumulant is the first momentum, i.e. the expectation value of ξ. The 2nd cumulant is also called variance and is a measure for the width of the distribution. We have

$$K_2 = \left\langle (\xi - \langle \xi \rangle)^2 \right\rangle = \left\langle \xi^2 \right\rangle - \langle \xi \rangle^2. \tag{1.3.17}$$

Let us look on two simple examples of distributions.

The first one is defined by $0 = K_2 = K_3 = \ldots$ and $K_1 = M_1$ arbitrary real. Thus

$$C_\xi(u) = \exp(iuK_1) \Rightarrow W_\xi(x) = \frac{1}{2\pi} \int du \, C_\xi(u) \exp(-iux) = \delta(x - K_1). \tag{1.3.18}$$

This means that for vanishing 2nd and higher cumulants the stochastic experiment is determined because comparison with (1.3.2) shows that for sure the value of ξ is K_1.

The next example is defined by $0 = K_3 = K_4 = \ldots$. Then the distribution is given by

$$C_\xi(u) = \exp\left(iuK_1 - K_2 \frac{u^2}{2}\right) \Rightarrow W_\xi(x) = \frac{1}{\sqrt{2\pi K_2}} \exp\left[-\frac{(x - K_1)^2}{2K_2}\right] \tag{1.3.19}$$

which shows that K_2 is a measure for the width of the distribution while the higher cumulants are a measure of deviation from a Gaussian distribution (1.3.19).

1.4   Vector valued probability variables

In statistical physics the most realistic problems are described by more than one probability variable. In this section we shall treat the relevant mathematical foundations for such problems.

Clearly in principle there is nothing new. We might think about a stochastic experiment and define a function $\xi : \mathscr{P}\Omega \to \mathbb{R}^n$ and a probability distribution as in the case of one probability variable:

$$W_\xi : \mathbb{R}^n \to \mathbb{R} : W_\xi(x) = \left\langle \delta^{(r)}(\xi - x) \right\rangle \tag{1.4.1}$$

and the expectation value for a function f defined on the range of ξ:

$$\langle f(\xi) \rangle = \int_{\mathbb{R}^r} d^r x \, f(x) W_\xi(x). \tag{1.4.2}$$

The distribution for a subspace of ξ may be calculated as follows. Provided this subspace lies in the vector space spanned by the first i basis vectors then the distribution for the variables $\xi_1, \ldots, \xi_i$ is given by

$$W_i(x_1, \ldots, x_i) = \int_{\mathbb{R}^{r-i}} dx_{i+1} \cdots dx_r \, W_r(x_1, \ldots, x_r). \tag{1.4.3}$$

To understand this we remember

$$\delta^{(r)}(x - \xi) = \prod_{k=1}^{r} \delta(x_k - \xi_k) \tag{1.4.4}$$

and thus

$$W_i(x_1, \ldots, x_i) = \left\langle \delta^{(i)}(x_1 - \xi_1, \ldots, x_i - \xi_i) \right\rangle =$$
$$= \int_{\mathbb{R}^{r-i}} dx_{i+1} \cdots dx_r \left\langle \delta^{(r)}(x_1 - \xi_1, \ldots, x_r - \xi_r) \right\rangle \tag{1.4.5}$$

and together with (1.4.1) this is (1.4.3).

Again the Fourier transform of the distribution is the *characteristic* function:

$$C_r(u) = \langle \exp(iu\xi) \rangle = \int_{\mathbb{R}^r} \exp(iux) W_r(x). \tag{1.4.6}$$

The momenta are again the derivatives with respect to the probability variables

$$M_{n_1 \ldots n_r} = \left\langle \prod_{j=1}^{r} \xi_j^{n_j} \right\rangle = \left(\frac{1}{i} \right)^{\sum_{j=1}^{r} n_j} \prod_{j=1}^{r} \left(\frac{\partial}{\partial u_j} \right)^{n_j} C_r(u) \Bigg|_{u=0} \tag{1.4.7}$$

and we can reconstruct the characteristic function out of the momenta with help of Taylor expansion

$$C_r(u) = \sum_{n_1 \ldots n_r = 1}^{\infty} M_{n_1 \ldots n_r} \prod_{j=1}^{r} \frac{(iu_j)^{n_j}}{n_j!}. \tag{1.4.8}$$

Transforming back we may reconstruct the probability distribution:

$$W_r(x) = \left(\frac{1}{2\pi} \right)^r \int_{\mathbb{R}^r} d^r u \, \exp(-iux) C_r(u). \tag{1.4.9}$$

The cumulants are given by

$$K_{n_1 \ldots n_r} = \left(\frac{1}{i} \right)^{\sum_{j=1}^{r} n_j} \sum_{j=1}^{r} \prod_{j=1}^{r} \left(\frac{\partial}{\partial u_j} \right)^{n_j} \ln[C_r(u)] \Bigg|_{u=0}. \tag{1.4.10}$$

The characteristic function is given in terms of the cumulants by

$$C_r(u) = \exp \left[\sum_{n_1 \ldots n_r \in \mathbb{N}} K_{n_1 \ldots n_r} \prod_{j=1}^{r} \frac{(iu_j)^{n_j}}{n_j!} \right]. \tag{1.4.11}$$

The conditional probability distribution that $\xi_1 = x_1$ provided that the other $\xi_j = x_j$ ($j = 2, \ldots, r$) is given with help of (1.4.3) by:

$$P_{x_1 | x_2, \ldots, x_r} = \frac{W_r(x_1, \ldots, x_r)}{W_{r-1}(x_2, \ldots, x_r)} = \frac{W_r(x_1, \ldots, x_r)}{\int_{\mathbb{R}} dx_1 W_r(x_1, \ldots, x_r)}. \tag{1.4.12}$$

By definition the probability variables ξ_1 and ξ_2 are stochastically independent if the probability distribution factorizes like

$$W_2(x_1, x_2) = P(x_1 | x_2) \int_{\mathbb{R}} dx_1 W_2(x_1, x_2) =: W_1^{(1)}(x_1) W_2^{(2)}(x_2). \tag{1.4.13}$$

Thus we define probability variables ξ_1 and ξ_2 as *stochastically independent* if and only if the probability distribution is a product of two distributions for each of the probability variables alone.

The other extreme case of total determination of ξ_1 by measuring ξ_2, which means $\xi_1 = f(\xi_2)$, is given by

$$P(x_1 | x_2) = \delta[x_1 - f(\xi_2)]. \tag{1.4.14}$$

Then from (1.4.12) we find

$$W_2(x_1, x_2) = \delta[x_1 - f(x_2)] W_1(x_2). \tag{1.4.15}$$

There are different measures for stochastic dependence or correlations of variables. One of them is the *cross correlation*

$$\varkappa(\xi_1, \xi_2) = \langle \xi_1 \xi_2 \rangle - \langle \xi_1 \rangle \langle \xi_2 \rangle. \tag{1.4.16}$$

With (1.4.13) one sees immediately that for stochastically independent probability variables the cross correlation vanishes. But it is in general *not* true that one can conclude from the vanishing of the cross correlation that the probability variables are stochastically independent!

Another correlation measure is the *relative cross correlation*

$$R(\xi_1, \xi_2) = \frac{\varkappa(\xi_1, \xi_2)}{\sigma(\xi_1)\sigma(\xi_2)} \text{ with } \sigma(\xi_j) = \sqrt{\langle \xi_j^2 \rangle - \langle \xi_j \rangle^2}, \; j = 1, 2. \tag{1.4.17}$$

To generalize these ideas to more than two probability variables we observe that (1.4.16) is the 11-cumulant according to (1.4.11) (prove this as an exercise by calculating K_{11} with help of the definition (1.4.11)). Thus we define the cross correlation for r probability variables:

$$\varkappa(\xi_1, \ldots, \xi_r) = K_{1\ldots2} = \prod_{j=1}^{r} \left(\frac{1}{i} \frac{\partial}{\partial u_j} \right) \ln C_r(u_1, \ldots, u_r) \Bigg|_{u=0}. \tag{1.4.18}$$

Note that this function already vanishes if only one of the probability variable (for instance ξ_1) is stochastically independent of the rest of the variables, i.e., if

$$W_r(x) = W_1^{(1)}(x_1) W_{r-1}^{(2)}(x_2, \ldots, x_r), \tag{1.4.19}$$

where $W_{r-1}^{(2)}$ needs not to factorize.

1.5 Stochastic processes

Now we think about a typical physical question, namely how to describe the evolution of probability experiments in time. This means the situation the system the probability experiment is done with depends on time. This is called a *stochastic process*. For sake of simplicity take the case of only one probability observable. The time dependence of the system is then described by the time dependence of the probability variable ξ. The probability distribution at time t is then defined by (1.4.1):

$$W_1(x_1; t) = \langle \delta[x_1 - \xi(t)] \rangle. \tag{1.5.1}$$

The expectation value has to be taken over a big ensemble of systems prepared at the one time instant t.

Now it makes sense to ask after the probability for the event that ξ takes the values $x_1, \ldots, x_n$ at the times $t_1, \ldots, t_n$. By definition the probability distribution for this event is given by

$$W_n(x_n, t_n; \ldots; x_1, t_1) = \left\langle \prod_{j=1}^{n} \delta[x_j - \xi(t_j)] \right\rangle. \tag{1.5.2}$$

If W_n is independent of all times t_j then the process is called *stationary*.

We define the conditional probability that $\xi(t_n)$ takes on the value x_n provided it took the values $x_1, \ldots, x_{n-1}$ at the earlier instants of time $t_1 < t_2 < \cdots < t_n$:

$$P(x_n, t_n | x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = \frac{W_n(x_n, t_n; \ldots; x_1, t_1)}{\int_{\mathbb{R}} dx_n \, W_n(x_n, t_n; \ldots; x_1, t_1)} \tag{1.5.3}$$

where we have made use of (1.4.12).

Now one may classify the processes as follows.

1.5.1 Totally stochastic processes

A totally stochastic process is given if the result to a time is independent of all knowledge about the system in the past. This means that the probability distribution factorizes completely, i.e., the experiment done at time t is stochastically independent of all results obtained at earlier times which means by definition

$$W_n(x_n, t_n; \ldots; x_1, t_1) = \prod_{j=1}^{n} W_1(x_j, t_j). \tag{1.5.4}$$

This situation is not the one we expect to be valid in physics because we want to describe the processes as causal in time. Thus we like to calculate the behaviour of the system, even if we don't know its exact state, from the given behaviour in the past. It would be even nicer if it would be enough to know the probability contribution to one single instant in time and then it can be calculated by dynamical laws for later times.

1.5.2 Markovian processes

The situation described at last in the previous case is called a *Markovian process*. This means the conditional probability (1.5.3) depends only on the last time t_{n-1} before the experiment is done:

$$P(x_n, t_n | x_{n-1}, t_{n-1}; \ldots; x_1, t_1) = P(x_n, t_n | x_{n-1}, t_{n-1}). \tag{1.5.5}$$

Using (1.5.3) and (1.5.5) successively one finds or a Markovian process:

$$W_n(x_n, t_n; \ldots; x_1, t_1) = \left[\prod_{k=1}^{n-1} P(x_{k+1}, t_{k+1} | x_k, t_k) \right] W_1(x_1, t_1) \tag{1.5.6}$$

which shows that $P(x_{k+1}, t_{k+1} | x_k, t_k)$ is the transition probability for ξ to take on the value x_{k+1} at time t_{k+1} if it is known to have taken on the value x_k at time t_k and that for a Markovian process the time evolution of the system depends only on the result of the last known result at an earlier time, not on the whole history of knowledge about the system.

1.5.3 General processes

In the most general case (1.5.3) depends on all times t_k with $k = 1, \ldots, (n-1)$. This means that the system at time t_n has "memorized" the measurements done at all earlier times t_k.

There are examples in physics where this behaviour is really observed, for instance the so called non-Newtonian fluids which "remember" their whole history. Such processes cannot be described by an initial value partial differential equation problem but one has to find a "memory functional" to take care of the whole history of the system.

It is the main task of statistical physics to describe a system in terms of a stochastic process, if the complete knowledge about the system, which would be needed to describe it by the underlying deterministic theory, is not available because of practical reasons.[3]

1.6 Information

Now we come to the mathematical tool for specifying the probability distribution to a given probability experiment due to the given information we have about the system. The problem is to find an objective

---

[3]Note that quantum theory is meant to be deterministic in this context. The complete information we can have about the system is given by a state and the Hamiltonian of the system, if these are given we may calculate in principle the state and the Hamiltonian to any later instant of time by making use of the quantum mechanical time evolution equations (Schrödinger equation of motion).

idea how to determine the distribution without simulating more knowledge about the system than we have really about it. Thus we need a concept for preventing *prejudices* hidden in the wrong choice of a probability distribution.

The idea is to define a measure for the *missing information* about the system provided we define a probability distribution about the outcome of experiments on the system. Clearly this has to be defined relative to the complete knowledge about the system.

We start with the most simple case of a stochastic experiment with a finite number of possible results. The set of possible results may be written as $\Omega = \{1, 2, \ldots, n\}$ for sake of simplicity. All the elementary events might be equally probable. Thus we have the *Laplacian probability distribution*: $P(\{i\}) = 1/n$ for $i = 1, \ldots, n$.

Now comparing situations with $|\Omega| = m$ and $|\Omega| = n$ (where $|\Omega|$ denotes the number of elements contained in Ω) with $n > m$ we want our measure of lacking information to be less in the first case than in the second, because there are less possibilities for the experiment if there are less possible results:

$$I(n) > I(m) \text{ for } n, m \in N \text{ and } n > m. \tag{1.6.1}$$

For $n = 1$ there is no lack of information, because all experiments will give the same result, namely "1". Thus we set by definition

$$I(1) = 0. \tag{1.6.2}$$

Now we think about this as the situation that there is an object which can be in n cells. Each of the cells might be divided in m boxes. The total number of boxes is nm and the total lack of information when asking in which box the object may be is thus $I(nm)$. On the other hand we can locate the object equally well by first asking in which cell it is located (missing information $I(n)$) and then in which of the boxes this special cell is divided in (missing information $I(m)$). Now since this two questions are thought to be independent of each other it makes sense to define the complete lack of information to be the sum of the single ones $I(n) + I(m)$. This means that our measure has to fulfill:

$$I(nm) = I(n) + I(m). \tag{1.6.3}$$

If now $n/l \in \mathbb{N}$ this reads

$$I\left(\frac{n}{l}\right) = I(n) - I(l). \tag{1.6.4}$$

In order to define the measure for all positive rational numbers we take this equation to be valid for any pair of positive integers n and l. Further it is continued to all positive real arguments by continuous continuation.

Now from this assumptions we can determine the measure up to a positive constant factor. For this purpose we suppose that $x \in \mathbb{R}$ is chosen such that $\ln x = m/n$ with $m, n \in N$. Then we have

$$I(x^n) = I(e^m) \Rightarrow nI(x) = mI(e) \tag{1.6.5}$$

where we have made use of (1.6.3) which is valid also for positive real numbers because of (1.6.4) and continuity of the function. Thus we have

$$I(x) = \frac{m}{n}I(e) = I(e)\ln x. \tag{1.6.6}$$

Thus the measure for lacking information in the case of a finite number of equally probable possible results is given by $\ln x$ up to a positive factor $I(e)$ which is set to 1 by definition. Thus we have

$$I(n) = \ln n. \tag{1.6.7}$$

Now let us think about the definition of the measure for the lack of information provided for an experiment with n possible results we have given the probabilities for any elementary event P_i for $i = 1, \ldots, n$.

Then we can find the information measure by the following idea: We can make an experiment of equally distributed elementary events if we think about it in terms of repeating the experiment N times independently.

The first step is to show that the probability contribution for the single event k can be obtained nearly exactly by doing a large number of experiments, i.e., letting $N \Rightarrow \infty$. Let $N(k, N)$ be the number the result k occurs when doing N experiments.

The result of an experiment which consists of N independent experiments is given as an N-tuple of results. If $p = P_k$ then the probability that a certain N-tuple with $N(k, N)$ occurs is $p^{N(k,N)} q^{N-N(k,N)}$ because the experiments were supposed to be independent of each other. For the number $N(k, N)$ it is not important in which order the results show up and thus the distribution for this probability variable is given by

$$P[N(k,N)] = \binom{N}{N(k,N)} p^{N(k,N)} q^{N-N(k,N)}. \tag{1.6.8}$$

Now we want to calculate the expectation value and variance for $N(k, N)$. This can be done by defining the polynomial:

$$f(x) = \sum_{j=0}^{N} \binom{N}{j} p^j q^{N-j} x^j = (px + q)^N \tag{1.6.9}$$

where we have made use of the binomial theorem.

Then we can easily calculate the expectation value

$$\langle N(k,N) \rangle = \sum_{j=0}^{N} j \binom{N}{j} p^j q^{N-j} = f'(1) = Np \tag{1.6.10}$$

and the expectation value for the square

$$\langle N(k,N)^2 \rangle = [xf'(x)]'|_{x=1} = Np + N(N-1)p^2. \tag{1.6.11}$$

The expectation value for $N(k,N)/N$ and its variance

$$\left\langle \frac{N(k,N)}{N} \right\rangle = p, \ \sigma^2_{\left(\frac{N(k,N)}{N}\right)} = \frac{pq}{N}. \tag{1.6.12}$$

One can show that also all higher cumulants vanish for $N \to \infty$ which shows that the contribution approaches the one for a determined result for $N \to \infty$.

The number of possible results is the number of different N-tuples. There are $N!$ possible permutations of N elements but interchanging equal ones does not change the result. Thus the total number is for $N \to \infty$ given by

$$\frac{N!}{\prod_{j=1}^{n}(NP_j)!}. \tag{1.6.13}$$

By using (1.6.7) we find

$$I(N) = \ln \frac{N!}{\prod_{j=1}^{n}(NP_j)!} = \ln N! - \sum_{j=1}^{N} \ln[(NP_j)!]. \tag{1.6.14}$$

For $N \to \infty$ we can use Stirling's asymptotic formula

$$\ln N! \underset{N\to\infty}{\cong} N \ln N - N + O(\ln N). \tag{1.6.15}$$

Using this formula also for $(NP_j)!$ we find

$$I(N) \underset{N\to\infty}{\cong} -N \sum_{j=1}^{n} P_j \ln P_j + O(\ln N). \tag{1.6.16}$$

The missing information per experiment is thus given as

$$I[P] = \lim_{N \to \infty} \frac{I(N)}{N} = -\sum_{j=1}^{n} P_j \ln P_j = -\langle \ln P \rangle. \tag{1.6.17}$$

Clearly there is no problem in using this result also for a probability experiment with countable many possible results:

$$I[P] = -\sum_{j=1}^{\infty} P_j \ln P_j. \tag{1.6.18}$$

In the case of continuous distributions we have to take it as a limit of discrete distributions. For a d-dimensional probability vector ξ we can divide its definition range in cubes of volume $\Delta^d x$ resulting in a countable set of cubes. Then the discrete probability distribution is defined as

$$P_j = W_d(x_j) \Delta^d x \tag{1.6.19}$$

where x_j is an arbitrary value in the jth cube. The missing information is then calculable with help of (1.6.17) or (1.6.18):

$$I(W, \Delta^d x) = -\sum_j P_j \ln P_j = -\sum_j W_d(x_j) \Delta^d x \ln[W_d(x_j) \Delta^d x]. \tag{1.6.20}$$

In these notes we shall always use the technique of using a discrete version of the probability distribution and taking a limit to the continuous distribution after calculating the missing information for the discrete case.

Chapter 2

Basics of Quantum Statistics

In this chapter we shall describe how to formulate Quantum Theory for the case of incomplete knowledge about a system with help of statistics.

2.1 Basics of Quantum Theory

In this course we assume that the reader is familiar with quantum mechanics in terms of Dirac's bra- and ket formalism. We repeat the basic facts by giving some postulates about the structure of quantum theory which are valid in the non-relativistic case as well as in the relativistic one. In these notes we emphasize that quantum theory is a picture about physical systems in space and time. As we know this picture is in some sense valid for a wider range of phenomena than the classical picture of particles and fields.

Although it is an interesting topic we do not care about some problems with philosophy of quantum mechanics. In my opinion physicists have a well understood way in interpreting the formalism with respect to nature and the problem of measurement is not of practical physical importance. That sight seems to be settled by all experiments known so far: They all show that quantum theory is correct in predicting and explaining the outcome of experiments with systems and there is no (practical) problem in interpreting the results from calculating "physical properties of systems" with help of the formalism given by the mathematical tool "quantum theory". So let's begin with some formalism concerning the mathematical structure of quantum mechanics as it is formulated in Dirac's famous book.

- The state of a quantum system is described completely by a ray in a Hilbert space $\mathscr{H}$. A ray is defined as the following equivalence class of vectors:

$$[|\psi\rangle] = \{c\,|\psi\rangle\,|\,|\psi\rangle \in \mathscr{H}, c \in \mathbb{C} \setminus \{0\}\}. \tag{2.1.1}$$

- The observables of the system are represented by self-adjoint operators $\mathbf{O}$ which build together with the unity operator an algebra of operators acting in the Hilbert space. For instance in the case of a quantized classical point particle the *algebra of observables* is generated by the operators of the Cartesian components of configuration space and (canonical) momentum operators, which fulfill the *Heisenberg algebra*:

$$[\mathbf{x}_i, \mathbf{x}_k] = [\mathbf{p}_i, \mathbf{p}_k] = 0,\ [\mathbf{x}_i, \mathbf{p}_k] = \mathrm{i}\delta_{ik}\mathbf{1}. \tag{2.1.2}$$

Here and further on (except in cases when it is stated explicitly) we set (Planck's constant) $\hbar = 1$. In the next chapter when we look at relativity we shall set the velocity of light $c = 1$ too. In this so called natural system of units observables with dimension of an action are dimensionless. Space and time have

the same unit which is reciprocal to that of energy and momentum and convenient unities in particle physics are eV or MeV.

A possible result of a precise measurement of the observable O is necessarily an eigenvalue of the corresponding operator $\mathbf{O}$. Because $\mathbf{O}$ is self-adjoint its eigenvalues are real and the eigenvectors can be chosen such that they build a complete normalized set of kets.

- (**Born's Rule**) If $|o,j\rangle$ is a complete set of orthonormal eigenvectors of $\mathbf{O}$ to the eigenvalue o, the probability to find the value o when measuring the observable O is given by

$$P_\psi(o) = \sum_j |\langle o,j\,|\,\psi\rangle|^2. \qquad (2.1.3)$$

The most famous implication is Heisenberg's uncertainty relation which follows from the positive definiteness of the scalar product in Hilbert space:

$$\Delta A \Delta B \geq \frac{1}{2} |\langle [\mathbf{A}, \mathbf{B}] \rangle|. \qquad (2.1.4)$$

Two observables are simultaneously exactly measurable if and only if the corresponding operators commute. In this case both operators have the same eigenvectors. After a simultaneous measurement the system is in a corresponding simultaneous eigenstate.

A set of pairwise commuting observables is said to be complete if the simultaneous measurement of all these observables fixes the state of the system completely, i.e. if the simultaneous eigenspaces of these operators are one-dimensional (non-degenerate).

- Time is a real parameter. There is an self-adjoint operator $\mathbf{H}$ corresponding to the system such that if $\mathbf{O}$ is an observable then

$$\dot{\mathbf{O}} = \frac{1}{i} [\mathbf{O}, \mathbf{H}] + \partial_t \mathbf{O} \qquad (2.1.5)$$

is the operator of the time derivative of this observable.

The partial time derivative is needed only in the case of explicit time dependence. The fundamental operators like space and momentum operators, which form a complete generating system of the algebra of observables, are not explicitly time dependent (by definition!). It should be emphasized that $\dot{\mathbf{O}}$ is not the mathematical total derivative with respect to time. We shall see that the mathematical dependence on time is arbitrary in a wide sense, because if we have a description of quantum mechanics, then we are free to transform the operators and state kets by a time dependent (!) unitary transformation without changing any physical prediction (possibilities, mean values of observables etc.).

- Due to our first assumption the state of the quantum system is completely known if we know a state ket $|\psi\rangle$ lying in the ray $[|\psi\rangle]$, which is the state the system is prepared in, at an arbitrary initial time. This preparation of a system is possible by performing a precise simultaneous measurement of a complete complete set of observables.

It is more convenient to have a description of the state in terms of Hilbert space quantities than in terms of the projective space (built by the above defined rays). It is easy to see that the state is uniquely given by the projection operator

$$\mathbf{P}_{|\psi\rangle} = \frac{|\psi\rangle \langle\psi|}{\||\psi\||^2}, \qquad (2.1.6)$$

with $|\psi\rangle$ an arbitrary ket contained in the ray (i.e. the state the system is in).

- In general, for example if we like to describe macroscopic systems with quantum mechanics, we do not know the state of the system exactly. In this case we can describe the system by a statistical operator ρ which is positive semi definite (that means that for all kets $|\psi\rangle \in \mathscr{H}$ we have $\langle\psi|\rho|\psi\rangle \geq 0$) and fulfills the normalization condition $\mathrm{Tr}\rho = 1$. It is chosen such that it is consistent with the knowledge about the system we have and contains no more information than one really has. This concept will be explained in a later section.

The trace of an operator is defined with help of a complete set of orthonormal vectors $|n\rangle$ as $\mathrm{Tr}\rho = \sum_n \langle n|\rho|n\rangle$. The mean value of any operator $\mathbf{O}$ is given by $\langle\mathbf{O}\rangle = \mathrm{Tr}(\mathbf{O}\rho)$.

The meaning of the statistical operator is easily seen from this definitions. Since the operator $\mathbf{P}_{|n\rangle}$ answers the question if the system is in the state $[|n\rangle]$ we have $p_n = \mathrm{Tr}(\mathbf{P}_{|n\rangle}\rho) = \langle n|\rho|n\rangle$ as the probability that the system is in the state $[|n\rangle]$. If now $|n\rangle$ is given as the complete set of eigenvectors of an observable operator $\mathbf{O}$ for the eigenvalues O_n then the mean value of $\mathbf{O}$ is $\langle\mathbf{O}\rangle = \sum_n p_n O_n$ in agreement with the fundamental definition of the expectation value of a stochastic variable in dependence of the given probabilities for the outcome of a measurement of this variable.

The last assumption of quantum theory is that the statistical operator is given for the system at all times. This requires that

$$\dot{\rho} = \frac{1}{i}[\rho, \mathbf{H}] + \partial_t \rho = 0. \tag{2.1.7}$$

This equation is also valid for the special case if the system is in a pure state that means $\rho = \mathbf{P}_{|\psi\rangle}$.

2.1.1 Choice of the Picture

Now that we have shortly repeated how quantum mechanics works, we like to give the time evolution a mathematical content, i.e. we settle the time dependence of the operators and states describing the system. As mentioned above it is in a wide range arbitrary how this time dependence is chosen. The only observable facts about the system are expectation values of its observables, so they should have a unique time evolution. To keep the story short we formulate the result as a theorem and prove afterwards that it really gives the right answer. Each special choice of the mathematical time dependence consistent with the axioms of quantum mechanics given above is called a *picture* of quantum mechanics. Now we can state

Theorem 1. *The picture of quantum mechanics is uniquely determined by the choice of an arbitrary self-adjoint Operator* $\mathbf{X}$ *which can be a local function of time. Local means in this context that it depends only on one time, so to say the time point "now" and not (as could be consistent with the causality property of physical laws) on the whole past of the system.*

This operator is the generator of the time evolution of the fundamental operators of the system. This means that it determines the unitary time evolution operator $\mathbf{A}(t, t_0)$ *of the observables by the initial value problem*

$$i\partial_t \mathbf{A}(t, t_0) = -\mathbf{X}(t)\mathbf{A}(t, t_0), \ \mathbf{A}(t_0, t_0) = 1 \tag{2.1.8}$$

such that for all observables which do not depend explicitly on time

$$\mathbf{O}(t) = \mathbf{A}(t, t_0)\mathbf{O}(t_0)\mathbf{A}^\dagger(t, t_0). \tag{2.1.9}$$

Then the generator of the time evolution of the states is necessarily given by the self-adjoint operator $\mathbf{Y} = \mathbf{H} - \mathbf{X}$, *where* $\mathbf{H}$ *is the Hamiltonian of the system. This means the unitary time evolution operator of the states is given by*

$$i\partial_t \mathbf{C}(t, t_0) = +\mathbf{Y}(t)\mathbf{C}(t, t_0). \tag{2.1.10}$$

Proof. The proof of the theorem is not too difficult. At first one sees easily that all the laws given by the axioms like commutation rules (which are determined by the physical meaning of the observables due to symmetry requirements which will be shown later on) or the connection between states and probabilities is not changed by applying different unitary transformations to states and observables.

So there are only two statements to show: First we have to assure that the equation of motion for the time evolution operators is consistent with the time evolution of the entities themselves and second we have to show that this mathematics is consistent with the axioms concerning "physical time evolution" above, especially that the time evolution of expectation values of observables is unique and independent of the choice of the picture.

For the first task let us look on the time evolution of the operators. Because the properties of the algebra given by sums of products of the fundamental operators, especially their commutation rules, shouldn't change with time, the time evolution has to be a linear transformation of operators, i.e. $\mathbf{O} \to \mathbf{A}\mathbf{O}\mathbf{A}^{-1}$ with an invertible linear operator $\mathbf{A}$ on Hilbert space. Because the observables are represented by self-adjoint operators, this property has to be preserved during evolution with time leading to the constraint that $\mathbf{A}$ has to be unitary, i.e. $\mathbf{A}^{-1} = \mathbf{A}^{\dagger}$.

Now for $t > t_0$ the operator $\mathbf{A}$ should be a function of t and t_0 only. Now let us suppose the operators evolved with time from a given initial setting at t_0 to time $t_1 > t_0$ by the evolution operator $\mathbf{A}(t_0, t_1)$. Now we can take the status of this operators at time t_1 as a new initial condition for their further time development to a time t_2. This is given by the operator $\mathbf{A}(t_1, t_2)$. On the other hand the evolution of the operators from t_0 to t_2 should be given simply by direct transformation with the operator $\mathbf{A}(t_0, t_2)$. One can easily see that this long argument can be simply written mathematically as the consistency condition:

$$\forall t_0 < t_1 < t_2 \in \mathbb{R} : \mathbf{A}(t_2, t_1)\mathbf{A}(t_1, t_0) = \mathbf{A}(t_2, t_1), \tag{2.1.11}$$

i.e. in short words: The time evolution from t_0 to t_1 and then from t_1 to t_2 is the same as the evolution directly from t_0 to t_2.

Now from unitarity of $\mathbf{A}(t, t_0)$ one concludes:

$$\mathbf{A}\mathbf{A}^{\dagger} = 1 = \text{const.} \Rightarrow (\mathrm{i}\partial_t \mathbf{A})\mathbf{A}^{\dagger} = \mathbf{A}\partial_t(\mathrm{i}\mathbf{A})^{\dagger}, \tag{2.1.12}$$

so that the operator $\mathbf{X} = -\mathrm{i}(\partial_t \mathbf{A})\mathbf{A}^{\dagger}$ is indeed self-adjoint: $\mathbf{X}^{\dagger} = \mathbf{X}$. Now using eq. (2.1.11) one can immediately show that

$$[\mathrm{i}\partial_t \mathbf{A}(t, t_0)]\mathbf{A}^{\dagger}(t, t_0) = [\mathrm{i}\partial_t \mathbf{A}(t, t_1)]\mathbf{A}^{\dagger}(t, t_1) := -\mathbf{X}(t) \tag{2.1.13}$$

which shows that $\mathbf{X}(t)$ does not depend on the initial time t_0, i.e. it is really local in time as stated in the theorem. So the first task is done since the proof for the time evolution operator of the states is exactly the same: The assumption of a generator $\mathbf{X}(t)$ or. $\mathbf{Y}(t)$ which is local in time is consistent with the initial value problems defining the time evolution operators by their generator.

Now the second task, namely to show that this description of time evolution is consistent with the above mentioned axioms, is done without much sophistication. From (2.1.9) together with the definition (2.1.8) one obtains for an operator which may depend on time:

$$\frac{\mathrm{d}\mathbf{O}(t)}{\mathrm{d}t} = \frac{1}{\mathrm{i}}[\mathbf{O}(t), \mathbf{X}(t)] + \partial_t \mathbf{O}(t). \tag{2.1.14}$$

This equation can be written with help of the "physical time derivative" (2.1.5) in the following form:

$$\frac{\mathrm{d}\mathbf{O}(t)}{\mathrm{d}t} = \dot{\mathbf{O}} - \frac{1}{\mathrm{i}}[\mathbf{O}, \mathbf{H} - \mathbf{X}]. \tag{2.1.15}$$

One sees that the eqs. (2.1.14) and (2.1.15) together with given initial values for an operator $\mathbf{O}$ at time t_0 are uniquely solved by applying a unitary time evolution operator which fulfils the eq. (2.1.8).

Now the statistical operator ρ fulfils that equations of motion as any operator. But by the axiom (2.1.7) we conclude from eq. (2.1.15)

$$\frac{\mathrm{d}\rho(t)}{\mathrm{d}t} = -\frac{1}{\mathrm{i}}[\rho(t), \mathbf{Y}] \tag{2.1.16}$$

and that equation is solved uniquely by a unitary time evolution with the operator $\mathbf{C}$ fulfilling (2.1.10).

Q.E.D.

It should be emphasized that this evolution takes only into account the time dependence of the operators which comes from their dependence on the fundamental operators of the algebra of observables. It does not consider an explicit dependence in time! The statistical operator is always time dependent. The only very important exception is the case of thermodynamical equilibrium where the statistical operator is a function of the constants of motion (we'll come back to that later in our lectures).

Now we have to look at the special case that we have full quantum theoretical information about the system, so we know that this system is in a pure state given by $\rho = \mathbf{P}_{|\psi\rangle} = |\psi\rangle\langle\psi|$ (where $|\psi\rangle$ is normalized). It is clear, that for this special statistical operator the general eq. (2.1.16) and from that (2.1.10) is still valid. It follows immediately, that up to a phase factor the state ket evolves with time by the unitary transformation

$$|\psi, t\rangle = \mathbf{C}(t, t_0)|\psi, t_0\rangle. \tag{2.1.17}$$

From this one sees that the normalization of $|\psi, t\rangle$ is 1 if the ket was renormalised at the initial time t_0. The same holds for a general statistical operator, i.e. $\mathrm{Tr}\rho(t) = \mathrm{Tr}\rho(t_0)$ (exercise: show this by calculating the trace with help of a complete set of orthonormal vectors).

2.1.2 Formal Solution of the Equations of Motion

We now like to integrate the equations of motion for the time evolution operators formally. let us do this for the case of $\mathbf{A}$ introduced in (2.1.9). Its equation of motion which we like to solve now is given by (2.1.8).

The main problem comes from the fact that the self-adjoint operator $\mathbf{X}(t)$ generating the time evolution depends in general on the time t and operators at different times need not commute. Because of this fact we cant solve the equation of motion like the same equation with functions having values in $\mathbb{C}$.

At first we find by integration of (2.1.8) with help of the initial condition $\mathbf{A}(t_0, t_0) = \mathbf{1}$ an integral equation which is equivalent to the initial value problem (2.1.8):

$$\mathbf{A}(t, t_0) = \mathbf{1} + \mathrm{i}\int_{t_0}^{t} \mathrm{d}\tau \mathbf{X}(\tau)\mathbf{A}(\tau, t_0). \tag{2.1.18}$$

The form of this equation leads us to solve it by defining the following iteration scheme:

$$\mathbf{A}_n(t, t_0) = \mathbf{1} + \mathrm{i}\int_{t_0}^{t} \mathbf{X}(\tau)\mathbf{A}_{n-1}(\tau, t_0)\mathrm{d}\tau, \ \mathbf{A}_0(t, t_0) = \mathbf{1}. \tag{2.1.19}$$

The solution of the equation should be given by taking the limit $n \to \infty$. We will not think about the convergence because this is a rather difficult problem.

One can prove by induction that the formal solution is given by the series

$$\mathbf{A}(t, t_0) = \sum_{k=0}^{\infty} \mathbf{A}^{(k)}(t, t_0) \text{ with} \tag{2.1.20}$$

$$\mathbf{A}^{(k)}(t, t_0) = \int_{t_0}^{t} \mathrm{d}\tau_1 \int_{t_0}^{\tau_1} \mathrm{d}\tau_2 \ldots \int_{t_0}^{\tau_{k-1}} \mathrm{d}\tau_k \mathbf{X}(\tau_1)\mathbf{X}(\tau_2)\ldots\mathbf{X}(\tau_k).$$

To bring this series in a simpler form let us look at $\mathbf{A}^{(2)}(t, t_0)$:

$$\int_{t_0}^{t} d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \mathbf{X}(\tau_1)\mathbf{X}(\tau_2). \tag{2.1.21}$$

The range of the integration variables is the triangle in the $\tau_1\tau_2$-plane shown at figure 2.1:
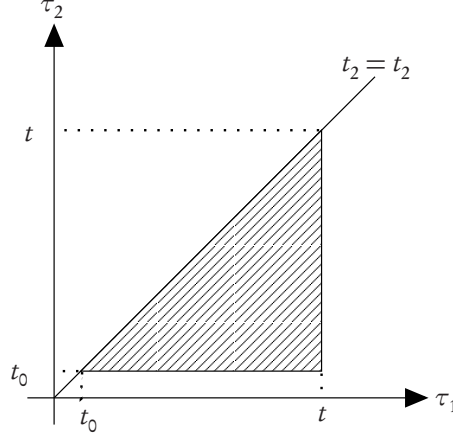

Figure 2.1: *Range of integration variables in (2.1.21)*

Using Fubini's theorem we can interchange the both integrations

$$\mathbf{A}^{(2)} = \int_{t_0}^{t} d\tau_1 \int_{\tau_1}^{t} d\tau_2 \mathbf{X}(\tau_2)\mathbf{X}(\tau_1). \tag{2.1.22}$$

A glance on the operator ordering in (2.1.21) and (2.1.22) shows that the operator ordering is such that the operator at the later time is on the left. For this one introduces the *causal time ordering operator* T_c invented by Dyson. With help of T_c one can add this both equations, leading to the result

$$2\mathbf{A}^{(2)}(t, t_0) = T_c \int_{t_0}^{t} d\tau_1 \int_{t_0}^{t} d\tau_2 \mathbf{X}(\tau_1)\mathbf{X}(\tau_2). \tag{2.1.23}$$

We state that this observation holds for the general case of an arbitrary summand in the series (2.1.20), i.e.

$$\mathbf{A}^{(k)}(t, t_0) = \frac{1}{k!} T_c \int_{t_0}^{t} d\tau_1 \cdots \int_{t_0}^{t} d\tau_n \mathbf{X}(\tau_1)\cdots\mathbf{X}(\tau_n). \tag{2.1.24}$$

To prove this assumption we apply induction. Assume the assumption is true for $k = n-1$ and look at the nth summand of the series. Because the assumption is true for $k = n-1$ we can apply it to the $n-1$ inner integrals:

$$\mathbf{A}^{(n)}(t, t_0) = \frac{1}{(n-1)!} T_c \int_{t_0}^{t} d\tau_1 \int_{t_0}^{\tau_1} d\tau_2 \cdots \int_{t_0}^{\tau_1} d\tau_n \mathbf{X}(\tau_1)\cdots\mathbf{X}(\tau_n). \tag{2.1.25}$$

Now we can do the same calculation as we did for $A^{(2)}$ with the outer integral and one of the inner ones. Adding all the possibilities of pairing and dividing by n one gets immediately

$$\mathbf{A}^{(n)}(t, t_0) = \frac{1}{n!} T_c \int_{t_0}^{t} d\tau_1 \cdots \int_{t_0}^{t} d\tau_n \mathbf{X}(\tau_1)\cdots\mathbf{X}(\tau_n), \tag{2.1.26}$$

and that is (2.1.24) for $k = n$. So our assumption is proven by induction.

With this little combinatorics we can write the series formally

$$\mathbf{A}(t, t_0) = T_c \exp\left[\mathrm{i} \int_{t_0}^{t} \mathrm{d}\tau \, \mathbf{X}(\tau) \right]. \tag{2.1.27}$$

This is the required solution of the equation of motion. For the operator $\mathbf{C}(t, t_0)$ one finds the solution by the same manipulations to be:

$$\mathbf{C}(t, t_0) = T_c \exp\left[-\mathrm{i} \int_{t_0}^{t} \mathrm{d}\tau \, \mathbf{Y}(\tau) \right]. \tag{2.1.28}$$

2.1.3 Example: The Free Particle

The most simple example is the free particle. For calculating the time development of quantum mechanical quantities we chose the Heisenberg picture defined in terms of the above introduced operators $\mathbf{X} = \mathbf{H}$ and $\mathbf{Y} = 0$. We take as an example a free point particle moving in one-dimensional space. The fundamental algebra is given by the space and the momentum operator which fulfil the Heisenberg algebra

$$\frac{1}{\mathrm{i}} [\mathbf{x}, \mathbf{p}] = \mathbf{1}, \tag{2.1.29}$$

which follows from the rules of canonical quantization from the Poisson bracket relation in Hamiltonian mechanics or from the fact that the momentum is defined as the generator of translations in space.

As said above in the Heisenberg picture only the operators representing observables depend on time and the states are time independent. To solve the problem of time evolution we can solve the operator equations of motion for the fundamental operators rather than solving the equation for the time evolution operator. The Hamiltonian for the free particle is given by

$$\mathbf{H} = \frac{\mathbf{p}^2}{2m}, \tag{2.1.30}$$

where m is the mass of the particle. The operator equations of motion can be obtained from the general rule (2.1.14) with $\mathbf{X} = \mathbf{H}$:

$$\frac{\mathrm{d}\mathbf{p}}{\mathrm{d}t} = \frac{1}{\mathrm{i}} [\mathbf{p}, \mathbf{H}] = 0, \quad \frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = \frac{1}{\mathrm{i}} [\mathbf{x}, \mathbf{H}] = \frac{\mathbf{p}}{m}. \tag{2.1.31}$$

That looks like the equation for the classical case but it is an operator equation. But in our case that doesn't effect the solution which is given in the same way as the classical one by

$$\mathbf{p}(t) = \mathbf{p}(0) = \text{const}, \quad \mathbf{x}(t) = \mathbf{x}(0) + \frac{\mathbf{p}(0)}{m} t. \tag{2.1.32}$$

Here we have set without loss of generality $t_0 = 0$.

Now let us look on the time evolution of the wave function given as the matrix elements of the state ket and a complete set of orthonormal eigenvectors of observables. We emphasize that the time evolution of such a wave function is up to a phase independent of the choice of the picture. So we may use any picture we like to get the answer. Here we use the Heisenberg picture where the state ket is time independent. The whole time dependence comes from the eigenvectors of the observables. As a first example we take the momentum eigenvectors and calculate the wave function in the momentum representation. From (2.1.31) we get up to a phase:

$$|p, t\rangle = \exp(\mathrm{i}\mathbf{H}t)|p, 0\rangle = \exp\left(\mathrm{i} \frac{p^2}{2m} t \right) |p, 0\rangle, \tag{2.1.33}$$

and the time evolution of the wave function is simply

$$\psi(p,t) = \langle p,t\,|\,\psi\rangle = \exp\left(-\mathrm{i}\frac{p^2}{2m}t\right)\psi(p,0). \tag{2.1.34}$$

This can be described by the operation of an integral operator in the form

$$\psi(p,t) = \int \mathrm{d}p'\,\underbrace{\langle p,t\,|\,p',0\rangle}_{U(t,p;0,p')}\langle p',0\,|\,\psi\rangle = \int \mathrm{d}p'\,U(t,p;0,p')\psi(p',0). \tag{2.1.35}$$

From (2.1.32) one finds

$$U(t,p,0,p') = \exp\left(-\mathrm{i}\frac{p^2}{2m}t\right)\delta(p-p'). \tag{2.1.36}$$

It should be kept in mind from this example that the time evolution kernels or *propagators* which define the time development of wave functions are in general distributions rather than functions.

The next task we like to solve is the propagator in the space representation of the wave function. We will give two approaches: First we start anew and calculate the space eigenvectors from the solution of the operator equations of motion (2.1.32). We have by definition:

$$\mathbf{x}(t)\,|x,t\rangle = \left(\mathbf{x}(0) + \frac{\mathbf{p}(0)}{m}t\right)|x,t\rangle = x\,|x,t\rangle. \tag{2.1.37}$$

Multiplying this with $\langle x',0|$ we find by using the representation of the momentum operator in space representation $\mathbf{p} = 1/\mathrm{i}\partial_x$:

$$(x'-x)\langle x',0\,|\,x,t\rangle = \frac{\mathrm{i}t}{m}\partial_{x'}\langle x',0\,|\,x,t\rangle \tag{2.1.38}$$

which is solved in a straightforward way:

$$U(t,x;0,x')^* = \langle x',0\,|\,x,t\rangle = N\exp\left[-\mathrm{i}\frac{m}{2t}(x'-x)^2\right]. \tag{2.1.39}$$

Now we have to find the normalization factor N. It is given by the initial condition

$$U(0,x;0,x') = \delta(x-x'). \tag{2.1.40}$$

Since the time evolution is unitary we get the normalization condition

$$\int \mathrm{d}x'\,U(0,x;t,x') = 1. \tag{2.1.41}$$

For calculating this integral from (2.1.39) we have to regularise the distribution to get it as a weak limit of a function. This is simply done by adding a small negative imaginary part to the time variable $t \to t-\mathrm{i}\epsilon$. After performing the normalization we may tend $\epsilon \to 0$ in the weak sense to get back the searched distribution. Then the problem reduces to calculate a Gaussian distribution. As the final result we obtain

$$U(t,x;0,x') = \sqrt{\frac{m}{2\pi\mathrm{i}t}}\exp\left[\mathrm{i}\frac{m}{2t}(x'-x)^2\right]. \tag{2.1.42}$$

An alternative possibility to obtain this result is to use the momentum space result and transform it to space representation. We leave this nice calculation as an exercise for the reader. For help we give the hint that again one has to regularise the distribution to give the resulting Fourier integral a proper meaning.

2.2 Mixed states

Now we want to elaborate further the above ad hoc given definition of mixed states. For this purpose we think about how to formulate the idea of statistical quantum theory in terms of probability theory given in chapter 1.

First let us give an example from physics. Suppose we have measured exactly an observable O (and we have read off its value o from the apparatus). Then we know from the above mentioned basics that the system is in a eigenstate of the corresponding self-adjoint operator $\mathbf{O}$ with eigenvalue o[1]. But if the measured eigenvalue o is degenerated we do not know any more than that. Thus it seems to be sensible to try a description of the situation in terms of probability theory on grounds of the known information.

The probabilities we look for correspond to a future experiment and the first task is to find out what are suitable experiments. This is not as easy as in classical physics where we might measure whatever we like and nothing happens to the system. The only thing we do is to realize a property of the system we have not known before.

In quantum theory this changes dramatically because the system has not all properties which can be measured but only the one given by the previous measurement of O. Thus the set of possible results Ω depends on which observable we are going to measure at next. Now suppose we measure an observable O' which is compatible with O, which means that their corresponding operators commute $\left[\mathbf{O}, \mathbf{O}'\right] = 0$.

In this case the possible result is an eigenvalue of o' but this might not be a complete description because it might be that the simultaneous eigenspace space of $\mathbf{O}$ and $\mathbf{O}'$ to the measured values o and o' respectively might be degenerated too. Thus the set Ω of possible results for the measurement of $\mathbf{O}$ is given by

$$\Omega = \{|o, r\rangle \langle o, r|\} \tag{2.2.1}$$

where $|o, r\rangle$ is an arbitrary complete set of orthonormalised eigenvectors of $\mathbf{O}$ with eigenvalue value o and r is labeling the degeneracy of the eigenspace. This basis can be chosen such that $|o, r\rangle$ is also a eigenbasis for $\mathbf{O}'$. Of course, measuring O' will result again in a eigenstate of $\mathbf{O}$ with the measured eigenvalue o because the observables were supposed to be compatible.

We do not know which will be the state the system is in completely and thus we can not know in which state it will go when measuring O'. The best we can do is to assume a probability distribution for these state. This distribution is completely defined by giving P_r with $0 \leq P_r \leq 1$ and $\sum_r P_r = 1$.

In general we can conclude that we can define the set of possible results by

$$\Omega = \{|n\rangle \langle n|\}_{n \in \mathbb{N}} \tag{2.2.2}$$

where the $|n\rangle$ build a complete orthonormal system (cons). The probability distribution is given by the $P(n)$, i.e., the probabilities that one finds the system in the state $|n\rangle$.

Now we like to find a tool to incorporate these statistics into the quantum theoretical formalism. If we like to calculate the expectation value of an observable O it is customary to take $|o, \alpha\rangle$, the cons. defined by the eigenvectors of the operator $\mathbf{O}$ corresponding to the observable O. Then by definition the expectation value of O due to the probability distribution $P(o, \alpha)$ is given by

$$\langle O \rangle = \sum_{o, \alpha} o P(o, \alpha) = \sum_{o, \alpha} \langle o, \alpha | \mathbf{O} | o, \alpha \rangle P(o, \alpha). \tag{2.2.3}$$

We like now to find a description without any reference to a special cons. For this purpose we introduce a new cons $\{|n\rangle\}_{n \in \mathbb{N}}$ into 2.2.3:

$$\langle O \rangle = \sum_{o, \alpha, n_1, n_2} \langle o, \alpha | n_1 \rangle \langle n_1 | \mathbf{O} | n_2 \rangle \langle n_2 | o, \alpha \rangle P(o, \alpha). \tag{2.2.4}$$

---

[1] We assume the operator to have a discrete spectrum.

Introducing the operator

$$\mathbf{R} = \sum_{o,\alpha} P(o,\alpha) |o,\alpha\rangle \langle o,\alpha| \tag{2.2.5}$$

we can write the sum over o, α as a matrix element with respect to the cons. $\{|n\rangle\}_{n\in\mathbb{N}}$:

$$\sum_{o,\alpha} \langle n_2|o,\alpha\rangle o,\alpha n_1 P(o,\alpha) = \langle n_2|\mathbf{R}|n_1\rangle \tag{2.2.6}$$

and thus (2.2.3) reads

$$\langle O\rangle = \sum_{n_1,n_2} \langle n_2|\mathbf{R}|n_1\rangle \langle n_1|\mathbf{O}|n_2\rangle = \sum_{n_2} \langle n_2|\mathbf{RO}|n_2\rangle = \mathrm{Tr}(\mathbf{RO}). \tag{2.2.7}$$

Herein the Tr of an operator $\mathbf{A}$ is defined with help of a cons. $\{|n\rangle\}_{n\in\mathbb{N}}$ by

$$\mathrm{Tr}\,\mathbf{A} = \sum_n \langle n|\mathbf{A}|n\rangle \tag{2.2.8}$$

from which one concludes that this is invariant under changing the cons. Thus (2.2.7) is the cons.-independent expression we are looking for.

The mixed state is thus described by an operator $\mathbf{R}$ which is called *the statistical operator*. By definition it fulfills the picture independent equation of motion (2.1.7), *the von Neumann equation of motion*:

$$\dot{\mathbf{R}} = \frac{1}{i}[\mathbf{R},\mathbf{H}] + \partial_t \mathbf{R} = 0. \tag{2.2.9}$$

This expresses the fact that, if the statistical operator is given at one instant of time, it is associated to the system at any later time.

The mathematical time evolution is given by the unitarian matrix $\mathbf{C}$ defined in eq. (2.1.10):

$$\mathbf{R}(\mathbf{F}(t),t) = \mathbf{C}(t,t_0)\mathbf{R}(\mathbf{F}(t_0),t)\mathbf{C}^\dagger(t,t_0) \tag{2.2.10}$$

where we had to write out the dependence on the fundamental operators $\mathbf{F}$ and the explicit time dependence. Of course the time evolution operator $\mathbf{C}$ takes account of the time dependence coming in through the fundamental operators, not the explicit time dependence.

From the representation (2.2.5) we can read off easily the coordinate free properties a statistical operator has to fulfill. Since the probabilities $P(o,\alpha)$ are real, it is self-adjoint:

$$\mathbf{R}^\dagger = \sum_{o,\alpha} P^*(o,\alpha)(|o,\alpha\rangle\langle o,\alpha|)^\dagger = \sum_{o,\alpha} P(o,\alpha)|o,\alpha\rangle\langle o,\alpha| = \mathbf{R}. \tag{2.2.11}$$

Further the $P(o,\alpha)$ are positive or 0 and thus $\mathbf{R}$ is positive semidefinite:

$$\langle\psi|\mathbf{R}|\psi\rangle = \sum_{o,\alpha} P(o,\alpha)|\langle\psi|o,\alpha\rangle|^2 \geq 0. \tag{2.2.12}$$

The statistical operator is normalized such that

$$\langle 1\rangle = \mathrm{Tr}\,\mathbf{R} = \sum_{o,\alpha} P(o,\alpha) = 1. \tag{2.2.13}$$

If we are given a self-adjoint positive semidefinite operator with trace 1 we can conclude from its eigenvalue properties that it is a statistical operator corresponding to a probability experiment with the set of possible results Ω given by the set of projectors to the eigenvectors of the operator.

A pure state is represented in this formalism by a projection operator $\mathbf{P}$, i.e. an operator with $\mathbf{P}^2 = \mathbf{P}$. The eigenvector with eigenvalue 1 is the state $|\psi\rangle$ represented by the projector and the probability distribution tells us that (seen from the point of view of a statistical operator) the system is known to be in the pure state $[|\psi\rangle]$.

All this shows that we have to determine the statistical operator with the properties (2.2.11-2.2.13) at an initial time which fulfills Jaynes' principle of least prejudice from (1.6.17-1.6.18) we read off that the measure for the missing information is given by

$$I[\mathbf{R}] = -\langle \ln R \rangle = -\text{Tr}(\mathbf{R}\ln\mathbf{R}) \tag{2.2.14}$$

which is the *von Neumann entropy* corresponding to the statistical operator $\mathbf{R}$.

As an important example we calculate the statistical operator fulfilling Jaynes' principle for the case that we know the expectation values of some observables O_k for $k = 1\ldots n$. These observables need not be compatible because only knowing their expectation values does not mean that we know them exactly. Thus it can make sense to determine expectation values for not compatible observables.

Thus we have to find the statistical operator $\mathbf{R}$ which maximizes the von Neumann entropy (2.2.14) out of all statistical operators which fulfill the constraints:

$$\mathscr{O}_k = \langle O_k \rangle = \text{Tr}(\mathbf{R}\mathbf{O}), \; \text{Tr}\,\mathbf{R} = 1. \tag{2.2.15}$$

This problem is solved as usual for extremal problems with constraints by introduction of Lagrange parameters λ_k and Ω:

$$I'[\mathbf{R},\Omega,\lambda] = -\text{Tr}\left[\mathbf{R}\left(\ln\mathbf{R} + \sum_{k=1}^{n}\lambda_k\mathbf{O}_k + \Omega - 1\right)\right]. \tag{2.2.16}$$

Herein we have chosen $\Omega - 1$ as the Lagrange parameter for the constraint $\text{Tr}\,\mathbf{R} = 1$.

Now we have to maximize I' by varying $\mathbf{R}$ over all self-adjoint positive semidefinite operators. This is done most conveniently by introducing the eigen-cons. of the self-adjoint operator

$$\mathbf{Q} = \sum_{k=1}^{n}\lambda_k\mathbf{O}_k \tag{2.2.17}$$

where the set of Lagrange parameters is fixed. Let us call this cons. $\{|\alpha\rangle\}_{\alpha\in\mathbb{N}}$. Then we write

$$\mathbf{R} = \sum_{\alpha}P_\alpha|\alpha\rangle\langle\alpha|, \quad \mathbf{Q}|\alpha\rangle = q_\alpha|\alpha\rangle. \tag{2.2.18}$$

Plugging this in (2.2.16) this results in

$$I'[\mathbf{R},\Omega,\lambda] = -\sum_{\alpha}P_\alpha(\ln P_\alpha + q_\alpha + \Omega - 1). \tag{2.2.19}$$

Now varying $\mathbf{R}$ over all positive semi-definite self-adjoint operators is given by variation of all $P_\alpha \geq 0$ independently:

$$\delta I' = -\sum_{\alpha}\delta P_\alpha(\Omega + q_\alpha + \ln P_\alpha) \overset{!}{=} 0. \tag{2.2.20}$$

Since the P_α are varied independently of each other, we find

$$P_\alpha = \exp(-\Omega - q_\alpha) \tag{2.2.21}$$

As we see the P_α are all positive.

Plugging this result into (2.2.18) and using (2.2.17) we find

$$\mathbf{R} = \exp\left(-\Omega\mathbf{1} - \sum_{k=1}^{n} \lambda_k \mathbf{O}_k\right). \tag{2.2.22}$$

The Lagrange parameters have to be chosen such that the constraints (2.2.15) are fulfilled. We shall prove in an appendix of this chapter that this is really the (unique) statistical operator fulfilling Jaynes' principle, i.e., which maximizes the von Neumann entropy (2.2.14) under all statistical operators respecting the constraints (2.2.15).

Especially we have

$$Z = \exp\Omega = \mathrm{Tr}\left[\exp\left(-\sum_{k=1}^{n} \lambda_k \mathbf{O}_k\right)\right]. \tag{2.2.23}$$

Z is called the partition sum. Taking the derivative of Ω with respect to the Lagrange parameters we obtain the expectation values of the corresponding observables:

$$\frac{\partial \Omega}{\partial \lambda_k} = \frac{1}{Z}\frac{\partial Z}{\partial \lambda_k} = \mathrm{Tr}[\mathbf{O}_k \mathbf{R}] = \langle O_k \rangle = \mathcal{O}_k. \tag{2.2.24}$$

The proof for this conjecture is given in an appendix not to interrupt our line of arguments[2].

2.2.1 Example: Position and Momentum

As an example we like to calculate the statistical operator according to Jaynes' principle given the average position and momentum and their standard deviations of a one dimensional system. The set of operators $\mathbf{O}_k$ is then given by

$$\mathbf{O}_1 = \mathbf{x}, \ \mathbf{O}_2 = \mathbf{p}, \ \mathbf{O}_3 = \mathbf{x}^2 - \mathcal{X}^2, \ \mathbf{O}_4 = \mathbf{p}^2 - \mathcal{P}^2 \text{ with } \mathcal{X} = \langle x \rangle, \ \mathcal{P} = \langle p \rangle. \tag{2.2.25}$$

This we can plug directly in (2.2.22) to obtain the statistical operator. Without loss of generality we can write this as

$$\mathbf{R} = \exp(-\Omega)\exp\left[-\frac{\lambda_1}{2}(\mathbf{x} - x_0)^2 - \frac{\lambda_2}{2}(\mathbf{p} - p_0)^2\right] \tag{2.2.26}$$

where x_0, p_0, λ_1 and λ_2 are substitutes for the Lagrange parameters.

This problem can be solved analytically because we may set

$$\xi = \mathbf{x} - x_0, \ \pi = \mathbf{p} - p_0 \tag{2.2.27}$$

which fulfill the commutator relations of the Heisenberg algebra

$$\frac{1}{i}[\xi, \pi] = 1. \tag{2.2.28}$$

With them the statistical operator reads

$$\mathbf{R} = \exp(-\Omega)\exp\left(-\frac{\lambda_1}{2}\xi^2 - \frac{\lambda_2}{2}\pi^2\right). \tag{2.2.29}$$

This has the form of exponential of the Hamiltonian of a harmonic oscillator. The solution for the corresponding eigenvalue problem is given in any basic quantum mechanics text book. Thus we summarize only the results.

---

[2]Note that this is not trivial because generally the operators $\mathbf{O}_k$ need not commute

The first step is to define the *annihilation operator for an oscillator quantum*

$$\mathbf{a} = (\lambda_1 \lambda_2)^{1/4} \left(\frac{1}{\sqrt{2\lambda_2}} \xi + \mathrm{i} \frac{1}{\sqrt{2\lambda_1}} \pi \right). \tag{2.2.30}$$

The Heisenberg algebra (2.2.28) gives the crucial commutator relation

$$\left[\mathbf{a}^\dagger, \mathbf{a} \right] = 1. \tag{2.2.31}$$

This relation is used to show that a basis for an irreducible representation of the Heisenberg algebra is given by the eigenvectors of the *number operator*

$$\mathbf{n} = \mathbf{a}^\dagger \mathbf{a}, \ \mathbf{n} |n\rangle = n |n\rangle \ \text{ with } n \in \mathbb{N} := \{0, 1, 2, \ldots\}. \tag{2.2.32}$$

The eigenvectors are built by successive application of the creation operator to *the vacuum state* $|0\rangle$:

$$\mathbf{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle \tag{2.2.33}$$

where the $|n\rangle$ are orthonormal vectors. With the choice of phases defined by (2.2.33) operating with $\mathbf{a}$ leads to

$$\mathbf{a} |0\rangle = 0, \ \mathbf{a} |n\rangle = \sqrt{n} |n-1\rangle \ \text{ for } n \in \mathbb{N}_{>0}. \tag{2.2.34}$$

The statistical operator reads

$$\mathbf{R} = \exp(-\Omega) \exp\left[-\sqrt{\lambda_1 \lambda_2} \left(\mathbf{n} + \frac{1}{2} \right) \right]. \tag{2.2.35}$$

The partition sum reads

$$Z = \exp(\Omega) = \exp\left(-\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right) \sum_{n=0}^\infty \exp(-n\sqrt{\lambda_1 \lambda_2}) = \frac{\exp\left(-\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right)}{1 - \exp(-\sqrt{\lambda_1 \lambda_2})} = \frac{1}{2 \sinh\left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right)}. \tag{2.2.36}$$

The explicit dependence on x_0 and p_0 is lost but it is simple to calculate directly that

$$\langle \xi \rangle = \mathrm{Tr}\, \xi \mathbf{R} = 0, \ \langle \pi \rangle = \mathrm{Tr}\, \pi \mathbf{R} = 0 \tag{2.2.37}$$

and thus

$$\langle x \rangle = \mathrm{Tr}\, \mathbf{x} \mathbf{R} = x_0, \ \langle p \rangle = \mathrm{Tr}\, \mathbf{p} \mathbf{R} = p_0. \tag{2.2.38}$$

The uncertainties (i.e. standard deviations) are calculated by differentiating Z with respect to λ_1 and λ_2:

$$\begin{aligned} \langle \xi^2 \rangle = \Delta x^2 = -\frac{2}{Z} \frac{\partial Z}{\partial \lambda_1} = \sqrt{\frac{\lambda_1}{4\lambda_2}} \coth\left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right), \\ \langle \pi^2 \rangle = \Delta p^2 = -\frac{2}{Z} \frac{\partial Z}{\partial \lambda_2} = \sqrt{\frac{\lambda_2}{4\lambda_1}} \coth\left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right). \end{aligned} \tag{2.2.39}$$

This shows that $\lambda_{1,2} \geq 0$ and

$$\Delta x \Delta p = \frac{1}{2} \coth\left(\frac{\sqrt{\lambda_1 \lambda_2}}{2} \right) \geq 1/2 \tag{2.2.40}$$

which shows that giving Δx and Δp values violating the *Heisenberg uncertainty relation* does not allow to chose a statistical operator which fulfills Jaynes' principle of least prejudice. This should of course be so

because we cannot know more than quantum theory allows us to know and since we resign information we can give position and momentum only with lower precision than by a pure state. Thus there cannot exist any statistical operator violating the Heisenberg uncertainty relation!

Using (2.2.36) in (2.2.35) we find that the statistical operator is given by

$$\mathbf{R} = \frac{1}{1 - \exp(-\sqrt{\lambda_1 \lambda_2})} \exp\left(-\sqrt{\lambda_1 \lambda_2} \mathbf{n}\right). \tag{2.2.41}$$

As (2.2.40) shows the state with minimal uncertainty, $\Delta x \Delta p = 1/2$, is given by the limit $\lambda_1 \lambda_2 \to \infty$. To see the meaning of this limit for $\mathbf{R}$ we evaluate its matrix elements with respect to the occupation-number basis, $|n\rangle$, which is the eigenbasis of $\mathbf{R}$:

$$R(n_1, n_2) = \langle n_1 | \mathbf{R} | n_2 \rangle = \frac{\exp(-\sqrt{\lambda_1 \lambda_2} n_2)}{1 - \exp(-\sqrt{\lambda_1 \lambda_2})} \delta_{n_1 n_2}. \tag{2.2.42}$$

We find that for $\lambda_1 \lambda_2 \to \infty$ we get a non-vanishin matrix element only for $n_2 = 0$, and thus in this limit $\mathbf{R}$ represents the pure state,

$$\mathbf{R}_{\Delta x \Delta p = 1/2} = |n = 0\rangle\langle n = 0|, \tag{2.2.43}$$

i.e., only the ground state of the Hamiltonian of a (shifted) harmonic oscillator leads to a state of minimal space-momentum uncertainty.

2.3   Appendix: Proof of two important theorems

In this appendix we shall give the proof for two important theorems used in the last section.

Theorem 2 (Uniqueness of the statistical operator). *Let $\{\mathbf{O}_j\}_{j=1,\dots,n}$ be a set of (not necessarily compatible) operators. Then the statistical operator*

$$\mathbf{R} = \exp(-\Omega - \sum_{j=1}^{n} \lambda_j \mathbf{O}_j) \tag{2.3.1}$$

is the only statistical operator under all operators which fulfill the constraints

$$\langle O \rangle_j = \mathcal{O}_j \tag{2.3.2}$$

with given expectation values $\mathcal{O}_j$ which maximizes the von Neumann entropy

$$I[\rho] = -\langle \ln \rho \rangle = -\mathrm{Tr}(\rho \ln \rho) \tag{2.3.3}$$

provided there exists a set of Lagrange parameters λ_j ($j = 1 \dots n$) such that the constraints (2.3.2) are fulfilled. Then this set of Lagrange parameters are unique, i.e., the maximum-entropy principle provides a unique statistical operator of minimal prejudice, given the constraints (2.3.2).

Proof. Let ρ be an arbitrary statistical operator fulfilling (2.3.2). Then we like to prove

$$S[\rho] - S[\mathbf{R}] \leq 0. \tag{2.3.4}$$

To this end we show that

$$\forall x \in \mathbb{R}_{>0} : \ln x \geq x - 1 \quad \text{and} \quad x - 1 = \ln x \Leftrightarrow x = 1. \tag{2.3.5}$$

We have only to investigate the function $f : \mathbb{R}_{>0} \to \mathbb{R}$ with $f(x) = x - 1 - \ln x$. Its derivative is given by $f'(x) = 1 - 1/x$ which is strictly monotone ascending and has the only 0 at $x = 1$. Thus f has the global minimum at $x = 1$ and $f(1) = 0$. This shows $f(x) \geq 0$ for all $x \in \mathbb{R}_{>0}$ and $\ln x = x - 1$ only for $x = 1$.

To use this equation we observe with help of the eigen-cons. of $\mathbf{R}$

$$\mathrm{Tr}[\mathbf{R} \ln \mathbf{R}] = -\sum_n R_n \left(\sum \lambda_j o_{nn}^{(j)} + \Omega \right) = -\Omega - \sum \lambda_j \mathcal{O}_j$$
$$= -\mathrm{Tr}\left[\rho \left(-\Omega - \sum_j \lambda_j \mathbf{O}_j \right) \right] = \mathrm{Tr}(\rho \ln \mathbf{R})$$

(2.3.6)

where we have used that ρ and $\mathbf{R}$ fulfill both the constraints (2.3.2). Herein we have used the following notation: R_n are the eigenvalues of $\mathbf{R}$ with the eigenvector $|n\rangle$ and $o_{nn}^{(j)} = \langle n | \mathbf{O}_j | n \rangle$.

Using now the eigen-cons. of ρ labeled by $|\alpha\rangle$ we find

$$S[\rho] - S[\mathbf{R}] = \sum_\alpha \rho_\alpha \ln\left(\frac{R_{\alpha\alpha}}{\rho_\alpha} \right) \leq \sum_\alpha (\rho_\alpha - R_{\alpha\alpha}) = 0$$

(2.3.7)

where we have used the inequality (2.3.5) and $\mathrm{Tr}\,\rho = \mathrm{Tr}\,\mathbf{R} = 1$.

This shows that, if a set of Lagrange parameters Ω and λ_j exists such that (2.3.2) is fullfilled for the statistical operator of the form (2.3.1) any other statistical operator ρ fullfilling (2.3.2) can not have a larger entropy than $\mathbf{R}$. Since the necessary condition for ρ to maximize the entropy is that it is of the form (2.3.1). Thus to prove the uniqueness of the statistical operator we have to show that the solution of the constraints (2.3.2) and $\mathrm{Tr}\,\mathbf{R} = 1$ leads to a unique solution for Ω and the λ_j.

Q.E.D.

Theorem 3 (Derivatives of exponential functions)**.** *Let* $\mathbf{Y} : D \to \mathrm{Hom}(\mathcal{H})$ *where D is an open subset of* $\mathbb{R}$*. Then we define*

$$\partial_a \mathbf{Y}(a) = \lim_{\Delta a \to 0} \frac{\mathbf{Y}(a + \Delta a) - \mathbf{Y}(a)}{\Delta a}.$$

(2.3.8)

If this limit exists then we call $\mathbf{Y}$ *differentiable in a.*

Supposed this is the case then also the operator $\exp[\mathbf{Y}(a)]$ *is differentiable in a and we have*

$$\partial_a \exp[\mathbf{Y}(a)] = \exp[\mathbf{Y}(a)] \int_0^1 \mathrm{d}\tau \exp[-\tau \mathbf{Y}(a)] \partial_a \mathbf{Y}(a) \exp[\tau \mathbf{Y}(a)].$$

(2.3.9)

Proof. Let $\mathbf{X} : D \to \mathrm{Hom}(\mathcal{H})$ (with D an open subset of $\mathbb{R}$) and $\mathbf{X}(t) = \mathbf{X}_1(t) + \mathbf{X}_2(t)$. Define operators $\mathbf{U}$ and $\mathbf{U}_j$ with ($j = 1, 2$) by the initial value problem

$$\partial_t \mathbf{U}(t, t_0) = \mathbf{X}(t)\mathbf{U}(t, t_0), \ \mathbf{U}(t_0, t_0) = 1;$$
$$\partial_t \mathbf{U}_j(t, t_0) = \mathbf{X}_j(t)\mathbf{U}_j(t, t_0), \ \mathbf{U}_j(t_0, t_0) = 1$$

(2.3.10)

which is solved in section 2.1 for the problem of time evolution. Now we define the operator $\mathbf{V}$

$$\mathbf{V}(t, t_0) = \mathbf{U}_1(t, t_0) + \int_{t_0}^t \mathrm{d}t' \mathbf{U}(t, t') \mathbf{X}_2(t') \mathbf{U}_1(t, t_0).$$

(2.3.11)

Differentiating with respect to t and using (2.3.10) and $\mathbf{U}(t, t) = 1$ for $t \in D$ we find that

$$\partial_t V(t, t_0) = \mathbf{X}(t) V(t, t_0).$$

(2.3.12)

Further it can be seen immediately from (2.3.10) that $\mathbf{V}(t_0, t_0) = 1$. Thus $\mathbf{V} \equiv \mathbf{U}$. Thus we have the identity

$$\mathbf{U}(t, t_0) = \mathbf{U}_1(t, t_0) + \int_{t_0}^{t} dt' \mathbf{U}(t, t') \mathbf{X}_2(t') \mathbf{U}_1(t, t_0). \tag{2.3.13}$$

Now we apply this to the case that $\mathbf{X}_1$ and $\mathbf{X}_2$ and thus $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2$ are independent of t:

$$\exp[(t - t_0)\mathbf{X}] = \exp[(t - t_0)\mathbf{X}_1] + \int_{t_0}^{t} dt' \exp[(t - t')\mathbf{X}] \mathbf{X}_2 \exp[(t' - t_0)\mathbf{X}_1]. \tag{2.3.14}$$

Now set $t_0 = 0$ and $t = 1$ to obtain

$$\exp[\mathbf{X}_1 + \mathbf{X}_2] = \exp[\mathbf{X}_1] + \exp(\mathbf{X}) \int_0^1 d\tau \exp[\tau(\mathbf{X}_1 + \mathbf{X}_2)] \mathbf{X}_2 \exp(\tau \mathbf{X}_1). \tag{2.3.15}$$

Setting $\mathbf{X}_1 = \mathbf{Y}(a)$ and $\mathbf{X}_2 = \Delta a \partial_a \mathbf{Y}(a)$ and letting $\Delta a \to 0$ we find the claim of the theorem.[3]   Q.E.D.

Theorem 4 (Calculating expectation values). *The partition sum for a statistical operator fulfilling Jaynes' principle with the given constraints (2.3.2) is defined as*

$$Z = \exp \Omega = \operatorname{Tr} \exp\left(-\sum_{j=1}^{n} \lambda_j \mathbf{O}_j\right). \tag{2.3.16}$$

Then the expectation values for the observable O_j are given by the derivatives of Z as follows:

$$\left\langle O_j \right\rangle = -\frac{1}{Z} \frac{\partial Z}{\partial \lambda_j} = -\frac{\partial \Omega}{\partial \lambda_j}. \tag{2.3.17}$$

Proof. Using

$$\mathbf{Y} = -\sum_{j=1}^{n} \lambda_j \mathbf{O}_j \text{ and } a = \lambda_k \text{ with } k = 1, \dots, n \tag{2.3.18}$$

in (2.3.9) we have

$$\frac{\partial}{\partial \lambda_k} \exp\left[-\sum_{j=1}^{n} \lambda_j \mathbf{O}_j\right] = -\exp\left[-\sum_{j=1}^{n} \lambda_j \mathbf{O}_j\right] \int_0^1 d\tau \exp\left[-\tau \sum_{j=1}^{n} \lambda_j \mathbf{O}_j\right] \mathbf{O}_k \exp\left[\tau \sum_{j=1}^{n} \lambda_j \mathbf{O}_j\right] \tag{2.3.19}$$

Now given two operators $\mathbf{X}$ and $\mathbf{Y}$ and provided the trace of their product exists we have

$$\operatorname{Tr}(\mathbf{XY}) = \operatorname{Tr}(\mathbf{YX}). \tag{2.3.20}$$

Using a cons. $\{|n\rangle\}_{n \in \mathbb{N}}$ this is proven by

$$\operatorname{Tr}(\mathbf{XY}) = \sum_{n_1, n_2} \langle n_1 |\mathbf{X}| n_2 \rangle \langle n_2 |\mathbf{Y}| n_1 \rangle = \sum_{n_1, n_2} \langle n_2 |\mathbf{Y}| n_1 \rangle \langle n_1 |\mathbf{X}| n_2 \rangle = \operatorname{Tr}(\mathbf{YX}). \tag{2.3.21}$$

Taking the trace of (2.3.19) we can write the rightmost exponential operator to the very left (taking the first under the integral). Then we can use the simple equation

$$[X, Y] = 0 \Rightarrow \exp(\mathbf{X} + \mathbf{Y}) = \exp(\mathbf{X}) \exp(\mathbf{Y}) \tag{2.3.22}$$

---

[3]Mention that this is not a rigorous proof because we have interchanged taking the integral and the limit.

which is proven with help of Cauchy's product law for series like in the case of the exponential of a complex number.[4] Then the integral becomes trivial and we find

$$\mathrm{Tr}\left[\frac{\partial}{\partial\lambda_k}\exp\left(\sum_{j=1}^{n}\lambda_j\mathbf{O}_j\right)\right]=-\mathrm{Tr}\left[\mathbf{O}_k\exp\left(\sum_{j=1}^{n}\lambda_j\mathbf{O}_j\right)\right]. \tag{2.3.23}$$

To finish the proof we have only to show that we can interchange differentiation with respect to λ_k with taking the trace operator $\mathbf{X}(a)$:

$$\partial_a\,\mathrm{Tr}\,\mathbf{X}(a)=\partial_a\sum_n\langle n\,|\mathbf{X}(a)|\,n\rangle. \tag{2.3.24}$$

To interchange the derivative with the summation we have to claim that the series is uniformly convergent in a neighbourhood of a and that the same is true for the series built by the derivatives of its summands. Then we can interchange the derivative with taking the trace and the theorem is proven. Q.E.D.

---

[4] If the operators in the exponential do not commute then there exists a more general formula known as Baker Campbell Hausdorff formula.

Equilibrium Thermodynamics

In this chapter we want to derive the results of phenomenological thermodynamics which deals with equilibrium, i.e., macroscopically static situations.

Due to the ideas developed in the previous section the stationarity of the statistical operator fixes the possible set of given expectation values as the known information about the system to the global conserved quantities of the system.

In order to obtain the full thermodynamic content of the phenomenological theory we have to take into account adiabatic changes of the external parameters which will be discussed in the first section. From these considerations one can deduce the whole thermodynamic content in terms of the thermodynamic potentials.

Beginning with the next section we shall treat some physical examples. Starting with ideal gases we shall introduce the perturbative methods needed for treating interacting systems.

3.1 Adiabatic changes of external parameters

In this section we shall prove the important adiabatic theorem:

Theorem 5. *Suppose the Hamiltonian* $\mathbf{H}(\chi)$ *of the system depends on an external parameter* χ.[1]. *We assume that the Hamiltonian has a non degenerate spectrum.*

Now let the external parameter change with time according to

$$\chi(t) = \chi_1 + \frac{t}{\tau}(\chi_2 - \chi_1). \tag{3.1.1}$$

Then in the limit of infinitesimal slow change, i.e., $\tau \to \infty$ *the time evolution from* $t = 0$ *to* $t = \tau$ *of a conserved observable* $O(\chi_1)$ *leads to a conserved observable with external parameter* χ_2.

Proof. For sake of simplicity we do the calculations in the Heisenberg picture. Then the full time dependence is at the operators. The state $|\psi\rangle$ the system is prepared in stays constant while the eigenstate $|\alpha, \chi\rangle$ of the Hamiltonian is time dependent through the time dependence of the external parameter χ. To fix the notation we write down the eigenvalue equation for $\mathbf{H}$:

$$\mathbf{H}(\chi)|\alpha, \chi\rangle = E_\alpha(\chi)|\alpha, \chi\rangle. \tag{3.1.2}$$

The Heisenberg picture equation of motion is

$$-\mathrm{i}\frac{\mathrm{d}}{\mathrm{d}t}|\alpha, \chi\rangle = \mathbf{H}(\chi)|\alpha\chi\rangle - \mathrm{i}\frac{\chi_2 - \chi_1}{\tau}\frac{\partial}{\partial \chi}|\alpha, \chi\rangle \tag{3.1.3}$$

---

[1]This might be an external field or as most often discussed in elementary applications the volume of the container of a gas

which leads to the Schrödinger equation for the energy representation

$$i\frac{d}{dt}\langle\alpha,\chi\,|\psi\rangle=\langle\alpha,\chi\,|\mathbf{H}(\chi)|\psi\rangle+i\frac{\chi_2-\chi_1}{\tau}\frac{\partial}{\partial\chi}\langle\alpha,\chi\,|\psi\rangle. \tag{3.1.4}$$

Making use of the operator

$$\mathbf{G}(\chi)=i\sum_\alpha\left(\frac{\partial}{\partial\chi}|\alpha,\chi\rangle\right)\langle\alpha,\chi|=-i\sum_\alpha|\alpha,\chi\rangle\frac{\partial}{\partial\chi}\langle\alpha,\chi| \tag{3.1.5}$$

leads to

$$i\frac{d}{dt}\langle\alpha,\chi\,|\psi\rangle=E_\alpha(\chi)\langle\alpha,\chi\,|\psi\rangle-\frac{\chi_2-\chi_1}{\tau}\sum_{\alpha'}G_{\alpha\alpha'}(\chi)\langle\alpha',\chi\,|\psi\rangle \tag{3.1.6}$$

with the matrix element

$$G_{\alpha\alpha'}(\chi)=\langle\alpha,\chi\,|\mathbf{G}(\chi)|\alpha',\chi\rangle. \tag{3.1.7}$$

Now we define

$$\langle\alpha,\chi\,|\psi\rangle=\exp\left[-i\int_0^t dt'E_\alpha(\chi(t'))\right]\phi_\alpha(t) \tag{3.1.8}$$

and substitute it in the equation of motion (3.1.6) using the integration

$$\int_0^t dt'E_\alpha(\chi(t'))=\frac{\tau}{\chi_2-\chi_1}\int_{\chi_1}^{\chi(t)}d\chi\,E_\alpha(\chi) \tag{3.1.9}$$

which leads after simple calculations to

$$i\frac{d\phi_{\alpha(t)}}{dt}=-\frac{\chi_2-\chi_1}{\tau}\sum_{\alpha'}G_{\alpha\alpha'}(\chi)\exp\left[\frac{i\tau}{\chi_2-\chi_1}\int_{\chi_1}^{\chi(t)}d\chi[E_\alpha(\chi)-E_{\alpha'}(\chi)]\right] \tag{3.1.10}$$

Now redefining the phase of the eigenvectors $|\alpha,\chi\rangle$ by

$$|\alpha',\chi\rangle=\exp[i\sigma_\alpha(\chi)]|\alpha,\chi\rangle \tag{3.1.11}$$

the operator $\mathbf{G}$ changes to

$$\mathbf{G}'(\chi)=\mathbf{G}\chi-\sum_\alpha|\alpha,\chi\rangle\frac{d\sigma_\alpha(\chi)}{d\chi}\langle\alpha,\chi| \tag{3.1.12}$$

which means

$$G'_{\alpha\alpha}(\chi)=\langle\alpha',\chi\,|\mathbf{G}(\chi)|\alpha',\chi\rangle-\frac{d\sigma_\alpha(\chi)}{d\chi}. \tag{3.1.13}$$

This means that we can make the diagonal elements of $\mathbf{G}$ vanish by proper choice of the arbitrary phases σ_α. From now on we assume that this is already the case for $\mathbf{G}$.

Integrating (3.1.10) over time we find

$$i[\phi_\alpha(\tau)-\phi_\alpha(0)]=-\frac{\chi_2-\chi_1}{\tau}\sum_{\alpha'}\int_0^\tau dt\,G_{\alpha\alpha'}(\chi)\phi_{\alpha'}\exp\left[\frac{i\tau}{\chi_2-\chi_1}\int_{\chi_1}^{\chi(t)}d\chi[E_\alpha(\chi)-E_{\alpha'}]\right]. \tag{3.1.14}$$

By our choice of phases such that $G_{\alpha\alpha}=0$ for all α in the sum of (3.1.10) only non-vanishing energy differences occur in the exponential because we assumed the Hamiltonian to have a non-degenerate spectrum. Thus the integral in the argument of the exponential is increasing with time and for $\tau\to\infty$ starts to oscillate more and

more rapidly while all the other functions vary little with time. Thus the integration over t tends to cancel the sum completely. This leads to the asymptotic statement of our theorem, namely

$$\phi_\alpha(\tau) \underset{\tau\to\infty}{\cong} \phi_\alpha(0). \tag{3.1.15}$$

Now let $\mathbf{A}(\chi)$ be a constant of motion with respect to the parameter dependent Hamiltonian $\mathbf{H}(\chi)$. Then it is diagonal with respect to the eigenbasis of the Hamiltonian:

$$\mathbf{A} = \sum_\alpha a_{\alpha\alpha}(\chi)|\alpha,\chi\rangle\langle\alpha,\chi| \tag{3.1.16}$$

and (3.1.15) shows that the adiabatic change of the parameter χ with time moves this from $\mathbf{A}(\chi_1)$ which is a constant of motion with respect to $\mathbf{H}(\chi_1)$ to $\mathbf{A}(\chi_2)$ which is a constant of motion with respect to $\mathbf{H}(\chi_2)$.

Q.E.D.

3.2 Phenomenological Thermodynamics revisited

In this section we shall briefly show that with the development of general statistical quantum physics one can derive all the knowledge of phenomenological thermodynamics from microscopic quantum statistics.

To this end we have to remember that phenomenological thermodynamics deals with stationary, i.e., macroscopic equilibrium states. This situation is completely defined by a statistic operator $\mathbf{R}$ which does not depend on time explicitly. Using the von Neumann equation of motion (2.2.9) this means

$$\frac{1}{i}[\mathbf{R},\mathbf{H}] = 0 \tag{3.2.1}$$

and this means for the operator fulfilling Jaynes' principle of least prejudice (2.2.22) must depend only on constants of motion. Thus the most general equilibrium statistic operator has to be built with given expectation values of conserved quantities. Denote an independent set of conserved quantities by $\{\mathbf{O}_i\}_{i=1\ldots n}$ any equilibrium operator has to be of the form

$$\mathbf{R} = \exp\left(-\Omega - \sum_{i=1}^{n}\lambda_i\mathbf{O}_i\right) \tag{3.2.2}$$

The most important statistical operator the *canonical statistical operator* where only the mean energy of the system is given

$$\mathbf{R} = \exp(-\Omega - \beta\mathbf{H}) \tag{3.2.3}$$

and the *grand canonical statistical operator* where the mean energy and numbers of particles are given[2]

$$\mathbf{R} = \exp(-\Omega - \beta\mathbf{H} - \alpha\mathbf{N}). \tag{3.2.4}$$

As we know from theorem 4 of section 2.3 the expectation values can be calculated by taking the derivatives of the logarithm of the partition sum Ω with respect to the Lagrange parameters λ_i:

$$\langle O_i\rangle = -\frac{\partial\Omega}{\partial\lambda_i}. \tag{3.2.5}$$

Due to the adiabatic theorem of the last section under an adiabatic change of the external parameters the statistical operator remains an equilibrium one with respect to the conserved quantities dependent on the

---

[2]Keep in mind that we are working in non-relativistic quantum theory now where the particle number is conserved provided we have particles which don't build bound states.

external parameters. On the other hand the adiabatic change is described as a time evolution in the limit of changing the external parameters very slowly. Since the von Neumann equation of motion tells us that the statistical operator $\mathbf{R}$ stays constant with time together with Ehrenfest's theorem we conclude that the von Neumann entropy $I = -\langle \ln \mathbf{R} \rangle$ stays constant within an adiabatic change of the external parameters.

On the other hand we have

$$I = -\langle \ln \mathbf{R} \rangle = \Omega + \sum_j \lambda_j o_j \text{ with } o_j = \langle \mathbf{O}_j \rangle. \tag{3.2.6}$$

This shows that Ω is a Legendre transform of the entropy, also known as *Massieu functions* in the literature.

Thus an adiabatic change of the parameters means that the system stays in the most general thermal equilibrium and for a closed system we have

$$\mathrm{d}I = \left(\frac{\partial I}{\partial \chi} \right)_{o_j = \text{const.}} \mathrm{d}\chi + \sum_j \lambda_j \mathrm{d}o_j = 0. \tag{3.2.7}$$

This we call the *Thermodynamic adiabatic theorem*.

Now using (3.2.6) and $\partial \Omega / \partial \lambda_j = -o_j$ this can be written as

$$\mathrm{d}I = \sum_\chi \left(\frac{\partial \Omega}{\partial \chi} \right)_{\lambda_j = \text{const.}} \mathrm{d}\chi + \sum_j \lambda_j \mathrm{d}o_j = 0 \tag{3.2.8}$$

because

$$\left(\frac{\partial I}{\partial \chi} \right)_{o_j = \text{const.}} = \left(\frac{\partial \Omega}{\partial \chi} \right)_{\lambda_j = \text{const.}}. \tag{3.2.9}$$

Now in general we can never know the average values of all conserved quantities (as we shall see in the next section this would be an infinite number of occupation numbers in the usual grand canonical approach!). In the usually explored cases of classical phenomenological thermodynamics we have only a view parameters like energy and particle numbers. But one looks at open systems. This means that we have a big macroscopic system (in general one may think of it as the whole universe) and we look on a little subsystem of also macroscopic dimensions.

We suppose now that the whole system is in thermodynamical and chemical equilibrium, i.e., the statistical operator describing the whole system is the grand canonical operator

$$\mathbf{R} = \exp(-\Omega - \beta \mathbf{H} - \alpha \mathbf{N}) \tag{3.2.10}$$

where $\mathbf{N}$ is the operator of the total particle number.[3]

Now we suppose that we make an adiabatic change on the smaller subsystem and that the whole system is so big that under that change the system stays at thermodynamical and chemical equilibrium, i.e., after the change of state the whole system is again described by a grand canonical statistical operator (3.2.10) with changed Lagrange parameters β and α and thus also changed average particle number and energy.

From the thermal adiabatic theorem (3.2.7) we obtain under these circumstances

$$\mathrm{d}I_1 = \beta \mathrm{d}\mathscr{E}_1 + \alpha \mathrm{d}\mathscr{N}_1 + \beta \sum_\chi \left(\frac{\partial \Omega}{\partial \chi} \right)_{\lambda_j = \text{const.}} = -\mathrm{d}I_2. \tag{3.2.11}$$

---

[3]Keep in mind that this is a equilibrium operator in the non-relativistic case only, because in relativistic physics the particle number is not conserved. In this case one has to substitute conserved charge operators (for instance electric charge, lepton number or baryon number) for $\mathbf{N}$.

Now we take especially the volume V of the small system as an external parameter. From (3.2.11) we obtain

$$d\mathcal{E}_1 = \frac{1}{\beta}dI_1 - \frac{1}{\beta}\left(\frac{\partial\Omega}{\partial V}\right)_{\alpha,\beta}dV - \frac{\alpha}{\beta}d\mathcal{N}_1. \tag{3.2.12}$$

If we identify our Lagrange parameters and the derivative of the Massieu function Ω with the phenomenological bulk properties of the thermodynamic system in the following way

$$I_1 = S, \ \beta = \frac{1}{T}, \ p = \frac{1}{\beta}\left(\frac{\partial\Omega}{\partial V}\right)_{\alpha,\beta=\text{const.}}, \ \mu = -\frac{\alpha}{\beta} \tag{3.2.13}$$

then (3.2.12) reads

$$d\mathcal{E}_1 = TdS - pdV + \mu d\mathcal{N}_1 \tag{3.2.14}$$

which is nothing else than the *first law of thermodynamics* namely the *macroscopic energy conservation*: The change of energy of the subsystem is given by the energy TdS, *the heat*, coming from the bigger subsystem (which takes account of thermal and chemical equilibrium and therefore often is called *the heat bath*), the mechanical work $-pdV$ and chemical work $\mu d\mathcal{N}_1$.

We also see that Shannon's measure of missing information for an open system which is in thermal and chemical equilibrium with a heat bath, which fulfills Jaynes' principle of least prejudice is nothing else than the macroscopic entropy invented by Clausius and Planck.

All this shows that phenomenological thermodynamics can be re-derived using the quantum statistics for the equilibrium in the grand canonical ensemble. The main advantage of the statistical approach is that one is (at least in principle) able to derive the properties of macroscopic systems the underlying dynamics of the very many particles it is built of.

It should only be mentioned that our quantum statistical approach avoids a lot of difficulties the classical statistics has with deriving phenomenological thermodynamics.

3.3   Ideal gases in the grand canonical ensemble

In this section we shall use the formalism developed so far for calculating the properties of one of the most important examples for quantum statistics, the *ideal gases*. From the microscopic point of view these are the most simple systems built by non interacting particles.

From the very beginning we shall use the methods of quantum field theory to calculate all the thermodynamic properties. This is done to introduce the formalism of canonical field quantization. In my notes about quantum field theory you may read the same ideas from the particle point of view while the canonical quantization formalism is more from the field point of view.

3.3.1   Non-relativistic ideal gases

In these notes we prefer the canonical field quantization formalism against the particle point of view of the same mathematics. This has the advantage of leading immediately to the interesting physical results but has the pedagogical disadvantage to hide the microscopic picture of particles behind the whole philosophy of statistics. The reader my look on my notes about quantum field theory for a complete treatment of the particle point of view.

We start with the classical action for the one particle Schrödinger equation:

$$A[\psi^*, \psi] = \int d^4x \underbrace{\left[i\psi^*\partial_t\psi - \frac{1}{2m}(\nabla\psi)^*(\nabla\psi)\right]}_{\mathscr{L}}. \tag{3.3.1}$$

Thereby we used the abbreviation $x = (t, \vec{x})$ within the integral.[4] The classical equations of motion are given by the *Hamiltonian principle* which tells us to look for the fields ψ^* and ψ which minimize the action functional. The result are the *Euler-Lagrange equations of motion*:

$$\frac{\delta A}{\delta \psi^*} = \frac{d}{dt} \frac{\partial \mathscr{L}}{\partial \partial_t \psi^*} + \nabla \frac{\partial \mathscr{L}}{\partial (\nabla \psi^*)} - \frac{\partial \mathscr{L}}{\partial \psi^*} = 0. \tag{3.3.2}$$

The variation with respect to ψ leads to the conjugate complex equation. Taking the Lagrangian from eq. (3.3.1) we immediately find the *Single particle Schrödinger equation for free particles*[5]:

$$i\partial_t \psi = -\frac{\Delta \psi}{2m}. \tag{3.3.3}$$

The important point of this field point of view is that it is very easy to find the expressions for total field energy, momenta, angular momenta and single particle number because these are conserved quantities due to Noether's theorem (see my notes on quantum field theory ch. 1). We shall not derive these well known expressions here which should be done by the reader as a simple exercise.

We aim now to quantize this "field formalism" canonically. For this purpose the first step is to transform into the Hamiltonian formalism which is a Legendre transformation with respect to the time derivatives of the fields. The canonical field momenta[6] are given by

$$\Pi = \frac{\partial \mathscr{L}}{\partial (\partial_t \psi)} = i\psi^* m, \Pi^* = \frac{\partial \mathscr{L}}{\partial (\partial_t \psi^*)} = 0 \tag{3.3.4}$$

and the Hamiltonian is given by making use of its definition

$$\mathscr{H} = \frac{1}{2m} (\nabla \psi^*)(\nabla \psi) = -\frac{i}{2m} (\nabla \Pi)(\nabla \psi). \tag{3.3.5}$$

In this case the vanishing of Π^* does not lead to any difficulties for the Hamiltonian formalism because in this we do not need to eliminate the time derivatives out of the Hamiltonian. The reason is that the Schrödinger field Lagrangian depends linearly on these derivatives and thus they vanish completely in the Hamiltonian. Nevertheless it is important to mention that the independent field degrees of freedom are given by ψ and Π and thus only those have to be used in the canonical quantization formalism. The reader should verify that the Hamiltonian formalism of course leads to the correct equations of motion for ψ and $\pi = i\psi^*$, namely the Schrödinger equation of motion and its complex conjugate.

As already mentioned above additionally to the space-time symmetries of Newtonian mechanics the Lagrangian for the Schrödinger equation also respects the symmetry transformation of a global phase transformation:

$$\psi'(x) = \exp(-i\alpha)\psi(x), \ \psi'^*(x) = \exp(+i\alpha)\psi^*(x). \tag{3.3.6}$$

According to Noether's theorem the conserved quantity following from this symmetry is the normalization integral

$$\mathbf{N} = \int d^3\vec{x} \, \psi^*(x)\psi(x) \text{ with } x = (t, \vec{x}). \tag{3.3.7}$$

---

[4]The reader should keep in mind that this has nothing to do with the relativistic Minkowski space notation because here we treat the non-relativistic case.

[5]For sake of simplicity we take spin-0-particles. As we know from the spin-statistics theorem of relativistic quantum field theory these particles are necessarily bosons. As we shall see for the case of the calculation of non interacting particles extension of the results for arbitrary spin particles is trivial

[6]Keep in mind that these have nothing to do with physical momenta. It is just a naming convention which comes to the theory from classical non-relativistic point mechanics without magnetic fields where the canonical momenta are of course identical with the physical momenta mv.

This should be proven also by the reader as a simple example using the notes about quantum field theory.

Now we can quantize the fields. As is discussed in the particle picture (see quantum field theory ch. 2) there are two possible quantization procedures which take account on the indistinguishability of identical particles: Quantization with commutator relations (the particles are then called *bosons*) or with anti-commutator relations (the particles are then called *fermions*).

Due to the canonical quantization formalism we introduce field operators ψ and Π instead of the classical "c-number"-fields and define the anti-commutator- or commutator relations according to the analogue *Poisson brackets*. If not mentioned otherwise in the following we treat fermions and bosons in parallel. The most differences are signs and the upper signs belong to fermions, the lower to bosons:

$$[\psi(t,\vec{x}),\psi(t,\vec{y})]_{\pm} = [\Pi(t,\vec{x}),\Pi(t,\vec{y})]_{\pm} = 0, \ \frac{1}{i}[\psi(t,\vec{x}),\Pi(t,\vec{y})]_{\pm} = \delta^{(3)}(\vec{x}-\vec{y}). \tag{3.3.8}$$

Now we like to calculate the thermodynamic properties of the ideal gas within a cubic box of length L. We assume that L is big compared to the De Broglie wave lengths of the typical momenta of the particles within the gas. Then we might neglect boundary effects which means that we may chose any boundary conditions which are customary for the calculation. In the end of the calculation we shall look on the limit $L \to \infty$. Then in equilibrium we expect the gas to be homogeneous and it is thus natural to use periodic boundary conditions, i.e.,

$$\psi(t,\vec{x}+L\vec{e}_i) = \psi(t,\vec{x}) \text{ for } i = 1,2,3. \tag{3.3.9}$$

Then we can expand the the field operator in a Fourier series

$$\psi(t,\vec{x}) = \sum_{\vec{n}\in\mathbb{Z}^3} \frac{1}{\sqrt{V}} \mathbf{a}(\vec{n}) \exp[-i\omega(\vec{n})+i\vec{p}(\vec{n})\vec{x}]$$

$$\mathbf{a}(\vec{n}) = \int_V \frac{d^3\vec{x}}{\sqrt{V}} \psi(t,\vec{x}) \exp[i\omega(\vec{n})t - i\vec{p}(\vec{n})\vec{x}] \tag{3.3.10}$$

$$\text{with } \vec{p}(\vec{n}) = \frac{2\pi}{L}\vec{n}, \ \vec{n}\in\mathbb{Z}^3.$$

$\omega(\vec{n})$ will be determined from the *Heisenberg picture equations of motion* for the field operators. To find these we have firstly to determine the (anti-) commutator relations for the $\mathbf{a}$ out of the canonical relations (3.3.8). From (3.3.10) we obtain

$$\left[\mathbf{a}(\vec{n}),\mathbf{a}^\dagger(\vec{n}')\right]_{\pm} = \int_V \frac{d^3\vec{x}'}{\sqrt{V}} \int_V \frac{d^3\vec{x}}{\sqrt{V}} \left[\psi(t,\vec{x}),\psi^\dagger(t,\vec{x}')\right]_{\pm} \times$$

$$\times \exp\{-i[\omega(\vec{n})-\omega(\vec{n}')]t + i[\vec{p}(\vec{n})-\vec{p}(\vec{n}')]\} = \tag{3.3.11}$$

$$= \int_V \frac{d^3\vec{x}}{V} \exp\{i[\vec{p}(\vec{n})-\vec{p}(\vec{n}')]\} = \delta^{(3)}(\vec{n}-\vec{n}') := \begin{cases} 0 & \text{for } \vec{n}\neq\vec{n}' \\ 1 & \text{for } \vec{n}=\vec{n}'. \end{cases}$$

Now it is easy to express the various quantities in terms of the $\mathbf{a}(\vec{n})$. One just has to apply (3.3.10) to the quantum counterparts of the classical expressions and to integrate out the $\vec{x}$-integral[7]. For the Hamiltonian and the "normalization" we find

$$\mathbf{H} = \frac{1}{2m}\sum_{\vec{n}\in\mathbb{Z}^3} \vec{p}^2(\vec{n})\mathbf{N}(\vec{n}), \ \mathbf{N} = \sum_{\vec{n}\in\mathbb{Z}^3} \mathbf{N}(\vec{n}) \text{ with } \mathbf{N}(\vec{n}) = \mathbf{a}^\dagger(\vec{n})\mathbf{a}(\vec{n}). \tag{3.3.12}$$

---

[7]The operator-ordering problem is in our case of non-relativistic quantum field theory not as serious as in the relativistic case. As we shall see in both cases problem is solved by choosing the lowest energy value at 0 and the particle number to 0 for the same state. At the end the operators have to be written in normal order, i.e., all $\mathbf{a}^\dagger$ have to be written to the left of all $\mathbf{a}$

To find the physical meaning of the operators $\mathbf{a}$, $\mathbf{a}^\dagger$ and $\mathbf{N}(\vec{n})$ we have to solve the eigenvector problem for the $\mathbf{N}(\vec{n})$. The first step is to show that the $\mathbf{N}(\vec{n})$ commute. Since they are self-adjoint operators this tells us that we can span the Hilbert space the field operators act on by a set of simultaneous eigenvectors of the $\mathbf{N}(\vec{n})$.

This is done by direct evaluation of the commutator

$$
\begin{aligned}
\left[\mathbf{N}(\vec{n}),\mathbf{N}(\vec{n}')\right] &= \left[\mathbf{a}^\dagger(\vec{n})\mathbf{a}(\vec{n}),\mathbf{a}^\dagger(\vec{n}')\mathbf{a}(\vec{n}')\right] = \\
&= \mathbf{a}^\dagger(\vec{n})\left[\mathbf{a}(\vec{n}),\mathbf{a}^\dagger(\vec{n}')\mathbf{a}(\mathbf{n}')\right] + \left[\mathbf{a}^\dagger(\vec{n}),\mathbf{a}^\dagger(\vec{n}')\mathbf{a}(\vec{n}')\right]\mathbf{a}(\vec{n}) = \\
&= \mathbf{a}^\dagger(\vec{n})\left\{\left[\mathbf{a}(\vec{n}),\mathbf{a}^\dagger(\vec{n}')\right]_\pm \mathbf{a}(\vec{n}) \mp \mathbf{a}^\dagger(\vec{n}')\left[\mathbf{a}(\vec{n}),\mathbf{a}(\vec{n}')\right]_\pm\right\} + \\
&\quad + \left\{\left[\mathbf{a}^\dagger(\vec{n}),\mathbf{a}^\dagger(\vec{n}')\right]_\pm \mathbf{a}(\vec{n}') \mp \mathbf{a}^\dagger(\vec{n}')\left[\mathbf{a}^\dagger(n),\mathbf{a}(\vec{n}')\right]_\pm\right\}\mathbf{a}(\vec{n}) = 0.
\end{aligned}
\tag{3.3.13}
$$

Thereby we made use of the identities

$$
[\mathbf{A},\mathbf{BC}]_\pm = [\mathbf{A},\mathbf{B}]_\pm \mathbf{C} \mp \mathbf{B}[\mathbf{A},\mathbf{C}]_\pm \text{ and } [\mathbf{AB},\mathbf{C}]_\pm = \mathbf{A}[\mathbf{B},\mathbf{C}]_\pm \mp [\mathbf{A},\mathbf{B}]_\pm \mathbf{C}.
\tag{3.3.14}
$$

Thus the $\mathbf{N}(\vec{n})$ commute and thus we can solve the eigenvalue problem for each $\mathbf{N}(\vec{n})$ separately. For this purpose we write simply $\mathbf{a}$, $\mathbf{a}^\dagger$ and $\mathbf{N}$.

We shall solve firstly the problem for fermions. In this case $\mathbf{N}$ is a projection operator:

$$
\mathbf{N}^2 = \mathbf{a}^\dagger \mathbf{a}\mathbf{a}^\dagger \mathbf{a} = \left\{\{\mathbf{a}^\dagger,\mathbf{a}\} - \mathbf{a}\mathbf{a}^\dagger\right\}\mathbf{a}^\dagger \mathbf{a} = \mathbf{N},
\tag{3.3.15}
$$

where we have used $(\mathbf{a}^\dagger)^2 = 0$. Suppose $|\alpha\rangle$ is an eigenvector of $\mathbf{N}$ with eigenvalue value α this means that $\alpha^2 = \alpha$ and thus $\alpha = 0$ or $\alpha = 1$. Now we suppose that the eigenspace of $\mathbf{N}$ is not degenerate. We shall show below that this conjecture is equivalent to the irreducibility of the representation of the field operators on the Hilbert space.[8] We have just to find how $\mathbf{a}$ and $\mathbf{a}^\dagger$ act on the eigenvectors $|\alpha\rangle$. For this we calculate the commutators

$$
\left[\mathbf{N},\mathbf{a}^\dagger\right] = \mathbf{a}^\dagger, \quad [\mathbf{N},\mathbf{a}] = -\mathbf{a}.
\tag{3.3.16}
$$

From this we find

$$
\mathbf{N}\mathbf{a}|\alpha\rangle = \{[\mathbf{N},\mathbf{a}] + \mathbf{a}\mathbf{N}\}|\alpha\rangle = (\alpha - 1)\mathbf{a}|\alpha\rangle
\tag{3.3.17}
$$

which means that $\mathbf{a}|\alpha\rangle$ is either a eigenvector of $\mathbf{N}$ with eigenvector $\alpha - 1$ or it is 0. From our conjecture that the eigenspaces of $\mathbf{N}$ are not degenerate we conclude

$$
\mathbf{a}|1\rangle = |0\rangle, \quad \mathbf{a}|0\rangle = 0
\tag{3.3.18}
$$

because $\mathbf{a}|0\rangle$ cannot be different from 0. Otherwise it would be an eigenvector of $\mathbf{N}$ with eigenvalue -1 which contradicts the projector properties of $\mathbf{N}$. Then $\mathbf{a}|1\rangle$ cannot be 0 because otherwise $\mathbf{a}$ would be 0 which is not compatible with the anti commutator relations of $\mathbf{a}$ and this leads together with the conjecture that the eigenspaces of $\mathbf{N}$ are non degenerate to the first equation (of course up to an arbitrary phase factor).

The correctness of the normalization for the first equation in (3.3.18) is also easily seen:

$$
\langle \mathbf{a}1 | \mathbf{a}1\rangle = \left\langle 1\left|\mathbf{a}^\dagger \mathbf{a}\right|1\right\rangle = \langle 1|\mathbf{N}|1\rangle = \langle 1|1\rangle = 1.
\tag{3.3.19}
$$

Thus a proper choice of the arbitrary phase of $|0\rangle$ leads to the first equation in (3.3.18).

Exactly the same arguments give

$$
\mathbf{a}^\dagger |0\rangle = |1\rangle, \quad \mathbf{a}^\dagger |1\rangle = 0.
\tag{3.3.20}
$$

---

[8]We want only mention that there are other approaches to quantum field theory which use explicitly a reducible representation of the algebra known as Thermo Field Dynamics (TFD). In our approach (named Schwinger-Keldysh real-time formalism) we shall find the "doubling of field degrees of freedom" on other grounds. Both formalism result in the same results in practice.

This shows that the $\mathbf{a}(\vec{n})$ are destruction and $\mathbf{a}^{\dagger}(\vec{n})$ creation operators for a fermion in a state of definite momentum $\vec{p}(\vec{n})$. The whole Hilbert space is spanned by

$$|N(\vec{n})\rangle = \prod_{\vec{n}\in\mathbb{Z}^3}[\mathbf{a}^{\dagger}(\vec{n})]^{N(\vec{n})}|0\rangle \tag{3.3.21}$$

where $N(\vec{n}) = 0, 1$ and $|0\rangle$ is the "vacuum state", i.e., the state describing the situation that no particle is present. We have also to fix the order of creation operators which means only the choice of a certain phase for the states. The states are antisymmetric with respect to interchange of two one-particle states.

The physical meaning of all this is now clear: From the canonical quantization with anti-commutators there cannot be more than 1 particle in a certain quantum state, i.e., fermions obey *Pauli's exclusion principle*. The states (3.3.21) may contain any total number of particles and the conserved quantity (3.3.7) is nothing else than the total number of particles described by the state. This Hilbert space, describing particles with an arbitrary number of particles, is known as the *Fock space for fermions*. The formalism above shows that it splits in the orthogonal sum of the subspaces of a fixed number of particles. Each of these subspaces is conserved within the time evolution because the Hamiltonian commutes with the number operator. This shows that our formalism is completely equivalent with the 1st quantization formalism for fermions.

We shall not repeat the same construction for bosons. The only difference is that in any state may be any number of particles and that the number states spanning the *Fock space for bosons* are given by:

$$|N(\vec{n})\rangle = \prod_{\vec{n}} \frac{1}{\sqrt{N(\vec{n})!}}[\mathbf{a}^{\dagger}(\vec{n})]^{N(\vec{n})}|0\rangle \;\; \text{with } N(\vec{n})\in\mathbb{N} := \{0, 1, 2, \ldots\}. \tag{3.3.22}$$

Now it is also easy to find the Heisenberg picture equation of motion for our free particle case:

$$\partial_t\psi(t,\vec{x}) = \mathrm{i}[\mathbf{H}, \psi(t,\vec{x})]. \tag{3.3.23}$$

On the right-hand-side we write the Hamiltonian and the field in terms of annihilation and creation operators according to eq. (3.3.12) and eq. (3.3.10) which using (3.3.16)[9] leads to:

$$[\mathbf{H}, \psi(t,\vec{x})] = -\sum_{\vec{n}\in\mathbb{Z}^3} \frac{1}{\sqrt{V}} \frac{\vec{p}^{\,2}(\vec{n})}{2m} \mathbf{a}(\vec{n}) \exp[-\mathrm{i}\omega(\vec{n}) + \mathrm{i}\vec{p}(\vec{n})\vec{x}]\psi(t,\vec{x}). \tag{3.3.24}$$

Writing also the left-hand-side of eq. (3.3.23) in terms of the Fourier series (3.3.10) this gives the *dispersion relation for free Schrödinger fields*:

$$\omega(\vec{n}) = \frac{\vec{p}^{\,2}(\vec{n})}{2m} \tag{3.3.25}$$

which shows also that our ansatz with time-independent annihilation operators in (3.3.10) is correct within the Heisenberg picture we are using here.

Now we want to derive the macroscopic properties of ideal gases in equilibrium. For this purpose we chose the grand canonical ensemble which is defined in terms of our information theoretical language as the choice of the statistical operator fulfilling Jaynes' principle of least prejudice consistent with the given average energy $\mathscr{E} = \langle H\rangle$ and the average total number of particles $\mathscr{N} = \langle N\rangle$. As we shall see this is a very good approximation for a closed system of $\mathscr{N}$ particles at energy $\mathscr{E}$ within a box of volume $V = L^3$. Here the approximation is "good" in the sense that for a macroscopic number of particles the standard deviations (i.e. uncertainties) for total energy and particle number are negligible compared to their values itself.

The grand canonical statistical operator is given by

$$\mathbf{R} = \exp[-\Omega - \beta\mathbf{H} - \alpha\mathbf{N}] \tag{3.3.26}$$

---

[9] $\mathbf{N}(\vec{n}')$ commutes with $\mathbf{a}(\vec{n})$ for $\vec{n}\neq\vec{n}'$

where we have made use of our general solution of the Jaynes' principle (2.2.22). As we have proven in section 2.3 we can calculate the average energy and particle number if we know the grand canonical partition sum as derivatives with respect to the according Lagrange parameters β and α. Thus our aim is to calculate this partition sum which is done most conveniently with evaluating the trace using the particle number eigenstates (3.3.21) for fermions or (3.3.22) for bosons

$$Z = \text{Tr}\left[\exp(-\beta \mathbf{H} - \alpha \mathbf{N})\right] = \prod_{\vec{n} \in \mathbb{Z}^3} \sum_{N(\vec{n})} \exp\left[\left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha\right) N(\vec{n})\right]. \tag{3.3.27}$$

The sum over $N(\vec{n})$ is simply done. In the case of fermions each $N(\vec{n})$ runs only over 0 and 1 giving always a finite result while for bosons these numbers run from 0 to infinity which gives a geometric series which is only convergent if the exponent is negative. As we shall see this fact gives rise to a completely different behaviour of boson gases compared to fermion gases in the quantum regime.

Assuming that the mentioned convergence restriction for the bosonic case is fulfilled we can write the result for both cases in one equation:

$$Z = \prod_{\vec{n} \in \mathbb{Z}^3} \left[1 \pm \exp\left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha\right)\right]^{\pm 1}. \tag{3.3.28}$$

For the following it is more convenient to discuss the physical properties with help of the Massieu function (the grand canonical potential)

$$\Omega = \ln Z = \pm \sum_{\vec{n} \in \mathbb{Z}^3} \ln\left[1 \pm \exp\left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha\right)\right]. \tag{3.3.29}$$

For both cases it is a necessary condition for this series to converge that $\beta > 0$. The range of α is not restricted for fermions but for bosons one has to pay attention that the exponent does not vanish. In the limit of an infinite volume we have to force $\alpha > 0$ and as we shall see below this has the very interesting physical consequence of the appearance of *Bose-Einstein condensation* in the quantum limit. This means that a macroscopic number of the bosons contained in the gas is in the lowest energy state. This leads to the highly impressive super-fluid behaviour of bosonic fluids as is ^{4}He which is a macroscopic quantum effect.

For sake of completeness we add also spin degrees of freedom. The whole formalism explained above still applies. The only difference is that the creation- and annihilation-operators contain an additional label σ which runs over $\{-s, -s+1, \ldots, s-1, s\}$ for particles of spin s. For instance for electrons $s = 1/2$.

Clearly we have to sum over σ in (3.3.29) additionally to the sum over $\vec{n}$. In the case of free particles, i.e., ideal gases the Hamiltonian does not contain the spin and thus this summation over the spin degrees of freedom gives just a multiplication with $g = 2s + 1$. Thus the correct grand potential for ideal gases including the spin degrees of freedom is given by

$$\Omega = \ln Z = \pm g \sum_{\vec{n} \in \mathbb{Z}^3} \ln\left[1 \pm \exp\left(-\beta \frac{\vec{p}^2(\vec{n})}{2m} - \alpha\right)\right]. \tag{3.3.30}$$

3.3.2 The ideal Fermi gas

In the case of Fermi gases the sum in (3.3.29) exists for $\alpha \in \mathbb{R}$ and $\beta \in \mathbb{R}_{>0}$. Thus we can go without further difficulties to the limit of an infinite box, i.e., $L \to \infty$. For large but finite L there are a lot of levels within a small volume $\mathrm{d}^3 \vec{p}$ in momentum space:

$$\vec{p}(\vec{n}) = \frac{2\pi}{L}\vec{n}, \ \vec{n} \in \mathbb{Z}^3, \ \mathrm{d}n_i = \frac{L}{2\pi}\mathrm{d}p_i \Rightarrow D(\vec{p})\mathrm{d}^3\vec{p} = \frac{V}{(2\pi)^3}\mathrm{d}^3\vec{p} \tag{3.3.31}$$

where $D(\vec{p})$ is the density of quantum states at momentum $\vec{p}$ (shortly called the *level density*).
This shows that in the limit of a very big box containing the gas we can write the sum (3.3.30)

$$\Omega = g \int d^3\vec{p} \frac{V}{(2\pi)^3} \ln\left[1 + \exp\left(-\beta\frac{\vec{p}^2}{2m} - \alpha\right)\right] \qquad (3.3.32)$$

This integral can be evaluated up to a single integral by introducing spherical coordinates:

$$\Omega = \frac{gV}{2\pi^2} \int_0^\infty dp\, p^2 \ln\left[1 + \exp\left(-\beta\frac{p^2}{2m} - \alpha\right)\right] \qquad (3.3.33)$$

Unfortunately this integral cannot be calculated analytically with elementary functions. Nevertheless we can calculate the behaviour of the gas in two important limits.

3.3.3 The classical limit

The most important quantum feature of the Fermi gas is Pauli's exclusion principle which forbids two particles to occupy the same single-particle state. Thus we conclude that the gas should show classical behaviour if statistically only a small number of particles is in a single state. This is the case either for a small number of particles which is not true for macroscopic gases or at given particle number at high temperatures, i.e., small β.

Formally the classical limit is obtained if $\exp(\alpha) \gg 1$.[10] Integrating (3.3.33) by parts leads to

$$\Omega = \frac{gV\beta}{6\pi^2 m} \int_0^\infty dp \frac{p^4}{\exp\left(\beta\frac{p^2}{2m} + \alpha\right) + 1} \qquad (3.3.34)$$

and in the classical limit we can neglect the 1 in the denominator against the exponential function:

$$\Omega = \frac{gV\beta}{6\pi^2 m} \exp(-\alpha) \int_0^\infty dp\, p^4 \exp\left(-\beta\frac{p^2}{2m}\right). \qquad (3.3.35)$$

This integral can be evaluated by using again the trick of "generating functions":

$$f(x) = \int_0^\infty dp \exp(-xp^2) = \frac{\sqrt{\pi}}{2} x^{-1/2}. \qquad (3.3.36)$$

Then we have

$$f'(x) = -\int_0^\infty dp\, p^2 \exp(-xp^2) = -\frac{\sqrt{\pi}}{4} x^{-3/2},$$
$$f''(x) = \int_0^\infty dp\, p^4 \exp(-xp^2) = \frac{3\sqrt{\pi}}{8} x^{-5/2}. \qquad (3.3.37)$$

For $x = \beta/(2m)$ we find for the integral (3.3.35):

$$\Omega = gV\left(\frac{m}{2\pi\beta}\right)^{3/2} \exp(-\alpha). \qquad (3.3.38)$$

---

[10]This can be shown by calculating the distribution of the particles over the states within the grand canonical ensemble. This task is devoted to an exercise for the reader.

From this we find the average particle number and energy by differentiation with respect to β or α respectively, i.e.

$$\mathcal{N} = -\frac{\partial \Omega}{\partial \alpha} = g V \left(\frac{m}{2\pi\beta} \right)^{3/2} \exp(-\alpha), \quad \mathcal{E} = -\frac{\partial \Omega}{\partial \beta} = \frac{3 g V}{2} \left(\frac{m}{2\pi\beta} \right)^{3/2} \frac{1}{\beta} \exp(-\alpha). \tag{3.3.39}$$

The first equation shows of course that our considerations are valid only in the low density limit because $\exp(-\alpha) \ll 1$.

Combining both equations (3.3.39) leads to the well known caloric equation of state for the ideal gas, namely

$$\mathcal{E} = \frac{3}{2} \frac{\mathcal{N}}{\beta}. \tag{3.3.40}$$

According to (3.2.13) the statistical quantities are connected with the macroscopic quantities by

$$T = \frac{1}{\beta}, \quad \alpha = -\mu\beta \tag{3.3.41}$$

where T is the temperature (measured in energy units which is due to our choice of $k_B = 1$) and μ is the chemical potential of the gas.

From (3.2.6) we obtain the entropy of the ideal gas

$$S = \Omega + \beta\mathcal{E} + \alpha\mathcal{N}. \tag{3.3.42}$$

According to (3.2.7) the natural independent quantities for the entropy are E, N and V:

$$S = \frac{5}{2}\mathcal{N} + \mathcal{N} \ln\left[\frac{g V}{\mathcal{N}} \left(\frac{m\mathcal{N}}{3\pi\mathcal{E}} \right)^{3/2} \right]. \tag{3.3.43}$$

This gives for the pressure according to (3.2.13) or (3.2.14):

$$p = \frac{1}{\beta} \left(\frac{\partial S}{\partial V} \right)_{\mathcal{N},\mathcal{E}=\text{const.}} = \frac{1}{\beta} \left(\frac{\partial \Omega}{\partial V} \right)_{\alpha,\beta=\text{const.}} = \frac{\mathcal{N}}{V\beta} \tag{3.3.44}$$

which may look more familiar in the form

$$pV = \mathcal{N}T. \tag{3.3.45}$$

This is nothing else than the equation of state for an ideal gas.

3.3.4 The Quantum Limit for a Fermi Gas

The quantum nature of the gas becomes most "visible" in the limit $T \to +0$, which means $\beta \to \infty$.

So let us first come to the case $T = 0$ itself. Then we have to write from the very beginning $\alpha = -\beta\mu$ and we have to keep μ finite while letting $\beta \to \infty$. The quantum nature of the Fermi gas becomes most clear if we calculate the distribution of the particles over the momentum states for arbitrary temperatures:

$$\rho(\vec{p}) = \frac{1}{\beta V} \frac{\delta\Omega}{\delta\omega(\vec{p})} = \frac{g}{1 + \exp\left[\beta\left(\frac{\vec{p}^2}{2m} - \mu \right) \right]}. \tag{3.3.46}$$

This is the *Fermi distribution* with a factor g taking into account the degenaracy of the free particle's state due to the spin degrees of freedom[11].

---

[11]For instance for electrons we have $g = 2$. The free electron gas theory applies for instance to the conducting electrons within a metal. The electrons can approximately be seen to move freely due to the positive background they are situated.

For $T \to 0$, i.e. $\beta \to \infty$ this becomes

$$\rho_0(\vec{p}) = g\Theta(\mu_0 - \frac{\vec{p}^2}{2m}) \tag{3.3.47}$$

which shows clearly the fermionic character of the particles. Because only g of them can join the same momentum state (due to different spin degrees of freedom) at 0 temperature they fill up all levels beginning with the lowest until all particles are put in the box.

The average total number of particles is

$$\mathscr{N}_0 = -\left(\frac{\partial\Omega}{\partial\alpha}\right)_{\beta\to\infty} = -gV\int \frac{\mathrm{d}^3\vec{p}}{(2\pi)^3}\rho_0(\vec{p}) = \frac{gV}{6\pi^2}(2m\mu_0)^{3/2} \tag{3.3.48}$$

which gives the chemical potential μ_0 at $T = 0$ as a function of the total number of particles and shows that μ_0 is rather big for macroscopic amounts of gases.

The mean total energy of the gas at 0 temperature is

$$\mathscr{E}_0 = \left(\frac{\partial\Omega}{\partial\beta}\right)_{\beta\to\infty} = \frac{gV}{20\pi^2}(2m\mu_0)^{5/2}. \tag{3.3.49}$$

At not too high temperatures we can expand the thermodynamic quantities with respect to the small parameter $1/(\mu\beta)$ which can be seen as follows. We have to calculate integrals of the form

$$F[f] = \frac{gV}{2\pi^2}\int_0^\infty \mathrm{d}P P^2 f(P)\frac{1}{1+\exp\left[\beta\left(\frac{P^2}{2m}-\mu\right)\right]}. \tag{3.3.50}$$

Now we substitute $x = \beta[P^2/(2m)-\mu]$:

$$F[f] = \frac{gV}{4\pi^2}\left(\frac{2m}{\beta}\right)^{3/2}\int_{-\mu\beta}^\infty \mathrm{d}x\,\sqrt{x+\beta\mu}f\left(\sqrt{\frac{2m(x+\mu\beta)}{\beta}}\right)\frac{1}{1+\exp x}. \tag{3.3.51}$$

Now we calculate the particle number:

$$\mathscr{N} = F[1] = \frac{gV}{4\pi^2}\left(\frac{2m}{\beta}\right)^{3/2}\int_{-\mu\beta}^\infty \mathrm{d}x\,\frac{\sqrt{x+\mu\beta}}{1+\exp x}. \tag{3.3.52}$$

Integration by parts gives

$$\mathscr{N} = \frac{gV}{6\pi^2}\left(\frac{2m}{\beta}\right)^{3/2}\int_{-\mu\beta}^\infty \mathrm{d}x\,\frac{\exp x}{(1+\exp x)^2}(x+\mu\beta)^{3/2}. \tag{3.3.53}$$

Now the first factor of the integrand is an even function in x which is exponentially damped for $|x| \to \infty$ while the second factor is a slowly varying function where the first factor is big for large $\mu\beta$. Thus on the one hand we can take without much loss of precision the lower boundary of the integral to $-\infty$ and on the other hand we can expand the second factor (und thus the whole integral) in powers of $1/(\mu\beta)$:

$$(x+\mu\beta)^{3/2} = (\mu\beta)^{3/2}\left(1+\frac{3}{2}\frac{x}{\mu\beta}+\frac{3}{8}\frac{x^2}{\mu^2\beta^2}+\cdots\right). \tag{3.3.54}$$

Using

$$\int_{-\infty}^{\infty} dx \frac{\exp x}{(1+\exp x)^2} = -\left.\frac{1}{1+\exp x}\right|_{-\infty}^{\infty} = 1,$$

$$\int_{-\infty}^{\infty} dx\, x^2 \frac{\exp x}{(1+\exp x)^2} = \frac{\pi^2}{3},$$

(3.3.55)

where the last integral is proven in appendix A, in (3.3.53) gives

$$\mathcal{N} = \frac{gV}{6\pi^2}(2m\mu)^{3/2}\left[1 + \frac{\pi^2}{8}\frac{1}{\mu^2\beta^2} + O[(\mu\beta)^{-4}]\right].$$

(3.3.56)

The mean energy is calculated in the same way as the particle number with the result

$$\mathcal{E} = \frac{gV}{20m\pi^2}(2m\mu)^{5/2}\left[1 + \frac{5}{8}\left(\frac{\pi}{\mu\beta}\right)^2 + O[(\mu\beta)^{-4}]\right].$$

(3.3.57)

An important result for the free electron gas model for the metal electrons was the specific heat for the Fermi gas

$$C = \frac{1}{V}\frac{\partial E}{\partial T} = \frac{g}{16m\mu^2}(2m\mu)^{5/2}T \to 0 \text{ for } T \to 0.$$

(3.3.58)

Thus it is a direct consequence of the quantum character of the electron that explains why the electrons do not contribute to specific heat of a metal at low temperatures.

We have to quantify what means "low temperatures". As we have seen for $T = 0$ the electrons fill the energy levels up to the fermi surface. For the specific heat only the excited electrons count and this means that the smallness of C is a measure for the validity of the approximations done.

The exact evaluation of $\mathcal{N}$ from (3.3.46) gives after the substitution $x = \beta p^2/(2m)$

$$\mathcal{N} = \frac{gV}{4\pi^2}\left(\frac{2m}{\beta}\right)^{3/2}\int_0^{\infty} dx \frac{\sqrt{x}}{1+\exp(x+\alpha)}$$

(3.3.59)

and thus the quantity

$$Q = \frac{4\pi^2\mathcal{N}}{gV}\left(\frac{\beta}{2m}\right)^{3/2} = \int_0^{\infty} dx \frac{\sqrt{x}}{1+\exp(x+\alpha)}$$

(3.3.60)

is a measure for the quantum behaviour of the gas, because the classical limit was shown to be given for $\alpha \to \infty$ and then Q is small while in the quantum limit, i.e., $\alpha \to -\infty$ Q becomes big. One can estimate that in the case of the conducting electrons of metals the temperature when the statistics becomes classical, i.e., for $Q \approx 1$, is $\Theta \approx 10^5 \text{K} \approx 8.6\text{eV}$.

3.3.5 The ideal Bose gas

The grand partition sum for the ideal Bose gas was calculated in (3.3.30) with the result

$$\Omega = -\sum_{\vec{n}\in\mathbb{Z}^3} \ln\left[1 - \exp\left(-\beta\frac{\vec{p}^2(\vec{n})}{2m} - \alpha\right)\right]$$

(3.3.61)

which is only sensible for $\alpha = -\mu\beta > 0$[12].

---

[12]This is true at least in the limit of an infinite volume because in this case for $\alpha < 0$ or $\mu > 0$ the integral corresponding to the sum would not exist because of the singularity at $p^2/(2m) = \mu$

As we shall see soon this restriction of the chemical potential μ to negative values gives rise to the fascinating phenomenon of *Bose-Einstein condensation*.

For sake of completeness we note that again for $\alpha \gg 0$ we find again the same classical limit as in the case of an ideal Fermi gas, namely the Boltzmann grand canonical sum (3.3.35) with the same conclusions already drawn there. This is not surprising since the classical limit means that the quantum features of being bosons or fermions become negligible because the statistical number of particles occupying one quantum state is small.

Thus we can come immediately to the quantum limit of the ideal Bose gas. For the mean particle number one finds by taking the derivative of (3.3.61)

$$\mathscr{N} = -\frac{\partial \Omega}{\partial \alpha} = g \sum_{\vec{n} \in \mathbb{Z}^3} \frac{1}{\exp\left(\beta \frac{\vec{p}^2(\vec{n})}{2m} + \alpha\right) - 1}. \tag{3.3.62}$$

Now we take the naive large volume limit where it seems that we could write an integral instead of the sum as we did in the fermionic case:

$$\mathscr{N}_{\int} = \frac{gV}{2\pi^2} \int_0^\infty \frac{P^2 \mathrm{d}P}{\exp\left(\beta \frac{P^2}{2m} + \alpha\right) - 1}. \tag{3.3.63}$$

Substitution of $x = \beta P^2/(2m)$ gives

$$\mathscr{N}_{\int} = \frac{gV}{4\pi^2} \left(\frac{(2m)}{\beta}\right)^{3/2} \int_0^\infty \frac{\mathrm{d}x \sqrt{x}}{\exp(x + \alpha) - 1} \tag{3.3.64}$$

We find that the integral exists for all $\alpha > 0$ and is monotoneously decreasing with α. Even for $\alpha = 0$ the integral exists because the integrand is going like $1/\sqrt{x}$ for $x \approx 0$ and thus the with respect to the lower boundary improper integral exists.

If $N_{\int}$ would be really the mean particle number of our gas this would mean that for a given temperature the particle number is bounded

$$\mathscr{N}_{\int} \leq \frac{gV}{4\pi^2} \left(\frac{(2m)}{\beta}\right)^{3/2} \int_0^\infty \frac{\mathrm{d}x \sqrt{x}}{\exp x - 1}. \tag{3.3.65}$$

In other words this would mean that we were not able to cool down a certain given number of particles below a certain temperature $T > 0$ which is not observed. The real bound for the temperature is $T = 0$!

To find the answer to this question we might look at the extreme limit $T = 0$. This means we have to think about the problem to put N particles in a box such that the energy becomes a minimum. But this is simple to do in the case of bosons, because all N particles are allowed to occupy the lowest energy state. This means that even a "macroscopic number" of particles can be together in this state. This also means that in the limit $T = 0$ a macroscopic system is in a pure quantum state.

Now for small $T > 0$ there also can be a huge number of particles in the lowest one particle state which is not negligible against the fluctuations of the particle number. On the other hand the only restriction for the approximation of the partition sum as an integral in the large volume limit was that the energy levels are close together. Thus we conclude that $\mathscr{N}_{\int}$ is the expectation value of the number of particles being *not* in the ground state. Thus the correct mean number of particles is

$$\mathscr{N} = \mathscr{N}_c + \mathscr{N}_{\int} = \frac{g}{\exp \alpha - 1} + \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int \mathrm{d}x \frac{\sqrt{x}}{\exp(x + \alpha) - 1} \tag{3.3.66}$$

where we have taken the contribution for the mean number of particles in the ground state $\mathscr{N}_c$ from the sum (3.3.62). This also shows that for $\alpha \gg 0$ the integral alone is a good enough approximation because then the

first summand on the right hand side, i.e., the mean number of particles in the ground state (shortly named the *condensate*) becomes negligible. So in the classical limit the condensation is negligible. On the other hand for small temperatures now the mean particle number is not bounded because the condensate term can become big for $\alpha \approx 0$ so that there is no restriction to the number of particles at low temperature.

For the grand canonical potential of course we have also to take the large volume limit with taking into account the condensate of particles in the ground state:

$$\Omega = -g \ln[1 - \exp(-\alpha)] - \frac{gV}{4\pi^2} \left(\frac{2m}{\beta}\right)^{3/2} \int_0^\infty \sqrt{x} \ln[1 - \exp(-x - \alpha)]. \tag{3.3.67}$$

3.4 Perturbation theory in the real-time formalism

In the previous section we have investigated in detail the equilibrium properties of ideal gases, i.e., non-interacting non-relativistic and relativistic identical particles in canonical or, if there are conserved charges, grand canonical ensembles.

As we know from relativistic vacuum quantum field theory the most interesting physical models are not solvable exactly. The great success of quantum field theory is thus based on perturbation theory (which has of course to be completed with renormalization theory in order to obtain finite results). Now we want to derive the perturbation theory for the case of quantum statistics.

3.4.1 The Schwinger-Keldysh time contour

In order to have something simple at hand we shall use ϕ^4-theory as a toy model. This is defined in terms of the Lagrangian

$$\mathscr{L} = \underbrace{\frac{1}{2}(\partial_\mu \phi)(\partial^\mu \phi) - \frac{m^2}{2}\phi^2}_{\mathscr{L}_0} - \underbrace{\frac{\lambda}{4!}\phi^4}_{-\mathscr{H}_I} \tag{3.4.1}$$

where ϕ is a real scalar field. As in vacuum quantum field theory we shall use the interaction picture, where the operators evolve in time with respect to $\mathbf{H}_0$. Thus the field operator's equation of motion can be solved exactly because these evolve like free fields and $\mathbf{H}_0$ is time independent. In the previous section we have shown that

$$\mathbf{H}_0 = \int \mathrm{d}^3\vec{p}\,\omega(\vec{p})\mathbf{N}(\vec{p}) \text{ with } \omega(\vec{p}) = \sqrt{\vec{p}^2 + m^2}. \tag{3.4.2}$$

The interaction part is time dependent in the interaction picture and in our case of ϕ^4-theory given as

$$\mathbf{H}_I(t) = \int \mathrm{d}^3\vec{x}\,\frac{\lambda}{4!} : \phi^4(x) : \tag{3.4.3}$$

where we have implied normal-ordering in order to renormalize the vacuum energy to 0.

Using theorem 1 of chapter 2 the statistical operator obeys the time evolution according to

$$\mathbf{R}(t) = \mathbf{C}(t, t_0)\mathbf{R}(t_0)\mathbf{C}^\dagger(t, t_0) \text{ with } \mathbf{C}(t, t_0) = T_c \exp[-\mathrm{i}\int_{t_0}^t \mathrm{d}\tau\mathbf{H}_I(\tau)]. \tag{3.4.4}$$

Now we want to calculate the expectation value of an arbitrary quantity which depends on a single time argument. This can also be a field operator, but for our arguments the dependence of these on $\vec{x}$ does not matter at all. Since such quantities are built with help of field operators we know this operator including its time dependence exactly. By definition we have

$$\langle O(t)\rangle = \mathrm{Tr}[\mathbf{R}(t)\mathbf{O}(t)]. \tag{3.4.5}$$

Using the time evolution operator (3.4.4) this reads

$$\langle O(t) \rangle = \text{Tr}[\mathbf{C}(t, t_0)\mathbf{R}(t_0)\mathbf{C}^\dagger(t, t_0)\mathbf{O}(t)]. \tag{3.4.6}$$

Since under the trace we can write the first time evolution operator to the very right this can be read as

$$\langle O(t) \rangle = \text{Tr}[\mathbf{R}(t_0)\mathbf{C}^\dagger(t, t_0)\mathbf{O}(t)\mathbf{C}(t, t_0)]. \tag{3.4.7}$$

Using the hermiticity of $\mathbf{H}_I$ and the fact that the ordering of operator products is just changed under self-adjoint conjugation we have

$$\mathbf{C}^\dagger(t, t_0) = T_a \exp[\mathrm{i}\int_{t_0}^{t} \mathrm{d}\tau \mathbf{H}_I(\tau)]. \tag{3.4.8}$$

Thereby T_a denotes the anti-causal time-ordering which orders the operators with increasing time arguments from left to right. Inserting this into (3.4.7) one sees that it can be written as a *path-ordered* product using the *Schwinger-Keldysh path* shown in figure 3.1. With this time path the (3.4.7) reads
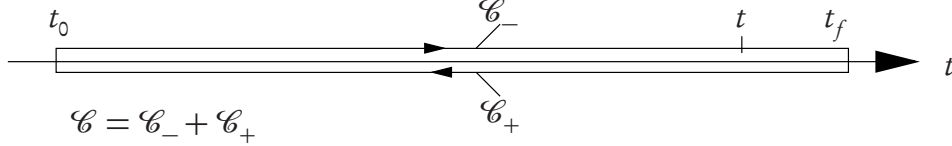

Figure 3.1: *The Schwinger-Keldysh time contour*

$$\langle O(t) \rangle = \text{Tr}\left\{ \mathbf{R}(t_0)T_{\mathscr{C}} \exp[-\mathrm{i}\int_{\mathscr{C}} \mathrm{d}\tau \mathbf{H}_I(\tau)]\mathbf{O}(t) \right\}. \tag{3.4.9}$$

Thereby $T_{\mathscr{C}}$ is the *contour ordering parameter* which orders the operators according to the time arguments on the contour from right to left. In the figure we have split the contour to the —- and +-branch. On the former the operators are time-ordered (causal) on the latter anti-time-ordered (anti-causal). All times on the upper branch are earlier than these on the lower. We have only to assure that t_0 is less and t_f is greater than the "external time" t.

In order to clarify we describe in detail how one has to read (3.4.9). Since we write from left to right we have to begin at the end of the contour. Using usual times (not the contour parameter) we have to take into account that the integration along the lower part is backwards in time. Thus (3.4.9) can be split in an integration along the lower branch where the operators are anti-time-ordered and we have an additional sign from the direction of the integration. Then we integrate in usual direction from t to t_f with time ordered operators. At this point we have to insert $\mathbf{O}(t)$ and then integrate again with time-ordered operator products from t_0 to t. As shown above the integrals along the +-branch is simply $\mathbf{C}^\dagger(t_f, t_0)$ while along the —-branch they are the $\mathbf{C}$-operator itself with the appropriate time arguments:

$$T_{\mathscr{C}} \exp[-\mathrm{i}\int_{\mathscr{C}} \mathrm{d}\tau \mathbf{H}_I(\tau)]\mathbf{O}(t) = \mathbf{C}^\dagger(t_f, t_0)\mathbf{C}(t_f, t)\mathbf{O}(t)\mathbf{C}(t, t_0). \tag{3.4.10}$$

Now we can use the composition property for the time evolution operator to cancel $\mathbf{C}^\dagger(t_f, t)$ from the +-branch against $\mathbf{C}(t_f, t)$ from the —-branch so that we really obtain (3.4.9) as claimed.

It is also clear that one can extend the contour to the left by the same arguments as before. To calculate expectation values for operator products with arbitrary time arguments it is customary to extend the contour to the whole real axis.

So far all is well defined for an arbitrary statistical operator. This first step has shown how to calculate expectation values in the interaction picture as expectation values with respect to the initial state (i.e. the statistical

operator $\mathbf{R}(t_0)$). One obtains the perturbation theory by power expansion the contour-ordered exponential function in (3.4.9) leading to diagram techniques comparable with the Feynman rules for the vacuum theory. However for a general initial state $\mathbf{R}(t_0)$ these diagram rules are very complicated since Wick's theorem does not reduce the problem to calculate contractions but one has to use a big set of free correlation functions to describe all the correlations contained in the initial statistical operator.

Fortunately this is different for the equilibrium case. Now we shall go further for this case. For our real scalar fields (i.e. self-adjoint field operators) there are no conserved charges. So we have to use the canonical ensemble. The statistical operator is

$$\mathbf{R}(t) = \frac{1}{Z}\exp[-\beta\mathbf{H}(t)] \text{ with } Z = \operatorname{Tr}\exp[-\beta\mathbf{H}(t)]. \tag{3.4.11}$$

Contrary to the Heisenberg picture description used for the ideal gases in the previous section in the here used interaction picture the statistical operator is time-dependent.

To go further we need the following

Theorem 6. *We define the operator*

$$\mathbf{U}(t, t_0) = \mathbf{A}^{\dagger}(t, t_0)\mathbf{C}(t, t_0) \tag{3.4.12}$$

where $\mathbf{A}$ *is the time evolution operator for the operators and* $\mathbf{C}$ *that for the states in the interaction picture, i.e.*

$$\mathbf{A}(t, t_0) = \exp[+\mathrm{i}(t - t_0)\mathbf{H}_0], \ \mathbf{C}(t, t_0) = T_c \exp[-\mathrm{i}\int_{t_0}^{t} \mathrm{d}\tau\mathbf{H}_I(\tau)]. \tag{3.4.13}$$

Thereby we have used the fact that $\mathbf{H}_0$, *the Hamiltonian for the free fields, is time-independent in the interaction picture.*

If $\mathbf{H}$ *is not explicitly time-dependent, then for the operator* $\mathbf{U}$ *the following holds*

$$\mathbf{U}(t, t_0) = \exp[-\mathrm{i}(t - t_0)\mathbf{H}(t_0)] \tag{3.4.14}$$

and thus the initial statistical operator is given by the analytic continuation of $\mathbf{U}$ *to complex times with negative imaginary part* $-\mathrm{i}\beta$:

$$\mathbf{R}(t_0) = \frac{1}{Z}\exp[-\beta\mathbf{H}(t_0)] = \frac{1}{Z}\mathbf{U}(t_0 - \mathrm{i}\beta, t_0) = \frac{1}{Z}\exp[-\beta\mathbf{H}_0]\mathbf{C}(t_0 - \mathrm{i}\beta, t_0). \tag{3.4.15}$$

Proof. We use theorem 1 in chapter 2 to calculate the time derivative of $\mathbf{U}$ defined in (3.4.14), the arguments in all operators are (t, t_0) so that we can omit them for sake of simplicity:

$$\partial_t \mathbf{U} = (\partial_t \mathbf{A})\mathbf{C} + \mathbf{A}\partial_t \mathbf{C} = -\mathrm{i}\mathbf{A}^{\dagger}\mathbf{H}_0\mathbf{C} - \mathrm{i}\mathbf{A}^{\dagger}\mathbf{H}_I(t)\mathbf{C}. \tag{3.4.16}$$

Inserting $\mathbf{1} = \mathbf{A}\mathbf{A}^{\dagger}$ between $\mathbf{H}_I(t)$ and $\mathbf{C}$ on the right term and using the fact that $\mathbf{A}$ is the time evolution operator for the operators in the interaction picture, i.e.,

$$\mathbf{O}(t) = \mathbf{A}\mathbf{O}(t_0)\mathbf{A}^{\dagger}. \tag{3.4.17}$$

Applied to $\mathbf{H}_I(t)$ we obtain finally

$$\partial_t \mathbf{U} = -\mathrm{i}\mathbf{H}(t_0)\mathbf{U}. \tag{3.4.18}$$

By definition of $\mathbf{A}$ and $\mathbf{C}$ the initial condition for $\mathbf{U}$ is $\mathbf{U}(t_0, t_0) = 1$. Since $\mathbf{H}$ is supposed to be not explicitly time-dependent $\mathbf{H}(t_0)$ is constant in time and thus the solution of the initial value problem for $\mathbf{U}$ given by (3.4.18) is simply

$$\mathbf{U}(t, t_0) = \exp[-\mathrm{i}(t - t_0)\mathbf{H}(t_0)] \tag{3.4.19}$$

and inserting $t = t_0 - \mathrm{i}\beta$ yields (3.4.15) which is crucial for perturbation theory at finite temperature in both the imaginary- and the real-time formalism. Q.E.D.

Inserting (3.4.15) into (3.4.9) we obtain

$$\langle O(t) \rangle = \text{Tr} \left\{ \exp[-\beta \mathbf{H}_0] \mathbf{C}(t_0 - \mathrm{i}\beta, t_0) T_{\mathscr{C}} \exp[-\mathrm{i} \int \mathrm{d}\tau \mathbf{H}_I(\tau)] \mathbf{O}(t) \right\} \tag{3.4.20}$$

This formula shows immediately that we can write this in terms of the expectation value of a path-ordered integral along the Schwinger-Keldysh-path extended with a piece running vertically down from t_0 to $t_0 - \mathrm{i}\beta$ in the complex time plane (see 3.2). The expectation value of this path-ordered integral has to be taken *with respect to the canonical statistical operator for an ideal gas*. As we shall see because of the fact that $\mathbf{H}_0$ is bilinear
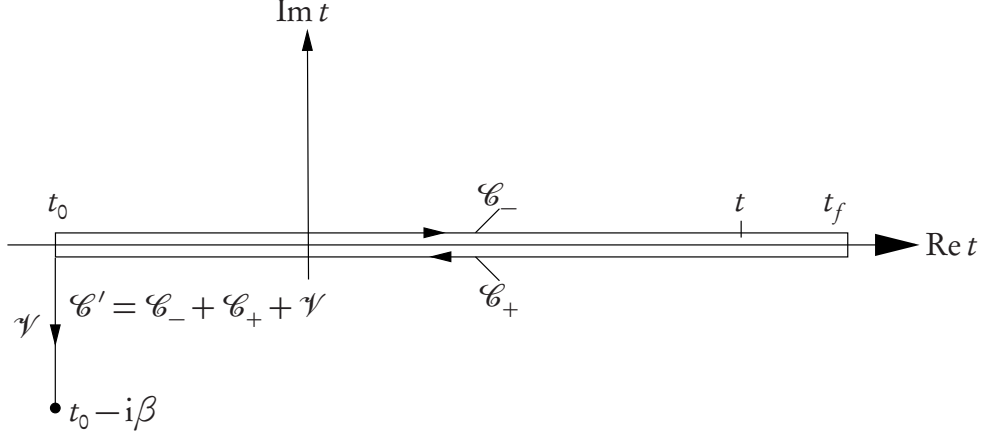

Figure 3.2: *The extended Schwinger-Keldysh time contour for thermal equilibrium*

in the fields (i.e. a one-particle operator) for this statistical operator Wick's theorem holds for the expectation values in the usual form. This means that the Feynman rules can be taken from the vacuum case with the following extensions: One has to use integrals along the extended Schwinger-Keldysh time path instead of time integrals in the vacuum theory and instead of contractions with the vacuum state one has to use thermal averages of contour-ordered products of two field operators with respect to the ideal gas statistical operator. Different to the vacuum case is of course that normal-ordered products do not vanish when calculating these averages.

We shall also see further on that in many cases the vertical part of the contour drops in the calculations and we have to calculate the integrals only along the real-time path (3.1).

To show all this we shall use the rest of this section.

3.4.2 The Dyson-Wick series

As we shall see in the following chapters physical information about an interacting system can completely be extracted with help of the Green's functions as is also the case for the vacuum theory. The main difference to vacuum theory is that in the case of a thermal background it does not make any sense to speak about *asymptotic states*. Thus the physical quantities are not S-matrix elements but, for instance, the reaction of the system to an external perturbation or the production rate of particles out of a heat bath (to speak in equilibrium language).

The n-point Green's functions we need to calculate such quantities perturbatively are defined as the thermal averages of time contour-ordered field operator products:

$$\mathrm{i}G^{(n)}(x_1, x_2, \ldots, x_n) = \langle T_{\mathscr{C}'} \phi(x_1) \phi(x_2) \ldots \phi(x_n) \rangle. \tag{3.4.21}$$

Using (3.4.20) we can write this as

$$\mathrm{i}G^{(n)}(x_1, \ldots, x_n) = \frac{1}{Z} \text{Tr} \left\{ \exp(-\beta \mathbf{H}_0) T_{\mathscr{C}'} \exp[-\mathrm{i} \int_{\mathscr{C}'} \mathrm{d}\tau \mathbf{H}_I(\tau)] \phi(x_1) \cdots \phi(x_n) \right\}. \tag{3.4.22}$$

These are the exact Green's functions and cannot be calculated for the most of the physically interesting interacting field theories as is even the case for vacuum quantum field theory. Thus we have to power expand the exponential function in order to obtain a perturbation theory in powers of the coupling constant.

Doing so we see that the main task is to calculate the statistical average of time-ordered products of normal ordered local field operator products[13] with respect to the free Green's function.

In order to find the Feynman rules to manage this calculations in a pictorial way (as in the vacuum case) we have to apply Wick's theorem. The operator version is proven in QFT 3.8. The only difference at the operator level is that in the definition of contractions instead of time ordered field operator products our contour ordered products enter:

$$\overset{\centerdot}{\phi}(x_1)\overset{\centerdot}{\phi}(x_2) = \langle 0 | T_{\mathscr{C}'} \phi(x_1)\phi(x_2) | 0 \rangle. \tag{3.4.23}$$

With this redefinition we can write down Wick's theorem in the same form as for the vacuum case:

$$
\begin{aligned}
T_{\mathscr{C}} \mathbf{UV} \cdots \mathbf{XYZ} \;=\; & \underbrace{: \mathbf{UV} \cdots \mathbf{XYZ} :}_{\text{normal-ordered product without contractions}} & + \\[2mm]
+ \;\; & \underbrace{: \mathbf{U}^{\centerdot}\mathbf{V}^{\centerdot}\mathbf{W} \cdots \mathbf{XYZ} : + : \mathbf{U}^{\centerdot}\mathbf{V}\mathbf{W}^{\centerdot} \cdots \mathbf{XYZ} : + \cdots}_{\text{sum of all normal-ordered products together with one contraction}} & + \\[2mm]
+ \;\; & \;\vdots & \\
+ \;\; & \underbrace{\mathbf{U}^{\centerdot}\mathbf{V}^{\centerdot\centerdot}\mathbf{W}^{\centerdot\centerdot\centerdot} \cdots \mathbf{X}^{\centerdot\centerdot\centerdot}\mathbf{Y}^{\centerdot\centerdot}\mathbf{Z}^{\centerdot} + \cdots}_{\text{sum over all possible total contracted operator pairs}} & .
\end{aligned} \tag{3.4.24}
$$

It is clear that in the case that we have a product with an odd number of operators in the last line there is just one operator left, in the case of an even number of operators the last line is the product of contracted pairs, which is a c-number.

Now the problem is to show that the average of a normal-ordered product of field operators with respect to the free Hamiltonian is a sum of products of averages for the normal-ordered product of two field operators. For this it is sufficient to show that this is the case for the normal-ordered product of annihilation and creation operators. But this is proven by writing out the trace in terms of the number operator basis for bosons (3.3.22):

$$\mathrm{Tr} \exp(-\beta \mathbf{H}_0) : \mathbf{AB} \cdots := \sum_{N(\vec{p})=0}^{\infty} \exp[-\beta \sum_{\vec{p}} N(\vec{p}) \omega(\vec{p})] \langle N(\vec{p}) | : \mathbf{ABC} \cdots : | N(\vec{p}) \rangle. \tag{3.4.25}$$

Thereby $\mathbf{A}$, $\mathbf{B}$, ... are arbitrary annihilation and creation operators, and we have used (3.4.2) for the case of a finite quantization volume with periodic boundary conditions (of course we shall go to the infinite volume limit at the end of the calculation).

Due to the definition of the trace operator the normal ordered product is "sandwiched" between the same states, i.e. we have to calculate the diagonal elements only. Since we have a normal ordered product we act on the ket of this matrix element with annihilation operators lowering the particle number with the given momentum. The creation operators can be seen to act to the left on the bra as annihilation operators also lowering the particle numbers. Since bra and ket are the same when calculating the trace the result is only different from 0 if one has the same creation and annihilation operators.

---

[13]The normal-ordering comes from the renormalization of the free vacuum explained in detail in QFT section 3.4. For ϕ^4-theory the interaction is $\mathbf{H}_I = \int \mathrm{d}^3 \vec{x} \, \lambda/4! : \phi^4(x):$

The most simple example is the thermal average of a normal ordered product of two operators:

$$Z_0 \left\langle \mathbf{a}^\dagger(\vec{p}_1)\mathbf{a}(\vec{p}_2) \right\rangle_0 = \sum_{N(\vec{p})=0}^{\infty} \exp\left[-\beta \sum_{\vec{p}} N(\vec{p})\omega(\vec{p})\right] \left\langle N(\vec{p}) \left| \mathbf{a}^\dagger(\vec{p}_1)\mathbf{a}(\vec{p}_2) \right| N(\vec{p}) \right\rangle$$

$$= \delta^{(3)}(\vec{p}_1 - \vec{p}_2) \sum_{N(\vec{p})=0} N(\vec{p}_1) \exp\left[-\beta \sum_{\vec{p}} N(\vec{p})\omega(\vec{p})\right]. \tag{3.4.26}$$

This sum can be calculated if we read the bosonic partition sum as a functional of $\omega(\vec{p})$:

$$\ln Z_0[\omega(\vec{p})] = \sum_{\vec{p}} \ln\left\{1 - \exp[-\beta\omega(\vec{p})]\right\}. \tag{3.4.27}$$

Then we find

$$\left\langle \mathbf{a}^\dagger(\vec{p}_1)\mathbf{a}(\vec{p}_2) \right\rangle_0 = -\delta^{(3)}(\vec{p}_1 - \vec{p}_2)\frac{1}{\beta}\frac{\delta}{\delta\omega(\vec{p}_1)}\ln Z_0 = \delta^{(3)}(\vec{p}_1 - \vec{p}_2)n[\omega(\vec{p}_1)] \tag{3.4.28}$$

where

$$n(x) = \frac{1}{\exp(\beta|x|) - 1} \tag{3.4.29}$$

is the *Bose distribution function*.

From the orthonormality of $\{|N(\vec{p})\rangle\}$ we deduce immediately that of course the expectation values of normal-ordered products exhibit also a pairing structure like the vacuum expectation values. Instead of the contractions one has to write expressions of the form (3.4.28). Thus the vacuum expectation value of a normal-ordered product of field operators is the sum over all possible "thermal pairings". But a glance on (3.4.24) shows that these sum combines with the contractions.

This means that for calculating n-point functions we can apply Wick's theorem as in vacuum quantum field theory but we have to use the free contour Green's function

$$i\Delta_{\mathscr{C}'}(x_1, x_2) = \left\langle T_{\mathscr{C}'} \phi(x_1)\phi(x_2) \right\rangle_0 \tag{3.4.30}$$

instead of the Feynman propagator. To prove this one has only to apply the thermal Wick theorem to (3.4.22) and use the arguments of QFT 3.9 to show that the partition sum Z drops and one has only to sum over all connected and disconnected diagrams which are linked to the external points (*linked cluster theorem*), i.e., the Dyson-Wick series for n-point functions is given by

$$iG^{(n)}(x_1, \ldots, x_n) =$$

$$= \sum_{j=0}^{\infty} \frac{1}{j!} \left\langle T_{\mathscr{C}'} \int_{\mathscr{C}'} \mathrm{d}^4 y_1 \cdots \mathrm{d}^4 y_j \left(\frac{-i\lambda}{4!}\right)^j : \phi^4(y_1) : \ldots : \phi^4(y_j) : \phi(x_1) \cdots \phi(x_n) \right\rangle_0^{(1)} \tag{3.4.31}$$

where we have used the notation of QFT section 3.9. $\langle \cdots \rangle_0^{(1)}$ means that one has to take the expectation value according to the Feynman rules but leaving out all diagrams which contain closed sub-diagrams (i.e. such sub-diagrams which are not linked to at least one of the external points $x_1, \ldots, x_n$).

For sake of clarity we summarize the *thermal Feynman rules for ϕ^4-theory*:

There are the following graphical elements:

With these graphical elements one calculates a contribution to $iG^{(n)}(x_1, x_2, \ldots, x_n)$ as follows: A diagram of such a contribution consists of n external points labelled with $x_1, \ldots, x_n$ and k vertices. Only such diagrams contribute where each vertex is linked at least to one external point. To obtain the full contribution to the n-point function one has to sum over all possible topologies of such diagrams if one obeys the following rules for extracting the analytical expression out of a given diagram:
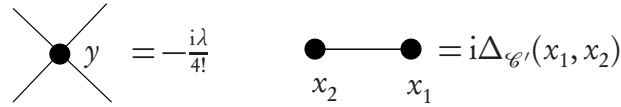
$$\times \!\!\!\! \bullet \, y \;=\; -\frac{\mathrm{i}\lambda}{4!} \qquad\qquad \underset{x_2 \qquad\;\; x_1}{\bullet\!\!-\!\!\!-\!\!\bullet} \;=\; \mathrm{i}\Delta_{\mathscr{C}'}(x_1, x_2)$$

Figure 3.3: *Graphical elements for the Feynman diagrams of ϕ^4-theory for the quantum field theory at finite temperature*

1. Write down a factor $-\mathrm{i}\lambda/4!$ for each vertex and a factor $1/k!$ for a diagram with k vertices.[14]

2. Count the number of ways to connect the vertices with the external points and the internal points giving the diagram and multiply the analytical expression with this number.

3. Each line connecting two points, internal or external, stands for a contour Green's function $\mathrm{i}\Delta_{\mathscr{C}'}$. For a line beginning and ending at the same point in space and time (*tadpole contribution*) there is an extra rule to avoid tadpole singularities due to normal-ordering given in the next subsection.

4. Integrate over the internal points. Thereby each time-integral stands for a integration along the contour $\mathscr{C}'$.

3.4.3 The free contour Green's functions

Now we derive the free thermal propagator which we need for calculating perturbatively the n-point functions.

There are two different possibilities to find this propagator. One is to use directly the plane wave representation of the field operator (see QFT (3.74)):

$$\phi(x) = \int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{p}}{\sqrt{2\omega(\vec{p})(2\pi)^3}} [\mathbf{a}(\vec{p})\exp(-\mathrm{i}px) + \mathbf{a}^\dagger(\vec{p})\exp(\mathrm{i}px)]_{p_0=\omega(\vec{p})} \text{ with } \omega(\vec{p}) = \sqrt{\vec{p}^2 + m^2} \qquad (3.4.32)$$

and the number eigenbasis (3.3.22) to calculate the trace in

$$\mathrm{i}\Delta_{\mathscr{C}'}(x, y) = \langle T_{\mathscr{C}'}\phi(x)\phi(y)\rangle_0. \qquad (3.4.33)$$

The reader is warmly invited to do this calculation himself as an exercise.

Here we will use the other method: Firstly we prove some properties of the propagator (3.4.33) then we shall use the free Klein-Gordon equation for the field operator

$$(\Box + m^2)\phi(x) = 0 \qquad (3.4.34)$$

and the equal-time commutation relation from canonical quantization

$$\phi(t, \vec{x})\overset{\leftrightarrow}{\partial}_t \phi(t, \vec{y}) = \mathrm{i}\delta^{(3)}(\vec{x} - \vec{y}) \qquad (3.4.35)$$

to derive equations of motion for the free thermal Green's function (3.4.33) and solve these with help of the properties we shall derive now directly from its definition.

From translation invariance of the equilibrium state and the Lagrangian one sees immediately that the Green's function depends only on the difference of its arguments. Thus we write

$$\mathrm{i}\Delta_{\mathscr{C}'}(x) = \langle T_{\mathscr{C}'}\phi(x)\phi(0)\rangle, \; \Delta_{\mathscr{C}'}(x, y) = \Delta_{\mathscr{C}'}(x - y). \qquad (3.4.36)$$

---

[14]The first factor comes from the interaction Lagrangian, the second from the Dyson-Wick series

To prove this one has just to use (3.4.32) and take into account that $\langle \mathbf{aa} \rangle_0 = \langle \mathbf{a^\dagger a^\dagger} \rangle_0 = 0$ for any momentum arguments and that the expectation value of a product of an annihilation with a creation operator contains always a δ-distribution for the momenta.

The next property of the propagator is known as the *Kubo-Martin-Schwinger condition for bosons* often abbreviated as *KMS-condition*:

$$\Delta_\mathscr{C}(t - \mathrm{i}\beta, \vec{x}) = \Delta_{\mathscr{C}'}(t, \vec{x}). \tag{3.4.37}$$

This is proven by using the cyclic commutability of operator products under the trace and also under the contour-ordering operator:

$$
\begin{aligned}
\mathrm{i}\Delta_{\mathscr{C}'}(t - \mathrm{i}\beta, \vec{x}) &= \langle T_{\mathscr{C}'} \boldsymbol{\phi}(t - \mathrm{i}\beta, \vec{x}) \boldsymbol{\phi}(0) \rangle_0 = \\
&= \langle T_{\mathscr{C}'} \exp(\beta \mathbf{H}_0) \boldsymbol{\phi}(t, \vec{x}) \exp(-\beta \mathbf{H}_0) \boldsymbol{\phi}(0) \rangle_0 = \\
&= \frac{1}{Z_0} \mathrm{Tr}\{\exp(-\beta \mathbf{H}_0) T_{\mathscr{C}'} \boldsymbol{\phi}(t, \vec{x}) \boldsymbol{\phi}(0)\} = \mathrm{i}\Delta_{\mathscr{C}'}(x).
\end{aligned}
\tag{3.4.38}
$$

This extends the definition range of the Green's function from $0 \leq -\mathrm{Im}\, t \leq \beta$ to the whole complex plane, but as we shall see the Green's function is an analytic function only in the open strip given by the original range. It is these analytic structure we are after when calculating the Green's functions with the equation of motion method!

For this purpose it is more convenient to treat the time argument on various parts of the contour separately. We start with the vertical part. In the following we write $-$, $+$ or V to label where the time arguments are located. The point $t = 0$, which we chose as the t_0 of the general derivation of the contour method, lies on all three parts of the contour by definition.

We start with times on the vertical part of the contour. It is convenient to switch to the parameterization $t = -\mathrm{i}\tau$, where $\tau \in (0, \beta)$. This means switching to the *imaginary-time formalism*. Now the box operator reads

$$\Box = -\partial_\tau^2 - \nabla^2 = -\Box_E. \tag{3.4.39}$$

Herein $\Box_E$ stands for the four dimensional Laplacian. The E refers to *Euclidean space time*. But one should keep in mind that the τ is restricted to $(0, \beta)$ and the Green's function has to be continued periodically due to the KMS-condition (3.4.37).

Now we write for $\tau \in \mathbb{R}$:

$$\mathrm{i}\Delta_V(x_E) = \langle T_\tau \boldsymbol{\phi}(x_E) \boldsymbol{\phi}(0) \rangle_0 \quad \text{with } x_E = (-\mathrm{i}\tau, \vec{x}). \tag{3.4.40}$$

This can be written in the form

$$\mathrm{i}\Delta_V(x_E) = \langle \boldsymbol{\phi}(0)\boldsymbol{\phi}(x_E) + [\boldsymbol{\phi}(x_E), \boldsymbol{\phi}(0)]\Theta(\tau) \rangle_0. \tag{3.4.41}$$

Applying (3.4.34) and $\mathrm{d}_t\Theta(\tau) = \delta(\tau)$ and the equal-real-time commutation relation (3.4.35) to this equation we find the equation of motion for the Green's function

$$(\Box_E - m^2)\Delta_V(x_E) = +\mathrm{i}\delta^{(4)}(x_E). \tag{3.4.42}$$

This is together with the boundary conditions the Green's function for a free Klein-Gordon field in Euclidean space-time. From this equation follows that $\Delta_V(x_E)$ is continuous at $x_E = 0$, while integration over a small τ-interval $(-\epsilon, \epsilon)$ and letting $\epsilon \to 0$ gives $\partial_\tau \Delta(+0, \vec{x}) - \partial_\tau(-0, \vec{x}) = \mathrm{i}$.

In the interval $(0, \beta)$ the function is analytic in τ because there the operators have a fixed order in (3.4.41). Thus we shall integrate (3.4.42) in this interval. Due to the KMS-Condition the boundary conditions for this region read

$$\Delta_V(+0, \vec{x}) = \Delta_V(\beta - 0, \vec{x}), \quad \partial_\tau \Delta_V(+0, \vec{x}) - \partial_\tau \Delta_V(\beta - 0, \vec{x}) = \mathrm{i}. \tag{3.4.43}$$

This equations can be simply integrated taking the Fourier transform with respect to $\vec{x}$:

$$\Delta_V(x_E) = \int \frac{d^3\vec{p}}{(2\pi)^3} \exp(i\vec{p}\vec{x}) \tilde{\Delta}_V(\tau, \vec{p}) \tag{3.4.44}$$

and (3.4.42) reads

$$[\partial_\tau^2 - \omega^2(\vec{p})]\tilde{\Delta}_V(\tau, \vec{p}) = i\delta(\tau). \tag{3.4.45}$$

For $\tau \in (0, \beta)$ the right-hand-side vanishes and the general solution of the equation is

$$\tilde{\Delta}_V(\tau, \vec{p}) = A\exp[\omega(\vec{p})\tau] + B\exp[-\omega(\vec{p}\tau)]. \tag{3.4.46}$$

The boundary conditions (3.4.43) finally determine the constants A and B:

$$\tilde{\Delta}_V(\tau, \vec{p}) = -\frac{i}{2\omega(\vec{p})}\left\{ n[\omega(\vec{p})]\exp[\omega(\vec{p})\tau] + [1 + n[\omega(\vec{p})]]\exp[-\omega(\vec{p})\tau] \right\} \text{ for } \tau = it \in (0, \beta). \tag{3.4.47}$$

This representation is known as the *Mills representation* for the propagator.

The periodic continuation can be obtained using a *Fourier series representation*:

$$\tilde{\Delta}_V(\tau, \vec{p}) = \frac{1}{\beta} \sum_{-\infty}^{\infty} \exp(-i\omega_n\tau) \Delta_V(\omega_n, \vec{p}) \text{ with } \omega_n = \frac{2\pi n}{\beta}. \tag{3.4.48}$$

The ω_n are known as the *Matsubara frequencies for bosons*. With help of (3.4.47) one finds

$$\Delta_V(\omega_n, \vec{p}) = \int_0^\beta d\tau \exp(i\omega_n\tau)\tilde{\Delta}_V(\tau, \vec{p}) = -\frac{i}{\omega_n^2 + \vec{p}^2 + m^2}. \tag{3.4.49}$$

This looks like a Euclidean propagator but with discrete energies ω_n.

For time arguments on the real time axis we start with the analytic Wightman functions

$$i\Delta^{+-}(x) = \langle\phi(x)\phi(0)\rangle, \; i\Delta^{-+}(x) = \langle\phi(x)\phi(0)\rangle \tag{3.4.50}$$

which shows that Δ^{+-} is the analytic continuation of Δ_V to real values for the time argument. Thus for the Mills representation we have only to set $\tau = it$ in (3.4.47):

$$i\tilde{\Delta}(t, \vec{p}) = \frac{1}{2\omega}\left\{ n(\omega)\exp(i\omega t) + [1 + n(\omega)]\exp(-i\omega t) \right\} \tag{3.4.51}$$

where we use ω as abbreviation for $\omega(\vec{p}) = \sqrt{\vec{p}^2 + m^2}$.

The Fourier transform of this is

$$\begin{aligned} i\Delta^{+-}(p) &= \int dt\, i\tilde{\Delta}^{+-}(t, \vec{p})\exp(ip_0 t) = \\ &= \frac{2\pi}{2\omega}\left\{ n(\omega)\delta(\omega + p_0) + [1 + n(\omega)]\delta(\omega - p_0) \right\} = \\ &= 2\pi\delta(p^2 - m^2)[\Theta(p_0) + n(\omega)]. \end{aligned} \tag{3.4.52}$$

From $\Delta^{-+}(x) = \Delta^{+-}(-x)$ we find immediately

$$i\Delta^{-+}(p) = 2\pi\delta(p^2 - m^2)[\Theta(-p_0) + n(\omega)]. \tag{3.4.53}$$

For the calculation of the time-ordered propagator

$$\tilde{\Delta}^{--}(t,\vec{p}) = \Theta(t)\tilde{\Delta}^{+-}(t,\vec{p}) + \Theta(-t)\tilde{\Delta}^{-+}(t,\vec{p}) \tag{3.4.54}$$

we use the Fourier transform of products

$$\int \mathrm{d}t\, f(t)g(t)\exp(\mathrm{i}t\,p_0) = [f * g](p_0) \tag{3.4.55}$$

which is a convolution integral:

$$[f * g](l_0) = \int \mathrm{d}t \int \frac{\mathrm{d}p_0}{2\pi} \int \frac{\mathrm{d}k_0}{2\pi} \tilde{f}(p_0)\tilde{g}(k_0)\exp[-\mathrm{i}t(p_0 + k_0 - l_0)] = \int \frac{\mathrm{d}k_0}{2\pi}\tilde{f}(l_0 - k_0)\tilde{g}(k_0). \tag{3.4.56}$$

Formal integration by parts (using $\Theta' = \delta$ in the sense of distributions) leads to the Fourier transform of the Θ-distribution:

$$\tilde{\Theta}(p_0) = \int \mathrm{d}t\,\Theta(t)\exp(\mathrm{i}t\,p_0) = \frac{\mathrm{i}}{p_0} \tag{3.4.57}$$

but this has not a proper meaning without defining what is meant by integrating this over p_0 with a test function. To define this properly it is most convenient to see it as the weak limit of an analytic function. For that purpose define

$$\tilde{\Theta}_\epsilon(p_0) = \frac{\mathrm{i}}{p_0 + \mathrm{i}\epsilon} \text{ with } \epsilon > 0. \tag{3.4.58}$$

The Fourier transform of this function, which is for each $\epsilon > 0$ a regular function in a neighbourhood of the real axis, can be obtained by closing the integration path with a big half circle in the lower (for $t > 0$) or the upper (for $t < 0$) half of the complex p_0-plane. Using the residuum theorem one finds

$$\Theta_\epsilon(t) = \mathrm{i}\int \frac{\mathrm{d}p_0}{2\pi}\Theta(p_0)\exp(-\mathrm{i}t\,p_0) = \Theta(t) \tag{3.4.59}$$

which shows that for $\epsilon \to +0$ in the weak sense the regularization (3.4.58) for (3.4.57) gives the correct result. We write (3.4.54) in the form

$$\tilde{\Delta}^{--}(t,\vec{p}) = \underbrace{\Theta(t)[\tilde{\Delta}^{+-}(t,\vec{p}) - \tilde{\Delta}^{-+}(t,\vec{p})]}_{\tilde{\Delta}^R(t,\vec{p})} + \tilde{\Delta}^{-+}(t,\vec{p}) \tag{3.4.60}$$

and perform the Fourier transformation of the *retarded propagator* with help of (3.4.56), (3.4.52-3.4.53) and (3.4.58):

$$\mathrm{i}\Delta^R(p) = \mathrm{i}[\Theta * (\Delta^{+-} - \Delta^{-+})](p) = \frac{\mathrm{i}}{p^2 - m^2 + \mathrm{i}\epsilon\, \mathrm{sign}\, p_0}. \tag{3.4.61}$$

Thus we have finally

$$\mathrm{i}\Delta^{--}(p) = \frac{\mathrm{i}}{p^2 - m^2 + \mathrm{i}\epsilon\, \mathrm{sign}\, p_0} + 2\pi\delta(p^2 - m^2)[\Theta(-p_0) + n(\omega)]. \tag{3.4.62}$$

The last propagator is

$$\tilde{\Delta}^{++}(t,\vec{p}) = \Theta(-t)\tilde{\Delta}^{+-}(t,\vec{p}) + \Theta(t)\tilde{\Delta}^{-+}(t,\vec{p}) = \tilde{\Delta}^{+-}(t,\vec{p}) - \tilde{\Delta}^R(t,\vec{p}) \tag{3.4.63}$$

and its Fourier transform

$$\mathrm{i}\Delta^{++}(p) = -\frac{\mathrm{i}}{p^2 - m^2 + \mathrm{i}\epsilon\, \mathrm{sign}\, p_0} + 2\pi\delta(p^2 - m^2)[\Theta(p_0) + n(\omega)]. \tag{3.4.64}$$

It should be emphasized that for some calculations it is important to be careful with the δ-distribution arising in the free Green's functions. The problems become obvious when a product of δ-distributions at the same point occur in a calculation. In such cases our calculations show that one should use the following regularization for the δ-distribution:

$$2\pi \mathrm{i}\delta_\epsilon(x) = \frac{1}{x-\mathrm{i}\epsilon} - \frac{1}{x+\mathrm{i}\epsilon} = \frac{2\mathrm{i}\epsilon}{x^2+\epsilon^2}, \text{ with } \epsilon > 0. \tag{3.4.65}$$

The reader should show that the weak limit $\epsilon \to +0$ is really the δ-distribution.

To see that this is the correct regularization for the Green's functions the reader should calculate $\Delta^{+-}(p)$ for $t = x_0 - \mathrm{i}\epsilon$ with $\epsilon > 0$ which uses the fact that Δ^{+-} is the analytic continuation of Δ_V as done above. This calculation shows that one also has to substitute p_0 instead of ω in the Bose factors n.

The regulated free Green's functions read:

$$
\begin{aligned}
\mathrm{i}\Delta^{--}(p) &= \frac{\mathrm{i}}{p^2-m^2+\mathrm{i}\epsilon\,\mathrm{sign}\,p_0} + 2\pi\delta_\epsilon(p^2-m^2)[\Theta(-p_0)+n(p_0)], \\
\mathrm{i}\Delta^{++}(p) &= \frac{-\mathrm{i}}{p^2-m^2+\mathrm{i}\epsilon\,\mathrm{sign}\,p_0} + 2\pi\delta_\epsilon(p^2-m^2)[\Theta(p_0)+n(p_0)], \\
\mathrm{i}\Delta^{-+}(p) &= 2\pi\delta_\epsilon(p^2-m^2)[\Theta(-p_0)+n(p_0)], \\
\mathrm{i}\Delta^{+-}(p) &= 2\pi\delta_\epsilon(p^2-m^2)[\Theta(p_0)+n(p_0)].
\end{aligned}
\tag{3.4.66}
$$

The last problem we have to solve is to calculate the thermal contraction of two fields at the same space-time point. Because this problem arises in the Feynman rules if we contract two field operators as they are written in the Hamiltonian we have to normal-order them:

$$\Delta_{\mathrm{TP}} = \left\langle :\boldsymbol{\phi}^2(x): \right\rangle_0. \tag{3.4.67}$$

For this purpose we have to use (3.4.32) because normal-ordering is defined in terms of annihilation and creation operators. The calculation is straight-forward and left as an exercise for the reader. The last step is to use (3.4.28) to obtain

$$\mathrm{i}\Delta_{\mathrm{TP}} = 2\int_{\mathbb{R}^3} \frac{\mathrm{d}^3\vec{p}}{\sqrt{2\omega(\vec{p})(2\pi)^3}} n[\omega(\vec{p})] = \int \frac{\mathrm{d}^4x}{(2\pi)^4} n(p_0) 2\pi\delta(p^2-m^2). \tag{3.4.68}$$

Thus the regulated form in Fourier space is

$$\mathrm{i}\Delta_{\mathrm{TP}} = 2\pi\delta_\epsilon(p^2-m^2)n(p_0). \tag{3.4.69}$$

3.5 The generating functional

The perturbation theory can also be formulated in terms of generating functionals. This is an important tool for quantum field theory. In QFT chapter 4.5 we have used the path integral methods to find the generating functional. Now we shall derive the same for thermal quantum field theory with help of the operator formalism. The functional method has the advantage that one can derive certain classes of diagrams (for instance the connected Green's functions or the generating functional for the proper vertex functions).

The idea bases on theorem 4 in chapter 2.3. We introduce a c-number external current J and define the vacuum expectation value

$$Z[J] = Z\left\langle T_{\mathscr{C}'} \exp\left[\mathrm{i}\int_{\mathscr{C}'} \mathrm{d}^4 J(x)\boldsymbol{\phi}(x)\right]\right\rangle. \tag{3.5.1}$$

With help of theorem 4 (ch. 2.3) and the definition of the Green's function (3.4.21) it is immediately clear that we can express these functions as functional derivatives with respect to J:

$$iG^{(n)}(x_1,\ldots,x_n) = \langle T_{\mathscr{C}'}\phi(x_1)\cdots\phi(\mathbf{x_n})\rangle = \frac{1}{i^n}\frac{1}{Z}\frac{\delta^n Z[J]}{\delta J(x_1)\cdots\delta J(x_n)}\bigg|_{J=0} \tag{3.5.2}$$

and the Dyson-Wick series (3.4.22) can be summarized in the important formula:

$$Z[J] = \exp\left\{-i\int d^4x\,\mathscr{H}_I\left[\frac{\delta}{\delta J(x)}\right]\right\}Z_0[J] \tag{3.5.3}$$

where $Z_0[J]$ is the generating functional for the free theory.

Fortunately we can calculate this generating functional exactly. It is given by

$$Z_0[J] = \text{Tr}\left\{\exp(-\beta\mathbf{H}_0)T_{\mathscr{C}'}\exp\left[i\int_{\mathscr{C}'} d^4x J(x)\phi(x)\right]\right\}. \tag{3.5.4}$$

We can apply directly the thermal Wick's theorem.

First we expand the exponential

$$Z_0[J] = Z_0\left\langle T_{\mathscr{C}'}\exp\left[i\int_{\mathscr{C}'}d^4x J(x)\phi(x)\right]\right\rangle_0 =$$
$$= Z_0\left\langle T_{\mathscr{C}'}\left(1 + \frac{i^2}{2!}\langle J_1 J_2\phi_1\phi_2\rangle_{12} + \frac{i^4}{4!}\langle J_1\cdots J_4\phi_1\cdots\phi_4\rangle_{1\ldots4} + \cdots\right)\right\rangle_0 \tag{3.5.5}$$

where we have used the fact that the thermal expectation value of an odd number of field operators vanishes due to Wick's theorem. We have also introduced the shorthand notation

$$\int_{\mathscr{C}'} d^4x_1 d^4x_2\cdots d^4x_n f(x_1,x_2,\ldots,x_n) = \langle f_{12\ldots n}\rangle_{12\ldots n} \tag{3.5.6}$$

which is convenient for the functional methods.

Since the expression in the expectation value of each term is completely symmetric in the space-time points we have to integrate over we have just to count the number of contractions:

$$iG_0^{(2k)}(x_1,\ldots,x_{2k}) = \langle T_{\mathscr{C}'}\phi_1\phi_2\cdots\phi_{2k}\rangle_0. \tag{3.5.7}$$

We denote the number of possible contractions of $2k$ field operators with $P(2k)$. For contracting ϕ_1 there are $2k-1$ possible partners. For each of these possibilities there are $P[2(k-1)]$ contractions for the other operators. Thus we get

$$P(2k) = (2k-1)P[2(k-1)] \Rightarrow P(2k) = (2k-1)!! := 3\cdot5\cdot7\cdots(2k-1). \tag{3.5.8}$$

Inserting this into (3.5.5) we obtain

$$Z_0[J] = Z_0\sum_{k=0}^{\infty}\left[\frac{i^{2k}}{(2k)!}(2k-1)!!(i\langle\Delta_{12}J_1 J_2\rangle_{12})^k\right] \tag{3.5.9}$$

where we have written $\Delta_{12} = \Delta(x_1-x_2)$ for convenience. Now the series can be rewritten in the form

$$Z_0[J] = Z_0\sum_{k=0}^{\infty}\left[\frac{1}{k!}\left(-\frac{i}{2}\langle\Delta_{12}J_1 J_2\rangle_{12}\right)^k\right] = Z_0\exp\left(-\frac{i}{2}\langle J_1\Delta_{12}J_2\rangle_{12}\right). \tag{3.5.10}$$

The reader is warmly invited to re-derive the Feynman rules for ϕ^4-theory with help of this result making use of (3.5.3).

Comparing this with the case of vacuum quantum field theory in QFT chapter 4.5 we see that the theorems 2 and 3 about the generating functionals for connected Green's functions and the one-particle irreducible (1PI) amputated diagrams (also called proper vertex functions) hold because these were derived independent from the use of path integral methods. The only input is the generating functional Z which is only different from the vacuum case by the fact that instead of time-ordering we use the ordering along the extended Schwinger-Keldysh path $\mathscr{C}'$.

Thus we can immediately write down the results of these theorems. The interested reader may look in QFT chapter 4.6 for the proofs!

Theorem 7 (Connected Green's Functions). *The functional*

$$\mathrm{i}W[J] = \ln\{Z[J]\} \Longleftrightarrow Z[J] = \exp\{\mathrm{i}W[J]\} \tag{3.5.11}$$

is the generating functional for connected Green's functions:

$$G_c^{(n)}(x_1,\ldots,x_n) = \left(\frac{1}{\mathrm{i}}\right)^n \left.\frac{\delta^n W[J]}{\delta J(x_1)\cdots\delta J(x_n)}\right|_{J=0}. \tag{3.5.12}$$

These are depicted as diagrams with n external points which do not split in the product of disconnected pieces.

Clearly this is a sensible subclass of diagrams since the Green's functions can be written as sums over products of connected Green's functions.

Theorem 8 (The Effective Action). *By performing a functional Legendre transformation of the generating functional of the connected Green's functions W*

$$\Gamma[\varphi_J] = W[J] - \int_{\mathscr{C}'} \mathrm{d}^4x\, \varphi_J(x)J(x) \ \text{with} \ \varphi_J(x) = \frac{\delta W[J]}{\delta J(x)} \tag{3.5.13}$$

one obtains a generating functional for the proper vertex functions.

Proper vertex functions are defined as 1PI connected diagrams where the propagators connecting the external points with the rest of the diagram are omitted. 1PI means that the diagram keeps to be connected if one cuts any single line.

Especially the negative functional inverse of the two-point connected Green's function is given by

$$\Gamma_J^{(2)}(x_1,x_2) = -\frac{\delta^2\Gamma[\varphi_J]}{\delta\varphi_J(x_1)\delta\varphi_J(x_2)}. \tag{3.5.14}$$

The convention for the n-point proper vertex functions is as follows

$$\Gamma_J^{(n)}(x_1,\ldots,x_n) = \mathrm{i}\frac{\delta^n\Gamma[\varphi_J]}{\delta\varphi_J(x_1)\cdots\delta\varphi_J(x_n)}. \tag{3.5.15}$$

Finally we mention that the original theory with $J=0$ is defined by the stationary condition

$$\frac{\delta\Gamma\varphi}{\delta\varphi(x)} = -J \stackrel{!}{=} 0. \tag{3.5.16}$$

Now we can show one of the important features of the real-time formalism, namely that the generating functional (3.5.10) factorizes in a part containing the integrals along the imaginary and one along the real time path.

For this purpose we note

$$\langle J_1 J_2 \Delta_{12} \rangle_{12} = \left(\int_{\mathscr{C}} d^4 x_1 \int_{\mathscr{C}} d^4 x_2 + 2 \int_{\mathscr{C}} d^4 x_1 \int_{\mathscr{V}} d^4 x_2 + \int_{\mathscr{V}} d^4 x_1 \int_{\mathscr{V}} d^4 x_2 \right) \Delta_{12} J_1 J_2. \tag{3.5.17}$$

We have to show that the mixed integral vanishes. But this is clear because if one time is on the vertical part but the other is on the real part we integrate an analytic function along the closed real path which means that the whole integral vanishes. Thus we have

$$Z_0[J] = Z_0 Z_0^{\mathscr{C}}[J] Z_0^{\mathscr{V}}[J] \text{ with}$$

$$Z_0^{\mathscr{C}}[J] = \exp\left(-\frac{i}{2} \int_{\mathscr{C}} d^4 x_1 d^4 x_2 \Delta_{12} J_1 J_2 \right) \text{ and} \tag{3.5.18}$$

$$Z_0^{\mathscr{V}}[J] = \exp\left(-\frac{i}{2} \int_{\mathscr{V}} d^4 x_1 d^4 x_2 \Delta_{12} J_1 J_2 \right).$$

This shows that for *connected Green's functions with at least one external real time one has to take into account only the real-time closed time path, i.e., the original Schwinger-Keldysh contour*[15].

This is so important because then one can use the usual momentum representation for the real-time formalism. That means if one likes to calculate real-time propagators one can use all the techniques of vacuum quantum field theory.

On the other hand it is clear that if one wants to calculate equilibrium bulk properties which means to calculate perturbative corrections to the potential of a free gas $\Omega_0 = \ln Z_0$ only the imaginary time path plays a role because then one has to calculate closed diagrams an for $J = 0$ the theory is translation invariant, i.e. the last integral is over a analytic one point function which is independent of x. Thus the real time path cancels. This tells us that we may use the techniques of the imaginary-time formalism which is Euclidean quantum field theory with the difference that one has to take sums over the Matsubara frequencies instead of energy integrals.

3.6 Real-time Feynman rules in momentum space

In this section we shall give the real-time Feynman rules customary for practical purposes. The Schwinger-Keldysh contour $\mathscr{C}$ is extended over the whole axis as explained above and we go to the momentum space for both time and space. For this purpose we rewrite the integrals in the space-time version of the Feynman rules in terms of usual time integrals, i.e., we use the four Green's functions given in (3.4.66) in their momentum space version. Generally an integration over a function defined along the contour is given by

$$\int_{\mathscr{C}} d^4 x f(x) = \int d^4 x f_-(x) - \int d^4 x f_+(x). \tag{3.6.1}$$

This shows how to reformulate the Feynman rules in terms of the "matrix Green's functions" in the space-time version. Firstly we have to introduce a sign on the inner space-time point in order to define on which part of the contour it lies. Due to the change of the direction of the time integration along the +-part we have to define the —-vertex as usual to be $-i\lambda/4!$ and the +-vertex as $+i\lambda/4!$ (this explains also the use of $\pm$ for the parts of the contour).

---

[15]Proof: Using theorem 7 we realize that the generating functional iW is given by $iW = \ln Z + \ln Z^{\mathscr{C}}[J] + \ln Z^{\mathscr{V}}[J]$

For the propagator lines we have to introduce an arrow to define which Green's function Δ^{ij} is meant where $i,j=\pm$. The graphs are calculated in the same way as before but one uses $\pm$-signs for inner vertex points and the external points and the n-point functions split in the according 2^n parts labelled by n indices $\pm$. Of course now the contour integrals along $\mathscr{C}'$ in the Feynman rules have to be substituted by usual integrals over space and time. But this has to be paid with the price that one has to sum over $\pm$ for the inner vertex points.

Since in the equilibrium case the free Green's functions depend only on the difference of their space time arguments, one can use the theorem about the Fourier transform of convolution integrals

$$F(x)=\int \mathrm{d}^4 y f(x-y)g(y)\Leftrightarrow \tilde{F}(p)=\tilde{f}(p)\tilde{g}(p) \text{ with } \tilde{f}(p)=\int \mathrm{d}^4 x f(x)\exp(\mathrm{i}px). \tag{3.6.2}$$

and a simple calculation shows that the Feynman rules in momentum space are given by the basic diagrammatical building blocks shown in figure 3.4:

1. Calculate the symmetry factor of the diagram as explained above for the space-time Feynman rules.

2. Write down the analytic expression according to the rules given in figure 3.4. On each vertex take care of four momentum conservation and integrate over the residual internal momenta, i.e., *the loop momenta*.

3. Sum over the $\pm$-labels of all vertices.

Figure 3.4: *Graphical elements for the real-time Feynman diagrams of ϕ^4-theory for the quantum field theory at finite temperature.*

3.7   Self-energies

The self energy is defined in terms of the real-time matrix Green's functions

$$\hat{G}=\begin{pmatrix} G^{--} & G^{-+} \\ G^{+-} & G^{++} \end{pmatrix}, \hat{\Delta}=\begin{pmatrix} \Delta^{--} & \Delta^{-+} \\ \Delta^{+-} & \Delta^{++} \end{pmatrix} \tag{3.7.1}$$

where $\hat{G}$ denotes the matrix of the exact Green's functions and Δ the same for the free Green's functions (both in momentum space).

The self energy is defined as

$$\hat{\Sigma}=\hat{\Delta}^{-1}-\hat{G}^{-1} \tag{3.7.2}$$

and this means it is the sum over all truncated 1PI diagrams with two external points containing at least one loop.

Now for the exact as well as for the free Green's functions the following relations hold

$$G^{+-}+G^{-+}=G^{--}+G^{++}, \Delta^{+-}+\Delta^{-+}=\Delta^{--}+\Delta^{++}. \tag{3.7.3}$$

The proof follows directly from the definition of the exact or free Green's functions as expectations values of $\mathscr{C}$-contour ordered products of field operators.

Using this relation we can write the Green's functions in "retarded representation" as follows

$$\hat{G}'=\hat{R}^{-1}\hat{G}\hat{R}=\begin{pmatrix} 0 & G^A \\ G^R & F \end{pmatrix} \text{ with } \hat{R}=\frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \hat{R}^{-1}=\hat{R}^\dagger. \tag{3.7.4}$$

Herein the matrix elements are defined as

$$G^R = G^{--} - G^{-+} = G^{+-} - G^{++},$$
$$G^A = G^{--} - G^{+-} = G^{-+} - G^{++}, \tag{3.7.5}$$
$$F = G^{+-} + G^{-+} = G^{--} + G^{++}.$$

The same holds of course for the free propagators. Thereby G^R denotes the retarded, G^A the advanced propagator and F a symmetrized correlator.

Direct inversion of the transformed Green's function matrices on the one hand and the transformation of the self-energy matrix on the other gives

$$\hat{\Sigma}' = \hat{R}^{-1}\hat{\Sigma}\hat{R} = \begin{pmatrix} \Omega & \Sigma^R \\ \Sigma^A & 0 \end{pmatrix} \tag{3.7.6}$$

with the relations

$$\Sigma^{-+} + \Sigma^{+-} = -\Sigma^{--} - \Sigma^{++}, \; \Sigma^R = \Sigma^{--} + \Sigma^{-+}, \; \Sigma^A = \Sigma^{--} + \Sigma^{+-}, \; \Omega = -\Sigma^{-+} - \Sigma^{+-}. \tag{3.7.7}$$

From the definition of the self energy (3.7.2) one derives the Dyson-Schwinger equation and transformation with $\hat{R}$

$$\hat{G}' = \hat{\Delta}' + \hat{\Delta}'\hat{\Sigma}'\hat{G}'. \tag{3.7.8}$$

Writing down this explicitly one can read off that the retarded and advanced Green's functions are built with the retarded and advanced self energies Σ^R and Σ^A. This must be the case because a retarded function cannot contain advanced parts! Thus we have usual Dyson-Schwinger equations for the retarded and advanced functions and a little more complicated relation for the correlator

$$G^R = \frac{\Delta^R}{1 - \Delta^R\Sigma^R} = \frac{1}{p^2 - m^2 - \Sigma^R + i\epsilon \, \text{sign} \, p_0},$$
$$G^A = \frac{\Delta^A}{1 - \Delta^A\Sigma^A} = \frac{1}{p^2 - m^2 - \Sigma^A - i\epsilon \, \text{sign} \, p_0}, \tag{3.7.9}$$
$$F = \frac{F_0 + \Delta^R G^A \Omega + F_0 G^A \Sigma^A}{1 - \Delta^R\Sigma^R}.$$

We see that it is much more convenient to solve the matrix Dyson-Schwinger equation in terms of retarded and advanced Green's functions and non-causal correlators than to solve the matrix equation directly. For later use we list the according free Green's functions with the appropriate ϵ-regulators:

$$\Delta^R(p) = \frac{1}{p^2 - m^2 + i\epsilon \, \text{sign} \, p_0},$$
$$\Delta^A(p) = \frac{1}{p^2 - m^2 - i\epsilon \, \text{sign} \, p_0}, \tag{3.7.10}$$
$$F_0(p) = 2\pi[1 + n(p_0)]\delta_\epsilon(p^2 - m^2).$$

This result is obtained by using the definitions (3.7.5) and the free Green's functions (3.4.66) with the regulated δ-distribution (3.4.65). Further we have used the important formula

$$\frac{1}{x \pm i\epsilon} = \mathscr{P}\frac{1}{x} \mp i\pi\delta(x) \tag{3.7.11}$$

where $\mathscr{P}$ tells us to take the Cauchy's principal value when one applies this *distribution* to a test function defined along the real axis. The proof is left to the reader as an exercise. (Hint: One has to close the integration

path along the real axis by a big semi-circle). This formula plays a very important role in physics. For instance in classical electrodynamics it relates the fraction index to the absorption coefficient for electro-magnetic waves in a medium in terms of the so called *Kramers-Kronig relations*. The deeper reason for its applicability in physics is of course the analytic structure of retarded functions. We shall investigate this structure in detail later.

3.8 Tadpole self-energy and the Dyson-Schwinger equation

Now we come to a first example which is trivial only from a naive point of view but turns out to be a little involved if one is not careful in the use of the distributions contained in the Green's functions. As will turn out the correct regularization of the δ-distribution arising in the free propagators is crucial in order to get rid of undefined products of these.

We just calculate the most simple one-loop contribution to the self energy in ϕ^4-theory, which is given by the tadpole shown in figure 3.5.

$$= -\mathrm{i}\Sigma_{\mathrm{TP}}^{ij}$$

Figure 3.5: *The 1-loop tadpole contribution to the self-energy. The ij run over $+$ and $-$. The bars on the external legs reflect the fact that the self-energy is truncated, i.e., the propagators according to the external legs have to be omitted.*

To calculate this diagram we first look after its symmetry factor: Since it is of first order in the perturbation (or which is in our case equivalent to being first order in λ) the factor from the power expansion of the exponential function is $1/1! = 1$ and there is a factor $1/4!$ from the interaction Hamiltonian. Now we have to count the number of possible contraction leading to the very diagram. The contraction is made over two fields with external points times four fields in the interaction Hamiltonian. There are 4 ways to contract the field $\phi(x_1)$ with one internal field $\phi(y)$. Now there are 3 possibilities left to do the same with $\phi(x_2)$. The remaining two fields have to be contracted with each other (thus there is only one possibility for that). The symmetry factor is thus given by $4 \cdot 3/4! = 1/2$.

Taking into account that the contractions of the external fields giving propagators for the external legs have to be omitted and that the two fields building the loop are normal-ordered the Feynman rules give

$$-\mathrm{i}\Sigma^{--} = -\frac{\mathrm{i}\lambda}{2}\int \mathrm{d}^4 p\, \mathrm{i}\Delta_{\mathrm{TP}}(p). \tag{3.8.1}$$

Now we have just to insert (3.4.67) where we can use the non-regulated δ-distribution directly because there are no problems with products. Integrating over the δ-distribution and a little use of algebra results in

$$\Sigma^{--} = \frac{\lambda}{16\pi^3}\int \mathrm{d}^3\vec{p}\, \frac{n(\omega)}{\omega} \text{ with } \omega = \sqrt{\vec{p}^2 + m^2}. \tag{3.8.2}$$

Thanks to the exponential from the Bose-distribution this result is perfectly finite. This is the case because we have used the normal-ordered form of the propagator for the tadpole. This is equivalent to a mass renormalization in the vacuum. If we omit normal-ordering (for instance when using path integrals) and use the naive Feynman rule with the propagator Δ^{--} this is reflected by the fact that we can write the $---$-propagator in

terms of the vacuum Feynman propagator and the tadpole propagator

$$i\Delta^{--}(p) = \frac{i}{p^2 - m^2 + i\epsilon \operatorname{sign} p_0} + 2\pi[\Theta(-p_0) + n(p_0)]\delta_\epsilon(p^2 - m^2) =$$
$$= \frac{i}{p^2 - m^2 + i\epsilon} + 2\pi n(p_0)\delta_\epsilon(p^2 - m^2) \tag{3.8.3}$$

where we have used (3.7.11) in order to express the free propagator in terms of the free *Feynman propagator* (which is the time-ordered propagator) and the normal-ordered piece due to the heat bath which is proportional to the Bose distribution n.

If we use this propagator instead of the special tadpole Feynman rule we have to renormalize the vacuum contribution which leads to a quadratically divergent term. This can be done with help of the usual vacuum techniques for regularization[16] and renormalization of the mass term (here we adopt the physical renormalization scheme where the whole vacuum contribution has to be absorbed into the physical mass in order to keep the pole of the propagator at $p^2 = m^2$, i.e., to keep m the physical mass). It is important that the associated counter terms in the Lagrangian are temperature-independent. The temperature-dependent part is finite. We end with the result we also have found using the normal-ordering description where the vacuum piece is renormalized away from the very beginning.

This shows that the normal-ordering description is equivalent to a renormalization of the vacuum. On the other hand we can use Feynman rules without using the normal-ordering concept because there in higher orders infinities for the mass arise and have to be absorbed into the bare mass leading to a finite result for the physical mass. The physically more important result of this little exercise shows that the mass is substituted by an effective mass, in other words *the effective physical parameters of the particles change due to the interaction of these with the particles which are in the heat bath.*

Clearly the off-diagonal elements of the tadpole self-energy vanish and the $++$-component is the negative of the $--$-component. Thus we can simply write

$$\hat{\Sigma}_{TP} = \begin{pmatrix} \mu^2 & 0 \\ 0 & -\mu^2 \end{pmatrix} \text{ with } \mu^2 = \Sigma_{TP}^{--} = \frac{\lambda}{16\pi^3} \int d^3\vec{p} \, \frac{n(\omega)}{\omega}. \tag{3.8.4}$$

The integral can be simplified by introduction of spherical coordinates. The angular part gives just a factor 4π because the integral is independent of the angles. The $P = |\vec{p}|$-integration can be substituted by integration over ω finally resulting in

$$\mu^2 = \frac{\lambda}{4\pi^2} \int_m^\infty d\omega \, \frac{\sqrt{\omega^2 - m^2}}{\exp(\beta\omega) - 1}. \tag{3.8.5}$$

Only in the case of massless bosons, i.e., $m = 0$ this can be calculated analytically giving the result

$$\mu^2|_{m=0} = \frac{\lambda}{24\beta^2} = \frac{\lambda}{24} T^2. \tag{3.8.6}$$

Now we want to solve the Dyson-Schwinger equation (3.7.8). We have

$$\hat{\Sigma}'_{TP} = \begin{pmatrix} 0 & \mu^2 \\ \mu^2 & 0 \end{pmatrix} \Rightarrow G^R(p) = \frac{1}{p^2 - m^2 - \mu^2 + i\epsilon \operatorname{sign} p_0}, \quad G^A(p) = G^{R*}(p). \tag{3.8.7}$$

Using this result in the equation for F and plugging in the definition for δ_ϵ (3.4.65) we find after a little rearrangement of the various terms:

$$F = -2\pi i[1 + n(p_0)]\delta_\epsilon(p^2 - m^2 - \mu^2) \tag{3.8.8}$$

---

[16]The most convenient and mathematically convincing one is the dimensional regularization invented by 't Hooft and Veltman (1972).

Comparing (3.8.7) and (3.8.8) with the vacuum results (3.7.10) we see that indeed the tadpole contribution to the self-energy has the physical meaning of a (temperature-dependent!) effective mass term: The particles at finite temperature are dressed with heat bath particles and obtain (among other effects beyond the tadpole contribution) from this cloud a bigger effective mass compared to the mass in vacuum.

3.9   Absence of ambiguities

Due to the δ-distributions in the free thermal propagators on the first glance one could think that there may arise some trouble applying the Feynman rules. As we shall show now this is not the case provided one is working with the regularised expressions (3.4.66) with the "causal regularisation" (3.4.65) for the δ-distribution. We shall show this on two simple examples.

3.9.1   Dyson series with tadpole self-energy revisited

The first is the already treated resummation of the Dyson series with a tadpole self-energy. We did this in the previous section by using the analytic properties and the fact that retarded and advanced Green's functions are resummed by the retarded and advanced functions themselves. This reduced the summation of the Dyson series to a simple geometric series for the retarded or advanced function, and we could reconstruct the other components of the matrix Green's function by using the analytic properties. The result was that the resummed Green's function in this case is simply the free one with an effective temperature dependent mass $M^2 = m^2 + \Theta$, where $\Theta = \Sigma^{--}$ is the tadpole self-energy.

Now we want to rederive this simple fact from the original matrix formalism in $\pm$-representation. Here it is important first to sum the series and then to take the ϵ-regulator to 0 in the sense of a weak limit.

For that purpose it is more convenient to write (3.4.65-3.4.66) in the form

$$
\begin{aligned}
\delta_\epsilon(p^2 - m^2) &= \frac{1}{2\pi i}[\Delta_F^*(p) - \Delta_F(p)], \\
\Delta^{--}(p) &= \Delta_F(p) - 2\pi i n(p_0)\delta_\epsilon(p^2 - m^2), \\
\Delta^{++}(p) &= -\Delta_F^*(p) - 2\pi i n(p_0)\delta_\epsilon(p^2 - m^2) \\
\Delta^{-+}(p) &= -2\pi i[\Theta(-p_0) + n(p_0)]\delta_\epsilon(p^2 - m^2) \\
\Delta^{+-}(p) &= -2\pi i[\Theta(p_0) + n(p_0)]\delta_\epsilon(p^2 - m^2)
\end{aligned}
\tag{3.9.1}
$$

with the Feynman propagator

$$
\Delta_F(p) = \frac{1}{p^2 - m^2 + i\epsilon}.
\tag{3.9.2}
$$

The first contribution to the Dyson series reads

$$
\hat{\Delta}_1 = \hat{\Delta}\hat{\Sigma}\hat{G} = \mu^2 \begin{pmatrix} \Delta_F^2 + n(p_0)(\Delta_F^2 - \Delta_F^{*2}) & (\Delta_F^2 - \Delta_F^{*2})[\Theta(-p_0) + n(p_0)] \\ (\Delta_F^2 - \Delta_F^{*2})[\Theta(p_0) + n(p_0)] & -\Delta_F^{*2} + (\Delta_F^2 - \Delta_F^{*2})n(p_0) \end{pmatrix}.
\tag{3.9.3}
$$

Since

$$
\Delta_F^2 = \frac{\partial}{\partial m^2}\Delta_F
\tag{3.9.4}
$$

we have according to (3.9.1)

$$
\Delta_F^2 - \Delta_F^{*2} = 2\pi i \delta_\epsilon'(p^2 - m^2).
\tag{3.9.5}
$$

Via induction one proves quickly

$$\hat{G}_n = \hat{\Delta}(\hat{\Sigma}\hat{\Delta})^n =$$
$$= \mu^{2n} \begin{pmatrix} \Delta_F^{n+1} + n(p_0)(\Delta_F^{n+1} - \Delta_F^{*n+1}) & (\Delta_F^{n+1} - \Delta_F^{*n+1})[\Theta(-p_0) + n(p_0)] \\ (\Delta_F^{n+1} - \Delta_F^{*n+1})[\Theta(p_0) + n(p_0)] & -\Delta_F^{*n+1} + (\Delta_F^{n+1} - \Delta_F^{*n+1})n(p_0) \end{pmatrix} \qquad (3.9.6)$$

and

$$\Delta_F^{n+1} = \frac{1}{n!} \frac{\partial}{\partial m^2} \Delta_F. \qquad (3.9.7)$$

Damit ergibt die Summe

$$\sum_{n=0}^{n} \hat{G}_n = \sum_{n=0}^{n} \frac{\mu^n}{n!} \frac{\partial}{\partial_{m^2}} \hat{\Delta} = \begin{pmatrix} G^{--} & G^{-+} \\ G^{+-} & G^{++} \end{pmatrix}, \qquad (3.9.8)$$

where the G^{ij} are defined as Δ^{ij} cf. (3.9.1) but with m^2 substituted by $M^2 = m^2 + \mu^2$ as we have expected. This shows that there are no ambiguities with products δ-distributions and Feynman propagators or powers of δ-distributions. It is only important that one uses the causally regularized δ-distributions and takes $n(p_0)$ and *not* $n(\sqrt{\vec{p}^2 + m^2})$ in the Feynman rules.

3.9.2 Two-loop tadpole

Our next example is closely related to the previous one but shows that the same cancellation of badly defined expressions takes place also within loop integrals.

3.10 Φ-derivable Approximations

As we have seen in the previous section on hand of the example of the tadpole contribution for the self energy perturbation theory breaks down for high temperatures because this contribution goes like $g\,T^2$ which is large for large T. Thus a resummation scheme is recommended. Now in literature there is known a variational method which keeps track of nice features like conservation laws and thermodynamic consistency.

This functional method starts with the addition of a bilocal classical source term to the classical action. In terms of path integrals (see QFT chapter 1 and 4 for an introduction in the vacuum case) the generating functional for non-connected Green's functions, which is identical with the partition sum when taking our extended Keldysh contour, reads:

$$Z[J,K] = N\int \mathrm{D}\phi \exp\left[\mathrm{i}S[\phi] + \mathrm{i}\langle J\phi\rangle + \frac{\mathrm{i}}{2}\langle K_{12}\phi_1\phi_2\rangle_{12}\right].\tag{3.10.1}$$

Defining $W[J,K]$ via $Z = \exp(\mathrm{i}W)$ we know that this is the generating functional of connected Green's functions with respect to J for a theory which contains also the bilocal source K. The definition of the effective action is generalised by not only taking the Legendre transformation with respect to φ and J but also with respect to G and φ:

$$\Gamma[\varphi,G] = W[J,K] - \langle J_1\varphi_1\rangle_1 - \frac{1}{2}\langle(\varphi_1\varphi_2 - \mathrm{i}G_{12})K_{12}\rangle_{12}.\tag{3.10.2}$$

Now we like to derive the general diagrammatic properties for this *Baym functional* with help of the path integral formalism.

From the general features of the path integral one derives

$$\frac{\delta^2 Z}{\delta J_1\delta J_2} = -\langle T_{\mathscr{C}'}\phi_1\phi_1\rangle = 2\mathrm{i}\frac{\delta Z}{\delta K_{12}}.\tag{3.10.3}$$

Substituting $Z = \exp(\mathrm{i}W)$ we obtain

$$\frac{\delta W}{\delta K_{12}} = \frac{1}{2}[\varphi_1\varphi_2 + \mathrm{i}G_{12}] \text{ with } \varphi_1 = \frac{\delta W}{\delta J_1},\ G_{12} = -\frac{\delta^2 W}{\delta J_1\delta J_2}.\tag{3.10.4}$$

To obtain a diagrammatical analysis we derive saddle-point expansion of the path integral. To this end we shift the integration variable by an arbitrary field φ:

$$Z[J,K] = \exp\left[\mathrm{i}S[\varphi] + \mathrm{i}\langle j_1\varphi_1\rangle_1 + \frac{\mathrm{i}}{2}\langle K_{12}\varphi_1\varphi_2\rangle\right]\times$$
$$\times\underbrace{\int \mathrm{D}\phi'\exp\left[\frac{\mathrm{i}}{2}\langle G_{12}^{-1}\phi_1'\phi_2'\rangle_{12} + \mathrm{i}S_I[\phi',\varphi] + \mathrm{i}\langle j_1\phi_1'\rangle\right]}_{Z_1[j]}.\tag{3.10.5}$$

Here we used the following abbreviations:

$$G_{12}^{-1} = \frac{\delta^2 S[\varphi]}{\delta\varphi_1\delta\varphi_2} + K_{12} := \mathscr{D}_{12}^{-1} + K_{12},$$
$$S_I[\phi',\varphi] = S[\phi'+\varphi] - S[\varphi] - \left\langle\frac{\delta S[\varphi]}{\delta\varphi_1}\right\rangle_1 - \frac{1}{2}\left\langle\frac{\delta^2 S[\varphi]}{\delta\varphi_1\delta\varphi_2}\phi_1'\phi_2'\right\rangle_{12},\tag{3.10.6}$$
$$j_1 = J_1 + \frac{\delta S[\varphi]}{\delta\varphi_1} + \langle K_{12}\varphi_2\rangle_2.$$

Now we chose J and K such that φ and G become the exact mean field and propagator cf. (3.10.4). From the mean field definition it follows that we have to chose $j = j'$ such that

$$\frac{\delta Z_1[j']}{\delta j'} = 0. \tag{3.10.7}$$

This means that the expectation value of the field ϕ' for the field theory defined by vanishes by definition. In this way $Z_1[j]$ determines completely $Z[J,K]$ and can be treated with well-known standard techniques (see QFT section 4.6.5) leading to loop expansions or perturbative expansions as a formal power series with respect to the coupling constant.

We do not need to repeat the steps leading to the loop expansion. This derivation is given in QFT section 4.6.5. We simply repeat the result up to order $\hbar$, i.e., up to one loop:

$$W_1[j]|_{\varphi'=0} := -\mathrm{i}\ln Z_1[j]|_{\varphi=0} = \Gamma_1[\varphi'=0] = \frac{\mathrm{i}}{2}\operatorname{Tr}\ln(M^2 G^{-1}) + \Phi[\varphi, G]. \tag{3.10.8}$$

This is an exact result which defines the functional Φ. Up to now the only fact we know about it is that within the diagrammatic interpretation Φ contains point vertices defined by the classical interaction action functional $S_I[\varphi, \phi]$ and lines representing the operator G. All diagrams in Γ_1 are one-particle irreducible, i.e., they cannot be disjoined by cutting only one line. The parameter M in (3.10.8) is an arbitrary mass scale which will be eliminated at the end when we are renormalising the mean energy to 0 for the vacuum case.

All we have to do now is to use (3.10.8) in (3.10.5) together with the definition (3.10.6) to obtain the final result for the Baym functional (3.10.2):

$$\Gamma[\varphi, G] = S[\varphi] + \frac{\mathrm{i}}{2}\left\langle \mathscr{D}_{12}^{-1}(G_{12} - \mathscr{D}_{12})\right\rangle + \frac{\mathrm{i}}{2}\operatorname{Tr}\ln(M^2 G) + \Phi[\varphi, G]. \tag{3.10.9}$$

Now we like to investigate our theory at vanishing external sources J and K. From the definition of the functional by the double Legendre transformation (3.10.2) we obtain immediately the *equations of motion* for the mean field φ and the exact propagator G:

$$\frac{\delta \Gamma[\varphi, G]}{\delta \varphi_1} = J_1 \overset{!}{=} 0,$$
$$\frac{\delta \Gamma[\varphi, G]}{\delta G_{12}} = \mathrm{i} K_{12} \overset{!}{=} 0. \tag{3.10.10}$$

With help of (3.10.9) this can be expressed in terms of the Φ-functional:

$$\frac{\delta S[\varphi]}{\delta \varphi} + \frac{1}{2}\left\langle \frac{\delta \mathscr{D}_{1'2'}^{-1}}{\delta \varphi_1} G_{1'2'}\right\rangle_{1'2'} + \frac{\delta \Phi[\varphi, G]}{\delta \varphi_1} \overset{!}{=} 0.$$
$$G_{12}^{-1} - \mathscr{D}_{12}^{-1} \overset{!}{=} 2\mathrm{i}\frac{\delta \Phi}{\delta G_{12}}. \tag{3.10.11}$$

The first line is the equation of motion for the mean field. While the first term provides the classical part of the equation (in our case of a scalar field it is of the Klein-Gordon type), the second term is a local one-loop quantum correction depending on the exact propagator in form of a tadpole contribution. Both contribution together build the mean field level. The third contribution consists of diagrams with at least two loops and goes in general beyond the mean field level, i.e., it generates the non-local quantum contributions to the field equation of motion.

The second line shows another important property of the Φ-functional: Since G is the full propagator and $\mathscr{D}$ is the free propagator at presence of the mean field the right hand side of the equations is the self-energy

Figure 3.6: *The Hartree-Fock approximation in non-relativistic quantum theory. The dashed line represents the interaction potential. Note that this is not a propagator and thus the diagrams for $i\Gamma_2$ need not be 2PI with respect to cutting these lines. The thick lines represent exact propagators iG and the thin eliminated lines stand for amputated legs in the corresponding self-energy diagrams. Using only the first diagram one obtains the Hartree approximation. Both together give the Hartree-Fock approximation. In literature one finds often the names direct and exchange term for these both self-energy diagrams*

expressed as functional of the mean fields and the exact propagator. In the corresponding diagrams lines represent *exact* rather than perturbative propagators and thus cannot contain any self-energy insertions in any propagator line. This class of diagrams is known as *skeleton diagrams*. The Φ-functional thus generates all (amputated) skeleton self-energy diagrams. Now the functional derivation (up to a factor i) of a diagram contained in Φ means to open any propagator line contained in this diagram, which leaves always a amputated self-energy diagram, and to add all these contributions. As we have argued above all these diagrams must be pure skeletons and thus must not disjoin if one cuts two lines. This is only possible if this is also the case for the closed diagrams which constitute Φ itself. Thus we can summarise our path integral analysis as follows:

Theorem 9. *The Φ-functional is diagrammatically represented by the sum of all closed two-particle irreducible (2PI) diagrams with at least two loops. While the vertices are given by the interaction part $S_I[\phi, \varphi]$ of the classical action at presence of a background field φ cf. (3.10.6) the propagator lines stand for exact propagators iG.*

The self-energy is obtained from the Φ-functional by opening any line of the diagrams which contribute to Φ in the described way. All these diagrams belong to the skeleton class of self-energy diagrams and expresses the self-energy in terms of exact propagators. Together with the Dyson equation (second line of (3.10.11)) this results in a closed self-consistent equation of motion for the exact propagator while the mean field is determined by the first line of (3.10.11).

Now it is clear that we are not able to solve the so far exact self-consistent equations of motion. But instead of taking an arbitrary approximation for a self-consistent calculation we approximate Γ_2 by just truncating the series of diagrams contributing to it at a *certain loop order*. Since we know that for the solutions of the equations of motion (3.10.11) this means an approximation for the effective action we can be sure that our self-consistent approximation fulfils the conservation laws which result from Noether's theorem from a symmetry, which is linearly realized on the fields (see QFT section 4.6). This is the case for the space-time symmetries and the most global gauge symmetries (leading to the conservation of charges) but not for the local gauge theories because the quantum analogue of gauge invariance for the gauge fixed classical action is BRST-invariance and this is a non-linearly realized symmetry.

We should finally mention that in the original works about the above explained formalism due to Luttinger and Ward and Baym and Kadanoff a slightly modified functional Γ_2 is named Φ and this gave the method the name Φ-*derivable approximations*.

As a last remark we report the fact that for the non-relativistic case the well known *self consistent field methods* in atomic and nuclear physics (also known as Hartree and Hartree-Fock approximations) is equivalent with taking the approximation depicted in figure 3.6 for the generating functional.
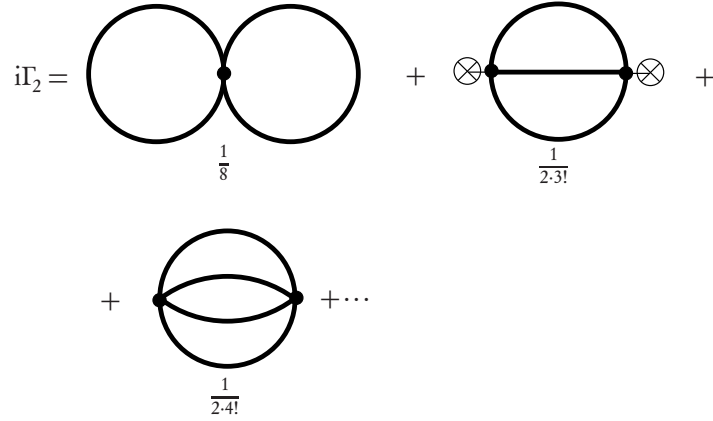
Figure 3.7: *The first diagrams contributing to the 2PI generating functional Γ_2. The bold lines represent again the exact propagator iG and the needles the mean field φ*

3.11   A simple example: Gap-equations for ϕ^4-Theory

Now we come to a first simple example of the above derived self-consistent scheme, namely ϕ^4-theory with the most simple approximation we consider for the functional Γ_2. We just keep the "eight"-diagram shown as the first drawing in fig. 3.7.

From this a straight forward calculation give the equations of motion for the mean field φ and the self-energy. In our case we assume thermal equilibrium and from homogeneity it follows $\varphi = $ const.:

$$\varphi\left(m^2 + \Sigma + \frac{\lambda}{6}\varphi^2\right) = 0, \quad \Sigma = \frac{i\lambda}{2}\int \frac{d^d l}{(2\pi)^d} G(l) + \Sigma_{CT}. \tag{3.11.1}$$

Here and in the following we can restrict ourselves to calculate the ——-self-energy alone because the off-diagonal elements of the self-energy vanish for our tadpole-approximation and thus the Green's functions never mix within Dyson's equation. Of course the self-energy is also independent of p but diverges quadratically due to UV-divergences.

The aim of the following is to show how to renormalise this theory for different cases with temperature independent counter terms. For that purpose we look at different cases.

3.11.1   Massive particles

This is the most simple case. The theory exhibits an unbroken symmetry against "field reflection" $\phi \rightarrow -\phi$ and thus we have $\varphi = 0$. The self-consistent propagator is of the form

$$G(p) = \frac{1}{p^2 - M^2 + i\epsilon} - 2i\pi n(p_0)\delta(p^2 - M^2) \text{ with } M^2 = m^2 + \Sigma. \tag{3.11.2}$$

Since the propagator is of the same form as a free one (but with a temperature dependent effective mass M) it is easy to calculate the self energy with dimensionally regularised standard integrals (see QFT chapter 6). The result is

$$\Sigma = -\frac{\lambda}{32\pi^2}M^2\left[\frac{1}{\epsilon} - \gamma + 1 + \ln\left(\frac{4\pi\mu^2}{M^2}\right) + O(\epsilon)\right] + \frac{\lambda}{2}\int \frac{d^4 l}{2\pi^4}\delta(p^2 - M^2)2\pi n(p_0) + \Sigma_{CT} \tag{3.11.3}$$

where $2\epsilon = 4 - d$ and "Σ_{CT}" stands for the counter terms which renormalise this divergent result.

Now the problem is to show that we can chose the counter terms independent of the state of the system, i.e. independent of the temperature, although the infinite term $\propto M^2/\epsilon$ depends on temperature. Since $M^2 = m^2 + \Sigma$ one part is independent of temperature and can be subtracted as a mass counter term in the Lagrangian. In ϕ^4-theory we have also a coupling constant renormalisation. From the viewpoint of perturbation theory it is clear that there must occur vertex counter terms because of sub-divergences. The question is now how to determine the vertex counter term within the self-consistent scheme.

To see this we use again our generating functional. In principle the equations of motion are given by setting the external auxiliary sources J and K to 0 resulting in a closed set of equations of motion for the self-energy and the mean field. Nevertheless these conditions also dictate consistency conditions for the higher order vertex functions.

For this purpose we use the defining equations (3.10.1) and (3.10.2) to rearrange the identity

$$\frac{\delta K_{12}}{\delta K_{34}} = \frac{1}{2}(\delta_{13}\delta_{24} + \delta_{14}\delta_{23}) := I_{12;34} \tag{3.11.4}$$

with help of the functional chain rule as follows:

$$-\frac{1}{4}I_{12;34} = \left\langle \frac{\delta^2\Gamma}{\delta G_{12}\delta G_{56}} \frac{\delta^2 W}{\delta K_{56}\delta K_{34}} \right\rangle_{56}. \tag{3.11.5}$$

From the definition of the generating functional (3.10.1) we find using the Feynman-Kac formula for path ordered products:

$$\frac{1}{Z}\frac{\delta^2 Z}{\delta K_{12}\delta K_{34}} = -\frac{1}{4}\langle T_{\mathscr{C}}\phi_1\phi_2\phi_3\phi_4\rangle = -\frac{i}{4}G_{34;56}^{(c,4)} \tag{3.11.6}$$

which is the sum of all connected four-point functions without vacuum sub-diagrams.

Using the definition for W in (3.10.4) we find

$$i\frac{\delta^2 W}{\delta K_{12}\delta K_{34}} = \frac{1}{4}\Big[G_{14}G_{23} + G_{24}G_{13} - iG_{12;34}^{(c)}\Big] \tag{3.11.7}$$

where $G^{(c)}$ is the connected four point Green's function.

Now we like to express this in terms of the 2PI functional $\Gamma^{(2)}$. For this purpose we have to differentiate the definition (3.10.9) two times with respect to G. As an intermediate step we need the derivative of G^{-1}. This is simply done as follows:

$$\langle G_{12}^{-1}G_{23}\rangle_2 = \delta_{13} \Rightarrow \frac{\delta G_{12}^{-1}}{\delta G_{56}} = -\langle G_{12}^{-1}G_{34}^{-1}I_{23;56}\rangle_{23}. \tag{3.11.8}$$

With help of this result we find

$$\frac{\delta^2\Gamma}{\delta G_{12}\delta G_{56}} = -\frac{i}{2}\langle G_{14}^{-1}G_{23}^{-1}I_{34;56}\rangle_{34} + \underbrace{\frac{\delta^2\Gamma_2}{\delta G_{12}\delta G_{56}}}_{:=\frac{1}{4}\tilde{\Gamma}_{12;56}}. \tag{3.11.9}$$

Putting this into (3.11.5) one obtains

$$iG_{1234}^{(c)} = i\langle G_{15}G_{26}G_{37}G_{48}\tilde{\Gamma}_{5678}\rangle_{5678} + \frac{1}{2}\langle G_{15}G_{26}\tilde{\Gamma}_{5678}G_{7834}^{(c)}\rangle_{5678}. \tag{3.11.10}$$

This is one of various forms of the *Bethe-Salpeter-equation* which is depicted in figure 3.8.
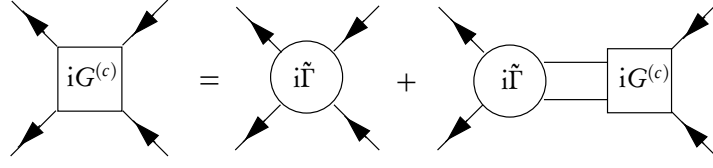
Figure 3.8: *The consistency condition of Bethe-Salpeter type. Herein bold lines stand for the self-consistent propagator, the box for the connected four-point function and the circle for the effective Bethe-Salpeter kernel which is the sum of connected amputated diagrams which are 1PI in the channel under consideration. It should be kept in mind that $G^{(c)}$ is not amputated, i.e. external legs stand for self-consistent propagators.*

$$ p = p_1 + p_2 = p_3 + p_4 $$

Figure 3.9: *The dinosaur-diagram chain determining the connected four-point function consistent with the lowest order Γ_2-functional for ϕ^4-theory.*

For our choice of the Γ_2-functional the four-point vertex $G_4^{(c)}$ is given by the chain of "dinosaur-diagrams" shown in figure 3.9.

We define λ to be the coupling at zero momentum transfer in the scattering amplitude given by $G_4^{(c)}$ in the vacuum. This can be obtained by renormalising the vacuum dinosaur diagram at $p = 0$ to 0. Since it is logarithmitically divergent (which is seen immediately from power counting and Weinberg's theorem) this is a sufficient condition to render it finite.

Fortunately we need to calculate it only for $T = 0$. Looking at (3.11.3) in that case the unique solution of our self-consistent tadpole approximation is $\Sigma = \varphi = 0$ since the explicitly temperature dependent second integral vanishes in the limit $T \to 0$. So for the vacuum we have to calculate the dinosaur diagram with the free Feynman propagators. For our purposes we need to calculate it only for $p = 0$:

$$ \Gamma_4^{\text{vac}}(p = 0) = -\delta\lambda = \frac{\lambda^2}{32\pi^2}\left[\frac{1}{\epsilon} - \gamma - \ln\left(\frac{m^2}{4\pi\mu^2}\right)\right]. \tag{3.11.11} $$

The full counter term structure for the self-energy is depicted in figure 3.10.

Plugging this into (3.11.3) we immediately find the finite self-consistent equation

$$ M^2 = m^2 + \Sigma = m^2 + \frac{\lambda}{32\pi^2}\left[M^2\ln\left(\frac{M^2}{m^2}\right) + m^2 - M^2\right] + \frac{\lambda}{4\pi^2}\int_M^\infty \sqrt{\omega^2 - M^2}\,n(\omega). \tag{3.11.12} $$

The numerical result for this equation is shown in figure 3.11. The approximation breaks down at big λ and/or T due to the first terms on the right hand side of eq. (3.11.12) originating from the quantum fluctuations because then there exists no real solution for M^2.

3.12 Massless particles

This case is only slightly different from the massive case in the previous section. Here an arbitrary scale enters the game because there is no natural one there except the temperature itself. The perturbative result is finite in dimensional renormalisation because the temperature independent part of the tadpole vanishes

$$\Sigma_{\mathrm{CT}} = \underset{\delta m^2}{\times} + \underset{-\frac{\Sigma}{\lambda}\delta\lambda}{\bigcirc}$$

Figure 3.10: *The counter terms for the self-consistent self-energy. In the second diagram we have written the loop integral in terms of the self-energy.*
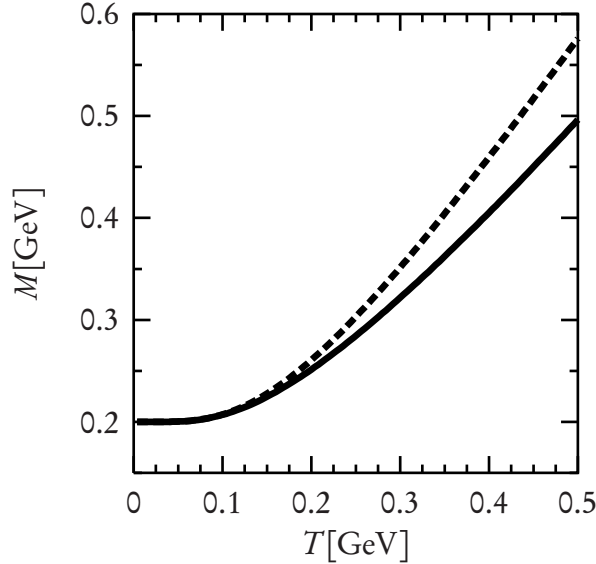

Figure 3.11: *The numerical solution for (3.11.12). Both the perturbative (dashed line) and the self-consistent (solid line) are shown. As parameters $m = 200 MeV$ and $\lambda = 40$ were chosen.*

for $m \to 0$ as can be seen from (3.11.3). Thus only the coupling constant renormalisation remains. This cannot be taken at the renormalisation point $p = 0$ because the branch cut of the function starts exactly at $p^2 = 0$. In addition the real part goes $\propto \ln p^2$ for $p^2 \approx 0$. Thus we have to chose an arbitrary momentum scale $p^2 = -\Lambda^2 < 0$ as the renormalisation point. Nevertheless there is no ambiguity in this result because a change of the renormalisation scale from Λ to Λ' is nothing than a finite change of the coupling constant and this has to be adapted to experiment in any case. The finite equation obtained from this calculation is

$$\Sigma = M^2 = \frac{\lambda}{32\pi^2}M^2\left[1 + \ln\left(\frac{M^2}{\Lambda^2}\right)\right] + \frac{\lambda}{4\pi^2}\int_M^\infty \sqrt{\omega^2 - M^2}\,n(\omega). \tag{3.12.1}$$

The numerical solution for this case is shown in figure 3.12.

Also in this case for too high couplings and/or temperatures the approximation breaks down because there is no real solution for M^2.

3.12.1 The spontaneously broken case

As our last example we look on the case that our mass parameter is $m^2 = -\tilde{m}^2 < 0$. Then the solution $\varphi = 0$ is not longer a stable solution in the vacuum. The renormalised self-consistent equations of motion are now
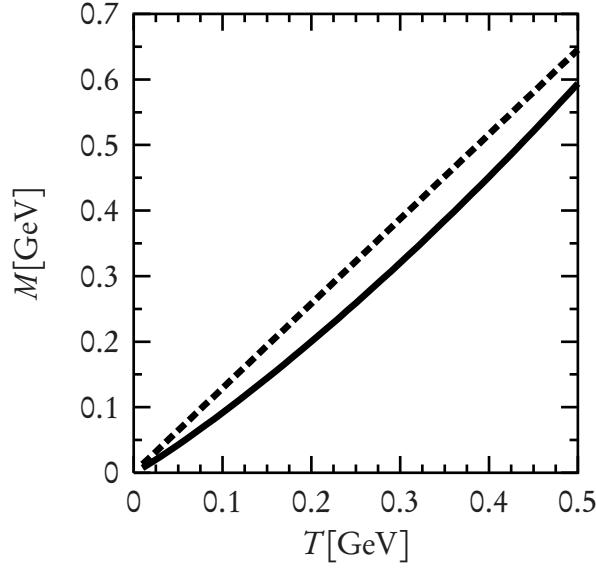
Figure 3.12: *The numerical solution for (3.12.1). Both the perturbative (dashed line) and the self-consistent (solid line) are shown. As parameters* $\Lambda = 100\,MeV$ *and* $\lambda = 40$ *were chosen.*

given by

$$
\varphi\left(\frac{\lambda}{3!}\varphi^2 - \tilde{m}^2 + \Sigma\right) = 0,
$$

$$
\Sigma = \frac{\lambda}{32\pi^2}\left[M^2\ln\left(\frac{M^2}{2\tilde{m}^2}\right) - M^2 + 2\tilde{m}^2\right] + \frac{\lambda}{4\pi^2}\int_M^\infty \mathrm{d}\omega\,\sqrt{\omega^2 - M^2}\,n(\omega) \tag{3.12.2}
$$

$$
\text{with } M^2 = \frac{\lambda}{2}\varphi^2 + \Sigma - \tilde{m}^2.
$$

Here again we only had to renormalise the vacuum case with the same techniques as in the unbroken cases. The only difference is that in addition to the tadpole subdivergence also the φ^2-term has to be renormalised due to the coupling constant renormalisation. To set the renormalisation scheme also in this case we define $\Sigma_{\mathrm{vac}} = 0$ which fixes the mass renormalisation. Then we have $\varphi_{\mathrm{vac}} = 3\tilde{m}^2/\lambda$ and the mass of the excitations from this condensate which now are the real particles of the theory is $2\tilde{m}^2 > 0$. The coupling constant again is defined at $s = 0$ to be λ in the vacuum.

3.12.2 The self-consistent effective potential

To see explicitly the symmetry restoration we also calculate the self-consistent effective action. We define this as the action which is calculated by iterating the gap equation (3.12.2) for the fixed mean field φ. As the numerical calculation shows for low temperatures the potential does not exist for too low mean fields φ because there is no stable real solution for the gap equation at all.

Usually for the calculation of bulk properties of a thermodynamical system, as is the effective potential, it is customary to use the imaginary time formalism. The reason is simply that when doing the last integration over a one-point function with respect to space-time variables the complete real-time contour is cancelled because the real part of the extended Schwinger-Keldysh contour is closed. In our case of thermal equilibrium the one-point functions are constant and by using the thermal limit of the finite volume definition used in the beginning of this chapter when we were calculating the properties of ideal gases, it is clear how to derive the expression for the effective action:

Using the definition (3.10.9) together with our tadpole approximation defined by the first-order "8-diagram" in figure 3.8 the only problem is to express the Matsubara frequency sums in terms of our already calculated real-time quantities. Thereby in any integral we have to omit the last integration over space time to get the effective *Lagrangian* instead of the effective action.

Let us start with the expression

$$\Gamma^{(1)} = -\frac{i}{2} \left\langle \mathscr{D}_{12}^{-1}(G_{12} - \mathscr{D}_{12}) \right\rangle_{12}.$$
(3.12.3)

We have to calculate only the integral over x_2 (for the time variable along the complete extended Schwinger-Keldysh-contour) while we fix x_1 with its time part on any point of the contour (the final integral does not depend on it because the integral must be a constant due to translation invariance). So we fix it on the vertical part of the contour and thus we have also to do the integration over x_2^0 only along the vertical part of the contour. Now we remember that in the imaginary time-formalism we have always "time-ordering" along the complex path and thus the Schwinger-Dyson-equation is solved as in the vacuum case which means that $\mathscr{D}^{-1} = \Sigma + G^{-1}$ which leads to

$$\mathscr{L}_{\text{eff}}^{(1)} = \frac{i}{2} \left\langle \Sigma_{12} G_{21} \right\rangle_2 = \frac{\Sigma}{2\beta} \sum_{\omega} \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{p_E^2 + M^2}$$
(3.12.4)

where we have used the Fourier representation for the imaginary time propagator (3.4.49) (remember that the self-consistent propagator G is a free one with the effective mass M) and the fact that in our tadpole approximation the self-energy is constant.

The remaining Matsubara sum and integral is proportional to our tadpole self-energy. Of course using the imaginary time Feynman rules we find that

$$\Sigma = \frac{\lambda}{2} \sum_{\omega} \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{p_E^2 + M^2}$$
(3.12.5)

and using this in (3.12.4) we get

$$\mathscr{L}_{\text{eff}}^{(1)} = \frac{\Sigma^2}{\lambda}.$$
(3.12.6)

Note that we have of course renormalised this result due to our physical renormalisation condition.

The next contribution we consider is the functional trace which comes from the logarithmic part of the effective action. In our case the calculation is very simple because from the path integral formulation for this trace we know that it is the thermodynamical potential $\Omega = \ln Z$ of an ideal gas consisting of particles with the effective mass M. This potential can be calculated most easily within the operator formalism with help of the number state basis as we have seen in the beginning of the chapter. Note that then also this contribution is renormalised due to the normal ordering description within the operator formalism which means not more than fixing the vacuum energy to 0. The final result for this contribution is thus

$$\mathscr{L}_{\text{eff}}^{(2)} = \frac{1}{2\beta} \sum_{p_0} \int \frac{d^3\vec{p}}{(2\pi)^3} \ln(p_E^2 + M^2).$$
(3.12.7)

The remaining part is only the "8-diagram" of figure 3.8 mentioned above. According to the imaginary time Feynman rules we have

$$i\mathscr{L}_{\text{eff}}^{(2)} = -\frac{i\lambda}{8\beta^2} \left[\sum_{p_0} \int \frac{d^3\vec{p}}{(2\pi)^3} \frac{1}{p_E^2 + M^2} \right]^2 + \text{CT}.$$
(3.12.8)

Again according to (3.12.5) this can be written in terms of the renormalised self-energy:

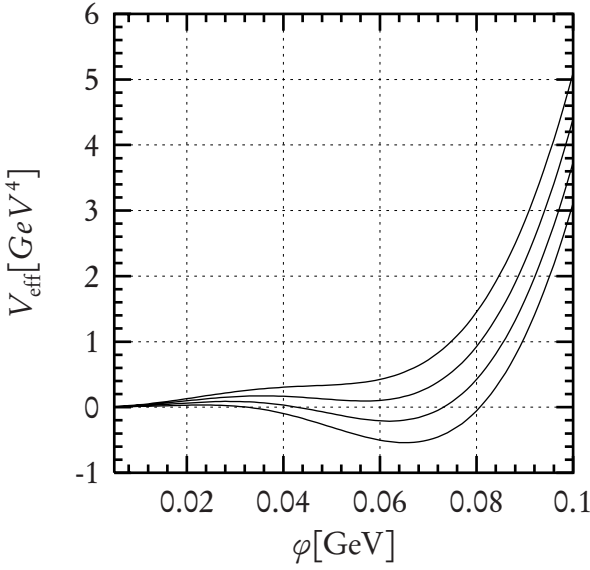$$\mathscr{L}_{\text{eff}}^{(2)} = -\frac{\Sigma^2}{2\lambda}.$$
(3.12.9)

Figure 3.13: *The evolution of the effective potential with temperatures between 150MeV and 180MeV. The potential is fixed to $V_{eff}(\varphi = 0) = 0$ for all temperatures.*

Finally we have to calculate the functional trace which is obtained most conveniently with help of the partition sum of an ideal gas consisting of particles with mass M:

$$Z_0(M) = \frac{1}{\sqrt{\mathrm{Det}(-iG^{-1})}} = N \int \mathrm{D}\phi \exp\left[\frac{i}{2}\langle G_{12}^{-1}\phi_1\phi_2\rangle_{12}\right] \tag{3.12.10}$$

where we have to integrate along the extended Schwinger-Keldysh time contour while the fields over which the path integral is taken have to obey periodic boundary conditions. According to our study of partition sums this can be solved most easily within the operator formalism where by implying normal ordering the same time we have plugged in the vacuum renormalisation which gives

$$\mathcal{L}_{\mathrm{eff}}^{(3)} = \frac{\Omega_0(M)}{\beta V} =$$

$$= -\int \frac{\mathrm{d}^3\vec{p}}{\beta(2\pi)^3}\ln[1-\exp(-\beta\omega)] = \tag{3.12.11}$$

$$= -\frac{1}{2\pi^2\beta}\int_M^\infty \mathrm{d}\omega(\omega^2-M^2)\ln[1-\exp(-\beta\omega)].$$

Now in order to get finally the *effective potential* we just have to write down our results (3.12.6), (3.12.9) and (3.12.11) together with the free Lagrangian:

$$V_{\mathrm{eff}}(\varphi) = -\frac{\mu^2}{2}\varphi^2 + \frac{\lambda}{4!}\varphi^4 - \frac{\Sigma^2}{2\lambda} + \frac{1}{2\pi^2\beta}\int_M^\infty \mathrm{d}\omega(\omega^2-M^2)\ln[1-\exp(-\beta\omega)]. \tag{3.12.12}$$

We have plotted this self-consistent effective action in figure 3.13

This shows explicitly how the first order phase transition comes about: For low temperatures (especially at $T = 0$) there is only one stable minimum at $\varphi \neq 0$ which is the broken phase. This minimum goes down until $\varphi = 0$ happens to become also a minimum above the first critical temperature T_{c1}. The minimum at $\varphi \neq 0$ remains the absolute minimum for a while. One can also see that there is a barrier between the two states showing up as a maximum in between these.

At a certain temperature (the second critical temperature T_{c2}) both minima are at the same level (here fixed to 0). Above this temperature the minimum at $\varphi = 0$ becomes the stable one but there remains an unstable minimum at $\varphi \neq 0$.
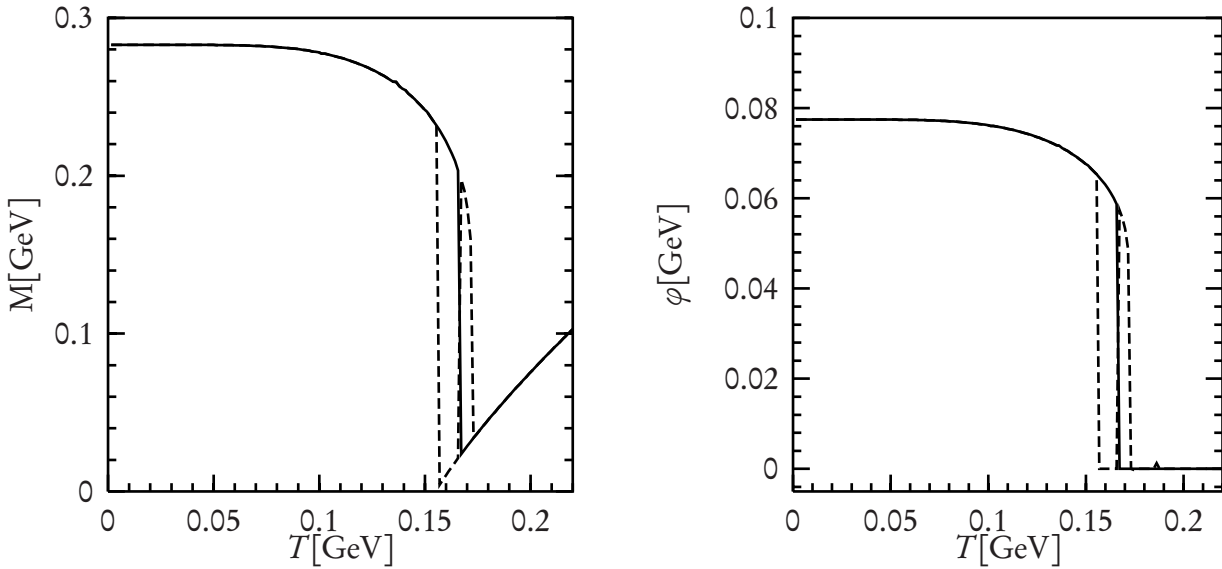
Figure 3.14: *The self-consistent solutions of the equations for M (left) and φ (right). The parameters were $\tilde{m} = 200MeV$ and $\lambda = 40$). The stable minimum of the effective potential is drawn with a solid, the unstable one with a dashed line. It shows that there is a first critical temperature at $T_{c1} \approx 158 MeV$, while the 2nd is $T_{c2} \approx 173 MeV$. For $T < T_{c1}$ only the symmetry-broken phase is an equilibrium state, for $T_{c1} < T < T_{c2}$ the symmetric phase ($\varphi = 0$) is another local minimum above the stable one. This unstable minimum is separated from the stable minimum at $\varphi \neq 0$ by a potential wall. In this region the system may be under-cooled or over-heated. For $T_{c2} < T < T_{c3}$ the symmetric phase becomes the stable one while the broken one remains an unstable minimum of the self-consistent effective potential. Above T_{c3} the 2nd minimum vanishes and only the symmetric phase is an equilibrium state.*

Finally above the third critical temperature T_{c3} the unstable minimum vanishes completely and the symmetric phase is the only stable equilibrium state.

This behaviour is depicted in figure 3.14 where the stable and the unstable minima of the effective potential are plotted against the temperature. These were calculated by finding numerically the minimum of the self-consistent effective potential.

Now we can also understand the qualitative physical behaviour of such a system: Suppose we start with an adiabatic heating[17] at a low temperature where only the broken phase (order parameter $\varphi \neq 0$) is an equilibrium state. Now coming in the range $T_{c1} < T < T_{c3}$ the system will stay in the broken phase. For $T_{c2} < T < T_{c3}$ this is the unstable phase. This phenomenon is known as *over-heating the system* and any small perturbation which makes it possible for the system to overcome the little gap showing up as the maximum in the effective action and will result a rapid change to the symmetric stable state $\varphi = 0$ [18]. Heating up the system above T_{c3} in any case the system will switch to the symmetric state.

Now suppose the opposite case of starting at temperatures above T_{c3} when the system is in the symmetric phase and lowering this temperature adiabatically. Cooling down to the range $T_{c1} < T < T_{c3}$ will keep the system in the symmetric state. While for $T_{c2} < T < T_{c3}$ this is the stable state for $T_{c1} < T < T_{c2}$ this describes an *under-cooled system* which will be brought to the stable broken state by a perturbation overcoming the potential wall separating the stable from the unstable minimum. Lowering the temperature below T_{c1} switches the system in any case to the broken phase.

---

[17]Remember that "adiabatic" is used in these notes in the sense that the change of state is done so slowly that the system always remains in an equilibrium state with increasing temperature.

[18]This behaviour is feared by chemists when boiling liquids because it may damage the whole apparatus!

As we have seen the transition from one to the other phase is characterised by a discontinuous change of the order parameter φ (and caused by that of course also the effective mass M). This is called a *1st-order phase transition*. As we have also discussed at length there are certain regions of the parameter space where both phases are possible equilibrium states separated by a potential wall but of course only one is stable.

Contrary to this in the case of a *2nd-order phase transition* the order parameter changes continuously and only the derivatives of the potential (for instance the specific heat) show discontinuities.

Appendix A

The imaginary time Feynman rules and other series summations

A.1 The Feynman rules

Although in these notes we use the real-time technique it is also important to know the imaginary time method. Especially if one likes to calculate bulk properties like pressure, entropy and other thermodynamical potentials the imaginary time formalism is much less complicated than the real-time formalism. On the other hand for calculating dynamical effects (for instance at first order linear response to an external perturbation or production and absorption rates) the real-time formalism is at place. Of course both formalisms are equivalent and to switch from one to the other is just a question of analytic continuation. But this can be a tedious task especially if one likes to do numerical calculations.

Our approach of the extended Schwinger-Keldysh contour unifies both points of view and the both formalisms can be split due to the factorization of the generating functional Z for disconnected Green's function.

It is very easy to obtain in the same way the imaginary time formalism by writing down the path integral formulation along the imaginary time path.

Let us use again the example of bosons to explain that. From the KMS-condition we know that we have to integrate over all fields with periodic boundary conditions $\phi(t + i\beta, \vec{x}) = \phi(t, \vec{x})$. Since the vertical part of the contour is not "doubled" like the real part we have no matrix structure and the Feynman rules look very similar to Euclidean vacuum quantum field theory. The only difference arises when we go to Fourier space: Instead of integrating over the energy variable p_0 we have to sum over the bosonic *Matsubara frequencies* $\omega_n = 2\pi kT$ with $k \in \mathbb{Z}$.

The pictorial elements of the Feynman rules are the same as in vacuum theory. In our conventional the four-vertex in ϕ^4-theory stands for $-i\lambda/24$ as in the vacuum theory but the line has to be substituted by $i\Delta_V$. Instead of integrating over $d^4 p/(2\pi)^4$ one has to sum over $1/\beta \sum_{\omega_n}$. The last integration in the case of closed diagrams gives the nonsense result $\beta(2\pi)^3 \delta(0)$ which is cured by taking a finite quantization volume with periodic boundary conditions as explained in chapter 3 in detail. This shows that the factor has to be substituted by βV and the corresponding intensive quantities (i.e. densities in this case) are well defined in the limit of infinite matter.

A.2 How to calculate Matsubara sums

The rough explanation in the previous section shows that within the imaginary-time formalism there arises the problem of summing over Matsubara frequencies. It is also important to show that the theory can be

rendered finite order by order in perturbation theory by just renormalizing the temperature-independent part of the diagrams, such that the counter terms in the Lagrangian remain independent of the state (in our this means independent of temperature).

Fortunately both tasks, calculating the sum in a practically way and how to split in the vacuum and the temperature part, can be solved by calculating the sum with help of complex contour integrals.

To this end suppose we like to sum a series of the form

$$s = \sum_{k=-\infty}^{\infty} f(2\pi i k). \tag{A.2.1}$$

We take the function f to be analytic in ranges defined precisely below and such that the series converges.

The idea of writing this as an complex integral is to multiply it with a meromorphic function which has simple poles at $2\pi i k$ with residuum one and use a contour enclosing these poles. Such a function is given by the *Bose distribution*:

$$g(z) = \frac{1}{\exp z - 1} \tag{A.2.2}$$

Of course this function has simple poles with residuum 1 at the desired points:

$$\operatorname*{Res}_{z \to 2\pi i k} g(z) = \lim_{\delta \to 0} \delta\, g(2\pi i k + \delta) = 1. \tag{A.2.3}$$

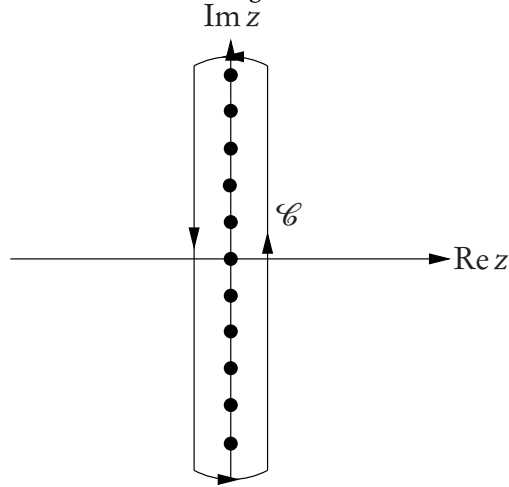Now we chose the integration contour as shown in figure A.1.

Figure A.1: The Matsubara contour. The dots show the simple poles of g.

If now the function f has no poles in an open strip around the imaginary axis from Cauchy's residuum theorem one obtains

$$\int_{\mathscr{C}} dz f(z) g(z) = 2\pi i \sum_{k=-\infty}^{\infty} f(2\pi i k). \tag{A.2.4}$$

Now we can rewrite this in a straight forward way as an integral along the imaginary axis:

$$
\begin{aligned}
\int_{\mathscr{C}} f(z) g(z) &= \int_{-i\infty+\epsilon}^{i\infty+\epsilon} dz \frac{f(z)}{\exp(z)-1} - \int_{-i\infty-\epsilon}^{i\infty-\epsilon} dz \frac{f(z)}{\exp(z)-1} = \\
&= \int_{-i\infty+\epsilon}^{i\infty+\epsilon} dz \left[\frac{f(z)}{\exp z - 1} + \frac{f(-z)}{\exp(-z)-1} \right].
\end{aligned}
\tag{A.2.5}
$$

This can be put in a somewhat more convenient form by the simple fact that the residuum theorem can be applied the same way for $g(-z)$ leading to

$$\int_{\mathscr{C}} dz f(z)g(x) = -\int_{\mathscr{C}} dz f(z)g(x \tag{A.2.6}$$

and so

$$\int_{\mathscr{C}} f(z)g(z) = \frac{1}{2}\int_{\mathscr{C}} f(z)[g(z) - g(-z)]. \tag{A.2.7}$$

Doing the same manipulations as in (A.2.5) this reads

$$\int_{\mathscr{C}} f(z)g(z) = \frac{1}{2}\int_{-i\infty+\epsilon}^{i\infty+\epsilon} [f(z) + f(-z)][g(z) - g(-z)]. \tag{A.2.8}$$

A little bit of algebra leads to

$$\int_{\mathscr{C}} dz f(z)g(z) = \frac{1}{2}\int_{-i\infty+\epsilon}^{i\infty+\epsilon} [f(z) + f(-z)] + \int_{-i\infty+\epsilon}^{i\infty+\epsilon} dz \frac{f(z) + f(-z)}{\exp z - 1}. \tag{A.2.9}$$

Together with (A.2.4) this gives the

$$\sum_{k=-\infty}^{\infty} f(2\pi in) = \frac{1}{2\pi i}\int_{-i\infty+\epsilon}^{i\infty+\epsilon} dz \frac{f(z) + f(-z)}{2} + \frac{1}{2\pi i}\int_{-i\infty+\epsilon}^{i\infty+\epsilon} dz \frac{f(z) + f(-z)}{\exp z - 1}. \tag{A.2.10}$$

To give the explicit form for the sum over Matsubara frequencies we just take $f(z) = h(z/\beta)$ and substitute $z = \beta x$ on the right hand side of (A.2.10):

$$\frac{1}{\beta}\sum_{k=-\infty}^{\infty} h(i\omega_k) = \frac{1}{2\pi i}\int_{-i\infty+\epsilon}^{i\infty+\epsilon} dx \frac{h(x) + h(-x)}{2} + \frac{1}{2\pi i}\int dx \frac{h(x) + h(-x)}{\exp(\beta x) - 1} \tag{A.2.11}$$

where we have written $\omega_k = 2\pi i k/\beta$ for the bosonic Matsubara frequencies.

The important point of this summation formula is, that it splits into a temperature independent and a dependent part. The first leads to the vacuum part of the diagrams and gives rise to the usual infinities of the diagram known from vacuum quantum field theory, and all the regularization and renormalization techniques can be taken from the vacuum case. The second one contains the Bose distribution function and as we shall see in the next section on a simple example this gives a finite contribution because the distribution function works like a cut-off for high momenta.

A.2.1 The tadpole self-energy revisited

A.3 Series summations

In this part we use the same technique as for summing over Matsubara frequencies for calculating some series we need for the treatment of ideal gases in chapter 3.

We need only to change the contour in the complex plane to sum series from 1 to ∞:

$$\sum_{n=1}^{\infty} f(2\pi in) = \frac{1}{4\pi i}\int_{\mathscr{C}'} dz \frac{f(z)}{\exp z - 1} \text{ with } f(z) = f(-z) \tag{A.3.1}$$

where the contour $\mathscr{C}'$ is given in figure A.2: Thereby f has to be regular in an open strip around the imaginary
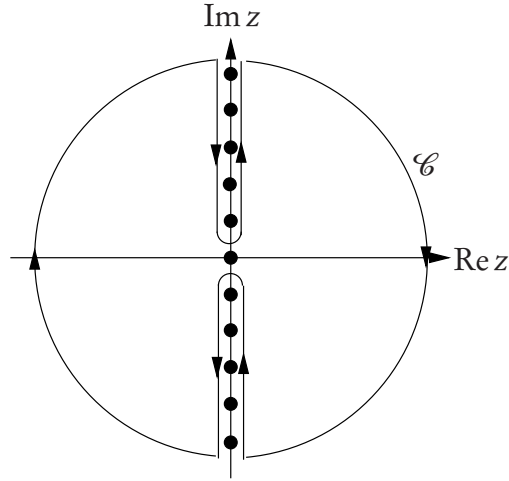
Figure A.2: *Integration path in (A.3.1).*

axes except in $z = 0$ where f may have a singularity. Clearly the semi-circles closing the path on both sides do not contribute to the integral if f is such that the series converges. From this analytic properties of f it is immediately clear that we can deform this path to a small circle containing only the singular point $z = 0$ oriented in mathematical negative direction (i.e., clock-wise). Thus using the residuum theorem we find

$$\sum_{n=1}^{\infty} f(2\pi i n) = -\frac{1}{2} \operatorname*{Res}_{z \to 0} \frac{f(z)}{\exp z - 1}. \tag{A.3.2}$$

We apply this formula to the series

$$S_k = \sum_{n=1}^{\infty} \frac{1}{n^k} \text{ for } k \in 2\mathbb{N}_{>0}. \tag{A.3.3}$$

Here we have

$$f(z) = \left(\frac{2\pi i}{z}\right)^k \tag{A.3.4}$$

and this function has only a pole of order k in $z = 0$ which means that it fulfills ther analyticity assumptions for (A.3.2).

The residuum can now be evalutated as follows:

$$\sum_{n=1}^{\infty} \frac{1}{n^k} = -\frac{(2\pi i)^k}{2} \frac{g^{(k)}(0)}{k!} \text{ with } g(z) = \frac{z}{\exp z - 1}. \tag{A.3.5}$$

Note that g is analytic in $z = 0$.

Here we list the result for some k[1]

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}, \quad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}, \quad \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^6}{945}. \tag{A.3.6}$$

The alternating sums can be obtained from that making use of

$$\sum_{n=1}^{\infty} \frac{1}{(2n)^k} = \frac{1}{2^k} \sum_{n=1}^{\infty} \frac{1}{n^k} \Rightarrow \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^k} = \left(1 - \frac{1}{2^{k-1}}\right) \sum_{n=1}^{\infty} \frac{1}{n^k}. \tag{A.3.7}$$

---

[1]To calculate the derivatives Mathematica was a helpful tool!

We apply these formulas to the calculation of some integrals we need in chapter 3 for ideal gases:

$$I_k = \int_{-\infty}^{\infty} dx \, x^k \frac{\exp x}{(1 + \exp x)^2} = 2 \int_0^{\infty} x^k \frac{\exp x}{(1 + \exp x)^2} \text{ for even postive integer } k. \tag{A.3.8}$$

Integration by parts gives

$$I_k = 2k \int_0^{\infty} dx \frac{x^{k-1}}{1 + \exp x} = 2k \int_0^{\infty} dx \frac{x^{k-1} \exp(-x)}{1 + \exp(-x)}. \tag{A.3.9}$$

Expanding in a geometric series and integrating gives

$$I_k = (2k)(k-1)! \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^k}. \tag{A.3.10}$$