Statistical Mechanics

Henri J.F. Jansen
Department of Physics
Oregon State University

October 12, 2008

II

Contents

List of Figures

Introduction.

Thermodynamics and statistical mechanics are two aspects of the study of large systems, where we cannot describe the majority of all details of that system. Thermodynamics approaches this problem from the observational side. We perform experiments on macroscopic systems and deduce collective properties of these systems, independent of the macroscopic details of the system. In thermodynamics we start with a few state variables, and discuss relations between these variables and between the state variables and experimental observations. Many theorists find this an irrelevant exercise, but I strongly disagree. Thermodynamics is the mathematical language of experimental physics, and the nature of this language itself puts constraints on what kind of experimental results are valid.

Nevertheless, true understanding needs the details on a microscopic level. We need to describe the behavior of all particles in the sample. That is an impossible task, and we need to resort to some kind of statistical treatment. Discussions on the foundations of statistical mechanics are very interesting, and help us understand the limitations of statistical mechanical treatments. This is analogue to the situation in quantum mechanics. Textbooks in quantum mechanics sue two approaches. One can simply start from Schrödinger's equation and discuss the mathematical techniques needed to solve this equation for important models. Or one can start with a discussion of the fundamentals of quantum mechanics and justify the existence of the Schrödinger equation. This is always done in less depth, because if one wants to go in great depth the discussion starts to touch on topics in philosophy. Again, I find that very interesting, but most practitioners do not need it.

In statistical mechanics one can do the same. The central role is played by the partition function, and the rules needed to obtain observable data from the partition function. If these procedures are done correctly, all limitations set by thermodynamics are obeyed. In real life, of course, we make approximations, and then things can go wrong. One very important aspect of that discussion is the concept of the thermodynamic limit. We never have infinitely large samples, and hence we see effects due to the finite size, both in space and in time. In such cases the statistical assumptions made to derive the concept of a partition function are not valid anymore, and we need alternative models. In many cases the modifications are small, but sometimes they are large. Quantification of small and large now depends on intrinsic time and length scales of our system.

The important difference between quantum mechanics and statistical mechanics is the fact that for all atomic systems quantum mechanics is obeyed, but for many systems the finite size of a sample is important. Therefore, in statistical mechanics it is much more important to understand what the assumptions are, and how they can be wrong. That is why we need to start with a discussion of the foundations. Again, we can go in greater depth that we

will do here, which leads to important discussions on the arrow of time, for example. We will skip that here, but give enough information to understand the basic ideas.

In order to introduce the partition function one needs to define the entropy. That can be done in many different ways. It is possible to start from chaos and Lyaponov exponents. It is possible to start with information theory and derive a fairness function. Although the latter approach is fundamentally more sounds, it does not illuminate the ideas one needs to consider when discussing validity of the theory. Therefore, I follow the old fashioned approach based on probability of states.

Working with statistical approaches is easier when the states of a system are labelled by an integer. Continuous variables need some extra prescriptions, which obscures the mathematics. Therefore we start with quantum mechanical systems. Counting is easiest. We follow the approach in the textbook by Kittel and Kroemer, an excellent textbook for further reference. It is aimed at the undergraduate level. Here we take their ideas, but expand on them. The toy model is a two state model, because counting states is simplest in that case. Almost every introduction to quantum statistical mechanics uses some form of this model, with states labelled up and down, plus and minus, black and white, occupied and empty, and so on. This basic model is simple enough to show mathematically how we use the law of large numbers (ergo, the thermodynamic limit) to arrive at Gaussian distributions. All conclusions follow from there.

The first few chapters discuss the definition of the entropy along these lines (chapter one), the partition function (chapter two), and the grand partition function, where the number of particles varies (chapter three). Remember that in thermodynamics we use N for the number of mols in a material, and use the constant R, the molar gas constant. In statistical mechanics we use N as the number of particles, and use k_B, the Boltzmann constant.

The first three chapters describe quantum states for the whole system. A very important application is when the energies of the particles are independent, in which case we can simplify the calculations enormously. This is done in chapter four. In chapter five we apply this formalism to the important examples of non-interacting fermions and bosons. The mathematical background needed is also explained in the textbook by Huang. This chapter is a must read for anybody who is going to work in solid state physics.

Chapters six and seven introduce different manners of describing the partition function. In chapter six we follow the density matrix formalism, which is very useful for all kinds of theoretical models. In chapter seven we go to classical statistical mechanics, which is simpler in terms of calculations, and is therefore often used in models where quantum effects do not play a role.

Life becomes more complicated when phase transitions show up. How do we know tat a system has a phase transition, how can we estimate the critical temperature? In chapter eight we use mean field theory and its cousins to relate values of the critical temperature to parameters in the microscopic model. We use the Ising model as our example, because it is very simple, yet realistic. We can actually perform all calculations in the Ising model. It is the best prototype.

In chapter nine we discuss methods that go beyond the mean field theory, and that allow us to calculate critical exponents as well. We give a very brief introduction to renormalization group theory, just enough to understand how it works. For more detail see the textbook by Plischke and Bergersen.

Missing from these notes is material on computational methods. The Monte Carlo method is extremely suitable for investigating phase transitions. In its simplest form it is very easy to implement. There have been many additions, though, and the Monte Carlo method is very well optimized. But we leave the discussion to a course on computational physics, since the Monte carlo method has applications in many other fields as well. Similarly, one can perform molecular dynamics calculations. Here, too, are many details that one needs to know in order to perform efficient and reliable calculations. We leave these discussions to the course on computational physics as well.

History of these notes:

1991 Original notes for first seven chapters written using the program EXP.

1992 Extra notes developed for chapters eight and nine.

2006 Notes of first six chapters converted to LaTeX, and significantly updated.

2008 Notes of the remaining chapters converted to LaTeX, and significantly updated. Corrections made.

Chapter 1

Foundation of statistical mechanics.

1.1   Introduction.

WHAT IS THE DIFFERENCE?

In our study of thermodynamics we have derived relations between state variables, response functions, and free energies. We have shown how a change of variables affects the choice of free energy. We derived limitations on response functions. Finally, we used equations of state to describe experiments. These equations of state were either derived from experimental results (i.e. a good guess of the functions) or from models of the free energy. At the heart of all our derivations was the thermodynamic limit. Our systems have to be large enough for fluctuations to be unimportant. In addition, we assume that we can introduce the idea of reservoirs, which are *external* systems much larger than the system we are investigating. In a manner of speaking, thermodynamics is the theory of experimental measurements. It defines what we can measure, and how results should be related.

In statistical mechanics we will try to derive these equations of state from *microscopic models*. It is here where we introduce our knowledge of physics. We might need a classical description of the motion of atoms in a gas, we could have to use a standard quantum mechanical theory of the motion of electrons in a crystal, or we might even have to introduce quantum electrodynamics for nuclei or quarks. Sometimes relativistic effects have to be included. Depending on the values of the state variables used to define the system different models could be needed. For example, at room temperature the electrons in some semiconductors behave like classical particles, but in other semiconductors quantum mechanical effects are essential.

Do we have to know why it works?

In a number of ways, statistical mechanics is like quantum mechanics. In the latter case, if we assume that a system is described by Schrödinger's equation, we can very often find the eigenstates and verify that these eigenstates give a good description of our experiments. Something similar is true in statistical mechanics. If we take the standard recipe to calculate the partition function and find all quantities of interest, often there is a good correspondence with experimental results. Most textbooks in quantum mechanics follow the approach mentioned above. They spend perhaps one chapter introducing Schrödinger's equation and mainly focus on solutions. Some text books covering statistical mechanics are based on the same philosophy. They essentially introduce the partition function and then show many techniques to obtain solutions for specific problems.

Older textbooks on quantum mechanics spent much more time discussing the fundamentals of quantum mechanics. There are still essential problems in this area and one often has to be very careful in interpreting derivations of Schrödinger's equation. On the other hand, even today most textbooks on statistical mechanics still try to justify and explain the fundamental assumptions in statistical mechanics. Again many problems are present, but not always discussed.

Traditionally, many textbooks start with classical statistical mechanics. It is certainly true that most of the development of statistical mechanics was based on classical physics. It turns out, however, that developing statistical mechanics is much easier in a discussion based on quantum mechanics. The main focus of this course is on equilibrium statistical mechanics, which means that we only need the time-independent Schrödinger equation and talk about stationary states. Non-equilibrium statistical mechanics and transport theory is in general a harder subject to study.

Extensive and intensive variables.

Thermodynamics is a study of the macroscopic variables of a large system. The equations typically contain quantities like the volume V of the system, the number of particles N, the magnetization $\vec{M}$, etc. As we have seen before, these quantities are extensive parameters. If the size of the system is increased by a factor of two, but kept in the same state, all these quantities have to become twice as large. In other words, if we combine two systems, initially we have to add the quantities of the separate systems. In the last case, it is of course possible for the total system to go through some process after contact is made, resulting in a change of the parameters. The combined system is not necessarily in equilibrium immediately after contact. Note the very important difference in the usage of N in thermodynamics and statistical mechanics. In thermodynamics we use N as the number of moles of material. Correspondingly, the chemical potential is an energy per mole, and the equations of state use the

gas constant R. In statistical mechanics we use N for the number of particles in the system. The chemical potential is now an energy per particle, and the equations contain the Boltzmann constant k_B. We have the following simple relation, $R = N_A k_B$, where N_A is Avogadro's number, the number of particles in a mole.

If we want to change the volume of a system we have to apply a pressure p. The work done during a volume change is of the form $p\Delta V$. In equilibrium, the pressure is the same in the whole system, unless other forces are present. If we take a system in equilibrium and separate it into two parts, both parts will have the same pressure. Therefore, the pressure is not additive and it is called an <u>intensive</u> parameter. There is an intensive parameter corresponding to most extensive parameters (the energy, an extensive parameter, is the exception). For example, the chemical potential μ is a partner of N, the magnetic field $\vec{H}$ pairs with $\vec{M}$, etc. These intensive parameters correspond in many cases to the external handles we have on a system and which can be used to define the state of a system. The energy contains terms like pV, μN, and $\vec{H} \cdot \vec{M}$.

How is temperature defined?

In thermodynamics we define the temperature T operationally, by how we can measure it. It clearly is an intensive parameter. The corresponding extensive parameter is called the entropy S and the energy contains a term of the form TS. The transport of this form of energy is called heat. Entropy cannot be measured directly like volume or number of particles, and this is often seen as a major problem in understanding entropy. But there are other extensive quantities that are often measured indirectly too! For example, the value for a magnetic moment follows from the response of a magnetic system to a small change in magnetic field. Measuring a magnetic moment this way is identical to measuring the entropy by submitting a system to a small change in temperature. It is true, however, that for all extensive quantities except entropy we are able to find some simple physical *macroscopic* picture which enables us to understand the meaning of that quantity.

What is the real entropy?

Entropy is associated with randomness or chaos, and these concepts are harder to put into a simple picture. Nevertheless, this is the path followed in statistical mechanics. Entropy is defined first, and then temperature simply follows as the intensive state variable conjugate to the entropy. But how to define entropy is not clear, and there are several approaches. In a technical sense, the quantities defined in statistical mechanics are only <u>entropy analogues</u>. For each proposed definition one needs to show that the resulting equations are equivalent to those derived in thermodynamics, which in the end describes experimental reality. We will only be able to do so in the thermodynamic limit, where the system becomes large. Also, we need to keep in mind that we always

have to be able to define an outside heat reservoir. Therefore, even though in statistical mechanics we will be able to evaluate the entropy analogue for very small systems or of the largest possible system, the universe, we cannot equate this calculated quantity with the entropy used in thermodynamics.

MICROSCOPIC ORIGIN OF MODEL PARAMETERS.

In thermodynamics one postulates relations between all the parameters (equations of state) and formulas for different forms of the energy (state functions). Typical equations of state are:

$$pV = NRT \tag{1.1}$$

for an ideal gas, or the van der Waals form

$$(p + a\frac{N^2}{V^2})(V - Nb) = Nk_BT \tag{1.2}$$

Based on macroscopic considerations, the parameters a and b in the second equation are simply parameters which could be fit to the experimental results. They can only be given a definite meaning in a discussion of the microscopic aspects of the models. For example, using experimental data one derives a value for b of about 1 $m^3/atom$. This is not a problem in thermodynamics, but it is in statistical mechanics, where the parameter b is equal to the volume of an atom. Having this statistical mechanics background available, one can exclude certain ranges of parameter values as unrealistic. Next, the equations of thermodynamics are used together with these equations of state to describe the properties of the system. Phase transitions are possible when using the van der Waals equation, but not with the ideal gas law.

1.2   Program of statistical mechanics.

WHAT DO WE DO?

Statistical mechanics aims to derive all these macroscopic relations from a microscopic point of view. An atomic model is essential in this case; thermodynamics can be derived without any reference to the atomic nature of matter. Since the number of particles is very large in many cases (we always have to take the limit $N \to \infty$), one can only discuss average quantities. The program of statistical mechanics can therefore be described by the following four steps:

1. Assume an atomic model.

2. Apply the equations of motion (Newton, Schrödinger, etc).

3. Calculate average quantities.

4. Take the thermodynamic limit and compare with experiment/thermodynamics.

COLLECTIVE VERSUS RANDOM BEHAVIOR.

For example, if one could solve all equations of motion, one would find $\vec{r}_i(t)$ and $\vec{p}_i(t)$ for all particles $i = 1, \cdots, N$. It is then easy to separate collective motion and random motion:

$$\vec{p}_i(t) = \vec{p}_{av}(t) + \delta\vec{p}_i(t) \tag{1.3}$$

Energy related to this average motion can easily be retrieved. For example, the energy stored in moving air can be used to drive a windmill. The energy stored in the random components $\delta\vec{p}_i(t)$ of the motion is much harder to get out. Heat engines have a well defined upper limit (Carnot limit) to their efficiency! Also, the amount of data needed to describe these random components is much too large. This is where entropy and temperature play a role. In some sense, entropy measures the amount of randomness in a system.

The energy stored in the random components of the motion, $\delta\vec{p}_i(t)$ is related to thermal energy. It is hard to recover, it is hard to measure. The only viable description we have available is a statistical one. we therefore define entropy as a measure of this randomness. Entropy measures everything we do not know easily, or therefore everything we do not know. Entropy measures non-information.

ENTROPY FIRST!

In statistical mechanics we define entropy first, and then derive temperature. This is the opposite from the situation in thermodynamics. Randomness involves the number of states available for a system, and states are much easier to count using quantum mechanics. This is the basic reason why quantum statistical mechanics is so much easier than classical statistical mechanics. In a quantum mechanical description we have to count numbers of states, in classical mechanics we have to integrate over regions of phase space, which is harder to describe exactly. Ideas similar to entropy are used in different fields like information theory where the amount of information in a message can also be related to a concept like entropy. In general, any system which contains many degrees of freedom which we cannot know and are not interested in can be described using a concept similar to entropy for these degrees of freedom!

1.3 States of a system.

ISING MODEL.

In order to show how to count states of a system it is very useful to have a definite example at hand. Although we derive basic results in this model, the final conclusions will be general. We will consider an Ising system (in one, two, or three dimensions) with atomic moments $\vec{\mu}_i$ of the form

$$\vec{\mu}_i = s_i \vec{\mu} \tag{1.4}$$

with $i = 1, \cdots, N$. The variables s_i can only take the values -1 or $+1$. A state of the system is given by a specific series of values of the set of numbers $\{s_i\}$, for example $\{-, -, +, +, +, \cdots\}$. Depending on our experimental abilities we can define a number of quantities of interest for this system. Here we will restrict ourselves to the total number of particles N and the total magnetic moment M, with $M = \sum_i s_i$. The real magnetic moment is, of course, $\vec{M} = M\vec{\mu}$. One could also include other quantities, like the number of domain walls (surfaces separating volumes of spins which are all up from regions where all spins are down). Or short range correlations could be included. Basically, anything that can be measured in an experiment could be included.

HOW MANY STATES?

How many states are there for a system with a given value of N? Obviously, 2^N. How many states are there for this system with specified values of N and M? The answer is g(N,M) and this function is called a multiplicity function. Obviously, the 2^N mentioned before is an example of a simpler multiplicity function g(N). If we define the number of spins up (with value $+1$) by $N_\uparrow$ and the number of spins down by $N_\downarrow$ it is easy to show that

$$g(N, M) = \binom{N}{N_\uparrow} \tag{1.5}$$

with $M = N_\uparrow - N_\downarrow$. One way of deriving this is to choose $N_\uparrow$ elements out of a total of N possibilities. For the first one we have N choices, for the second one $N - 1$, and for the last one $N - N_\uparrow + 1$. But the order in which we make the choices is not important, and hence we need to divide by $N_\uparrow!$. We then use the definition

$$\binom{N}{N_\uparrow} = \frac{N(N-1)\cdots(N - N_\uparrow + 1)}{N_\uparrow!} = \frac{N!}{N_\uparrow! N_\downarrow!} \tag{1.6}$$

where we have used $N = N_\uparrow + N_\downarrow$. The multiplicity function is properly normalized. Using the binomial expansion

$$(x + y)^N = \sum_{n=0}^{N} \binom{N}{n} x^n y^{N-n} \tag{1.7}$$

we find with $x = y = 1$ that

$$\sum_M g(N,M) = 2^N \tag{1.8}$$

Note that if we can only measure the values of N and M for our Ising system, all the states counted in g(N,M) are indistinguishable! Hence this multiplicity function is also a measure of the degeneracy of the states.

THERMODYNAMIC LIMIT AGAIN!

What happens to multiplicity functions when N becomes very large? It is in general a consequence of *the law of large numbers* that many multiplicity functions can be approximated by Gaussians in the thermodynamic limit. In our model Ising system this can be shown as follows. Define the relative magnetization x by $x = \frac{M}{N}$. The quantity x can take all values between -1 and $+1$, and when N is large x is almost a continuous variable. Since g(N,M) is maximal for $M = 0$, we want to find an expression for g(N,M) for $x \ll 1$. If N is large, and x is small, both $N_\uparrow$ and $N_\downarrow$ are large too. Stirling's formula now comes in handy:

$$N! \approx \sqrt{2\pi N} N^N e^{-N + \frac{1}{12N} + \mathcal{O}(\frac{1}{N^2})} \tag{1.9}$$

Leaving out terms of order $\frac{1}{N}$ and smaller one finds

$$\log(N!) \approx N\log(N) - N + \frac{1}{2}\log(2\pi N) \tag{1.10}$$

and with

$$\log(g(N,x)) = \log(N!) - \log(N_\uparrow!) - \log(N_\downarrow!) \tag{1.11}$$

we find

$$\log(g(N,x) \approx N\log(N) - N_\uparrow \log(N_\uparrow) - N_\downarrow \log(N_\downarrow)$$

$$+ \frac{1}{2}\left(\log(2\pi N) - \log(2\pi N_\uparrow) - \log(2\pi N \downarrow)\right) \tag{1.12}$$

$$\log(g(N,x) \approx (N_\uparrow + N_\downarrow)\log(N) - N_\uparrow \log(N_\uparrow) - N_\downarrow \log(N_\downarrow)$$

$$- \frac{1}{2}\log(2\pi N) + \frac{1}{2}\left(2\log(N) - \log(N_\uparrow) - \log(N_\downarrow)\right) \tag{1.13}$$

$$\log(g(N,x)) \approx -\frac{1}{2}\log(2\pi N) - (N_\uparrow + \frac{1}{2})\log(\frac{N_\uparrow}{N}) - (N_\downarrow + \frac{1}{2})\log(\frac{N_\downarrow}{N}) \tag{1.14}$$

Next we express $N_\uparrow$ and $N_\downarrow$ as a function of x and N via

$$x = \frac{N_\uparrow - N_\downarrow}{N} = \frac{2N_\uparrow - N}{N} \tag{1.15}$$

which gives

$$N_\uparrow = \frac{1}{2}N(1+x) \tag{1.16}$$

$$N_\downarrow = \frac{1}{2}N(1-x) \tag{1.17}$$

This leads to

$$\log(g(N,x)) \approx -\frac{1}{2}\log(2\pi N)$$

$$-\frac{1}{2}(N(1+x)+1)\log(\frac{1}{2}(1+x)) - \frac{1}{2}(N(1-x)+1)\log(\frac{1}{2}(1-x)) \tag{1.18}$$

Next we have to remember that this approximation is only valid for values of x close to zero. Hence we can use $\log(1+x) \approx x - \frac{1}{2}x^2$ to get

$$\log(g(N,x)) \approx -\frac{1}{2}\log(2\pi N) - \frac{1}{2}(N(1+x)+1)(\log(\frac{1}{2}) + x - \frac{1}{2}x^2)$$

$$-\frac{1}{2}(N(1-x)+1)(\log(\frac{1}{2}) - x - \frac{1}{2}x^2) \tag{1.19}$$

$$\log(g(N,x)) \approx -\frac{1}{2}\log(2\pi N) - (N+1)\log(\frac{1}{2}) - \frac{1}{2}((N+1)+Nx)(x - \frac{1}{2}x^2)$$

$$-\frac{1}{2}((N+1)-Nx)(-x - \frac{1}{2}x^2) \tag{1.20}$$

$$\log(g(N,x)) \approx -\frac{1}{2}\log(2\pi N) + (N+1)\log(2) - \frac{1}{2}(N-1)x^2 \tag{1.21}$$

which gives the final result

$$g(N,x) \approx \sqrt{\frac{2}{\pi N}} 2^N e^{-\frac{1}{2}x^2(N-1)} \tag{1.22}$$

Hence we have shown that the multiplicity function becomes Gaussian for large values of N and small values of x. The Gaussian form is reduced by a factor e from its maximum at $x = 0$ at a value $x = x_0$ with

$$x_0 = \sqrt{\frac{2}{N-1}} \tag{1.23}$$

which also is much smaller than one when N is very large. In other words, for all values where the Gaussian approximation to the multiplicity function has a large value it is close to the original function! In that sense we can replace the multiplicity function by a Gaussian form for all values of x. The relative errors are large when x is not small, but so is the value of g(N,x). In other words, when the *relative* difference between the actual multiplicity function and the Gaussian approximation is large, the values of the multiplicity function are very small and the *absolute* value of the error is very small. This result is general for all multiplicity functions. In the thermodynamic limit the deviations from the maximum value are always of a Gaussian nature in the area where the values are large. **This is one of the reasons that statistical mechanics works! It also reinforces the importance of the thermodynamic limit!** It is directly related to the law of large numbers in mathematics.

In order to keep the normalization correct we replace the factor $N-1$ in the exponent by N and will use later

$$g(N,x) \approx \sqrt{\frac{2}{\pi N}} 2^N e^{-\frac{1}{2}x^2 N} \tag{1.24}$$

Because we know that

$$\sum_x g(N,x) = 2^N \tag{1.25}$$

and that the stepsize for x is $\frac{2}{N}$ (which includes a factor of two since M can only change in units of two!), we can write

$$2^N \frac{2}{N} = \sum_x g(N,x)\Delta x \rightarrow \int_{-1}^1 g(N,x)dx \tag{1.26}$$

For the Gaussian approximation we find

$$\int_{-1}^1 \sqrt{\frac{2}{\pi N}} 2^N e^{-\frac{1}{2}x^2 N} dx = \int_{-\sqrt{\frac{N}{2}}}^{\sqrt{\frac{N}{2}}} \sqrt{\frac{2}{\pi N}} 2^N e^{-t^2} dt \sqrt{\frac{N}{2}} \approx$$

$$\int_{-\infty}^{\infty} \sqrt{\frac{2}{\pi N}} 2^N e^{-t^2} dt \sqrt{\frac{2}{N}} = 2^N \frac{2}{N} \tag{1.27}$$

which is the correct normalization indeed.

WHAT IS THE PHYSICAL MEANING?

At this point we start to wonder what the multiplicity function means. Suppose the parameter x in the previous example is an internal parameter of the system. This would be the case in a system in zero magnetic field, where there is no work involved when the magnetic moment changes. In thermodynamics we have seen that the equilibrium state is the state of maximal entropy. Here we see that the state with $x = 0$, which we expect to be the equilibrium state,

has the maximal multiplicity function. Therefore, the question can be asked if there is any relation between a multiplicity function and the entropy. We will return to this question in shortly.

How to use a multiplicity function?

Average values can also be calculated easily in a number of cases. If all states of the system are equally probable, the expected value for a quantity f is

$$< f >= \frac{\sum_{states} f(state)}{\sum_{states} 1} \tag{1.28}$$

If f only depends on the magnetization x in a given state, this reduces to

$$< f >= \int f(x)g(N,x)dx\frac{N}{2}2^{-N} \tag{1.29}$$

For the Ising model using the Gaussian approximation to the multiplicity function we find

$$< x >= 0 \tag{1.30}$$

$$< x^2 >= \frac{1}{N} \tag{1.31}$$

which means that there is no net magnetization and that the fluctuations are small. There is no net magnetization since for every state $\{s_i\}$ there is a corresponding state $\{s'_i\}$ with $s'_j = -s_j$ which has the opposite magnetization. One has to apply a magnetic field in order to produce a magnetic state in our simple Ising model. In realistic models to be discussed later there are interactions between neighboring spins that lead to a net magnetic moment. The fluctuations are small because the multiplicity function is very sharp as we have seen before.

1.4   Averages.

Once we have determined the eigenstates of a system, we have to describe in which of these states the system actually is. The basic assumption in statistical mechanics is:

All accessible quantum states are equally probable.

There is no a priori preference for any given state. One keyword in this statement is accessible. This means that we only have to consider those states which have properties consistent with the external constraints. For example, in a closed system we know that the number of particles is fixed. If we start with 27 particles, we only consider quantum states with 27 particles.

A different way of stating this basic assumption is that if we wait long enough, a system will go through all accessible quantum states. This seems to be in contradiction with ordinary quantum mechanics, which specifies that once a system is prepared in a given quantum state, it will remain in that quantum state unless there are external perturbations (i.e. a measurement). In statistical mechanics we therefore assume that our system is always in contact with a large outside world. This makes sense, since it is hard to shield a sample from all external influences. For example, try to counteract gravity! The atoms in a gas do collide with the walls of the container of that gas, and since the atoms in this wall move in a random way (even at zero temperature due to zero point motion), the wall adds a small random component to the motion of the atoms in the gas.

This scenario also requires that the system we are considering is chaotic. In a chaotic system a small difference in initial conditions growth exponentially fast, and the random, small perturbations by the walls of the container are very efficient in causing the system to cycle through all accessible states. An interesting question is if all systems are chaotic when they are very large (remember, we need to take the thermodynamic limit!). In addition, we do not know for sure what chaos means in quantum mechanics. In practice, however, all macroscopic systems seem to obey the laws of thermodynamics.

You run into a number of paradoxes when you want to calculate the entropy of the universe in the standard thermodynamical way. The discussion above shows you, however, that the entropy of the universe is not a conceptually sound quantity. There is no outside reservoir which can cause the universe to cycle through all possible states, and one can argue that the universe is in a given quantum state for all eternity.

TAKE YOUR TIME!

The argument made in the previous paragraphs allows for a statistical mechanical description of the solar system. It is an almost closed system, but there are weak interactions with the external world. Such a description is, of course, not very useful. The keywords here are *long enough*. A system can be trapped in a metastable state. Glass is a standard example of such a metastable state and in order to give a meaningful statistical mechanical description of the behavior of glass we have to exclude all crystalline states. If the duration of the measurement would be millions of years we have, of course, to include all those states. A solar system is actually a chaotic system and if we wait long enough all kinds of weird things can happen. A planet could leave and go off to infinity and beyond. Fortunately, the time scale of these phenomena is much larger than the lifetime of a typical star. Therefore, we do not worry about this happening. Should we?

TRULY ISOLATED?

Finally, we can ask what happens in a truly isolated system. Is the time evolution of such a system chaotic? This question is not easy to answer, since we need to make a measurement in order to find out, and this measurement does change the system! Another interesting theoretical question, but fortunately not of practical importance. Nevertheless fun to think about.

What can we measure?

The quantities one can measure are always average quantities. It is impossible to determine the positions as a function of time of molecules number two, three, and seven in a gas of distinguishable molecules. We can measure quantities like volume V and magnetization $\vec{M}$. These are average quantities, and they correspond to external handles on the system. The corresponding thermodynamic variables are the intensive variables pressure p and magnetic field $\vec{H}$.

Everything independent.

Hence, we have to know how to calculate averages. There are three ways of doing this. First, since all states are equally probable, we have that the average value of a function f is given by

$$< f >_{ensemble} = \frac{\sum_{states} f(state)}{\sum_{states} 1} \qquad (1.32)$$

In our simple Ising model this would mean a sum over 2^N states, if we assume that N is specified. If we can only measure N and M, this expression is too complicated. In this case f has to be a function of N and M only, otherwise it would imply the ability to measure an additional quantity. Consequently, the average reduces to

$$< f >= 2^{-N} \sum_M g(N,M) f(M) \qquad (1.33)$$

which only has $N + 1$ terms. We now have a sum over configurations of the system, where a specific configuration is defined by the values of N and M. The degeneracy of each configuration, or the number of states in each configuration, is measured by the multiplicity function.

Averages taken this way are called ensemble averages. The formal justification of this procedure in the old formulation of statistical mechanics is as follows. Consider a set of L identical systems, in our case Ising systems with N particles. The state of each system is chosen randomly from the accessible states. If $L \gg 2^N$, the number of systems with a configuration (N, M) is given by $Lg(N, M)2^{-N}$ and hence the average of a function $f(M)$ over all systems is given by

$$< f > = \frac{1}{L} \sum_{systems} f(system) = \frac{1}{L} \sum_M Lg(N,M)f(M)2^{-N} \qquad (1.34)$$

which reduces to our previous definition.

REALITY IS NOT UNCORRELATED!

In a real experiment we find average quantities in a very different way. If there are no external probes which vary as a function of time, one measures f over a certain time and takes the average:

$$< f >_{time} = \frac{1}{T} \int_0^T f(t)dt \qquad (1.35)$$

The basic assumption is that if we wait long enough a system will sample all accessible quantum states, in the correct ratio of times! Nobody has yet been able to prove that this is true, but it is very likely. The alternative formulation of our basic assumption tells us that all accessible states will be reached in this way with equal probability. This is also called the ergodic theorem. Hence we assume that, if we wait long enough,

$$< f >_{ensemble} = < f >_{time} \qquad (1.36)$$

In a measurement over time we construct a sequence of systems which are correlated due to the time evolution. The ergodic theorem therefore assumes that such correlations are not important if the time is sufficiently long. Again, beware of metastable states!

SPATIAL AVERAGES.

If the function f is homogeneous, or the same everywhere in the sample, it is also possible to define a spatial average. Divide the sample into many subsystems, each of which is still very large. Measure f in each subsystem and average over all values:

$$< f >_{space} = \frac{1}{V} \int f(\vec{r})d^3r \qquad (1.37)$$

The subsystems will also fluctuate randomly over all accessible states, and this spatial average reduces to the ensemble average too. In this formulation we use a sequence of systems which are again correlated. We assume that when the volume is large enough these correlations are not important.

The difference between the three descriptions is a matter of correlation between subsystems. In the ensemble average all subsystems are strictly uncorrelated. In the time average the system evolves and the subsystem at time $t+dt$ is

almost identical to the subsystem at time t. In a spatial average, a change in a given subsystem is directly related to a change in the neighboring systems. The time average only approaches the ensemble average when we measure very long, or $T \rightarrow \infty$. The spatial average only reduces to the ensemble average if we can divide the system in an infinite number of subsystems, which reduces the relative effects of the correlation. This requires $N \rightarrow \infty$. In the ensemble average we have to take $L \rightarrow \infty$, but that is not a problem for a theorist. The limit that the system becomes infinitely large is called the thermodynamic limit and is essential in a number of derivations in thermodynamics. Keep in mind, though, that we still need an even larger outside world to justify the basic assumption!

1.5   Thermal equilibrium.

Isolated systems are not that interesting. It is much more fun to take two systems and bring them in contact. Suppose we have two systems, A and B, which are in thermal contact. This means that only energy can flow back and forth between them. The number of particles in each system is fixed. Also, each system does not perform work on the other. For example, there are no changes in volume. Energy can only exchange because of the interactions of the two systems across their common boundary. Conservation of energy tells us that the total energy U of the combined system is constant. The big question is: **what determines the energy flow between A and B** and what is the condition for thermal equilibrium, e.g. no net energy flow. We invoke our basic assumption and note that *all accessible states of the total system are equally probable.* A configuration of the total system is specified by the distribution of energy over A and B, and hence by the value of U_A (since $U_B = U - U_A$ and U is constant). The most probable configuration, or most probable value of U_A , corresponds to the configuration which has the largest number of states available. The energy U_A of this most probably configuration will be well defined in the thermodynamic limit, since in that case multiplicity functions become very sharp on a relative scale.

EXPLANATION BY EXAMPLE.

As an example, consider the Ising model in the presence of a magnetic induction $\vec{B}$. Since the number of particles does not vary, we will drop the reference to N in the multiplicity functions in this paragraph. It is easy to show that

$$U = -\sum_i s_i \vec{\mu} \cdot \vec{B} = -M\vec{\mu} \cdot \vec{B} = -xN\vec{\mu} \cdot \vec{B} \qquad (1.38)$$

The energy of subsystem A is $U_A = -x_A N_A \vec{\mu} \cdot \vec{B}$ and of subsystem B is $U_B = -x_B N_B \vec{\mu} \cdot \vec{B}$. Because energy is conserved we have $xN = x_A N_A + x_B N_B$ and therefore the total relative magnetization x is the average of the values of

the magnetization of the subsystems. Since the total magnetization does not change, the only independent variable is x_A.

If we assume that the states of the subsystems are not correlated, we have the following result for the multiplicity functions:

$$g(N, x) = \sum_{x_A} g(N_A, x_A) g(N_B, x_B) \tag{1.39}$$

with x_A and x_B related as described above.

Next, we introduce a short hand notation for the individual terms:

$$t(x_A) = g(N_A, x_A) g(N_B, x_B) \tag{1.40}$$

and notice that this function is also sharply peaked around its maximum value when both N_A and N_B are large. For our model the multiplicity function is in that case

$$g(U) = \sqrt{\frac{2}{\pi N}} 2^N e^{-\frac{1}{2}x^2 N} \tag{1.41}$$

Therefore

$$t(x_A) = \frac{2}{\pi} \sqrt{\frac{1}{N_A N_B}} 2^N e^{-\frac{1}{2}(x_A^2 N_A + x_B^2 N_B)} = t_0 e^{-\frac{1}{2}(x_A^2 N_A + x_B^2 N_B)} \tag{1.42}$$

If the number of particles is very large, x_A is approximately a continuous variable and the number of states is maximal when the derivative is zero. It is easier to work with the log of the product in t, which is also maximal when the term is maximal. We find

$$\log(t(x_A)) = \log(t_0) - \frac{1}{2}(x_A^2 N_A + x_B^2 N_B) = \log(t_0) - \frac{1}{2}x_A^2 N_A - \frac{1}{2N_B}(xN - x_A N_A)^2) \tag{1.43}$$

The derivatives are straightforward (thank you, logs)

$$\left(\frac{\partial \log(t)}{\partial x_A} \right) = -x_A N_A - \frac{N_A}{N_B}(x_A N_A - xN) \tag{1.44}$$

$$\left(\frac{\partial^2 \log(t)}{\partial x_A^2} \right) = -N_A - \frac{N_A^2}{N_B} = -\frac{N N_A}{N_B} \tag{1.45}$$

The second derivative is negative, and hence we do have a maximum when the first derivative is zero. This is the case when

$$0 = -x_A N_B - (x_A N_A - xN) \tag{1.46}$$

or

$$x_A = x = x_B \tag{1.47}$$

The last statement is intuitively correct. In equilibrium, any part of a magnet has the same partial magnetization as the whole magnet.

Next, we approximate the form of the terms $t(x_A)$ in the case that both $N_A/gg1$ and $N_B/gg1$. In that case we can write

$$t(x_A) \approx t(x)e^{-\frac{N_A N}{2N_B}(x_A - x)^2}\tag{1.48}$$

which has the right value and curvature at the maximum. For large numbers of particles this form has a small relative error when the value is large, and a large relative error when the value is very small (and hence does not matter). This is again one of those cases where we rely on the thermodynamic limit. Statistical mechanics can be applied to small systems, but in such a case one cannot make a connection with thermodynamics.

The value of the term is reduced by a factor e for a fluctuation

$$\delta x_A = \sqrt{\frac{2N_B}{N_A N}}\tag{1.49}$$

For a given ratio of N_A and N_B this becomes very small in the limit $N \to \infty$. Again, if system A is small, the relative fluctuations in energy are large and one has to be careful in applying standard results of thermodynamics.

We are now able to do the summation over all terms:

$$g(N, x) = \sum_{x_A} t(x_A) \approx t(x) \sum_{x_A} e^{-\frac{N_A N}{2N_B}(x_A - x)^2}\tag{1.50}$$

or

$$g(N, x) \approx t(x) \int dx_A \frac{N_A}{2} e^{-\frac{N_A N}{2N_B}(x_A - x)^2} = t(x)\frac{1}{2}\sqrt{\frac{\pi N_A N_B}{N}}\tag{1.51}$$

Next, we use are familiar friends the logarithms again to get

$$\log(g(N, x)) \approx \log(t(x)) + \log(\frac{1}{2}\sqrt{\frac{\pi N_A N_B}{N}})\tag{1.52}$$

The first term on the right hand side contains terms proportional to N or N_A, while the second term contains only the logarithm of such numbers. Therefore, in the thermodynamic limit we can ignore the last terms and we find that the logarithm of the sum (or of g) is equal to the logarithm of the largest term! This might seem a surprising result, but it occurs in other places too. In a mathematical sense, it is again connected with the law of large numbers.

1.6   Entropy and temperature.

The statements in the previous section can be generalized for arbitrary systems. If two systems A and B, with multiplicity functions $g_A(N, U)$ and $g_B(N, U)$

are in thermal contact, the configuration with the maximum number of states available follows from finding the maximal value of

$$t(U_A) = g_A(N_A, U_A)g_B(N_B, U - U_A) \tag{1.53}$$

where U is the total energy of the combined system. We assume that the systems are very large, and that the energy is a continuous variable. Therefore, the condition for equilibrium is

$$0 = \frac{dt}{dU_A} = \left(\frac{\partial g_A}{\partial U}\right)_N (N_A, U_A)g_B(N_B, U - U_A) + g_A(N_A, U_A)\left(\frac{\partial g_B}{\partial U}\right)_N (N_B, U - U_A)(-1) \tag{1.54}$$

or

$$\frac{1}{g_A(N_A, U_A)}\left(\frac{\partial g_A}{\partial U}\right)_N (N_A, U_A) = \frac{1}{g_B(N_B, U - U_A)}\left(\frac{\partial g_B}{\partial U}\right)_N (N_B, U - U_A) \tag{1.55}$$

This leads us to define the entropy of a system by

$$S(U, N, V) = k_B log g(U, N, V) \tag{1.56}$$

where we have added the variable V. Other extensive variables might also be included. The symbol k_B denotes the Boltzmann factor. Note that the entropy is always positive. The entropy is a measure of the number of states available for a system with specified extensive parameters. We have to be careful, though. Strictly speaking, this definition gives us an entropy analogue. We need to show that we recover all the laws of thermodynamics using this definition. In this chapter we will consider changes in temperature. In the next chapter we discuss changes in volume and mechanical work. In the chapter after that changes in the particle number and the corresponding work term.

The temperature of a system is defined by

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{N,V} \tag{1.57}$$

where the partial derivative is taken with respect to U, keeping all other extensive parameters the same. Hence our criterion for thermal equilibrium is

$$T_A = T_B \tag{1.58}$$

which is not unexpected.

The multiplicity function for the total system g_{tot} is found by summing over all possible configurations of the combination A plus B, but as shown before the value of the logarithm of this sum is equal to the logarithm of the largest term. If the energy of system A in equilibrium is $\hat{U}_A$ we have

$$log(g_{tot})(U) = log(g_A(\hat{U}_A)) + log(g_B(U - \hat{U}_A)) \tag{1.59}$$

and hence the entropy of the total system is

$$S = S_A + S_B \tag{1.60}$$

which shows that for two systems in thermal equilibrium the entropy of the total system is equal to the sum of the entropies of subsystems. Hence the entropy is an extensive quantity. Since U is also extensive, it follows from the definition of T that T is intensive.

At this point we have two definitions of entropy. In thermodynamics we defined entropy by making the differential for heat flow exact. In statistical mechanics we define entropy related to the number of quantum states. In order for statistical mechanics to be a useful theory, these definitions have to be the same. This one can show as follows. The flow of energy between the two systems we considered was not due to work, hence must be the exchange of heat. Thermodynamics tells us that two systems are in equilibrium if they have the same temperature, when only heat can flow between them. Therefore, the quantity T defined in statistical mechanics must be a function of the real temperature only, since equilibrium is independent of volume and number of particles. Hence if we define

$$\frac{1}{T^{SM}} = \left(\frac{\partial S^{SM}}{\partial U} \right)_{N,V} \tag{1.61}$$

where the superscript SM denotes defined in statistical mechanics, we find

$$T^{SM} = f(T) \tag{1.62}$$

Also, by construction the entropy defined in statistical mechanics is a state function. For each value of U,V, and N we have a unique prescription to calculate the entropy. Therefore, the differential of this entropy is exact and we have

$$\bar{d}Q = TdS = f(T)dS^{SM} \tag{1.63}$$

As we saw in thermodynamics, the only freedom we have in making an exact differential out of the heat flow is a constant factor. Hence

$$T^{SM} = \alpha T \tag{1.64}$$

$$S^{SM} = \frac{1}{\alpha}S \tag{1.65}$$

The entropy is therefore essentially the same in both thermodynamics and statistical mechanics. The factor α is set equal to one, by choosing the correct value of the prefactor k_B in the statistical mechanical definition. A change in this value would correspond to a change in temperature scale. In order to show this, we have to consider a simple model system. A gas of independent particles is a good test case, and the calculations show that the equation of state derived from a microscopic model is the same as the experimental version, if we take $N_{Avogadro}k_B = R$.

1.7 Laws of thermodynamics.

Thermodynamics is based upon four laws or postulates. In statistical mechanics these laws follow immediately from our previous discussion.

Zeroth Law.

If two systems are in thermal equilibrium with a third system, they are in thermal equilibrium with each other. Trivial from $T_A = T_B \cup T_B = T_C \Rightarrow T_A = T_C$.

First Law.

Heat is a form of energy. Trivial, since we defined thermal contact by an exchange of energy and used this exchange of energy to define entropy.

Second Law.

Entropy always increases. Also trivial, since if the combined system starts with energies U_A^0 and U_B^0 we automatically have $g_A(U_A^0)g_B(U_B^0) \leqslant g_A(\hat{U}_A)g_B(\hat{U}_B)$, or after taking the logarithms $S_A^{init} + S_B^{init} \leqslant S_A^{final} + S_B^{final}$.

Third Law.

The entropy per particle is zero at zero temperature in the thermodynamic limit. This is more a statement about quantum mechanics. In general, the degeneracy of the ground state is small. Suppose the degeneracy of the ground state is N^p, which is unusually large. Then the entropy per particle in the ground state (and hence at $T = 0$) is $S(0) = k_B p \frac{\log(N)}{N}$, which is zero when $N \to \infty$. Only when the degeneracy of the ground state is on the order of e^N or $N!$ does one see deviations from this law. Such systems are possible, however, when we have very large molecules!

We also see that the local stability requirements derived in thermodynamics are fulfilled. Since the entropy is related to the maximal term in a product, we recover expressions like $\left(\frac{\partial^2 S}{\partial U^2}\right)_{N,V} \leqslant 0$.

In summary, the first law is an assumption made in both thermodynamics and statistical mechanics. The only difference is that the introduction of heat as a form of energy is more acceptable when based on a microscopic picture. The other three laws all follow from the definition of entropy in statistical mechanics. Entropy is a more tangible quantity in a microscopic theory. It is still useful, however, to be able to define entropy from a macroscopic point of view only. This approach leads to generalizations in systems theory. For any complicated system for which the state only depends on a small number of macroscopic variables one can define a quantity like entropy to represent all other variables which are not of interest. Such a system might not have a clear microscopic foundation and therefore a generalization of the microscopic definition of entropy in terms of states might be impossible.

1.8   Problems for chapter 1

PROBLEM 1.

In an Ising model the magnetic moment of an individual atom is given by $\vec{\mu}_i = s_i\vec{\mu}$ with $s_i = \pm 1$. A state of this system is given by a particular set of values $\{s_i\}$. The total number of spins is N, and the relative magnetization x is given by $\frac{1}{N}\sum_i s_i$. The number of states g(N,x) with a given value of N and x is approximated by

$$g(N, x) \approx \sqrt{\frac{2}{\pi N}} 2^N e^{-\frac{1}{2}Nx^2}$$

A magnetic induction $\vec{B}$ is applied to this system. The energy of the system is $U(\{s_i\})$.

(a) Calculate U(x).

(b) Calculate S(N,U).

(c) Calculate U(T,N).

PROBLEM 2.

Two Ising systems are in thermal contact. The number of atoms in system j is $N_j = 10^{24}$. The relative magnetization of system j is $x_j = \frac{1}{N_j}\sum_i s_{ij}$, where s_{ij} is the sign of the magnetic moment of the i-th site in system j. The average relative magnetization of the total system is fixed at a value x.

(a) What are the most probable values of x_1 and x_2?

Denote these most probable values by $\hat{x}_1$ and $\hat{x}_2$. Since only the total magnetization is fixed, it is possible to change these values by amounts δ_1 and δ_2.

(b) How are δ_1 and δ_2 related?

Consider the number of states for the total system as a function of δ_1: $g_{tot}(N, \delta_1) = g(N_1, \hat{x}_1 + \delta_1)g(N_2, \hat{x}_2 + \delta_2)$. Use the form given in problem 1 for the multiplicity functions of the subsystems.

(c) We change the spins of 10^{12} atoms in system 1 from -1 to $+1$. Calculate the factor by which g_{tot} is reduced from its maximal value $g_{tot}(N, x)$.

(d) Calculate the relative change in the entropy for this fluctuation.

PROBLEM 3.

A system has N sites. The probability that a given site is occupied by an atom is ξ. Only one atom at a time is able to occupy a site.

(a) Find the probability $p(M, N, \xi)$ that exactly M sites out of N are occupied.

Assume that on the average A sites are occupied, hence $\xi = \frac{A}{N}$.

(b) Find the probability p(M,A) that M sites are occupied when the average number of sites occupied is A in the limit $N \to \infty$. A and M remain finite!

(c) Show that $\sum\limits_{M=0}^{\infty} p(M, A) = 1$, $< M >= \sum\limits_{M=0}^{\infty} Mp(M, A) = A$, and that $< M^2 >= \sum\limits_{M=0}^{\infty} M^2 p(M, A) = A^2 + A$.

When A is very large, the distribution p(M,A) becomes very sharp around A.

(d) Show that in that case the distribution near A can be approximated by a Gaussian of width $\sqrt{A}$.

PROBLEM 4.

Use the following approximations derived from Stirling's formula:

(I) $\log(N!) \approx N \log(N)$

(II) $\log(N!) \approx N \log(N) - N$

(III) $\log(N!) \approx N \log(N) - N + \frac{1}{2} \log(N)$

(IV) $\log(N!) \approx N \log(N) - N + \frac{1}{2} \log(N) + \frac{1}{2} \log(2\pi)$

For each of these approximations find the smallest value of N for which the relative error in N! becomes less than 1%. (Relative error: $\frac{|x_{ap} - x_{ex}|}{|x_{ex}|}$.)

PROBLEM 5.

A standard game show problem is the following. You have to choose between three doors. Behind one of them is a car, behind the others is nothing. You point at one door, making your choice. Now the game show host opens another door, behind which there is nothing. He asks you if you want to change your choice of doors. Question: should you change? Answer: Yes!!!!! Question: Calculate the ratio of the probabilities of your original door hiding the car and the other, remaining door hiding the car.

PROBLEM 6.

The atomic spin on an atom can take 2S+1 different values, $s_i = -S, -S + 1, \cdots, +S - 1, +S$. The total magnetic moment is given by $M = \sum\limits_{i=1}^{N} s_i$ and the total number of atoms is N. Calculate the multiplicity function $g(N, M)$, counting all states of the system for these values N and M. Define $x = \frac{M}{N}$ and calculate $g(N, x)$ when N becomes very large.

PROBLEM 7.

Twelve physics graduate students go to the bookstore to buy textbooks. Eight students buy a copy of Jackson, six students buy a copy of Liboff, and two students buy no books at all. What is the probability that a student who bought a copy of Jackson also bought a copy of Liboff? What is the probability that a student who bought a copy of Liboff also bought a copy of Jackson?

PROBLEM 8.

For an ideal gas we have $U = \frac{3}{2} N k_B T$, where N is the number of particles. Use the relation between the entropy $S(U, N)$ and the multiplicity function $g(U, N)$ to determine how $g(U, N)$ depends on U.

PROBLEM 9.

The energy eigenstates of a harmonic oscillator are $\epsilon_n = \hbar\omega(n + \frac{1}{2})$ for $n = 0, 1, 2, \cdots$ Consider a system of N such oscillators. The total energy of this system in the state $\{n_1, n_2, \cdots, n_N\}$ is

$$U = \sum_{i=1}^{N} \epsilon_{n_i} = (M + \frac{1}{2}N)\hbar\omega$$

where we have defined

$$M = \sum_{i=1}^{N} n_i$$

Calculate the multiplicity function $g(M, N)$. Hint: relate $g(M, N)$ to $g(M, N+1)$ and use the identity

$$\sum_{k=0}^{m} \binom{n + k}{n} = \binom{n + 1 + m}{n + 1}$$

Show that for large integers $g(M, N)$ is a narrow Gaussian distribution in $x = \frac{M}{N}$.

Chapter 2

The canonical ensemble: a practical way for microscopic calculations.

2.1   Introduction.

WHICH STATE VARIABLES?

The state of a system depends on a number of parameters. In most measurements it is a good idea to change only one of those parameters and to observe the response of the system. Hence one needs a way to keep the other parameters at fixed, prescribed values. The same is true in theory. So far we have only considered the entropy S as a function of the internal energy U, the volume V, and the number of particles N. If we specify U we can find T. Of course, this is not what is happening in most measurements. It is easier to specify T and to measure U (or rather the specific heat in calorimetry). The same is true in our theoretical formalism. It is often easier to calculate all quantities at a given temperature.

The formal development in the previous chapter was necessary to establish that a thermodynamical and a statistical mechanical description of a large system are the same. We argued that the entropy as defined in thermodynamics and in statistical mechanics lead to the same equations describing the physics of a system, but only if we work in the thermodynamic limit. What the differences can be for small systems is not clear. The formal definition of entropy is, however, in many cases not the best way to attempt to obtain a quantitative description of a system. In this chapter we develop a more practical approach for statistical mechanical calculations.

Reservoirs are what we need!

The theoretical way to specify temperature is similar to the experimental set-up. We assume that the system $\mathcal{S}$ is in contact with and in thermal equilibrium with a very large reservoir $\mathcal{R}$ with temperature T. Only heat can flow back and forth between $\mathcal{R}$ and $\mathcal{S}$. Volumes do not change, and no work is done between the system and the reservoir. The number of particles in the reservoir is $N_{\mathcal{R}}$ and is much larger than the number of particles in the system $N_{\mathcal{S}}$. How much larger?? In thermal equilibrium the fluctuations of the energy per particle in the reservoir are

$$\frac{\Delta U_{\mathcal{R}}}{N_{\mathcal{R}}} \propto \frac{1}{\sqrt{N_{\mathcal{R}}}} \tag{2.1}$$

or with $U_{\mathcal{R}} = \epsilon_{\mathcal{R}} N_{\mathcal{R}}$ we can write this in the form $\Delta U_{\mathcal{R}} = \epsilon_{\mathcal{R}}\sqrt{N_{\mathcal{R}}}$. Suppose the maximum energy of the system is $U_{\mathcal{S}} = N_{\mathcal{S}}\epsilon_{\mathcal{S}}$. If we demand that

$$N_{\mathcal{R}} \gg N_{\mathcal{S}}^2 (\epsilon_{\mathcal{S}}/\epsilon_{\mathcal{R}})^2 \tag{2.2}$$

the system can take all of its energy from the reservoir without noticeably changing the state of the reservoir! In this case it is really allowed to assume that the reservoir is at a constant temperature. In practice, this leads to impossibly large reservoir sizes, but in theory that is, of course, not a problem. For example, if we assume that a typical reservoir has 10^{24} particles, we could only consider systems of about 10^{12} particles. This is too small. In reality, therefore, reservoirs do lead to finite size effects. These effects are very hard to calculate, and depend on the nature of the reservoir. We will therefore ignore them, and assume that the reservoirs are very large, and that their only effect is to set the temperature.

Probabilities.

We will now address the question: **what is the probability of finding the system $\mathcal{S}$ in a unique quantum state s if it is in thermal equilibrium with a reservoir at temperature T?** The energy U_0 of the total combination, system plus reservoir, cannot change. Hence if the system is in a state s with energy ϵ_s, the reservoir has to have energy $U_0 - \epsilon_s$. The total number of states available in this configuration for the combination of system and reservoir is $g_{\mathcal{R}}(U_0 - \epsilon_s)$. This does not depend on the multiplicity function of the system, since we specify the *state of the system*. The multiplicity function for the combined reservoir and system is

$$g_{\mathcal{S}+\mathcal{R}}(U_0) = \sum_s g_{\mathcal{R}}(U_0 - \epsilon_s) \tag{2.3}$$

All states of the combination of system and reservoir are equally probably, and hence the probability of finding the combination system plus reservoir in the configuration with the system in state s is

$$P_{\mathcal{S}}(s) \propto g_{\mathcal{R}}(U_0 - \epsilon_s) \tag{2.4}$$

Therefore the ratio of the probabilities of finding the system in state 1 and 2 is

$$\frac{P_{\mathcal{S}}(1)}{P_{\mathcal{S}}(2)} = \frac{g_{\mathcal{R}}(U_0 - \epsilon_1)}{g_{\mathcal{R}}(U_0 - \epsilon_2)} = e^{\frac{1}{k_B}(S_{\mathcal{R}}(U_0-\epsilon_1) - S_{\mathcal{R}}(U_0-\epsilon_2))} \tag{2.5}$$

Because the reservoir is very large, the difference between the entropies is

$$S_{\mathcal{R}}(U_0 - \epsilon_1) - S_{\mathcal{R}}(U_0 - \epsilon_2) \approx -(\epsilon_1 - \epsilon_2)\left(\frac{\partial S}{\partial U}\right)_{V,N}(U_0) = -\frac{\epsilon_1 - \epsilon_2}{T_{\mathcal{R}}} \tag{2.6}$$

The reservoir and the system are in thermal equilibrium, and have the same temperature T. We drop the subscript on the temperature because is is the same for both system and reservoir. We find

$$\frac{P_{\mathcal{S}}(1)}{P_{\mathcal{S}}(2)} = e^{-\frac{\epsilon_1-\epsilon_2}{k_B T}} \tag{2.7}$$

or in general

$$P_{\mathcal{S}}(s) \propto e^{-\frac{\epsilon_s}{k_B T}} \tag{2.8}$$

This is the well-known <u>Boltzmann factor</u>. The only reference to the reservoir is via the temperature T, no other details of the reservoir are involved. Normalization requires $\sum P(s) = 1$ and hence we define

$$\mathcal{Z}(T) = \sum_s e^{-\frac{\epsilon_s}{k_B T}} \tag{2.9}$$

$$P_{\mathcal{S}}(s) = \frac{1}{\mathcal{Z}} e^{-\frac{\epsilon_s}{k_B T}} \tag{2.10}$$

The probability $P_{\mathcal{S}}(s)$ depends on the temperature T of the system, and although there is no explicit reference to the reservoir there is still an implicit reservoir through T. Specifying T requires contact with an outside reservoir. Also, the probability only depends on the energy ϵ_s of the state and not on any other property of this state. All different states with the same energy have the same probability. The quantity $\mathcal{Z}(T)$ is called the <u>partition function</u>. We will see later that for a description of a system at temperature T the partition function is all we need to know. The probability distribution $P_{\mathcal{S}}(s)$ of this form is called the <u>Boltzmann distribution</u>. The partition function is also a function of volume and number of particles. The dependence on all extensive state variables, excluding the entropy, is via the energy of the quantum states. In general, the energy eigenvalues ϵ_s are functions of the volume V and the number of particles N. Note that the quantum states we consider here are *N-particle quantum states*. The energy $\epsilon_s(V, N)$ is the energy of a **collective state** and is not equal to the energy of a given particle!

At this point we seem to have a contradiction. The basic assumption of statistical mechanics tells us that all accessible quantum states are equally probably. The Boltzmann distribution on the other hand gives the states unequal probabilities! This will always occur when we specify some intensive parameter like T. In such a case one always needs a reservoir, and the basic assumption applies to the states of the reservoir plus system. If we specify the state of the system, this corresponds to a configuration, i.e. a set of states, of the combination reservoir plus system. The basic assumption only applies to states, not to configurations. For a configuration one always has to include a multiplicity function counting the number of states available for that configuration.

The Boltzmann factor is not correct in the limit $\epsilon_s \to \infty$, since that implies taking a large fraction of the energy of the reservoir and putting it in the system. The upper limit of ϵ_s for which the Boltzmann factor is valid can be derived from the considerations in the beginning of this section. This is not a practical problem, though, since we can make the reservoir very large. The probability of finding states with an energy which is too large for our formalism to be valid then becomes exponentially close to zero!

2.2 Energy and entropy and temperature.

The expectation value of the energy, or average energy, of the system is given by

$$U = \sum_s \epsilon_s P(s) \tag{2.11}$$

where we have dropped the subscript $\mathcal{S}$ for system in the probability function. This can be related to the partition function by

$$\left(\frac{\partial \log(\mathcal{Z})}{\partial T}\right)_{V,N} = \frac{1}{\mathcal{Z}}\left(\frac{\partial \mathcal{Z}}{\partial T}\right)_{V,N} = \frac{1}{\mathcal{Z}}\sum_s e^{-\frac{\epsilon_s}{k_B T}}\frac{\epsilon_s}{k_B T^2} \tag{2.12}$$

or

$$U(T,V,N) = k_B T^2 \left(\frac{\partial \log \mathcal{Z})}{\partial T}\right)_{V,N}(T,V,N) \tag{2.13}$$

The entropy can be obtained from U(T) through

$$\left(\frac{\partial S}{\partial T}\right)_{V,N} = \left(\frac{\partial S}{\partial U}\right)_{V,N}\left(\frac{\partial U}{\partial T}\right)_{V,N} = \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_{V,N} \tag{2.14}$$

which leads to

$$S(T,V,N) = S_{ref}(V,N) + \int_{T_{ref}}^{T} \frac{1}{T}\left(\frac{\partial U}{\partial T}\right)_{V,N} dT \tag{2.15}$$

since we assume that all other variables remain constant. It is useful to have an explicit form for the entropy, however. Based on

$$-\frac{U}{k_B T} = \sum_s \left(-\frac{\epsilon_s}{k_B T}\right) P(s) \qquad (2.16)$$

and

$$\log(P(s)) = -\frac{\epsilon_s}{k_B T} - \log(\mathcal{Z}) \qquad (2.17)$$

we find

$$-\frac{U}{k_B T} = \sum_s \left(\log(P(s)) + \log(\mathcal{Z})\right) P(s) = \sum_s P(s) \log(P(s)) + \log(\mathcal{Z}) \quad (2.18)$$

where we used $\sum P(s) = 1$. This leads to

$$\frac{\partial}{\partial T}\left(-\frac{U}{k_B T}\right) = \left(\frac{\partial \log(\mathcal{Z})}{\partial T}\right)_{V,N} + \frac{\partial}{\partial T} \sum_s P(s) \log(P(s)) \qquad (2.19)$$

and using the partial derivative relation for the entropy

$$\left(\frac{\partial S}{\partial T}\right)_{V,N} = \frac{\partial}{\partial T}\left(-k_B \sum_s P(s) \log(P(s))\right) \qquad (2.20)$$

The result of the integration is

$$S(T,V,N) = S_0(V,N) - k_B \sum_s P(s) \log(P(s)) \qquad (2.21)$$

where we still need to find the constant of integration $S_0(V,N)$. This is easiest if we take the limit that T goes to zero. In that case the system is in the ground state. Suppose the degeneracy of the ground state is g_0. Then $P(s) = g_0^{-1}$ for all states s with the ground state energy and $P(s) = 0$ for other states. The logarithm of zero is not finite, but the product of the form $x\log(x)$ is zero for $x = 0$. Hence we have

$$S(T = 0, V, N) = S_0(V,N) - k_B \sum_{sing} g_0^{-1} \log(g_0^{-1}) = S_0(V,N) + k_B \log(g_0)$$

$$(2.22)$$

From the formal definition of entropy we see that the second term is equal to the ground state entropy, and hence we find $S_0(V,N) \equiv 0$. As a result we have the useful formula

$$S(T,V,N) = -k_B \sum_s P(s) \log(P(s)) \qquad (2.23)$$

Up to now we have considered three variables describing our system: S, T, and U. These variables are not independent in thermal equilibrium. If we

specify one of them, the other two are determined. We have found two recipes for these functions. If the energy U is the independent variable, the entropy S(U) is constructed by counting states with $\epsilon_s = U$. The temperature T(U) follows from $\left(\frac{\partial S}{\partial U}\right)_{V,N} = T^{-1}$. On the other hand, if the temperature T is the independent variable, we calculate the Boltzmann factor and the partition function $\mathcal{Z}(T)$. The energy U(T) then follows from the derivative of $\log(\mathcal{Z})$, while the entropy follows from a calculation of the probabilities P(s). Note that in calculating the partition function we often use the multiplicity function, since

$$\mathcal{Z}(T) = \sum_E g(E) e^{-\frac{E}{k_B T}} \tag{2.24}$$

which is a simple Laplace transformation. This transformation can be inverted and used to obtain the multiplicity function $g(E)$ from the partition function, but that requires an integration in the complex temperature plane. Interesting, since complex temperature corresponds to time. We obtain factors like $e^{\imath E t}$. Field theory?? But that is for another course.

2.3   Work and pressure.

A system which remains in a given equilibrium state is rather boring. We would like to change the state of a system. This can be done by varying some of the external parameters. For example, if a system is in contact with a reservoir, we can change the temperature of the reservoir. The system will adjust itself until it is again in thermal equilibrium with the reservoir. This is an example of a process in which the system is brought from an equilibrium state with temperature T_i to an equilibrium state with temperature T_f. What is happening in between, though? If we suddenly switch the temperature of the reservoir, it will take some time before thermal equilibrium is restored again. The states in between are non-equilibrium states and the entropy of the system will increase. We can also change the temperature very slowly, in such a way that at each time the system is in thermal equilibrium with the reservoir. A process for which a system is always in thermal equilibrium in all of the intermediate states is called reversible. If we reverse the change in temperature of the reservoir, the system will go back to its initial state. If we suddenly switch the temperature back and forth, the system will end up in a different state with higher entropy. For the latter to happen we have to allow other variables to change too, if only T,S, and U can change the states are completely determined by T. Reversible processes are nice, since they can easily be described in mathematical terms. It is much harder to deal with irreversible processes.

   An important extensive parameter of a system is its volume. We will in this chapter still assume that the number of particles does not change. A change in volume will in general affect the energy eigenvalues of the quantum states in this system and we write $\epsilon_s(V)$. If we introduce changes in the volume of the system, we need a force to produce these changes. We assume that this force is homogeneous, and hence the only thing we need is an external pressure p. The

pressure p is the natural intensive parameter which has to be combined with the extensive volume. The work W performed on the system to reduce the volume of the system from V to $V - \Delta V$ (with ΔV small) is $p\Delta V$.

An example of a reversible process in which the volume of the system is reduced is a process where the system starts in a given state $\hat{s}$ and remains in that state during the whole process. It is clearly reversible, by definition. If we reverse the change in volume we will still be in the same quantum state, and hence return to the same state of the system at the original value of the volume. The energy of the system is $\epsilon_{\hat{s}}(V)$ and the change in this energy is equal to the work done on the system

$$p\Delta V = \Delta \epsilon_{\hat{s}}(V) = -\Delta V \left(\frac{\partial \epsilon_{\hat{s}}}{\partial V} \right)_N \tag{2.25}$$

Now consider an ensemble of systems, all with initial volume V and initial temperature T_i. The number of subsystems in a state s is proportional to the Boltzmann factor at T_i. We change the volume in a similar reversible process in which each subsystem remains in the same quantum state. The change in energy is given by

$$\Delta U = \sum_s \Delta \epsilon_s P(s) = p\Delta V \sum_s P(s) = p\Delta V \tag{2.26}$$

because the probabilities P(s) do not change. Since ΔV is a <u>reduction</u> in volume we find $p = -\left(\frac{\partial U}{\partial V} \right)$. We are working with two independent variables, however, and hence we have to specify which variable is constant. Since P(s) does not change, the entropy does not change according to the formula we derived before. The temperature, on the other hand, will have to change! Hence

$$p = -\left(\frac{\partial U}{\partial V} \right)_{S,N} \tag{2.27}$$

If we take other extensive variables into account we have to keep those the same too while taking this derivative. The last formula is identical to the formula for pressure derived in thermodynamics, showing that the thermodynamical and the statistical mechanical description of volume and pressure are the same. In thermodynamics we assumed that mechanical work performed on the outside world was given by $\bar{d}W = -pdV$. In general we have $dU = TdS - pdV$, but at constant entropy this reduces to $dU = -pdV$ and the formula for the pressure follows immediately. In our present statistical mechanical description we used the same formula, $-pdV$, for the work done by a system in a specific quantum state $\hat{s}$. But now we assume that the probabilities for the quantum states do not change, again leading to $dU = -pdV$. In thermodynamics the entropy is always used as a mathematical tool, and in order to arrive at the formula for the pressure we simply set $dS = 0$. In statistical mechanics the entropy is a quantity connected to the quantum states of a system. The fact that the entropy did not change in derivation for the formula for the pressure is a simple consequence of the requirement that the quantum states keep their original probabilities.

Another subtle point is that here we defined pressure via $p = -\left(\frac{\partial \epsilon_s}{\partial V}\right)_N$. This is a definition in a microscopic sense, and is the direct analogue of what we would expect based on macroscopic considerations. Because we derive the same work term as in thermodynamics, this microscopic definition is equivalent to the macroscopic, thermodynamic one.

VERY, VERY SIMPLE MODEL.

An example of a process outlined above is the following. This model is much too simple, though, and should not be taken too serious. Assume that the energy eigenvalues of our system depend on volume according to $\epsilon_s(V) = \epsilon_s(V_0)\frac{V}{V_0}$. During the process of changing the volume we also change the temperature and for each volume set the temperature by $T(V) = T_0 \frac{V}{V_0}$. In this case the probabilities are clearly unchanged because the ratio $\frac{\epsilon_s}{T}$ is constant. Hence in this very simple model we can describe exactly how the temperature will have to change with volume when the entropy is constant. The energy is given by $U(V, T(V)) = \sum \epsilon_s(V) P(s) = \frac{V}{V_0} U_0(T_0)$ and hence the pressure is $p = -\frac{U_0}{V_0}$. This is negative!

WHY THIS IS TOO SIMPLE.

The partition function in the previous example is given by

$$\mathcal{Z}(T, V) = sum_s e^{-\frac{V \epsilon_s(V_0)}{V_0 k_B T}} = F\left(\frac{V}{T}\right) \tag{2.28}$$

and hence the internal energy is

$$U = -k_B V \frac{F'\left(\frac{V}{T}\right)}{F\left(\frac{V}{T}\right)} \tag{2.29}$$

Also, from the formula for the entropy 2.23 we see that $S = G\left(\frac{V}{T}\right)$ and combining this with the formula for energy we get

$$U = V H(S) \tag{2.30}$$

which gives $p = -\frac{U}{V}$ and hence for the enthalpy $H = U + pV = 0$. This is exactly why this model is too simple. There are only two intensive state variables, and only one (T in this case) can be independent. We will need an extra dependence of the energy on another extensive state variable to construct a realistic system. The easiest way is to include N, the number of particles.

In general, beware of simple models! They can be nice, because they allow for quick calculations, and some conclusions are justified. In general, though, we need a dependence of the energy on at least two extensive state variables to obtain interesting results which are of practical importance.

VERY SIMPLE MODEL, REVISED.

Suppose that in the previous model we replace the energy relation by $\epsilon_s(V) = \epsilon_s(V_0) \left(\frac{V_0}{V}\right)^\alpha$. Then we would find $pV_0 = \alpha U_0$. This is positive when α is positive. For an ideal gas we have this form of relation between pressure and energy, with $\alpha = \frac{2}{3}$. This value of α is indeed consistent with the form of the energy eigenvalues for free particles. The energy relation is of the form $\epsilon = \frac{\hbar^2 k^2}{2m}$ and k is proportional to inverse length.

MORE ABOUT PRESSURE.

In summary, in the previous discussion we used the traditional definition of temperature to evaluate the work done on the system in a volume change. We linked this work to the change in eigenvalues in a particular process, in which the entropy remain the same, by using conservation of energy. In this way the pressure could be directly related to the thermodynamic functions we have available.

In order to find the pressure from the energy, we need U(S,V). This form is usually not available. Counting states we derived S(U,V) and hence we would like to relate the pressure to this function. A standard trick will do. A change in the entropy due to changes in the energy and volume is given by

$$dS = \left(\frac{\partial S}{\partial U}\right)_V dU + \left(\frac{\partial S}{\partial V}\right)_U dV \tag{2.31}$$

It is possible to change both the energy and the volume at the same time in such a way that the entropy does not change:

$$0 = \left(\frac{\partial S}{\partial U}\right)_V (\Delta U)_S + \left(\frac{\partial S}{\partial V}\right)_U (\Delta V)_S \tag{2.32}$$

After dividing by ΔV and taking the limit $\Delta V \to 0$ we get

$$0 = \left(\frac{\partial S}{\partial U}\right)_V \left(\frac{\partial U}{\partial V}\right)_S + \left(\frac{\partial S}{\partial V}\right)_U \tag{2.33}$$

or

$$p = T \left(\frac{\partial S}{\partial V}\right)_U \tag{2.34}$$

which is a more useful expression. In the next section we will also give the expression for the pressure if the temperature and the volume are specified. Of course, these tricks have been discussed extensively in thermodynamics, where changes of variables are the bread and butter of the game.

2.4 Helmholtz free energy.

The formula relating the differentials of the entropy, energy, and pressure is simplified by introducing the pressure and temperature. We find, as expected,

$$dU = TdS - pdV \qquad (2.35)$$

If other extensive parameters are allowed to change, we have to add more terms to the right hand side. The change in energy U, which is also called the internal energy, is equal to the work done on the system $-pdV$ plus the flow of heat to the system if the system is in contact with some reservoir, TdS. When the entropy of the system does not change, the complete change in internal energy can be used to do work. Therefore, the internal energy is a measure of the amount of work a system can do on the outside world in processes in which the entropy of the system does not change. Most processes are of a different nature, however, and in those cases the internal energy is not useful.

An important group of processes is described by changes in the volume at constant temperature. In order to describe these processes we introduce the Helmholtz free energy F:

$$F = U - TS \qquad (2.36)$$

which has the differential

$$dF = -SdT - pdV \qquad (2.37)$$

Therefore, in a process at constant temperature the work done on the system is equal to the change in the Helmholtz free energy. The Helmholtz free energy measures the amount of work a system can do on the outside world in processes at constant temperature. Since both the entropy and temperature are positive quantities, the Helmholtz free energy is always less than the internal energy. This is true in general. The amount of energy which can be retrieved from a system is usually smaller than the internal energy. Part of the internal energy is stored as heat and is hard to recover.

2.5   Changes in variables.

The variables we have introduced so far are S,T,U,V, and p. These variables are not independent, and if we specify two the other three are determined. We can choose, however, which two we want to specify. The cases we have considered either use (U,V) as independent variables and calculate all other quantities from S(U,V) or use (T,V) as independent variables and calculate all other quantities through the partition function $\mathcal{Z}(T,V)$ and U(T,V) as a derivative of $\log(\mathcal{Z})$. This implies restrictions on functions like S(U,V). For example, if we know S(U,V) the temperature follows from the partial derivative $\left(\frac{\partial S}{\partial U}\right)_V$. If we specify the temperature, the energy U is found by solving

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_V (U,V) \qquad (2.38)$$

Since we know that this equation has a unique solution, there are no two values of U for which S(U,V) has the same slope! Hence T is a monotonous

function of U and U is an increasing function of T! The ability to switch back and forth between different sets of independent variables implies constant signs for certain derivatives, and for response functions. For example, a solid always expands as a function of temperature. In thermodynamics we related these properties in a rigorous way to criteria for stability.

2.6 Properties of the Helmholtz free energy.

For a system at constant temperature and volume, the Helmholtz free energy is the state function determining all physical properties of the system. For example, consider a system in thermal contact with a large reservoir at temperature T. The volume of the system is V. Energy can flow back and forth between the system and the reservoir; the energy of the system is ϵ and the energy of the combination reservoir plus system is U_0. In order to find the value of ϵ which corresponds to the equilibrium value for the system, i.e. which is equal to U(T,V), we have to maximize the entropy of the combination system plus reservoir. Since

$$S_{tot}(\epsilon) = S_{\mathcal{R}}(U_0 - \epsilon) + S_{\mathcal{S}}(\epsilon) \approx S_{\mathcal{R}}(U_0) - \frac{1}{T}(\epsilon - TS_{\mathcal{S}}(\epsilon)) \qquad (2.39)$$

this means we have to minimize $\epsilon - TS_{\mathcal{S}}(\epsilon)$. In other words, if we specify the temperature and volume of a system, we can give the energy ϵ an arbitrary value, but that value will in general not correspond to a system in equilibrium. The equilibrium value of ϵ is found by minimizing the Helmholtz free energy $\epsilon - TS(\epsilon)$. This equilibrium value is called U. Very often, notation is sloppy and one tells you to minimize $U - TS(U)$. The Helmholtz energies in statistical mechanics and in thermodynamics therefore have identical properties, and again the statistical mechanical and thermodynamical descriptions are the same.

The pressure as a function of temperature and volume is related to the internal energy U and the entropy S by

$$p(T, V) = -\left(\frac{\partial U}{\partial V}\right)_T + T\left(\frac{\partial S}{\partial V}\right)_T \qquad (2.40)$$

Both terms can be important, although in general in a solid the first term is larger and in a gas the second. Beware, however, of the errors introduced by approximating the pressure by only one of the terms in this expression.

How are they related?

In thermodynamics we have seen that a system as a function of T and V (and N) is described by the Helmholtz energy $F(T, V, N)$. In statistical mechanics we needed the partition function $\mathcal{Z}(T, V, N)$. These two functions must therefore be related. It is possible to eliminate S from the definition of the Helmholtz free energy F, since $S = -\left(\frac{\partial F}{\partial T}\right)_{V,N}$. This gives

$$F = U + T \left(\frac{\partial F}{\partial T} \right)_{V,N} \tag{2.41}$$

which leads to the following differential equation for F

$$\frac{\partial}{\partial T} \left(\frac{F}{T} \right) = -\frac{U}{T^2} \tag{2.42}$$

Combined with the relation between U and $\log(\mathcal{Z})$ this leads to the simple equation

$$\left(\frac{\partial \frac{F}{T}}{\partial T} \right)_{V,N} = -k_B \left(\frac{\partial \log(\mathcal{Z})}{\partial T} \right)_{V,N} \tag{2.43}$$

or

$$F = -k_B T \ log(\mathcal{Z}) + c(V,N)T \tag{2.44}$$

The constant of integration c(V,N) is determined by the state of the system at low temperatures. Suppose the ground state has energy ϵ_0 and degeneracy g_0. The only terms playing a role in the partition function at low temperature correspond to the ground state and hence we have

$$\lim_{T \to 0} \mathcal{Z} = g_0 e^{-\frac{\epsilon_0}{k_B T}} \tag{2.45}$$

and hence in lowest order in T near T=0:

$$F(T,V,N) = -k_B T \ log(g_0) + \epsilon_0 + c(V,N)T \tag{2.46}$$

which gives for the entropy

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V,N} \approx k_B \log(g_0) - c(V,N) \tag{2.47}$$

and since we know that at $T = 0$ the entropy is given by the first term, we need to have $c(V,N) = 0$, and

$$F(T,V,N) = -k_B T \log(\mathcal{Z}(T,V,N)) \tag{2.48}$$

or

$$\mathcal{Z}(T,V,N) = e^{-\frac{F(T,V,N)}{k_B T}} \tag{2.49}$$

In other words, we not only know that F can be expressed as a function of T and V (and N) , we also have an explicit construction of this function. First we specify the volume V, then calculate all the quantum states $\epsilon_s(V)$ of the system. Next we specify the temperature T and evaluate the partition function $\mathcal{Z}(T,V)$, the Helmholtz free energy F(T,V) and all other quantities of interest.

The form of the relation between the partition function and the Helmholtz free energy is easily memorized by writing the equation in the following form,

$$e^{-\frac{F(T,V,N)}{k_B T}} = \sum_s e^{-\frac{\epsilon_s}{k_B T}} \tag{2.50}$$

which has Boltzmann like factors on each side.

The entropy follows from

$$S = \frac{1}{T}(U-F) = k_B \left(\sum_s \frac{\epsilon_s}{k_B T} P(s) + \log(\mathcal{Z}) \right) = k_B \sum_s \left(\frac{\epsilon_s}{k_B T} + \log(\mathcal{Z}) \right) P(s) \tag{2.51}$$

which gives

$$S = -k_B \sum_s P(s) \log(P(s)) \tag{2.52}$$

as before. This is just an easier way to derive the entropy as a function of volume and temperature in terms of the elementary probabilities.

2.7 Energy fluctuations.

A system in thermal contact with a large reservoir will have the temperature of this reservoir when equilibrium is reached. In thermal contact, energy is allowed to flow back and forth between the system and the reservoir and the energy of the system will not be equal to U at all times, only as an average. Thermal equilibrium implies only that there is no <u>net</u> flow of energy between the system and the reservoir.

ARE THE FLUCTUATIONS REAL?

So far we have made the connection between thermodynamics and statistical mechanics in the following sense. Thermodynamics is a macroscopic theory which is experimentally correct. Statistical mechanics is a microscopic theory, and the averages calculated in this theory seem to be the same as the corresponding quantities in thermodynamics. We have not addressed the concept of fluctuations, however. Fluctuations are much more difficult to describe in thermodynamics, although we have given some prescription how to do this. Fluctuations are easy to describe in statistical mechanics. But are the fluctuations described in statistical mechanics really what we observe experimentally? How does the thermodynamic limit play a role here? These are very difficult questions, but here we simply assume that the answer is that we have the correct description of fluctuations in thermodynamics, since the theory looks like the thermodynamic form. The fact that in the thermodynamic limit all distributions are Gaussian will have to play an important role.

In order to get the root-mean-square fluctuations $\Delta\epsilon$ in the energy, we have to calculate $< \epsilon_s^2 >$. These fluctuations are related to the heat capacity at constant volume, C_V, as can be shown in the following manner. By definition:

$$U =< \epsilon_s >= \frac{1}{\mathcal{Z}} \sum_s \epsilon_s e^{-\frac{\epsilon_s}{k_B T}} \tag{2.53}$$

$$C_V = \left(\frac{\partial U}{\partial T}\right)_{V,N} = \frac{1}{\mathcal{Z}} \frac{1}{k_B T^2} \sum_s \epsilon_s^2 e^{-\frac{\epsilon_s}{k_B T}} - \frac{1}{\mathcal{Z}^2} \sum_s \epsilon_s e^{-\frac{\epsilon_s}{k_B T}} \left(\frac{\partial \mathcal{Z}}{\partial T}\right)_{V,N} \tag{2.54}$$

$$C_V = \frac{1}{k_B T^2} < \epsilon_s^2 > -\frac{1}{\mathcal{Z}} \sum_s \epsilon_s e^{-\frac{\epsilon_s}{k_B T}} \left(\frac{\partial \log(\mathcal{Z})}{\partial T}\right)_{V,N} \tag{2.55}$$

$$C_V = \frac{1}{k_B T^2} < \epsilon_s^2 > - < \epsilon_s > \frac{U}{k_B T^2} \tag{2.56}$$

which gives with the relation $U =< \epsilon_s >$ that

$$k_B T^2 C_V =< \epsilon_s^2 > - < \epsilon_s >^2= (\Delta\epsilon)^2 \tag{2.57}$$

Two important conclusions are based on this formula. First of all, the right hand side is positive, and hence the heat capacity is positive. Hence the internal energy U is an increasing function of the temperature! Second, the heat capacity is an extensive quantity and hence $\Delta\epsilon \propto \sqrt{N}$. This tells us that the fluctuations in the energy increase with increasing system size. A more important quantity is the energy per particle, $\frac{U}{N}$. The fluctuations in the energy per particle are therefore proportional to $\frac{1}{\sqrt{N}}$. The energy per particle is therefore well defined in the thermodynamic limit, fluctuations are very small. Also, this formula for the energy fluctuations is very useful in numerical simulations. In these calculations one often follows the state of a system as a function of time at a given temperature, and the fluctuations in the energy are a direct result of the calculations. By varying the temperature, one also obtains U(T,V). Hence there are two independent ways of calculating the heat capacity, and this gives a very good test of the computer programs! Finally, the formula relating the heat capacity and the fluctuations in energy is an example of a general class of formulas. Response functions are always related to fluctuations in a system in equilibrium. In our case, if the fluctuations in energy of a given system at a given temperature are small, it is apparently difficult to change the energy of this system. Hence we expect also that we need a large change in temperature to make a small change in energy, and this implies that the heat capacity is small.

2.8 A simple example.

In order to illustrate the ideas developed in this chapter, we consider a simple example, even though simple examples are dangerous. The energy eigenstates of a system are given by two quantum numbers, n and l. The possible values for n are $1, 2, 3, \cdots, \infty$, but the values for l are limited by $l = 1, 2, \cdots, n$. The energy eigenvalues depend only on n:

$$\epsilon_{nl} = n\hbar\omega \tag{2.58}$$

where ω is a positive number. This problem is essentially a harmonic oscillator in *two* dimensions. The partition function follows from

$$\mathcal{Z}(T) = \sum_{n=1}^{\infty} n \left(e^{-\frac{\hbar\omega}{k_B T}} \right)^n \tag{2.59}$$

and since the factor in parenthesis is always between zero and one, this can be summed to give

$$\mathcal{Z}(T) = \frac{e^{-\frac{\hbar\omega}{k_B T}}}{(1 - e^{-\frac{\hbar\omega}{k_B T}})^2} \tag{2.60}$$

The Helmholtz free energy is

$$F(T) = \hbar\omega + 2k_B T \log(1 - e^{-\frac{\hbar\omega}{k_B T}}) \tag{2.61}$$

The derivative with respect to temperature yields the entropy

$$S(T) = \hbar\omega \frac{2}{T} \frac{1}{e^{\frac{\hbar\omega}{k_B T}} - 1} - 2k_B \log(1 - e^{-\frac{\hbar\omega}{k_B T}}) \tag{2.62}$$

The internal energy is $F + TS$ and is given by

$$U(T) = \hbar\omega \coth(\frac{\hbar\omega}{2k_B T}) \tag{2.63}$$

Note that the term with the logarithm disappeared in the formula for the energy U. This is characteristic of many problems. The Helmholtz free energy is $F = -k_B T \log(\mathcal{Z})$ and hence the entropy is $S = k_B \log \mathcal{Z} + k_B T \frac{1}{\mathcal{Z}} \left(\frac{\partial \mathcal{Z}}{\partial T} \right)_{V,N}$. Therefore the terms with the logarithms cancel in $F + TS$.

At this point it is always useful to consider the entropy and energy in two limiting cases, small and large temperatures. In our problem the natural energy scale is $\hbar\omega$ and the temperature is small if the thermal energy $k_B T$ is small compared with the spacing between the energy levels, $k_B T \ll \hbar\omega$. The temperature is large if the thermal energy $k_B T$ covers many energy levels, $k_B T \gg \hbar\omega$. For small values of the temperature we find

$$T \to 0 : U \approx \hbar\omega, S \approx 0 \tag{2.64}$$

as expected. The system is in the ground state, which is non-degenerate. For large values of the temperature we get

$$T \to \infty : U \approx 2k_BT, S \approx 2k_B - 2k_B \log(\frac{\hbar\omega}{k_BT}) \tag{2.65}$$

This makes sense too. When the temperature is large, the details of the quantum levels are unimportant and the internal energy should be proportional to k_BT. The number of quantum states at energy k_BT is $\frac{k_BT}{\hbar\omega}$ and the entropy should be proportional to $\log(\frac{k_BT}{\hbar\omega})$.

The heat capacity is also easy to get:

$$C(T) = \hbar\omega \frac{d}{dT} \coth(\frac{\hbar\omega}{2k_BT}) = \frac{(\hbar\omega)^2}{2k_BT^2} \frac{1}{sinh^2(\frac{\hbar\omega}{2k_BT})} \tag{2.66}$$

The heat capacity takes the simple limiting forms

$$T \to 0 : C \approx 0 \tag{2.67}$$

$$T \to \infty : C \approx 2k_B \tag{2.68}$$

In a later chapter we will derive the result that a classical harmonic oscillator in D dimensions has an internal energy of the form Dk_BT . If the temperature is large, our quantum system will mainly be in states with large quantum numbers, and the correspondence principle tells us that such a system behaves like a classical system. Our calculated value of the internal energy in the limit $T \to \infty$ is consistent with this idea.

In order to introduce other extensive variables, we have to make assumptions for ω. Suppose the frequency ω is a linear function of volume, $\omega = \alpha V$, with $\alpha > 0$. In that case we have F(T,V) and the pressure is given by

$$pV = -\hbar\omega \coth(\frac{\hbar\omega}{k_BT}) \tag{2.69}$$

This pressure is negative! This is an immediate consequence of our relation between frequency and volume, which implies that the energy <u>increases</u> with increasing volume, opposite to the common situation in a gas of volume V. The previous formula also shows that the pressure is directly related to the temperature, as it should be when there are only two sets of conjugate variables. The limiting values of the pressure are

$$T \to 0 : pV \approx -\hbar\omega \tag{2.70}$$

$$T \to \infty : pV \approx -2k_BT \tag{2.71}$$

We can compare these values with the formula for the pressure we derived before

$$p = -\left(\frac{\partial U}{\partial V}\right)_T + T\left(\frac{\partial S}{\partial V}\right)_T \tag{2.72}$$

If the temperature is small, only the first term is important. If the temperature is large, the second term is dominant. This is true in general. At low temperatures pressure is mainly related to changes in the internal energy, while at high temperature it is mainly due to changes in the entropy.

Finally, note that in our example $U = -pV$. The Gibbs-Duhem relation is $U = TS - pV$, and these two results are inconsistent. This is not surprising. The Gibbs-Duhem relation is based on the assumption that the energy is an extensive quantity. In our example this is not true. Neither the entropy nor the energy double their value when the volume is increased by a factor of two. This indicates that one has to be careful in introducing extensive parameters in a quantum-mechanical formalism. Our example is too simple, again!

2.9 Problems for chapter 2

PROBLEM 1.

The quantum states of a system depend on a single quantum number n, with $n = 0, 1, 2, \cdots, \infty$. The energy of state n is $\epsilon_n = n\epsilon$ (with $\epsilon > 0$). The system is at a temperature T.

(1) Calculate the partition function $\mathcal{Z}(T)$

(2) Calculate the energy U(T) and the entropy S(T)

(3) Calculate T(U)

(4) Calculate S(U) and check that $\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)$

PROBLEM 2.

The partition function of a system with quantum states $n = 0, 1, 2, \cdots, \infty$ and energies f(n) is given by $\mathcal{Z}_f(T)$ and for a system with energies g(n) by $\mathcal{Z}_g(T)$. The corresponding internal energies are $U_f(T)$ and $U_g(T)$; the corresponding entropies are $S_f(T)$ and $S_g(T)$. The quantum states of a composite system depend on two quantum numbers n and m, with $n, m = 0, 1, 2, \cdots, \infty$. The energy of a state n,m is $\epsilon_{n,m} = f(n) + g(m)$.

(1) Find the partition function $\mathcal{Z}(T)$ for the composite system in terms of $\mathcal{Z}_f(T)$ and $\mathcal{Z}_g(T)$

(2) Do the same for the energy and the entropy

PROBLEM 3.

A system has only two quantum states available. State number one has energy zero and state number two has energy $\epsilon > 0$.

(a) Calculate the partition function for this system.

(b) Calculate U(T) and S(T).

(c) Calculate the heat capacity as a function of temperature.

The heat capacity as a function of temperature peaks at a certain value of the temperature. This peak is called a Schottky anomaly.

(d) Is this peak associated with a phase transition? Give an argument justifying your answer.

PROBLEM 4.

(From Reif) A sample of mineral oil is placed in an external magnetic field H. Each proton has spin $\frac{1}{2}$ and a magnetic moment μ; it can, therefore, have two possible energies $\epsilon = \mp\mu H$, corresponding to the two possible orientations of its spin. An applied radio frequency field can induce transitions between these two energy levels if its frequency ν satisfies the Bohr condition $h\nu = 2\mu H$. The power absorbed from this radiation field is then proportional to the difference in the number of nuclei in these two energy levels. Assume that the protons in the mineral oil are in thermal equilibrium at a temperature T which is so high that $\mu H \ll kT$. How does the absorbed power depend on the temperature T of the sample?

PROBLEM 5.

(From Reif) A system consists of N weakly interacting particles, each of which can be in either of two states with respective energies ϵ_1 and ϵ_2, where $\epsilon_1 < \epsilon_2$.
    (a) Without explicit calculation, make a qualitative plot of the mean energy $\bar{E}$ of the system as a function of its temperature T. What is $\bar{E}$ in the limit of very low and very high temperatures? Roughly near what temperature does $\bar{E}$ change from its low to its high temperature limiting values?
    (b) Using the result of (a), make a qualitative plot of the heat capacity C_V (at constant volume) as a function of the temperature T.
    (c) Calculate explicitly the mean energy $\bar{E}(T)$ and heat capacity $C_V(T)$ of this system. Verify that your expressions exhibit the qualitative features discussed in (a) and (b).

PROBLEM 6.

(From Reif) The following describes a simple two-dimensional model of a situation of actual physical interest. A solid at absolute temperature T contains N negatively charged impurity atoms per cm^3; these ions replacing some of the ordinary atoms of a solid. The solid as a whole, of course, is electrically neutral. This is so because each negative ion with charge $-e$ has in its vicinity one positive ion with charge $+e$. The positive ion is small and thus free to move between lattice sites. In the absence of an electrical field it will therefore be found with equal probability in any one of the four equidistant sites surrounding the stationary negative ion (see diagram in hand-out).

If a small electrical field E is applied along the x direction, calculate the electric polarization, i.e., the mean electric dipole moment per unit volume along the x direction.

PROBLEM 7.

The probability of finding a system in a state s is P_s. In this case the entropy of the system is $S = -k_B \sum_s P_s \log(P_s)$. Assume that the system consists of two independent subsystems and that the state s is the combination of subsystem 1 being in state s_1 and system 2 being in state s_2. Use the formula for the entropy given above to show that $S = S_1 + S_2$.

PROBLEM 8.

The energy eigenvalues for a single particle in a given potential are ϵ_n. These energy values are independent of how many other particles are present. Show that the partition function $\mathcal{Z}(T, N)$ obeys the relation $\mathcal{Z}(T, N) = (\mathcal{Z}(T, 1))^N$.

PROBLEM 9.

The energy eigenvalues of a system are given by 0 and $\epsilon + n\Delta$ for $n = 0, 1, 2, \cdots$. We have both $\epsilon > 0$ and $\Delta > 0$. Calculate the partition function for this system. Calculate the internal energy and the heat capacity. Plot the heat capacity as a function of temperature for $0 < k_B T < \epsilon$ for (a) $\Delta \gg \epsilon$, (b) $\Delta = \epsilon$, and (c) $\Delta \ll \epsilon$.

Chapter 3

Systems with a variable number of particles.

3.1 Chemical potential.

MORE STATE VARIABLES?

The extensive system parameters we have studied are S,V,N, and U. Until now we have allowed for variations in S (defining temperature) and V (defining pressure). In many processes the number of particles is also allowed to change. The corresponding intensive parameter is the chemical potential, μ. In this section we will discuss the meaning of the chemical potential. In addition, the system could have other extensive variables. Changes in these variables lead to new intensive variables. The recipes in those cases are constructed in a straightforward manner by copying the steps in the development of T,p, and μ.

Two systems are in <u>diffusive contact</u> if particles can flow back and forth between them. Consider two systems with temperature T, and fixed volumes. The number of particles in system 1(2) is $N_1(N_2)$. The equilibrium configuration is found by minimizing the Helmholtz free energy of the total system, since the total system is at constant temperature and volume. If ΔN particles flow from 1 to 2, the change in free energy is

$$\Delta F = \left(\frac{\partial F_1}{\partial N}\right)_{T,V} (-\Delta N) + \left(\frac{\partial F_2}{\partial N}\right)_{T,V} (+\Delta N) \qquad (3.1)$$

We define the <u>chemical potential</u> by

$$\mu(T,V,N) = \left(\frac{\partial F}{\partial N}\right)_{T,V} (T,V,N) \qquad (3.2)$$

and the free energy is a minimum if the chemical potentials are equal, $\mu_1 = \mu_2$. This, of course, is similar to the conditions we found before for temperature and pressure. Although we could have defined the chemical potential as an arbitrary function of the partial derivative, the definition above is the only choice which is consistent with the experimentally derived chemical potential. The definition in terms of a derivative only holds for large values of N, in general one should write

$$\mu(T, V, N) = F(T, V, N) - F(T, V, N - 1) \tag{3.3}$$

which tells us that the chemical potential equals the free energy needed to subtract one particle to the system. This is another reminder that of a difference between thermodynamics and statistical mechanics. In thermodynamics, the number of moles N of a material is a continuous variable. In statistical mechanics, the number of particles N is a discrete variable, which can be approximated as continuous in the thermodynamic limit (showing up again!). Finally, if different types of particles are present, on specifies a chemical potential for each type according to

$$\mu_i(T, V, N_1, N_2, ...) = \left(\frac{\partial F}{\partial N_i} \right)_{T, V, N_{j \neq i}} \tag{3.4}$$

CHEMICAL POTENTIAL, WHAT DOES IT MEAN?

What is this chemical potential? The following example will help you to understand this important quantity. If the systems 1 and 2 are not in equilibrium, the Helmholtz free energy is reduced when particles flow from the system with a high value of the chemical potential to the system with the lower value, since $\Delta F = (\mu_2 - \mu_1)\Delta N$. We can also add some other potential energy to the systems. For example, the systems could be at different heights are have a different electric potential. This potential energy is Φ_i. If we bring ΔN particles from system 1 to system 2, we have to supply an amount of energy $\Delta E = (\Phi_2 - \Phi_1)\Delta N$ to make this transfer. Since this is a direct part of the energy, the total change in the Helmholtz free energy is given by

$$\Delta F = (\hat{\mu}_2 + \Phi_2 - \hat{\mu}_1 - \Phi_1)\Delta N \tag{3.5}$$

where $\hat{\mu}$ is the chemical potential evaluated *without* the additional potential energy. Therefore, if we choose Φ such that

$$\hat{\mu}_2 + \Phi_2 = \hat{\mu}_1 + \Phi_1 \tag{3.6}$$

the two systems will be in equilibrium. Of course, the presence of an external potential Φ leads to a term ΦN in the internal energy U and hence in F. As a result

$$\mu(T, V, N) = \left(\frac{\partial F}{\partial N}\right)_{T,V} = \hat{\mu} + \Phi \tag{3.7}$$

This is called the <u>total</u>chemical potential. **Two systems are in equilibrium when the total chemical potential is the same**. Ordinary, standard potentials therefore give a direct contribution to the chemical potential. But even without the presence of external potentials, the Helmholtz free energy will still depend on the number of particles N and we will find a value for the chemical potential. The value of the chemical potential without any external forces is called the <u>internal</u> chemical potential. Hence the total chemical potential is the sum of the internal chemical potential and the external potentials! Confusion arises when one uses the term chemical potential without specifying total or internal. Some textbooks assume that chemical potential means internal chemical potential, others imply total chemical potential. Keep this in mind when reading articles, books, and exam problems!! There is even more confusion, since we often call the quantity U the internal energy. If that would mean the energy without the effect of external potentials, the total energy would be the sum of these two quantities. We could define

$$U_{total} = U_{internal} + N\Phi \tag{3.8}$$

$$F_{internal} = U_{internal} - TS \tag{3.9}$$

$$F_{total} = F_{internal} + N\Phi \tag{3.10}$$

$$\mu = \hat{\mu} + \Phi \tag{3.11}$$

in order to be consistent. But I prefer to call U_{total} the internal energy and only use μ and not $\hat{\mu}$.

ORIGIN OF INTERNAL CHEMICAL POTENTIAL.

The chemical potential is related to the partial derivative of the Helmholtz free energy $F = U - TS$. Hence we have

$$\mu = \left(\frac{\partial U}{\partial N}\right)_{T,V} - T\left(\frac{\partial S}{\partial N}\right)_{T,V} \tag{3.12}$$

If the particles are independent, the internal energy is linear in N and the first contribution to the chemical potential is independent of N (but does depend on T and V). In real life, however, adding particles to a system changes the interactions between the particles that are already present. The energy depends on N in a non-linear manner. This is called correlation. But even if the particles are independent, the entropy will be non-linear in N, since adding particles will change the number of ways the particles already present can be distributed over

available states. Hence even for independent particles the chemical potential will depend on N.

Simple models of chemical potential effects often ignore any volume dependence. We already know from thermodynamics that this cannot be sufficient, because in that case the chemical potential and the temperature are not independent variables. This can be easily demonstrated, since in this case we need to have:

$$S = Nf(\frac{U}{N}) \tag{3.13}$$

for some function $f(x)$. This gives

$$\frac{1}{T} = f'(\frac{U}{N}) \tag{3.14}$$

or $U = Ng(T)$ and hence $S = Nf(g(T))$. In this case $F = Nh(T)$ is linear in N and $\mu = h(T)$. Even though such simple models are suspect, they can still be useful in some discussions.

BASIC FORMULATION.

In summary we have:

The internal chemical potential is a thermodynamic variable which is equivalent to a standard potential energy: the value of the difference in internal chemical potential between two systems is equal to the magnitude of the potential barrier needed to bring these two systems into diffusive equilibrium.

In other words, two systems are in diffusive contact are in equilibrium when the total chemical potentials are the same.

3.2   Examples of the use of the chemical potential.

ELECTRICITY IS ALWAYS USEFUL.

As an example we consider a standard, old-fashioned battery. This simple example is very useful to help us gain further understanding of what a chemical potential is. The negative electrode is made of Pb, the positive electrode of PbO_2 with a Pb core. The electrolyte is diluted H_2SO_4, in other words mixed with water . The chemical reaction at the negative electrode is

$$Pb + SO_4^{2-} \rightarrow PbSO_4 + 2e^- + 0.8eV \tag{3.15}$$

and at the positive electrode

$$PbO_2 + 2H^+ + H_2SO_4 + 2e^- \rightarrow PbSO_4 + 2H_2O + 3.2eV \qquad (3.16)$$
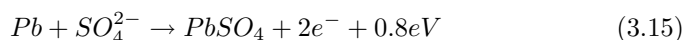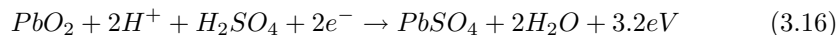
We ignore free electrons in the electrolyte, and hence the only place where we can take electrons or dump electrons is in the Pb of the electrodes. Therefore, if the terminals of the battery are not connected by an external wire, these reactions have as a net result a transport of electrons from the positive electrode to the negative electrode. The amount of energy gained by transferring **one** electron from the positive electrode to the negative electrode through these reactions is 2 eV. Therefore, the internal chemical potential is higher at the positive electrode, and $\hat{\mu}_+ - \hat{\mu}_- = 2eV$. Of course, after a number of electrons are transported from the positive electrode to the negative electrode there will be a difference in electrical potential energy between the electrodes. Suppose the potentials at these electrodes are V_+ and V_-. The flow of electrons in this unconnected battery will have to stop when the total chemical potential is constant, or

$$\hat{\mu}_+ + (-e)V_+ = \hat{\mu}_- + (-e)V_- \Rightarrow V_+ - V_- = 2V \qquad (3.17)$$

Do we know how much charge has been transferred? No, unless we know the capacitance of the battery, which depends on the shape and geometry of the plates and the space between them.

If we connect the terminals of the battery through an outside resistor, electrons will flow in the outside circuit. This will reduce the number of electrons in the negative electrode, and as a result the chemical reactions will continue to make up for the electrons leaving through the outside circuit. The actual difference in potential between the electrodes will be less than 2V and will depend on the current through the outside circuit. Of course, in real life the potential difference is always less than 2V because of internal currents through the electrolyte.

GRAVITY AND ATMOSPHERIC PRESSURE.

A second example is a very simplified study of the pressure in the atmosphere. Consider a column of gas in a gravitational field. There is only one type of molecule in this gas. The gravitational potential at a height h in the column is mgh, where m is the mass of a molecule. We assume that the temperature T is the same everywhere, in other words this is a model of a isothermal atmosphere. Probably not very realistic, but not a bad start. The density $n = \frac{N}{V}$ of the molecules is a function of h. Next, we focus on a slice of gas at a height h with a thickness dh. The value of dh is small enough that the density $n(h)$ does not vary appreciably within this slice. But it is also large enough that the slice of the gas contains many molecules. This is possible only because the gravitational potential doe not very much on the scale of the interatomic distance! We assume that the gas inside this slice through the column

is an ideal gas. This is again an idealization. The internal chemical potential for an ideal gas with density n and temperature T is

$$\hat{\mu}(n,T) = k_B T \log(n) + \mu_0(T) \tag{3.18}$$

which only depends on n via the logarithm. The gas in this column in the atmosphere is in diffusive contact, and equilibrium requires that the chemical potential is the same everywhere. Therefore we find:

$$k_B T \log(n(h)) + mgh = k_B T \log(n(0)) \tag{3.19}$$

which leads to

$$n(h) = n(0) e^{-\frac{mgh}{k_B T}} \tag{3.20}$$

Because the gas inside a small slice is ideal, we can use the ideal gas equation of state, $pV = NkT$, to obtain the isothermal atmospheric pressure relation

$$p(h) = p(0) e^{-\frac{mgh}{k_B T}} \tag{3.21}$$

3.3   Differential relations and grand potential.

The six most common basic variables describing a system are the extensive parameters entropy S, volume V, number of particles N, and the intensive parameters temperature T, pressure p, and chemical potential μ. Every new measurable quantity adds one extensive parameter and one intensive parameter to this list. The free energy is measures by the internal energy U or of a quantity derived from the internal energy by a Legendre transformation. Of the six basic variables only three are independent, the other three can be derived. As we have seen before we need at least one extensive parameter in the set of independent variables, otherwise the size of the system would be unknown.

If our choice of basic variables is T,V, and N, the way to construct the other functions is via the partition function and the Helmholtz free energy, using partial derivatives to get S,P, and μ and using $U = F + TS$ to get U. If our choice of basic variables is U,V, and N, we calculate the entropy first. Remember that

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U} \right)_{V,N} \tag{3.22}$$

and

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V} \right)_{U,N} \tag{3.23}$$

The third partial derivative can be obtained in the following way. A general change in the entropy as a result of changes in the basic variables is

$$dS = \left(\frac{\partial S}{\partial U}\right)_{V,N} dU + \left(\frac{\partial S}{\partial V}\right)_{U,N} dV + \left(\frac{\partial S}{\partial N}\right)_{U,V} dN \qquad (3.24)$$

Consider a *reversible* change in the state of the system in which we keep T and V constant but change S,U, and N:

$$(\Delta S)_{T,V} = \left(\frac{\partial S}{\partial U}\right)_{V,N} (\Delta U)_{T,V} + \left(\frac{\partial S}{\partial N}\right)_{U,V} (\Delta N)_{T,V} \qquad (3.25)$$

After dividing by $(\Delta N)_{T,V}$ and taking the limit of the variations to zero we find

$$\left(\frac{\partial S}{\partial N}\right)_{T,V} = \left(\frac{\partial S}{\partial U}\right)_{V,N} \left(\frac{\partial U}{\partial N}\right)_{T,V} + \left(\frac{\partial S}{\partial N}\right)_{U,V} \qquad (3.26)$$

$$\left(\frac{\partial S}{\partial N}\right)_{T,V} - \frac{1}{T}\left(\frac{\partial U}{\partial N}\right)_{T,V} = \left(\frac{\partial S}{\partial N}\right)_{U,V} \qquad (3.27)$$

$$\frac{1}{T}\left(\frac{\partial F}{\partial N}\right)_{T,V} = \left(\frac{\partial S}{\partial N}\right)_{U,V} \qquad (3.28)$$

from which we get, using the definition of the chemical potential,

$$\left(\frac{\partial S}{\partial N}\right)_{U,V} = -\frac{\mu}{T} \qquad (3.29)$$

The relation between the differential changes in dS and dU,dV, and dN is now very simple and is usually written in the form

$$dU = TdS - pdV + \mu dN \qquad (3.30)$$

This is the standard formulation of the first law, as expected. It confirms the point we learned in thermodynamics, that with S,V,N as independent variables the partial derivatives of U give the remaining dependent variables. The set S,V,N is therefore a natural set for U. Of course, if we for example use T,V, and N as independent variables we can also calculate U(T,V,N) and the relation between the differentials. In this case, however, the coefficients are not simple functions of the dependent variables, but more complicated response functions:

$$dU = C_V dT + \left(-p + T\left(\frac{\partial S}{\partial V}\right)_{T,N}\right) dV + \left(\frac{\partial U}{\partial N}\right)_{T,V} dN \qquad (3.31)$$

The Helmholtz free energy F was introduced exactly for that reason, because it has T,V,N as natural independent variables:

$$dF = -SdT - pdV + \mu dN \qquad (3.32)$$

One important class of processes are those which occur at constant T,V, and N. A second group keeps T,V, and μ constant. This is certainly true in many

chemical reactions! How do we describe the equilibrium state of a system where the values of μ, T, and V are specified. A simple observation is based on the relations between the differentials. If we define $\Omega = F - \mu N$ we find

$$d\Omega = -SdT - pdV - Nd\mu \qquad (3.33)$$

Hence Ω has as natural variables T,V, and μ. This is exactly what we need. Also, for a process at constant μ and T the change in Ω is equal to the work performed by the system. In other words, Ω measures the amount of energy available in a system to do work at constant chemical potential and temperature. The free energy Ω is called the grand potential.

3.4   Grand partition function.

One way of calculating Ω is via the Helmholtz free energy F(T,V,N), obtained from the partition function. If we solve $\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}$ for $N(\mu, T, V)$ we can construct $\Omega(\mu, T, V) = F(T, V, N(\mu, T, V)) - \mu N(\mu, T, V)$. This assumes that we are able to invert the formula for $\mu(T, V, N)$ and solve for N uniquely. That this is possible is by no means obvious from a mathematical point of view only. Of course, in thermodynamics we know that the relation between μ and N is never decreasing, and hence if we did thermodynamics there would be no problem. Here we are doing statistical mechanics, though, and we have to show that the statistical mechanical description is equivalent to the thermodynamical one.

We will now describe a procedure to calculate the grand potential directly, and hence in this procedure we find the inverse relation for N. The procedure is the analogue of what we did to construct the partition function. It is a very general construction, which works for all types of free energies! Consider a system $\mathcal{S}$ in thermal and diffusive equilibrium with a very large reservoir $\mathcal{R}$. Both energy and particles can be exchange. The total energy U_0 of $\mathcal{R} + \mathcal{S}$ and the total number of particles N_0 in $\mathcal{R} + \mathcal{S}$ are constant. Therefore, the probability of finding the system $\mathcal{S}$ in a particular quantum state s with energy ϵ_s and number of particles n_s is again proportional to the number of quantum states $g_\mathcal{R}(U_0 - \epsilon_s, N_0 - n_s)$ the reservoir has available with energy $U_0 - \epsilon_s$ and number of particles $N_0 - n_s$. Or

$$\frac{Prob(1)}{Prob(2)} = \frac{g_\mathcal{R}(U_0 - \epsilon_1, N_0 - n_1)}{g_\mathcal{R}(U_0 - \epsilon_2, N_0 - n_2)} \qquad (3.34)$$

The multiplicity function is the exponent of the entropy. If we assume the the reservoir is very large, the presence of the system causes only small perturbations, and in first order we find

$$\frac{Prob(1)}{Prob(2)} = \exp(\frac{1}{k_B}(n_2 - n_1)\left(\frac{\partial S_\mathcal{R}}{\partial N}\right)_{U,V} + \frac{1}{k_B}(\epsilon_2 - \epsilon_1)\left(\frac{\partial S_\mathcal{R}}{\partial U}\right)_{V,N}) \quad (3.35)$$

The partial derivatives are $-\frac{\mu_{\mathcal{R}}}{T_{\mathcal{R}}}$ and $\frac{1}{T_{\mathcal{R}}}$ for the reservoir. Since we assume thermal and diffusive equilibrium, both μ and T are the same for system and reservoir, $T_{\mathcal{R}} = T_{\mathcal{S}} = T$ and $\mu_{\mathcal{R}} = \mu_{\mathcal{S}} = \mu$, and hence we the probability of finding the system with temperature T and chemical potential μ in a quantum state s is equal to

$$Prob(s) = \frac{1}{3} e^{\frac{1}{k_B T}(\mu n_s - \epsilon_s)} \tag{3.36}$$

The exponential factor is called the Gibbs factor in this case and is an obvious generalization of the Boltzmann factor. The normalization factor $\mathfrak{Z}$ is called the grand partition function or grand sum and is defined by

$$\mathfrak{Z}(T, V, \mu) = \sum_{states} e^{\frac{1}{k_B T}(\mu n_s - \epsilon_s)} \tag{3.37}$$

The probabilities are again a ratio of an exponential factor and a normalization sum. We saw before that the partition function contains all information needed to describe a system. The grand partition function also contains all information needed. Sometimes it is easier to work with the partition function, sometimes with the grand partition function.

Since both partition functions contain all information, it must be possible to construct one from the other. The grand partition function is related to the partition function by a sum over particle numbers:

$$\mathfrak{Z}(T, V, \mu) = \sum_{\hat{N}} e^{\frac{\mu \hat{N}}{k_B T}} \mathcal{Z}(T, V, \hat{N}) \tag{3.38}$$

This follows when we split the summation over all states in the grand sum into two parts, a summation over all possible number of particles and for each term in that summation a sum over all states with the corresponding number of particles. The latter gives the partition function. In a later chapter we will give a more formal justification why this is allowed even in systems where the number of particles is not conserved. It is related to the invariance of the trace operator in quantum mechanics. The transformation above is a standard Laplace transform, and can be inverted. This requires, however, a analytic continuation of the grand partition function in the complex μ plane. Mathematically, that is not a problem, but what does that mean in terms of physics?

How do we extract information from the grand partition function? One case is easy. The average number of particles in the system, $<n_s> = \sum n_s Prob(s)$:

$$\left(\frac{\partial \mathfrak{Z}}{\partial \mu}\right)_{T,V} = \frac{1}{k_B T} \sum_s n_s e^{\frac{1}{k_B T}(\mu n_s - \epsilon_s)} = \frac{1}{k_B T} \mathfrak{Z} <n_s> \tag{3.39}$$

Therefore, the average number of particles (now described by the function $N(T, V, \mu)$), is directly related to the grand partition function by

$$N(T, V, \mu) = k_B T \left(\frac{\partial \log(\mathfrak{Z})}{\partial \mu}\right)_{T,V} \tag{3.40}$$

Since the temperature in many formulas appears as $\frac{1}{k_B T}$, this combination is often abbreviated by defining $\beta = \frac{1}{k_B T}$. Using β simplifies many superscripts to simple products in stead of ratios. Of course, this β should not be confused with the critical exponent describing the disappearance of the order parameter near the critical temperature! Also, the chemical potential often occurs in the combination $e^{\frac{\mu}{k_B T}}$. This exponential is called the <u>absolute activity</u> (often denoted by $\lambda = e^{\frac{\mu}{k_B T}}$) or the <u>fugacity</u> (often denoted by $z = e^{\frac{\mu}{k_B T}}$). The notation in the latter case is not the <u>same in</u> all books, unlike the conventional abbreviation β, which is used almost uniformly. In terms of β, one finds immediately that

$$\left(\frac{\partial \mathfrak{Z}}{\partial \beta}\right)_{\mu,V} = \sum_s (\mu n_s - \epsilon_s)e^{\beta(\mu n_s - \epsilon_s)} = \mathfrak{Z}(\mu N - U) \tag{3.41}$$

since $U = < \epsilon_s >$. This gives

$$U = \mu N - \left(\frac{\partial \log(\mathfrak{Z})}{\partial \beta}\right)_{\mu,V} = \left(\frac{\mu}{\beta}\frac{\partial}{\partial \mu} - \frac{\partial}{\partial \beta}\right) \log(\mathfrak{Z}) \tag{3.42}$$

The recipe for calculations at specified values of μ, T, and V is now clear. Calculate the grand partition function $\mathfrak{Z}(\mu, T, V)$ and then the number of particles, internal energy, etc. The construct the grand potential. There is a better and direct way, however, as one might expect.

Since the grand potential is the correct measure of energy available to do work at constant V, μ and T, it should be related to the grand partition function. The correct formula can be guessed directly when we remember that:

$$\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = -N \tag{3.43}$$

After comparing this partial derivative with the expression for N in 3.40 we conclude that:

$$\Omega = -k_B T \log(\mathfrak{Z}) + f(T,V) \tag{3.44}$$

where the function f(T,V) is still unknown. We will show that it is zero. This last formula can be inverted to give an expression for $\log(\mathfrak{Z})$. Inserting this expression in the formula for U leads to

$$U = \left(\frac{\mu}{\beta}\frac{\partial}{\partial \mu} - \frac{\partial}{\partial \beta}\right) \log(\mathfrak{Z}) = \left(\frac{\mu}{\beta}\frac{\partial}{\partial \mu} - \frac{\partial}{\partial \beta}\right)(\beta(f - \Omega)) \tag{3.45}$$

$$U = \Omega - f + \left(\mu\frac{\partial}{\partial \mu} - \beta\frac{\partial}{\partial \beta}\right)(f - \Omega) \tag{3.46}$$

$$U = \Omega - f - \mu\left(\frac{\partial \Omega}{\partial \mu}\right)_{\beta,V} - \beta\left(\frac{\partial f}{\partial \beta}\right)_{\mu,V} + \beta\left(\frac{\partial \Omega}{\partial \beta}\right)_{\mu,V} \tag{3.47}$$

Next we use the following thermodynamical relations, $\left(\frac{\partial\Omega}{\partial\mu}\right)_{V,T} = -N$ and $\left(\frac{\partial\Omega}{\partial\beta}\right)_{V,\mu} = -T\beta^{-1}\left(\frac{\partial\Omega}{\partial T}\right)_{V,\mu} = TS\beta^{-1}$ to get

$$U = \Omega - f + \mu N - \beta\left(\frac{\partial f}{\partial\beta}\right)_{\mu,V} + TS \tag{3.48}$$

and with $U = \Omega + TS + \mu N$ this gives

$$0 = f + \beta\left(\frac{\partial f}{\partial\beta}\right)_{\mu,V} = \left(\frac{\partial\beta f}{\partial\beta}\right)_{\mu,V} \tag{3.49}$$

or

$$f(T,V) = k_B T g(V) \tag{3.50}$$

The partial derivative of the grand potential with respect to T yields the entropy:

$$S = -\left(\frac{\partial\Omega}{\partial T}\right)_{\mu,V} = \left(\frac{\partial k_B T\log(\mathfrak{Z})}{\partial T}\right)_{V,\mu} + k_B g(V) \tag{3.51}$$

In the limit $T \to 0$ the grand partition function approaches $\log(g_0)$ where g_0 is the degeneracy of the ground state. Therefore, the first term on the right hand side approaches $k_B\log(g_0)$, which is the entropy at zero temperature. Therefore,

$$\lim_{T\to 0} g(V) = 0 \tag{3.52}$$

but, since $g(V)$ does not depend on T, this shows that $g(V) = 0$ and hence

$$\Omega = -k_B T\log(\mathfrak{Z}) \tag{3.53}$$

or

$$\mathfrak{Z} = e^{-\beta\Omega} \tag{3.54}$$

This form is very similar to the one we found for the relation between the canonical sum and the Helmholtz free energy.

Working with the grand canonical ensemble is necessary in every situation where the number of particles is allowed to vary. Besides for systems in diffusive contact, it is also the ensemble of choice in quantum field theory, where particles are destroyed and created via operators a and $a^\dagger$ and the quantum states are elements of a Fock space with variable numbers of particles.

Once we have found the grand potential, we are also able to evaluate the entropy. A simple formula for the entropy is again obtained in terms of the probabilities of the quantum states. We have

$$S = \frac{U - \Omega - \mu N}{T} = \frac{1}{T}\sum_s \epsilon_s Prob(s) + k_B\log(\mathfrak{Z}) - \frac{\mu}{T}\sum_s n_s Prob(s) \tag{3.55}$$

$$S = k_B \sum_s \frac{\epsilon_s - \mu n_s}{k_B T} Prob(s) + k_B \log(\mathfrak{Z}) \sum_s Prob(s) \qquad (3.56)$$

where we used the fact that the sum of all probabilities is one. This then leads to:

$$S = -k_B \sum_s \left(\frac{\mu n_s - \epsilon_s}{k_B T} - k_B \log(\mathfrak{Z}) \right) Prob(s) \qquad (3.57)$$

or

$$S = -k_B \sum_s Prob(s) \log(Prob(s)) \qquad (3.58)$$

just as we had in the canonical ensemble. Note, however, that the similarity is deceptive. The summation is now over a much larger set of states, including all possible values of N. So how can the result be the same? The answer is found easily when we introduce the independent variables. The result of 3.58 is $S(T,V,\mu)$, while for the canonical case we get $S(T,V,N)$. Like in thermo-dynamics, we are doing calculus of many variables, and keeping track of the independent variables is again important!

The fluctuations in the number of particles follow from a calculation of the average square of the number of particles, just as in the case of the energy in the canonical ensemble. We start with

$$< N^2 > = \sum_s n_s^2 Prob(s) \qquad (3.59)$$

which is equal to $\beta^{-2} \frac{1}{\mathfrak{Z}} \left(\frac{\partial^2 \mathfrak{Z}}{\partial \mu^2} \right)$. Using with $< N > = \sum_s n_s Prob(s) = \beta^{-1} \frac{1}{\mathfrak{Z}} \left(\frac{\partial \mathfrak{Z}}{\partial \mu} \right)$ we find

$$\left(\frac{\partial N}{\partial \mu} \right)_{T,V} = \beta^{-1} \frac{1}{\mathfrak{Z}} \left(\frac{\partial^2 \mathfrak{Z}}{\partial \mu^2} \right) - \beta^{-1} \frac{1}{\mathfrak{Z}^2} \left(\left(\frac{\partial \mathfrak{Z}}{\partial \mu} \right) \right)^2 \qquad (3.60)$$

or

$$k_B T \left(\frac{\partial N}{\partial \mu} \right)_{T,V} = < n_s^2 > - < n_s >^2 = (\Delta N)^2 \qquad (3.61)$$

All the partial derivatives are at constant T and V. This formula shows that N is a monotonically increasing function of μ, and hence μ a monotonically increasing function of N, at constant T and V. This also followed from the stability criteria in thermodynamics. The quantity $\left(\frac{\partial N}{\partial \mu} \right)_{T,V}$ is a response function, just like C_V. The formula above again shows that this type of response function is proportional to the square root of the fluctuations in the related state variable. If the fluctuations in the number of particles in a given system are large, we need only a small change in chemical potential to modify the average number of particles. In other words, if intrinsic variations in N at a given value of μ

occur easily, it is also easy to induce a change in the average by external means, e.g. changing μ. Finally, the fluctuations ΔN are proportional to $\sqrt{N}$. In the thermodynamic limit the fluctuations in the number of particles are infinitely large, but the fluctuations per particle become infinitesimally small.

3.5 Overview of calculation methods.

The development of the ideas used in statistical mechanics starts with the discussion of a closed system, with extensive parameters V, U, and N (and possibly more), which can be set at specified values. These extensive parameters are the independent parameters describing the system. The link between the microscopic world and thermodynamics is obtained by calculating the entropy. There are several ways of doing so (more about that in a later chapter), but here we start with the multiplicity function, $g(U, V, N)$. This function gives the number of microscopic states available to the system at given U, V, and N. Next, we define the entropy analogue by $S(U, V, N) = k_B \log(g(U, V, N))$. We then take the thermodynamic limit, making the system infinite, and all variables become continuous. The thermodynamic limit is always needed, no matter which formulation of statistical mechanics one is using. Sometimes people try to hide it, and calculate statistical mechanical functions for small systems. But those results are suspect, because only in the thermodynamic limit can we prove that the entropy analogue defined above is equivalent to the entropy defined in thermodynamics. Once we have shown that the entropy analogue is indeed equivalent to the thermodynamical entropy, we can use all results from thermodynamics, and we are in business.

The path followed above is a good path for theoretical development, but often not a good way for practical calculations. Using the temperature T as an independent variable makes life much easier (most of the time). In our model system we can now choose the following combinations of independent variables: (T, V, N), (T, V, μ) , and (T, p, N). Note that we need at least one extensive variable. Other combinations are possible to, but do not lead to useful free energies (see thermodynamics). Next, we construct the Hilbert space of all possible quantum states which are consistent with the values of the extensive parameters in the combination. This space is, for the three cases above, $\mathcal{S}(V, N)$, $\mathcal{S}(V)$, and $\mathcal{S}(N)$, respectively. Next, we calculate the partition sum for these situations:

$$\mathcal{Z}(T, V, N) = \sum_{s \in \mathcal{S}(V,N)} e^{-\beta \epsilon_s} \tag{3.62}$$

$$\mathfrak{Z}(T, V, \mu) = \sum_{s \in \mathcal{S}(V)} e^{-\beta(\epsilon_s - \mu n_s)} \tag{3.63}$$

$$\zeta(T, p, N) = \sum_{s \in \mathcal{S}(N)} e^{-\beta(\epsilon_s + p v_s)} \tag{3.64}$$

where in the last case (which we did not discuss here) the volume of the state is v_s. This last formulation is only included for completeness, and to show how other transformations are made. The argument of the exponent in the sum is easy to find if we think about the free energies that are used in the situations above. In the first case it is F, in the second case $\Omega = F - \mu N$, and in the last case $G = F + pV$. The relations with the free energy follow from:

$$F(T, V, N) = -k_B T \log(\mathcal{Z}(T, V, N)) \tag{3.65}$$

$$\Omega(T, V, \mu) = -k_B T \log(\mathfrak{Z}(T, V, \mu)) \tag{3.66}$$

$$G(T, p, N) = -k_B T \log(\zeta(T, p, N)) \tag{3.67}$$

or in exponential form, which resembles the sum in the partition functions,

$$\mathcal{Z}(T, V, N) = e^{-\beta F(T, V, N)} \tag{3.68}$$

$$\mathfrak{Z}(T, V, \mu) = e^{-\beta \Omega(T, V, \mu)} \tag{3.69}$$

$$\zeta(T, p, N) = e^{-\beta F(T, p, N)} \tag{3.70}$$

Once we have the free energies, we can revert to thermodynamics. We can also do more calculations in statistical mechanics, using the fact that the probability of finding a state is always the exponential factor in the sum divided by the sum. This is a generalization of the Boltzmann factor. We can then calculate fluctuations, and show that certain response functions are related to fluctuations. For example, we would find:

$$(\Delta V)^2 = k_B T \left(\frac{\partial V}{\partial p} \right)_{T, N} \tag{3.71}$$

The example with changing V to p is a bit contrived, but shows the general principle. A practical example involves magnetic systems, where we can either use the total magnetization M of the applied field H as an independent variable. There we would have:

$$\mathcal{Z}(T, V, N, M) = \sum_{s \in \mathcal{S}(V, N, M)} e^{-\beta \epsilon_s} = e^{-\beta F(T, V, N, M)} \tag{3.72}$$

or

$$\mathcal{Z}(T, V, N, H) = \sum_{s \in \mathcal{S}(V, N)} e^{-\beta(\epsilon_s - H m_s)} = e^{-\beta \mathcal{G}(T, V, N, H)} \tag{3.73}$$

where $\mathcal{G}$ is a generalized form of a Gibbs energy. We have used the same symbol for the partition function (sorry, I could not find a new and better one). This is also very commonly done in the literature. Fluctuations follow from

$$(\Delta M)^2 = k_B T \left(\frac{\partial M}{\partial H} \right)_{T,V,N} \tag{3.74}$$

Finally, the different partition sums are related by Laplace transformations. For example, we have

$$\mathfrak{Z}(T, V, \mu) = \sum_N e^{\beta \mu N} \mathcal{Z}(T, V, N) \tag{3.75}$$

We can also do the back transformation, but that involves an integration in the complex μ plane. Although we can relate complex temperature T to time, I have no idea what the physical meaning of complex μ would be.

3.6 A simple example.

A very simple example of the use of the grand canonical ensemble is a study of a one-dimensional photon gas where all the photons have frequency ω. The energy ϵ_n of the quantum state n is $n\hbar\omega$, while the number of particles in that state is $N_n = n$. The values for n are 0,1,2,.... This example is oversimplified, since one quantum number controls two quantum states. Also, there are only two sets of thermodynamic variables. We should therefore only use it as an illustration of how we sometimes can do calculations, and not draw strong conclusions from the final results. The grand partition function follows from

$$\mathfrak{Z}(T, \mu) = \sum_{n=0}^{\infty} e^{\beta(\mu - \hbar\omega)n} \tag{3.76}$$

This series can only be summed for $\mu < \hbar\omega$, leading to

$$\mathfrak{Z}(T, \mu) = \frac{1}{1 - e^{\beta(\mu - \hbar\omega)}} \tag{3.77}$$

All thermodynamic quantities can now be evaluated from the grand potential and we find

$$\Omega(T, \mu) = k_B T \log(1 - e^{\beta(\mu - \hbar\omega)}) \tag{3.78}$$

$$S(T, \mu) = -k_B \log(1 - e^{\beta(\mu - \hbar\omega)}) + \frac{1}{T} \frac{\hbar\omega - \mu}{1 - e^{\beta(\mu - \hbar\omega)}} \tag{3.79}$$

$$N(T, \mu) = frac1{e^{\beta(\hbar\omega - \mu)}} - 1 \tag{3.80}$$

$$U(T, \mu) = \Omega + TS + \mu N = N\hbar\omega \tag{3.81}$$

We have followed all steps in the statistical mechanical prescription for the grand canonical ensemble, and we obtained reasonable answers. Nevertheless, thermodynamics tells us that for a system with only T, S, μ, and N as state

variables $U = TS + \mu N$ (Gibbs-Duhem), and hence $\Omega = 0$. Also, $T = \left(\frac{\partial U}{\partial S}\right)_N$ which is zero according to the formula $U = N\hbar\omega$. What went wrong?

There are two problems. First of all, in a T,S,μ,N system we cannot choose T and μ as independent variables in thermodynamics. We have to introduce an extensive parameter. Second, we have to take the thermodynamic limit. The formula for N can be inverted to yield a formula for μ:

$$\mu = \hbar\omega - k_B T \log(1 + \frac{1}{N}) \tag{3.82}$$

which tells us that for a given value of N the chemical potential and the temperature are related. In the thermodynamic limit $N \to \infty$ we find $\mu = \hbar\omega$. For this value of the chemical potential the series for the grand partition function does not converge anymore, hence we have to perform all calculations for finite values of N and take the limit only in the end. The grand potential per particle is

$$\frac{\Omega(T, N)}{N} = -k_B T \frac{log(N+1)}{N} \tag{3.83}$$

and this quantity is zero in the thermodynamic limit. Hence we retrieved the result anticipated from thermodynamics. We also find

$$S(T, N) = k_B \log(N+1) + N k_B \log(1 + \frac{1}{N}) \tag{3.84}$$

which shows that in the thermodynamic limit $\frac{S}{N} = 0$. In other words, temperature is not well-defined for this system. This problem is independent of the thermodynamic limit, but is inherent to our simple model where we use only one quantum number for both energy and number of particles. For a given number of particles the energy is fixed! Hence we really only have two independent variables in our system and thermodynamically our system is completely characterized once we know the value of N. In realistic physical models the parameters N and U have to be independent. An example is found in the next section.

3.7   Ideal gas in first approximation.

The ideal gas is the first realistic and complete model we encounter. There are three sets of thermodynamic variables. We have volume and pressure, entropy and temperature, and chemical potential and number of particles. Therefore, we can choose T and μ independently, after which the value of p is fixed.

In the original experiments with a gas as a function of pressure, volume, and temperature it was noted that the relations between these thermodynamic quantities were very similar for many different gases. Therefore, one postulated the so-called ideal gas laws and tried to explain why a real gas shows small (or large) differences from these ideal gas laws. Of course, we now know that the most important parameter for ideal gas behavior is the density $n = \frac{N}{V}$, and that

the ideal gas laws are simply the limits of the correct functional dependencies to zero density. It is just fortunate that almost every gas is close to ideal, otherwise the development of thermal physics would have been completely different!

The quantum mechanical model of an ideal gas is simple. Take a box of dimensions $L \times L \times L$, and put one particle in that box. If L is large, this is certainly a low density approximation. The potential for this particle is defined as 0 inside the box and ∞ outside. The energy levels of this particle are given by

$$\epsilon(n_x, n_y, n_z) = \frac{\hbar^2}{2M} \left(\frac{\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2) \tag{3.85}$$

with n_x ,n_y, and n_z taking all values 1,2,3,..... The partition function is easy to calculate in this case

$$\mathcal{Z}_1 = \sum_{n_x n_y n_z} e^{-\beta\epsilon(n_x n_y n_z)} =$$

$$\sum_{n_x} e^{-\frac{\hbar^2}{2Mk_BT}\left(\frac{\pi}{L}\right)^2 n_x^2} \sum_{n_y} e^{-\frac{\hbar^2}{2Mk_BT}\left(\frac{\pi}{L}\right)^2 n_y^2} \sum_{n_z} e^{-\frac{\hbar^2}{2Mk_BT}\left(\frac{\pi}{L}\right)^2 n_z^2} \tag{3.86}$$

Define $\alpha^2 = \frac{\hbar^2\pi^2}{(2ML^2k_BT)}$ and realize that the three summations are independent. This gives

$$\mathcal{Z}_1 = \left(\sum_{n=1}^{\infty} e^{-\alpha^2 n^2}\right)^3 \tag{3.87}$$

We cannot evaluate this sum, but if the values of α are very small, we can use a standard trick to approximate the sum. Hence, what are the values of α? If we take $\hbar \approx 10^{-34}Js$, $M \approx 5 \times 10^{-26}kg$ (for A around 50), and $L = 1m$, we find that $\alpha^2 \approx \frac{10^{-38}J}{k_BT}$, and $\alpha \ll 1$ translates into $k_BT \gg 10^{-38}J$. With $k_B \approx 10^{-23}JK^{-1}$ this translates into $T \gg 10^{-15}K$, which from an experimental point of view means always. Keep in mind, though, that one particle per m^3 is an impossibly low density!

In order to show how we calculate the sum, we define $x_n = \alpha n$ and $\Delta x = \alpha$ and we obtain

$$\sum_{n=1}^{\infty} e^{-\alpha^2 n^2} = \frac{1}{\alpha}\sum_{n=1}^{\infty} e^{-x_n^2}\Delta x \tag{3.88}$$

In the limit that Δx is very small, numerical analysis shows that

$$\sum_{n=1}^{\infty} e^{-x_n^2}\Delta x \approx \int_0^{\infty} e^{-x^2} dx - \frac{1}{2}\Delta x + \mathcal{O}\left(e^{-\frac{1}{\Delta x}}\right) \tag{3.89}$$

where the integral is equal to $\frac{1}{2}\sqrt{\pi}$. The term linear in Δx is due to the fact that in a trapezoidal sum rule we need to include the first (n=0) term with a factor

$\frac{1}{2}$. The remainder of the error converges very rapidly in this case, because of the special nature of the integrant. Therefore, for $T \gg 10^{-15} K$ we can replace the sum by an integral and we find:

$$\mathcal{Z}_1 \approx \left(\frac{\sqrt{\pi}}{2\alpha} \right)^3 \tag{3.90}$$

Note that in the case $\alpha \ll 1$ we have $\mathcal{Z}_1 \gg 1$. Since this sum appears in the denominator of the probability for a state, the probability of being in a given state is very small.

The volume V of the box is L^3 and the quantum concentration $n_Q(T)$ is defined by

$$n_Q(T) = \left(\frac{Mk_BT}{2\pi\hbar^2} \right)^{\frac{3}{2}} \tag{3.91}$$

Combining these formulas leads to

$$\mathcal{Z}_1(T,V) = Vn_Q(T) \tag{3.92}$$

Since we have one particle only, the density is $n = \frac{1}{V}$ and the condition $\mathcal{Z}_1 \gg 1$ is equivalent to $n \ll n_Q(T)$. For a normal gas (which is very dilute), this is always obeyed, and we can replace sums by integrals. This is called the classical limit. Discontinuities due to the discrete energy spectrum in quantum mechanics are not important, and everything is smooth. On the other hand, if $n \gtrsim n_Q(T)$ we need to do the sums. This is the case when the density is large or the temperature is low.

We note that $\mathcal{Z}_1 \propto T^{\frac{3}{2}}$. The factor three is due to the presence of three summations, one for each spatial coordinate. If we would be working in D dimensions, this would be $\mathcal{Z}_1 \propto T^{\frac{D}{2}}$. Therefore, the internal energy is

$$U(T,V,N=1) = k_BT^2 \left(\frac{\partial \mathcal{Z}_1}{\partial T} \right)_V = \frac{D}{2} k_BT \tag{3.93}$$

as expected for one particle.

What do we do with N particles? We assume that the energy states of the particles are independent, and hence the energy of all particles together is the sum of all one particle energies. This would suggest that $\mathcal{Z}(T,V,N) = \mathcal{Z}_1^N(T,V)$. If the particles were distinguishable, this would be correct. But we know from quantum mechanics that the particles are not distinguishable. Hence we do not know which particle is in which state! Therefore, we propose that the partition function for N particles in this box is given by

$$\mathcal{Z}(T,V,N) = \frac{1}{N!} (\mathcal{Z}_1)^N \tag{3.94}$$

The presence of a power of N can be easily argued for, as we did above, but the factor $N!$ is not clear. If we assume that the particles are indistinguishable, it makes sense to divide by $N!$. This factor $N!$ did lead to a lot of discussion, and

the more rigorous treatment which will follow will indeed justify it. Before this more rigorous theory had been developed, however, experiments had already shown that one needed this factor! Using this formula we obtain the results for an ideal gas:

$$\mathcal{Z}(T, V, N) = \frac{1}{N!} \left(V n_Q(T) \right)^N \tag{3.95}$$

$$U = k_B T^2 \left(\frac{\partial \log(\mathcal{Z})}{\partial T} \right)_{V,N} = \frac{3}{2} N k_B T \tag{3.96}$$

$$F = -k_B T \log(\mathcal{Z}) = k_B T \log(N!) - N k_B T \log(n_Q V) \tag{3.97}$$

Using Stirling's formula and only keeping terms in F which are at least proportional to N, we obtain

$$F = N k_B T \left(\log(\frac{n}{n_Q(T)}) - 1 \right) \tag{3.98}$$

$$p = - \left(\frac{\partial F}{\partial V} \right)_{T,N} = \frac{N k_B T}{V} \tag{3.99}$$

$$S = - \left(\frac{\partial F}{\partial T} \right)_{V,N} = N k_B \left(\log(\frac{n_Q(T)}{n} + \frac{5}{2} \right) \tag{3.100}$$

This last formula is called the Sackur-Tetrode relation.

If one ignores the N! in the definition of the partition function for N particles, several problems appear. First of all, the argument of the logarithm in all formulas would be $n_Q(T)V$, and quantities like F and S would not be extensive, i.e. they would not be proportional to N (but to $N \log N$)! Also, the second term in S would be $\frac{3}{2} N k_B$. S can be measured precise enough to show that we really need $\frac{5}{2}$ Hence these experiments showed the need of a factor N! (or at least for an extra $N \log N - N$).

The formula for the entropy is valid for an ideal gas, or for a gas in the classical limit. Nevertheless, it contains factors $\hbar$ (through $n_Q(T)$). That is really surprising. In many cases the classical limit can be obtained by assuming that $\hbar$ is small. Here that does not work. So why does $\hbar$ show up in a classical formula? Interesting indeed.

The chemical potential follows from the Helmholtz free energy:

$$\mu = k_B T \log(\frac{n}{n_Q(T)}) \tag{3.101}$$

A calculation of the grand partition function is also straightforward. We find

$$\mathfrak{Z}(T, \mu, V) = \sum_{\hat{N}} z^{\hat{N}} \frac{1}{\hat{N}!} (\mathcal{Z}_1)^{\hat{N}} = e^{z \mathcal{Z}_1} \tag{3.102}$$

and the grand potential is

$$\Omega(T, \mu, V) = -k_B T z \mathcal{Z}_1 = -k_B T e^{\beta\mu} V n_Q(T) \tag{3.103}$$

The average number of particles is $-\left(\frac{\partial\Omega}{\partial\mu}\right)_{T,V}$, leading to

$$N = e^{\beta\mu} V n_Q(T) \tag{3.104}$$

which is consistent with 3.101.

We can also check the Gibbs-Duhem relation. We find

$$TS - pV + \mu N = Nk_B T \left(\log(\frac{n_Q(T)}{n} + \frac{5}{2}\right) - \frac{Nk_B T}{V}V + k_B T \log(\frac{n}{n_Q(T)})N \tag{3.105}$$

$$TS - pV + \mu N = \frac{3}{2}Nk_B T \tag{3.106}$$

which is equal to U indeed.

In the previous formulas we have assumed that we can replace the summations by integrations. Is that still valid for N particles? Could anything go wrong? In replacing the sum by an integral we essentially wrote

$$\mathcal{Z}_1 = I(1 + \varepsilon) \tag{3.107}$$

where the constant I represents the result obtained using the integral and ε is the relative error. This error is proportional to α, which is proportional to $(Vn_Q(T))^{-\frac{1}{3}}$. We can replace the summation by an integral is $Vn_Q(T) \gg 1$. For N particles we need to calculate

$$\mathcal{Z}_1^N = I^N(1 + \varepsilon)^N \approx I^N(1 + N\varepsilon) \tag{3.108}$$

as long as the error is small, or as long as $N\epsilon \ll 1$. This implies

$$N^3 \ll Vn_Q(T) \tag{3.109}$$

or

$$n \ll n_Q(T)N^{-2} \tag{3.110}$$

and in the thermodynamic limit this is never obeyed!! That is strange. The formulas that resulted from our partition function are clearly correct, but the error in the partition function is not small! What went wrong? The answer is that we are not interested in the error in the partition function itself, but in the error in the Helmholtz free energy, because that is the relevant quantity for all physics. In the Helmholtz free energy we have:

$$F = F_{ideal} - Nk_B T \log(1 + \varepsilon) \tag{3.111}$$

Our new condition is therefore that the second term is small compared to the first term, or

$$|\log(1 + \varepsilon)| \ll |\log(\frac{n}{n_Q(T)}) - 1| \tag{3.112}$$

We now take the same requirement as in the one particle case, and assume that

$$n \ll n_Q(T) \tag{3.113}$$

which also implies $V n_Q(T) \gg N$ or $\varepsilon \ll 1$. This means that the left hand side of 3.112 is much smaller than one. Since the right hand side is much larger than one (note the absolute sign, the logarithm itself is large and negative!), the Helmholtz free energy is indeed approximated by the formula we gave.

The inequality $n \ll n_Q(T)$ inspired the name quantum concentration. If the concentration is comparable to or larger than the quantum concentration, we have to use the series. Also, at that point quantum effects like the Pauli principle will start to play a role. If the condition above is fulfilled, we are in the classical limit

Hence everything seems OK. But there might be one last objection. We need to take derivatives of the free energy, and is the derivative of the error term small? The two terms in the free energy are:

$$F_{ideal} = -N k_B T \left(\log(\frac{n}{n_Q(T)}) - 1 \right) \tag{3.114}$$

and

$$F_{trunc} = -N k_B T \varepsilon = -N k_B T f(V, T)(V n_Q(T))^{-\frac{1}{3}} \tag{3.115}$$

where $f(V, T)$ is some function of order one. As an example, we take the derivative with respect to V, in order to get the pressure. The ideal gas result is $\frac{N k_B T}{V}$, and the derivative of the truncation error in the free energy has two terms. One part comes from the last factor, and gives

$$\frac{1}{3} \frac{N k_B T}{V} f(V, T)(V n_Q(T))^{-\frac{1}{3}} \tag{3.116}$$

which is again small if $V n_Q(T) \gg N$. The second part gives

$$-\frac{N k_B T}{V} \left[V \left(\frac{\partial f}{\partial V} \right)_T (V, T) \right] (V n_Q(T))^{-\frac{1}{3}} \tag{3.117}$$

and could be large if $f(V, T)$ is an oscillating function. This is fortunately not the case. Formula 3.89 tells us that $f(V, T)$ is equal to $-\frac{1}{2}$ plus terms that are exponentially small. The first part gives zero derivatives, and the second part gives derivatives which are exponentially small.

The error analysis above is typical for fundamental questions in statistical mechanics. In many cases it is ignored, and often without consequences. In

the next chapter we will see that Bose-Einstein condensation is an example of a natural phenomenon that would not be described at all if we ignored such an error analysis. In fact, in the description of many phase transitions one needs to be careful!

3.8   Problems for chapter 3

PROBLEM 1.

The Helmholtz free energy of a system at an appropriate temperature is given by $F(V, N) = N \log(\frac{N}{V}) - N$.
1. Calculate the pressure and chemical potential.
M such systems, all with volume V, are in diffusive contact and in equilibrium. In addition, there is a potential energy per particle Φ_i in system i. The total number of particles in the M combined systems is N.
2. Find the number of particles in each of the subsystems.

PROBLEM 2.

The quantum states of a given system have a variable volume. The energy of state n is ϵ_n, while the volume of state n is ν_n. This system is at a given pressure p.
a) Show that the probability of finding the system in state n is proportional to $e^{-\beta(\epsilon_n + p\nu_n)}$.
b). In this case we define a partition function $\mathbf{Z}(T, p, N) = \sum_n e^{-\beta(\epsilon_n + p\nu_n)}$. Show that the Gibbs energy $G = U - TS + pV$ is related to $\mathbf{Z}$ by $G = -k_B T \log(\mathbf{Z})$.
c) Find a formula relating the fluctuations in volume to the isothermal compressibility.

PROBLEM 3.

A system contains an ideal gas of atoms with spin $\frac{1}{2}$ in a magnetic field $B(\vec{r})$. The concentration of the spin up (down) particles is $n_\uparrow(\vec{r})$ ($n_\downarrow(\vec{r})$). The temperature is T.

(A) Evaluate the **total** chemical potential for the spin up and down particles.

(B) These two chemical potentials have to be the same and independent of $\vec{r}$. Explain why.

(C) Calculate the magnetic moment of this gas as a function of position.

(D) Show that the concentration of magnetic particles is high in regions with a large magnetic field.

PROBLEM 4.

The state of a many body system is characterized by two quantum numbers, n and m. The possible values of the quantum number n are $0, 1, 2, \cdots, \infty$, while the values of m are in the range $0, 1, \cdots, n$. The energy of the system in the state (n, m) is $n\hbar\omega$ and the number of particles is m. Evaluate the grand partition function for this system.

PROBLEM 5.

An ideal gas of atoms with mass m is contained in a cylinder that spins around with angular frequency ω. The system is in equilibrium. The distance to the axis of the cylinder is r. The radius of the cylinder is R. Calculate the density of the gas as a function of r.

PROBLEM 6.

Extremely relativistic particles obey the relation $E(\vec{k}) = \hbar c |\vec{k}|$. Assume we have a gas of these identical particles at low density, or $n \ll n_Q(T)$.

(A) Calculate the partition function $\mathcal{Z}_1(T, V)$ for N=1.

(B) Calculate $\mathcal{Z}(T, V, N)$.

(C) Calculate $p(T, V, N)$, $S(T, V, N)$, and $\mu(T, V, N)$.

PROBLEM 7.

The chemical potential of an ideal gas is given by 3.101. Suppose $n \ll n_Q(T)$. In this case we have $\mu < 0$. A bottle contains an ideal gas at such an extremely low density. We take one molecule out of the bottle. Since $\mu < 0$ the Helmholtz free energy will go up. The equilibrium of a system is reached when the Helmholtz free energy is minimal. This seems to favor molecules entering the bottle. Nevertheless, we all know that if we open the bottle in an environment where the density outside is lower than inside the bottle, molecules will flow out. What is wrong with the reasoning in this problem?

PROBLEM 8.

The only frequency of the radiation in a certain cavity is ω. The average number of photons in this cavity is equal to $\left[e^{\frac{\hbar\omega}{k_B T}} - 1 \right]^{-1}$ when the temperature is T. Calculate the chemical potential for this system.

PROBLEM 9.

A system can be in one of four states. These are

(1) No particles, zero energy

(2) One particle, energy ϵ_1

(3) One particle, energy ϵ_2

(4) Two particles, energy $\epsilon_1 + \epsilon_2 + I$

where $\epsilon_2 > \epsilon_1 > 0$ and $I < 0$. This sign of I is unusual. The temperature is T and the chemical potential is μ. This is a very simple model, and in order to make this realistic we should really consider many of these systems together. That is not asked in this question, though.

(A) Calculate the grand partition function

(B) Calculate $N(T, \mu)$

(C) Calculate $\left(\frac{\partial N}{\partial T}\right)_\mu$ and show that this can be negative. What does that mean?

Chapter 4

Statistics of independent particles.

4.1 Introduction.

If you think about the consequences of assuming a low density limit for a gas, it comes to mind immediately that the interactions between the particles must be very small. An ideal gas is defined as a gas of *non-interacting* particles in the low density limit. If we leave out the low density, but still assume that there are no interactions between the particles, quantum effects will start to play a role and in that case the gas is a Bose or Fermi gas. Although one might object that in this case it it not correct to ignore the inter-particle interactions, there are enough systems where these interactions are small and a non-interacting gas is a very good first approximation. This explains why it is useful at all to discuss Fermi or Bose gases.

The treatment of the ideal gas in the previous chapter is useful, but it is not a good starting point for more precise calculations. In this chapter we discuss an approach which is very versatile, and can be used to include the effects of particle statistics and interactions.

ORBITALS.

The energy levels used in the previous section were the energy levels of a single particle in a large box. What happens when we add a second particle to this box? If the particles are really independent, the presence of this second particle will not alter the energy levels of the first. Neither will the first particle influence the energy levels of the second, and hence both particle will have the same set

of energy levels available. These energy levels are called single particle states. In order to avoid using the word state in two different meanings, we will follow a standard approach and call these single particle states orbitals. One is used to the word orbital for systems of electrons, but here it is generalized to systems of arbitrary particles.

An example will help to describe this. Consider an atom, for which the states of the electrons are given by the quantum numbers n, l, m. We lump these quantum numbers together and use the symbol o, for orbital. The corresponding energies depend, however, on the states of all other electrons. In Helium, for example, the energy of an electron in the 1s state is different in the $1s^2 2s^0$ configuration and in the $1s^1 2s^1$ configuration. It is larger (more negative) in the second case, because the 2s electron does not screen the nucleus as well as the 1s electron. In this case there are many body effects, often described by the word correlation. The energy levels are of the form ϵ_o^s, where s stands for the state of the whole system.

ORBITAL ENERGY OF INDEPENDENT PARTICLES.

In an ideal gas we have a special situation. Here the particles do not interact, and the energy of a particle is independent of the states of the other particles. We say that there is no correlation, or no screening, and we can write

$$\epsilon_o^s = \epsilon_o \qquad (4.1)$$

for all orbitals o and in all states s of the total system. We say that the particles are independent.

This is a useful approximation to make, but is it realistic? In a Helium answer the answer is clearly no. But it turns out that for electrons in a metal the independent electron approximation is not a bad starting point. Well, more or less, as long as we replace the electrons by quasiparticles, that is electrons with a corresponding empty space around them, out of which other electrons have been repelled. For details, see books on solid state physics.

TOTAL ENERGY OF INDEPENDENT PARTICLES.

A general many body state of a system with N particles is completely known if we know how many particles are in each orbital. This defines a function of the single particle orbital index o for each many body state s, n_o^s. The energy of this many body state in terms of the single particle orbital energies ϵ_o is

$$E(state \quad s) = \sum_o n_o^s \epsilon_o \qquad (4.2)$$

and of course we have

$$N(state \quad s) = \sum_o n_o^s \qquad (4.3)$$

If we add one particle to the system, we assume that the orbital energies for this particle are the same as for the other particles. For most practical applications it is not necessary that this is true for an arbitrary number of particles in the system. Fluctuations in the number of particles are never very large and if the two formulas only hold for a certain range of values of N around the equilibrium value $< N >$, they are already useful and the subsequent discussion is valuable. This is the case in metals. The energy levels of the conduction electrons in a metal are in first approximation independent of the state of the electronic system for many changes in N which are of importance, but certainly not for changes starting with zero particles! Even for small changes in N one has to be aware of processes where the orbital levels do change because the correlation between the particles changes.

INCLUSION OF CORRELATION.

For atomic systems, as mentioned before, these correlation effects are always important and a typical formula for the energy of a rare earth atom with n_f electrons in the 4f shell is

$$E(n_f) = E(0) + n_f \epsilon_f + \frac{1}{2} n_f^2 U \qquad (4.4)$$

which introduces a <u>Coulomb interaction</u> U, which is of the same order of magnitude as ϵ_f (a few eV). This also shows why a starting point of independent particles can be very useful. The previous formula can be generalized to

$$E(state \quad s) = E(0) + \sum_o n_o^s \epsilon_o + \frac{1}{2} \sum_{o,o'} n_o^s n_{o'}^s U_{o,o'} \qquad (4.5)$$

If the Coulomb interactions are small, we can try to find solutions using perturbation theory, based on the starting point of the independent particle approximation. Of course, it is also possible to extend the equation above to include third and higher order terms!

INCLUSION OF QUANTUM STATISTICS.

How many particles can there be in one orbital? That depends on the nature of the particles! If the particles are **Fermions**, we can only have zero or one ($n_o = 0, 1$). If the particles are **Bosons**, any number of particles in a given orbital is allowed. Hence in the independent particle formalism the effects of quantum statistics are easy to include! This is another big advantage.

CALCULATION FOR INDEPENDENT SUBSYSTEMS.

Suppose a gas of non-interacting particles is in thermal and diffusive contact with a large reservoir. Hence both the temperature T and the chemical poten-

tial μ are specified. The volume V of the gas determines the orbital energies $\epsilon_o(V)$. We will now divide this gas in a large number of subsystems, each of which contains exactly one orbital. If all these subsystems are in thermal and diffusive equilibrium with the reservoir, they are in equilibrium with each other. Therefore, the properties of the total system can be obtained by adding the corresponding properties of all subsystems and the total system is really the sum of the subsystems! The properties of each subsystem, or each orbital o, follow from

$$3_o(T, \mu, V) = \sum_{n=0}^{?} e^{\frac{n(\mu - \epsilon_o)}{k_B T}} \tag{4.6}$$

because the energy of a state of the subsystem with n particles in the orbital o is simple $n\epsilon_o$. Quantum effects are still allowed to play a role here, and they determine the upper limit of the summation. The Pauli exclusion principle for fermions tells us that we can have at most one particle in each orbital and hence for fermions there are only two terms in the sum. For bosons, on the other hand, there is no such a restriction, and the upper limit is ∞.

FERMIONS.

We will first consider the case of Fermions. The grand partition function for an orbital is

$$3(T, \mu, V) = 1 + e^{\frac{\mu - \epsilon_o}{k_B T}} \tag{4.7}$$

Once we have the partition function for each orbital, we can calculate the average number of particles $< n_o >$ in that orbital. In terms of the probabilities P_n of finding n particles in this orbital we have $< n_o >= 0 P_0 + 1 P_1$. Hence we find

$$< n^o > = \frac{e^{\frac{\mu - \epsilon_o}{k_B T}}}{1 + e^{\frac{\mu - \epsilon_o}{k_B T}}} = \frac{1}{e^{\frac{\epsilon_o - \mu}{k_B T}} + 1} \tag{4.8}$$

Hence the average number of particles in a subsystem depends on T and μ and $\epsilon_o(V)$. The only quantity of the orbital we need to know to determine this average is the energy! No other aspects of the orbital play a role. In general, the function of the energy ϵ which yields the average number of particles in an orbital with energy ϵ is called a <u>distribution function</u> and we have in the case of Fermions:

$$f_{FD}(\epsilon; T, \mu) = \frac{1}{e^{\frac{\epsilon - \mu}{k_B T}} + 1} \tag{4.9}$$

This function is called the <u>Fermi-Dirac distribution function</u>. The shape of this function is well known, and is given in figure (4.1).
The Fermi-Dirac distribution function has the following general properties:
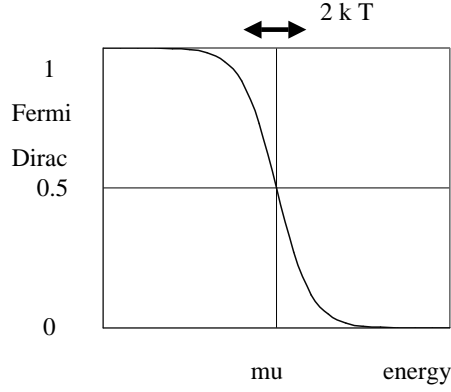
Figure 4.1: **Fermi Dirac distribution function.**

$$\lim_{\epsilon \to \infty} f_{FD}(\epsilon; T, \mu) = 0 \qquad (4.10)$$

$$\lim_{\epsilon \to -\infty} f_{FD}(\epsilon; T, \mu) = 1 \qquad (4.11)$$

$$f_{FD}(\epsilon = \mu; T, \mu) = \frac{1}{2} \qquad (4.12)$$

and the horizontal scale for this function is set by the product $k_B T$. In the limit $T \to 0$ the Fermi-Dirac function becomes a simple step function, with value 1 for $\epsilon < \mu$ and value 0 for $\epsilon > \mu$. Note that at $\epsilon = \mu$ the value remains $\frac{1}{2}$!

BOSONS.

In the case of Bosons the summation in the grand partition function goes from zero to infinity and we have

$$\mathfrak{Z}_o(T, \mu, V) = \sum_{n=0}^{\infty} e^{\frac{n(\mu - \epsilon_o)}{k_B T}} = \frac{1}{e^{\frac{\mu - \epsilon_o}{k_B T}} - 1} \qquad (4.13)$$

which is only valid when $\mu < \epsilon_o$, or else the series diverges. This is an important difference with the previous case, where μ could take all values. Here the possible values of the chemical potential are limited! The average number of particles follows from

$$< n_o >= k_B T \left(\frac{\partial \log(\mathfrak{Z}_o)}{\partial \mu} \right)_{T,V} \qquad (4.14)$$

which gives

$$< n_o > = \frac{1}{e^{\frac{\epsilon_o - \mu}{k_B T}} - 1} \qquad (4.15)$$

Again, this only depends on the values of the temperature and chemical potential (as given by the external reservoir) and the energy of the orbital. No other properties of the orbital play a role. The distribution function for Bosons is therefore

$$f_{BE}(\epsilon; T, \mu) = \frac{1}{e^{\frac{\epsilon - \mu}{k_B T}} - 1} \qquad (4.16)$$

This function is called the <u>Bose-Einstein distribution function</u>. The only difference with the Fermi-Dirac distribution function is the minus sign in the denominator. This is a small difference, but with large consequences. The Bose-Einstein distribution function has the following properties:

$$\lim_{\epsilon \to \infty} f_{FD}(\epsilon; T, \mu) = 0 \qquad (4.17)$$

$$\lim_{\epsilon \downarrow \mu} f_{FD}(\epsilon; T, \mu) = \infty \qquad (4.18)$$

and it is the last infinity that is the cause of all differences. The big difference between the two different distribution functions is the maximal value they can attain. For fermions the distribution function never exceeds the value of one, while for bosons there is no upper limit.

LIMIT OF SMALL OCCUPATION NUMBERS.

It is possible to choose T and μ in such a manner that the value of the distribution function is always much less than one for all orbital energies. In order to use this argument it is essential that the orbital energy has a lower limit (which is zero in our case of particles in a box). This requirement is quite natural. All realistic Hamiltonians are bounded from below, and the energy spectrum always has a minimum. The only exception is the Dirac Hamiltonian in first quantization, but in that case the existence of negative energy states is easily resolved in second quantization, where the negative energy states correspond to positrons with normal positive energies.

The requirement for small occupation numbers is that

$$\mu(T) \ll \epsilon_{min} - k_B T \qquad (4.19)$$

Note that this does not mean that in the limit $T \to \infty$ this inequality is not obeyed! The chemical potential also has a temperature dependence which needs to be taken into account. When the inequality above is valid, we have

$$e^{\frac{\epsilon - \mu}{k_B T}} \gg 1 \qquad (4.20)$$

and hence in the distribution functions the terms ± 1 can be ignored. In that case **both** distribution functions reduce to the same form

$$f_{MB}(\epsilon; T, \mu) = e^{\frac{\mu - \epsilon}{k_B T}} \tag{4.21}$$

which is the Maxwell-Boltzmann function. Quantum effects do not play a role anymore, since in this limit there is no distinction between Fermions and Bosons. This is therefore the classical limit and an ideal gas is also called a Boltzmann gas.

USE OF DISTRIBUTION FUNCTIONS.

Once we have chosen the appropriate distribution function for our system, it is easy to calculate thermodynamic variables. A general property Q of a independent gas as a function of T, μ, and V follows from the values of $Q_o(V)$ for all orbitals from

$$Q(T, V, \mu) = \sum_{orb} f(\epsilon_o(V); T, \mu) Q_o(V) \tag{4.22}$$

Two important examples are the number of particles and the internal energy

$$N(T, V, \mu) = \sum_{orb} f(\epsilon_o(V); T, \mu) \tag{4.23}$$

$$U(T, V, \mu) = \sum_{orb} f(\epsilon_o(V); T, \mu) \epsilon_o(V) \tag{4.24}$$

These two equation are sufficient to derive all of thermodynamics, although we later will see that there is a better way of doing that. From the first equation we can find by inversion $\mu(T, V, N)$ and then by integration $F(T, V, N)$. The second equation can be written as a function of T, V, N too when we substitute μ. But in many cases we do not need that, and the strength of this approach is really when we can use equation 4.22 directly. Note: in order for these sums to converge, we always need $\mu < \epsilon_o$. In other words we need $\mu < \epsilon_{min}$.

What have we gained by introducing this formulation of independent particles? The formulas we need still look very similar to those we used before. The important difference, however, is the number of terms involved. If the number of orbitals N_o is finite, the number of N particle states in the partition function is $N_o{}^N$ for bosons and $\binom{N_0}{N}$ for fermions. Since in practice $N_o \gg N$ the numbers are very similar for bosons and fermions. This indicates that the sum over many particle states in the general formalism is much more complicated than the sum over orbitals in the independent particle systems! In many cases the sum over orbitals reduces to an integral over a single variable, and these integrals are easy to analyze. Only for very small systems (e.g. 3 orbitals, 2 particles) is the complexity of the general approach similar to the independent particle formalism. But keep in mind that using distribution functions implies independent particles!

4.2   Boltzmann gas again.

An ideal gas is defined as a system of (1) free, (2) non-interacting particles in the (3) classical regime. Restriction (1) means that there are no external forces, beyond those needed to confine the particles in a box of volume V. Restriction (2) means that there are no internal forces. Restriction (3) means that we can ignore the quantum effects related to restrictions in occupation numbers. This means that we use the Maxwell-Boltzmann distribution function. Restrictions (1) and (3) are easy to relax. Most of solid state physics is based on such calculations, where electrons can only singly occupy quantum states in a periodic potential. Restriction (2) is much harder to relax, and many discussions in physics focus on these correlation effects.

IDEAL GAS AGAIN.

In this section we derive the results for the ideal gas again, but now from the independent particle point of view. This implicitly resolves the debate of the factor N! The energy levels of the particles in the box are the same as we have used before. The ideal gas follows the Boltzmann distribution function and hence we find

$$N = \sum_{orb} f_{MB}(\epsilon_o) = \sum_{orb} e^{\frac{\mu - \epsilon_o}{k_B T}} = e^{\frac{\mu}{k_B T}} \mathcal{Z}_1 \tag{4.25}$$

We again introduce the quantum concentration $n_Q(T)$ and solve for μ :

$$n_Q(T) = \left(\frac{M k_B T}{2 \pi \hbar^2} \right)^{\frac{3}{2}} \tag{4.26}$$

$$\mu = k_B T \log \left(\frac{n}{n_Q(T)} \right) \tag{4.27}$$

The requirement $\mu \ll -k_B T$ is indeed equivalent to $n \ll n_Q(T)$. The internal energy is also easy to obtain:

$$U = \sum_{orb} \epsilon_o e^{\frac{\mu - \epsilon_o}{k_B T}} = e^{\frac{\mu}{k_B T}} k_B T^2 \frac{\partial}{\partial T} \sum_{orb} e^{-\frac{\epsilon_o}{k_B T}} \tag{4.28}$$

$$U = \frac{N k_B T^2}{\mathcal{Z}_1} \left(\frac{\partial \mathcal{Z}_1}{\partial T} \right)_{N,V} = \frac{3}{2} N k_B T \tag{4.29}$$

as we had before. To get to the last step we used that $\mathcal{Z}_1$ is proportional to $T^{\frac{3}{2}}$.

ALWAYS THE SAME STRATEGY.

In many cases one uses the formula for N in the same way we did here. From it we solved for $\mu(T, N, V)$ and we also obtained U(N,T,V) accordingly.

Although the distribution functions depend on μ, it is often possible to use this approach to obtain functions of N,T,V. Also note that for the ideal gas $\left(\frac{\partial U}{\partial V}\right)_{T,N} = 0$ and hence that the pressure in an ideal gas is completely due to the derivative of the entropy, as we have stated before. This is due to the fact that

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = -\left(\frac{\partial U}{\partial V}\right)_{T,N} + T\left(\frac{\partial S}{\partial V}\right)_{T,N} \tag{4.30}$$

In a real gas system, of course, a change in volume will change the average distance between particles and hence the inter-particle interactions. This will give a small correction to the formula for U, and hence a small contribution to the pressure. In solids, on the other hand, the pressure is mainly due to the change in interaction energy.

SUMMATION OF THE CHEMICAL POTENTIAL.

We have transformed our independent variables to the set T, V, N and therefore need to find the corresponding free energy. The Helmholtz free energy is the important quantity and can be calculated as soon as $\mu(T, N, V)$ is found, because $\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V}$. The Helmholtz free energy for a system with zero particles is zero, and hence we have

$$F(N, T, V) = \int_0^N \mu dN' =$$

$$\int_0^N k_B T \log\left(\frac{N'}{V n_Q(T)}\right) dN' = N k_B T \left(\log(\frac{n}{n_Q(T)}) - 1\right) \tag{4.31}$$

where we used $\int_0^N \log(cN')dN' = N\log(c) + N\log(N) - N$.

Formally, there is an objection against using an integral over the number of particles when this number is small. Remember that the chemical potential is the free energy needed to add one particle to the system, and hence a correct expression for F is

$$F(N, T, V) = \sum_{\hat{N}=1}^{N} \mu(\hat{N}, T, V) =$$

$$\sum_{\hat{N}=1}^{N} k_B T \log\left(\frac{\hat{N}}{V n_Q(T)}\right) = k_B T \log(N!) - N\log(n_Q(T)V) \tag{4.32}$$

Here we used the fact that $\sum_{\hat{N}=1}^{N} \log(c\hat{N}) = N\log(C) + \log(N!)$. If N is very large we can again use Stirling's formula to show that this form really reduces

to the previous expression. We replace $\log(N!)$ by $N \log(N) - N$ only, since all other terms vanish in the thermodynamic limit.

Back to thermodynamics.

Once the Helmholtz free energy is known, the entropy and pressure can be calculated and we obtain again

$$p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} = \frac{Nk_BT}{V} \qquad (4.33)$$

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = Nk_B\left(\log(\frac{n_Q(T)}{n}) + \frac{5}{2}\right) \qquad (4.34)$$

Note that since $n \ll n_Q(T)$ the entropy is positive. This shows that we expect differences if the density becomes comparable to the quantum concentration. That is to be expected, because in that case there are states with occupation numbers that are not much smaller than one anymore. Also note that this classical formula for the entropy contains $\hbar$ via $n_Q(T)$. This is an example of a classical limit where one cannot take $\hbar = 0$ in all results! The entropy in our formalism is really defined quantum mechanically. It is possible to derive all of statistical mechanics through a classical approach (using phase space, etc), but in those cases $\hbar$ is also introduced as a factor normalizing the partition function! We will discuss this in a later chapter.

Check of Euler equation.

The Gibbs free energy is defined by $G = F + pV$ and is used for processes at constant pressure. For the ideal gas we find

$$G = \mu N \qquad (4.35)$$

This result is very important as we will see later on. We will show that it holds for all systems, not only for an ideal gas, and that it puts restrictions on the number of independent intensive variables. Of course, from thermodynamics we know that this has to be the case, it is a consequence of the Euler equation.

Heat capacities.

Important response functions for the ideal gas are the heat capacity at constant volume C_V and at constant pressure C_p. These functions measure the amount of heat ($T\Delta S$) one needs to add to a system to increase the temperature by an amount ΔT:

$$C_V = T\left(\frac{\partial S}{\partial T}\right)_{V,N} = \frac{3}{2}Nk_B \qquad (4.36)$$

$$C_p = T \left(\frac{\partial S}{\partial T} \right)_{p,N} = \frac{5}{2} N k_B \tag{4.37}$$

where we have used

$$\left(\frac{\partial S}{\partial T} \right)_{p,N} = \left(\frac{\partial S}{\partial T} \right)_{V,N} + \left(\frac{\partial S}{\partial V} \right)_{T,N} \left(\frac{\partial V}{\partial T} \right)_{p,N} \tag{4.38}$$

with $\left(\frac{\partial S}{\partial V} \right)_{T,N} = \frac{Nk_B}{V}$ and $\left(\frac{\partial V}{\partial T} \right)_{p,N} = \frac{Nk_B}{p} = \frac{V}{T}$.

It is obvious that we need $C_p \geqslant C_V$, since at constant pressure we have to do work against the outside world. This is an example of an inequality in thermal physics which relates two functions and holds in general for all systems. This relation was derived in thermodynamics based on equilibrium principles, and our current microscopic theory is in agreement with those general observations.

RATIO OF HEAT CAPACITIES.

The ratio of the heat capacities is often abbreviated by γ and is used in a number of formulas. For the ideal gas $\gamma = \frac{C_p}{C_V} = \frac{5}{3}$. From an experimental point of view this is a useful ratio, since it can be measured directly. From equation (4.38) we see that

$$\gamma = 1 + \left(\frac{\partial S}{\partial V} \right)_{T,N} \left(\frac{\partial V}{\partial T} \right)_{p,N} \left(\left(\frac{\partial S}{\partial T} \right)_{V,N} \right)^{-1} \tag{4.39}$$

$$\gamma = 1 + \left(\frac{\partial S}{\partial V} \right)_{T,N} \left(\frac{\partial V}{\partial T} \right)_{p,N} \left(\frac{\partial T}{\partial S} \right)_{V,N} \tag{4.40}$$

$$\gamma = 1 - \left(\frac{\partial T}{\partial V} \right)_{S,N} \left(\frac{\partial V}{\partial T} \right)_{p,N} \tag{4.41}$$

In any system where the ideal gas equation of state is $pV = Nk_BT$ obeyed we find

$$\left(\frac{\partial T}{\partial V} \right)_{S,N} = (1 - \gamma) \frac{T}{V} \tag{4.42}$$

or in regions where γ is constant we see that for adiabatic processes $T \propto V^{1-\gamma}$, which is equivalent to say that (again using the ideal gas law) for adiabatic processes pV^γ is constant.

4.3   Gas of poly-atomic molecules.

INTERNAL MOTION IS INDEPENDENT.

In our discussion of the ideal gas we have assumed that the orbitals are characterized by three quantum numbers n_x, n_y, and n_z. These three quantum numbers describe the motion of the center of mass of the particle, but did not include any internal degrees of freedom. Therefore, the previous formulas only are valid for a gas of mono-atomic molecules. A gas of poly-atomic molecules is easily treated if we assume that the energy associated with the internal degrees of freedom does not depend on (1) the presence of other particles and (2) the motion of the center of mass. Hence the rotation of a diatomic molecule is not hindered by neighboring molecules or changed by its own motion. In that case we can write, using *int* for the collective internal degrees of freedom:

$$\epsilon(n_x, n_y, n_z, int) = \frac{\hbar^2}{2M} \left(\frac{\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2) + \epsilon_{int} \qquad (4.43)$$

Then internal degrees of freedom represent the rotational quantum numbers (of which there are at most three) and the vibrational quantum numbers. If we have a molecule with N atoms, there are $3N$ internal degrees of freedom. Three are used for the center of mass, r for the rotations (r is 2 or 3) and hence $3(N-1) - r$ for the vibrational state.

CHANGES IN THE PARTITION FUNCTION.

In the classical regime $\mu \ll -k_B T$ and hence $\lambda = e^{\frac{\mu}{k_B T}} \ll 1$. In the partition function for a given orbital terms with λ^2 and higher powers can be neglected for bosons and in all cases we find

$$\mathfrak{Z}_o(T, \mu, V) = 1 + \lambda \sum_{int} e^{-\frac{\epsilon_o + \epsilon_{int}}{k_B T}} = 1 + \lambda \mathfrak{Z}_{int} e^{-\frac{\epsilon_o}{k_B T}} \qquad (4.44)$$

where the internal partition function is defined by

$$\mathfrak{Z}_{int}(T) = \sum_{int} e^{-\frac{\epsilon_{int}}{k_B T}} \qquad (4.45)$$

The average number of particles in orbital o, independent of its internal state, is therefore given by

$$< n_o >= \sum_{int} \frac{\lambda e^{-\frac{\epsilon_{int}}{k_B T}} e^{-\frac{\epsilon_o}{k_B T}}}{1 + \lambda \sum_{int} e^{-\frac{\epsilon_{int}}{k_B T}} e^{-\frac{\epsilon_o}{k_B T}}} \approx \lambda \mathfrak{Z}_{int} e^{-\frac{\epsilon_o}{k_B T}} \qquad (4.46)$$

where we used the fact that the denominator is approximately equal to one because λ is very small. Hence in the Boltzmann distribution function we have to replace λ by $\lambda \mathfrak{Z}_{int}(T)$ and we find

$$N = \lambda \mathfrak{Z}_{int}(T) n_Q(T) V \qquad (4.47)$$

$$\mu = k_B T \left(\log(\frac{n}{n_Q(T)}) - \log(\mathfrak{Z}_{int}) \right) \qquad (4.48)$$

$$F = Nk_BT \left(\log(\frac{n}{n_Q(T)}) - 1 \right) + F_{int}(T, N) \qquad (4.49)$$

$$F_{int}(T, N) = -Nk_BT \log \mathfrak{z}_{int}(T) \qquad (4.50)$$

EXPERIMENTAL ACCESS TO THE INTERNAL DEGREES OF FREEDOM.

For the entropy and internal energy we also find formulas that show that these quantities are the sum of the entropy/energy for the motion of the center of mass plus a term pertaining to the internal degrees of freedom. Since we have assumed that the energy of the internal motion does not depend on other particles, F_{int} does not depend on volume and the ideal gas law $pV = NkT$ remains valid! The heat capacities do change, however, since the internal energy is different. Therefore, the value of the ratio γ is different. As we will see later, the value of this ratio is a direct measure of the number of **active** degrees of freedom. Hence by measuring the pressure-volume relation in adiabatic expansion gives us direct information about the internal structure of the molecules!

DO WE HAVE THEM ALL?

The entropy is the sum of the entropy associated with the center of mass motion and the entropy of the internal state of the molecules. Hence the entropy is now larger than the entropy of mono-atomic molecules. We need more information to describe the state of large molecules, and hence we have more "un-knowledge", which translates into a larger entropy.

How do we know if we have all internal degrees of freedom accounted for? In principle, we do not know, because there could always be hidden degrees of freedom. For example, do we have to separate electrons and nucleï? Neutrons and protons? Quarks? Hence the real value of the entropy could even be larger as what we calculated here. We can also ask the question from an experimental point of view. In that case, the answer is simple, cool the system down to a low temperature. But what is low? Sometimes we see phase transitions at very low temperatures, and we are never sure if we went low enough. For example, the nuclear spins could order below a certain temperature, and the values of such temperatures are small indeed.

4.4 Degenerate gas.

The entropy of an ideal or Boltzmann gas is given by the Sackur-Tetrode formula. The temperature enters this equation through the quantum concentration. In the limit $T \to 0$, the entropy is approximately $\frac{3}{2}Nk_B \log(T)$ and approaches $-\infty$, which is obviously incorrect. The Sackur-Tetrode formula is not

valid anymore, because the condition $n \ll n_Q(T) \propto T^{1.5}$ will not be true for small temperatures. Every gas will show deviations from ideal gas behavior at low temperatures. A gas is called a quantum gas or degenerate gas when $n \approx n_Q(T)$. In a degenerate gas the differences between fermions and bosons become important.

At these higher densities the interactions between the particles also play a role, but the independent particle description is still a very good first approximation for most quantities. Inter-particle interaction can be included using perturbation theory. Of course, when $n \gg n_Q(T)$ that approach does not work anymore, and one really has to start with a model including inter-particle interactions.

An estimate of the temperature T_0 where these effects start to play a role comes from solving the equation $n = n_Q(T_0)$. We will calculate some typical examples of T_0. A gas of helium atoms has a density of about $2.5 \times 10^{25} m^{-3}$. With $M \approx 4 \times 10^{-27} kg$, $\hbar \approx 10^{-34} Js$, and $k_B \approx 10^{-23} JK^{-1}$ we get $T_0 \approx 1K$. Although helium is a liquid at these temperatures, special effects are seen in ^4He and also there are big differences between ^4He and ^3He at these temperatures.

Electrons in a metal have much higher densities that normal gases. A typical number is $n \approx 10^{29} m^{-3}$. With $M \approx 10^{-30} kg$ we find $T_0 \approx 10^5 K$. Such an electron gas is always degenerate, and the Pauli principle has to be taken into account. This is clearly seen in the physics of metals. The famous exchange energy for the electrons in a solid has to be taken into account. The conduction electrons in a semiconductor, on the other hand, have much lower densities. A typical range is $10^{23} \cdots 10^{26} m^{-3}$, and hence T_0 ranges from 10K to 1000K. At room temperature the electron gas in a semiconductor can be classical or degenerate, depending on the doping. This leads to interesting effects in these semiconductors.

4.5   Fermi gas.

Quantum gases come in two varieties, named after Fermi and Bose. There are some other cases of theoretical interest, but they are probably not important from a practical point of view. If the particles in a gas have integral spin we need Bose-Einstein statistics, if the spin is half-integral Fermi-Dirac statistics. We will first study a gas of identical, independent particles obeying Fermi-Dirac statistics. Kittel's book has a number of simple examples and is a good source of information for simple techniques. The same techniques are used in Sommerfeld's theory of the conduction electrons in a metal, and hence any text book on solid state physics will also be useful.

FERMI ENERGY.

The average number of particles in an orbital with energy ϵ is given by the Fermi-Dirac distribution function

$$f_{FD}(\epsilon) = \frac{1}{e^{\frac{\epsilon - \mu}{k_B T}} + 1} \tag{4.51}$$

The chemical potential μ depends on the temperature and the value of μ at $T = 0K$ is called the <u>Fermi energy</u> $\mu(T = 0) = \epsilon_F$. It is easy to show that in the limit $T \to 0$ the Fermi-Dirac function is 1 for $\epsilon < \mu$ and 0 for $\epsilon > \mu$. The chemical potential follows from

$$N = \sum_{orb} f_{FD}(\epsilon_o) = g \sum_{n_x n_y n_z} f_{FD}(\epsilon(n_x, n_y, n_z)) \tag{4.52}$$

where $g = 2S+1$ is the spin degeneracy; $g = 2$ for electrons. The spin factors out because the energy levels do not depend on spin, only on the spatial quantum numbers. Note that we use the fact that the particles are identical, we only specify the number of particles in each orbital, not which particle is in which orbital! The description is slightly more complicated when magnetic fields are included, adding a dependency of the energy on spin.

CONVERGENCE OF SERIES.

When we have an infinite sum, we always need to ask the question if this series converges. This means the following:

$$\sum_{n=1}^{\infty} x_n = S \tag{4.53}$$

if we can show that for $S_N = \sum_{n=1}^{N} x_n$ the following is true:

$$\lim_{N \to \infty} S_N = S \tag{4.54}$$

This means that for any value of $\epsilon > 0$ we can find a value N_ϵ such that

$$N > N_\epsilon \Rightarrow |S_N - S| < \epsilon \tag{4.55}$$

If the terms in the series are dependent on a variable like the temperature, we need to ask even more. Does the series converge uniformly? In general we want

$$\sum_{n=1}^{\infty} x_n(T) = S(T) \tag{4.56}$$

and this is true if for every $\epsilon > 0$ we can find a value $N_\epsilon(T)$ such that

$$N > N_\epsilon(T) \Rightarrow |S_N(T) - S(T)| < \epsilon \tag{4.57}$$

The problem is that the values of N_ϵ depend on T. What if, for example, $\lim_{T \to 0} N_\epsilon(T)$ does not exist? Than the sum of the series is not continuous at $T = 0$. That is bad. In order to be able to interchange limits and summations,

we need absolute convergence, that is we need to be able to find a minimum value of N for which the partial sum is a good approximation for all temperatures. Hence we need for every $\epsilon > 0$ we a value N_ϵ such that

$$N > N_\epsilon \Rightarrow |S_N(T) - S(T)| < \epsilon \tag{4.58}$$

In that case we can interchange summation and limits, and can integrate the sum of the series by integrating each term and sum the resulting integrals.

In our case, for large values of the quantum numbers the energy is very large and the distribution function can be approximated by

$$\lim_{\epsilon \to \infty} f_{FD}(\epsilon) \approx e^{\frac{\mu - \epsilon}{k_B T}} \tag{4.59}$$

Hence the terms in the series for N decay very rapidly. If we consider a temperature interval $[0, T_{max}]$ we always have that $e^{\frac{-\epsilon}{k_B T}} e^{\frac{-\epsilon}{k_B T_{max}}}$. Therefore we can always use the value at T_{max} to condition the convergence of the series. Hence we find that the series converges uniformly on the interval $[0, T_{max}]$, with endpoints included. We may interchange the limit $T \to 0$ and the sum and get

$$N = \sum_{orb} \Theta(\epsilon_F - \epsilon_o) \tag{4.60}$$

If we use free particle energies, which are always positive, we find immediately that $\epsilon_F > 0$ because $N > 0$. Note that the only problem is when we take $T \to \infty$. That limit needs to be analyzed separately, since we now have an infinite number of infinitesimally small terms to sum. We will see later how to deal with that limit. We already know the answer, though, from a previous section. At large temperatures we can replace the sum by an integral, if the orbital energies are free particle energies. This is always true for large quantum numbers, due to Bohr's correspondence principle.

GRAND PARTITION FUNCTION.

The grand partition function for a Fermi gas involves a sum over states. If we enumerate the orbitals for the independent particles, a state of the total system can be specified by the occupation of each orbital, or by a set of numbers $\{n_1, n_2, \ldots\}$. Here n_i denotes the number of particles in orbital i. Since we are dealing with fermions n_i is zero or one only. Examples of states are $\{0, 0, 0, \ldots\}$ for a state with no particles or $\{1, 1, 0, \ldots\}$ for a state with one particle in orbit 1 and in orbit 2.

The grand partition function is

$$3(T, \mu, V) = \sum_{\{n_1, n_2, \ldots\}} e^{\frac{1}{k_B T}(\mu N(\{n_1, n_2, \ldots\}) - E(\{n_1, n_2, \ldots\}))} \tag{4.61}$$

The total number of particles is easy to find:

$$N(\{n_1, n_2, \ldots\}) = \sum_o n_o \tag{4.62}$$

The energy is in general harder to find. But we now make again the assumption that the particles are independent. This allows us to write for the energy:

$$E(\{n_1, n_2, \ldots\}) = \sum_o n_o \epsilon_o \tag{4.63}$$

where the **many body energy** $E(\{n_1, n_2, \ldots\})$ simply is equal to the sum of the **single particle energies** ϵ_o corresponding to occupied states. The grand partition function can now be written in the form

$$\mathfrak{Z}(T, \mu, V) = \sum_{\{n_1, n_2, \ldots\}} e^{\frac{1}{k_B T} \sum_o (\mu - \epsilon_o) n_o} \tag{4.64}$$

$$\mathfrak{Z}(T, \mu, V) = \sum_{n_1=0}^{1} \sum_{n_2=0}^{1} \cdots \sum_{n_i=0}^{1} \cdots \prod_o e^{\frac{1}{k_B T}(\mu - \epsilon_o) n_o} \tag{4.65}$$

$$\mathfrak{Z}(T, \mu, V) = \left(\sum_{n_1=0}^{1} e^{\frac{1}{k_B T}(\mu - \epsilon_1) n_1} \right) \left(\sum_{n_2=0}^{1} e^{\frac{1}{k_B T}(\mu - \epsilon_2) n_2} \right) \cdots \tag{4.66}$$

Since the summation variables are just dummy variables, this is equal to

$$\mathfrak{Z}(T, \mu, V) = \prod_{orb} \left(\sum_{n=0}^{1} e^{\frac{1}{k_B T}(\mu - \epsilon_o) n} \right) = \prod_{orb} \mathfrak{Z}_o(T, \mu, V) \tag{4.67}$$

where we have defined the orbital grand partition function $\mathfrak{Z}_o$ as before by:

$$\mathfrak{Z}_o = \sum_{n=0}^{1} e^{n \frac{\mu - \epsilon_o}{k_B T}} \tag{4.68}$$

The grand partition function for a subsystem with a single orbital is therefore given by $\mathfrak{Z}_o = 1 + e^{\frac{\mu - \epsilon_o}{k_B T}}$.

Note that there is no factor N! in front of the product in (4.67), unlike we had before for the ideal gas partition function. The essential difference is that in the formula above the product is over **orbitals** and these are **distinguishable**. Previously we had a product over particles and those are identical.

GRAND ENERGY.

The grand energy follows from

$$\Omega(T, \mu, V) = -k_B T \log(\mathfrak{Z}_o) = -k_B T \sum_{orb} \log(\mathfrak{Z}_o(T, \mu, V)) \tag{4.69}$$

and the product has become a sum as expected. The energy should be the sum of the energies of the subsystems, and this implies a product in the grand partition function since the energy always needs the logarithm of this grand partition function.

Once the grand energy is available, all thermodynamic variables can be obtained. For example, the number of particles follows from

$$N = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = k_B T \sum_{orb} \frac{1}{\mathcal{3}_o(T,\mu,V)} \left(\frac{\partial \mathcal{3}_o(T,\mu,V)}{\partial \mu}\right)_{T,V} \tag{4.70}$$

or

$$N = \sum_{orb} \frac{e^{\frac{\mu-\epsilon_o}{k_B T}}}{1 + e^{\frac{\mu-\epsilon_o}{k_B T}}} = \sum_{orb} f_{FD}(\epsilon_o; T, \mu) \tag{4.71}$$

as expected.

ENTROPY OF A SYSTEM OF FERMIONS.

A useful formula for the entropy expresses the entropy in terms of the distribution functions.

$$S = -\left(\frac{\partial \Omega}{\partial T}\right)_{\mu,V} = k_B \sum_{orb} \log(\mathcal{3}_o(T,\mu,V)) + k_B T \sum_{orb} \left(\frac{\partial \log(\mathcal{3}_o(T,\mu,V))}{\partial T}\right)_{\mu,V} \tag{4.72}$$

$$S = k_B \sum_{orb} \log(1 + e^{\frac{\mu-\epsilon_o}{k_B T}}) + k_B T \sum_{orb} \frac{e^{\frac{\mu-\epsilon_o}{k_B T}}}{1 + e^{\frac{\mu-\epsilon_o}{k_B T}}} \frac{\mu-\epsilon_o}{-k_B T^2} \tag{4.73}$$

With the help of:

$$e^{\frac{\mu-\epsilon}{k_B T}} = \frac{1}{e^{\frac{\epsilon-\mu}{k_B T}}} = \frac{1}{e^{\frac{\epsilon-\mu}{k_B T}} + 1 - 1} =$$

$$\frac{1}{\frac{1}{f_{FD}} - 1} = \frac{f_{FD}}{1 - f_{FD}} \tag{4.74}$$

we get

$$S = k_B \sum_{orb} \log(1 + \frac{f_{FD}}{1 - f_{FD}}) - k_B \sum_{orb} \frac{\frac{f_{FD}}{1-f_{FD}}}{1 + \frac{f_{FD}}{1-f_{FD}}} \log(\frac{f_{FD}}{1 - f_{FD}}) \tag{4.75}$$

$$S = -k_B \sum_{orb} \log(1 - f_{FD}) - k_B \sum_{orb} f_{FD} \log(\frac{f_{FD}}{1 - f_{FD}}) \tag{4.76}$$

$$S = -k_B \sum_{orb} \left(f_{FD} \log(f_{FD}) + (1 - f_{FD}) \log(1 - f_{FD}) \right) \qquad (4.77)$$

This is very similar to the expression we had before in terms of probabilities. The big difference is again that this time we sum over orbitals. When we summed over many body states we had $S = -k_B \sum_s P_s \log(P_s)$. But now we sum over single particle orbitals, which is a much simpler sum. We can remember the formula above by the following analogy. For each orbital there are two states, it is either occupied or empty and the first term in the expression for S is related to the probability of an orbital being occupied, the second to the probability of the orbital being empty.

4.6 Boson gas.

ONLY A MINUS SIGN DIFFERENT!

The treatment of a gas of identical, independent bosons is almost identical to that for fermions. There is one important exception, though. The distribution function for bosons is

$$f_{BE}(\epsilon) = \frac{1}{e^{\frac{\epsilon - \mu}{k_B T}} - 1} \qquad (4.78)$$

for which we need that $\mu < \min(\epsilon)$, otherwise the number of particles in a given orbital with energy below μ would be negative! The fermion distribution function is always less than one, but for bosons the distribution function can take any positive value depending on how close μ is to ϵ in units of $k_B T$. Obviously, since $f(\epsilon)$ is a monotonically decreasing function of ϵ the orbital with the lowest energy will have the largest population and this orbital will cause problems in the limit $\mu \to \min(\epsilon)$.

GRAND PARTITION FUNCTION.

The grand partition function is calculated in a similar way as for fermions, with the important difference that the number of particles in each orbital can be between zero and infinity

$$\mathfrak{Z}(T, \mu, V) = \sum_{n=0}^{\infty} \sum_{n2=0}^{\infty} \sum_{n_i} e^{\frac{1}{k_B T} \left(\mu \sum_o n_o - \sum_o n_o \epsilon_o \right)} \qquad (4.79)$$

$$\mathfrak{Z}(T, \mu, V) = \sum_{n=0}^{\infty} \sum_{n2=0}^{\infty} \sum_{n_i} \prod_{orb} e^{\frac{n_o}{k_B T} (\mu - \epsilon_o)} \qquad (4.80)$$

$$3(T, \mu, V) = \prod_{orb} \left[\sum_{n=0}^{\infty} e^{\frac{n(\mu - \epsilon_o)}{k_B T}} \right] = \prod_{orb} 3_o(T, \mu, V) \tag{4.81}$$

The only difference is the limit of the summation now going to infinity. The partition function for one orbital is in this case

$$3_o(T, \mu, V) = \frac{1}{1 - e^{\frac{\mu - \epsilon_o}{k_B T}}} \tag{4.82}$$

and is always positive since $\mu < \min(\epsilon)$.

The grand potential follows from

$$\Omega(T, \mu, V) = -k_B T \sum_{orb} \log(3_o(T, \mu, V)) \tag{4.83}$$

in the same way we found for fermions. The total number of particles is equal to minus the derivative of Ω with respect to μ, and is, of course:

$$N = - \left(\frac{\partial \Omega}{\partial \mu} \right)_{T,V} = \sum_{orb} f_{BE}(\epsilon_o) \tag{4.84}$$

because

$$\left(\frac{\partial \log 3_o}{\partial \mu} \right)_{T,V} = \frac{1}{k_B T} 3_o^{-1} \frac{e^{\frac{\mu - \epsilon_o}{k_B T}}}{(1 - e^{\frac{\mu - \epsilon_o}{k_B T}})^2} \tag{4.85}$$

which is equal to the distribution function divided by $k_B T$.

A very useful formula for the entropy, again relating the entropy to probabilities, is

$$S = - \left(\frac{\partial \Omega}{\partial T} \right)_{V,\mu} =$$

$$-k_B \sum_{orb} \left(f_{BE}(\epsilon_o) \log(f_{BE}(\epsilon_o)) - (1 + f_{BE}(\epsilon_o)) \log(1 + f_{BE}(\epsilon_o)) \right) \tag{4.86}$$

The second term always gives a positive contribution, but the first term is negative if $f_{BE}(\epsilon_o) > 1$. But since we can combine terms according to

$$S = k_B \sum_{orb} \left(f_{BE}(\epsilon_o) \log(\frac{1 + f_{BE}(\epsilon_o)}{f_{BE}(\epsilon_o)}) + \log(1 + f_{BE}(\epsilon_o)) \right) \tag{4.87}$$

we see that each negative term in equation (4.86) is cancelled by a larger positive term, because the first term in this expansion is always positive. The second term in equation (4.86) does not have a simple physical interpretation, but is directly related to the physical phenomenon of stimulated emission.

4.7 Problems for chapter 4

PROBLEM 1.

Imagineons have weird statistics. The number of particles in a given orbital can be 0,1, or2.

(A.) Calculate the Imagineon distribution function $f_I(\epsilon)$.

(B.) Sketch the shape of this function with ϵ in units of $k_B T$.

Fermions with spin $\frac{1}{2}$ can also have two particles in one orbital (in the traditional sense the word orbital does not include a spin quantum number, like the 1s orbital in an atom).

(C.) Calculate the distribution function for the number of particles in each orbital.

(D.) Why is this result the same/different from the result of B?

PROBLEM 2.

Assume that for a system of N fermions the Fermi level coincides with the energy level of M orbitals.

(A.) Calculate the entropy at $T = 0$.

(B.) In the thermodynamic limit $N \to \infty$ this entropy is non-zero. What do you know about M in this case?

PROBLEM 3.

Starting with $\Omega(T, \mu, V) = -k_B T \sum_{orb} \log \mathfrak{Z}_{orb}(T, \mu, V)$ for a system of independent, identical bosons, show that the entropy is given by

$$S = -k_B \sum_{orb} \left(f_{BE}(\epsilon_o) \log(f_{BE}(\epsilon_o)) - (1 + f_{BE}(\epsilon_o)) \log(1 + f_{BE}(\epsilon_o)) \right) \quad (4.88)$$

PROBLEM 4.

The Maxwell distribution function f_M is given by $f_M(\epsilon; T, \mu) = e^{\frac{1}{k_B T}(\mu - \epsilon)}$. Show that

$$S(T, \mu, V) = N k_B - \sum_o f_M(\epsilon_o; T, \mu) \log(f_M(\epsilon_o; T, \mu))$$

where the sum is over orbitals.

PROBLEM 5.

Consider a system of independent particles. The number of orbitals with energy between E and $E + dE$ is given by $N(E)dE$. The function $N(E)$ is called the density of states. One measures the expectation value of a certain operator $\mathcal{O}$. For a particle in an orbital o the value of the operator depends only on the energy of the orbital, or $\mathcal{O}_o = O(\epsilon_o)$. Show that in the thermodynamic limit the ensemble average of the operator is given by

$$< \mathcal{O} >= \int_{-\infty}^{\infty} O(E)N(E)f(E; T, \mu)dE$$

where $f(E; T, \mu)$ is the distribution function for these particles.

PROBLEM 6.

The orbital energies of a system of Fermions are given by $\epsilon_i = i\Delta$, with $\Delta > 0$ and $i = 1, 2, 3, \cdots, \infty$. These energies are non-degenerate. If the system has N particles, show that the low temperature limit of the chemical potential gives $\epsilon_F = (N + \frac{1}{2})\Delta$.

PROBLEM 7.

The entropy for a system of independent Fermions is given by

$$S = -k_B \sum_o (f_{FD} \log(f_{FD}) + (1 - f_{FD}) \log(1 - f_{FD}))$$

Calculate $\lim_{T \to 0} f_{FD}(\epsilon, T, \mu)$ for $\epsilon < \mu$,$\epsilon = \mu$, and $\epsilon > \mu$.

The number of orbitals with energy ϵ_o equal to the Fermi energy ϵ_F is M. Calculate the entropy at $T = 0$ in this case.

Explain your answer in terms of a multiplicity function. Pay close attention to the issue of dependent and independent variables.

Chapter 5

Fermi and Bose systems of free, independent particles.

5.1 Fermions in a box.

All the formulas we have derived for the thermodynamic variables of an independent Fermi gas contain a sum over orbitals. One would like to convert these sums to integrals in order to use standard analytical techniques. In order to do this we need to know the details of the single particle energies. Without such further knowledge we cannot derive results in more detail.

Free, independent particles.

We still assume that the particles are independent. Also, in the simplest case we assume that there are no external forces acting on the particles. This is easy when we assume that the energy of the orbital is given by the energy levels of a particle in a box. Of course, there are the implicit forces due to the pressure on the sides of the box, but these forces are taken into account by our boundary conditions. When we are done with our calculations, the pressure needed to keep the box at a constant volume will follow automatically. We also assume that the box is a cube. This is not essential, but makes the mathematics easier. The sides of the cubic box are of length L and the energy levels are

$$\epsilon(n_x, n_y, n_z) = \frac{\hbar^2}{2M} \left(\frac{\pi}{L}\right)^2 (n_x^2 + n_y^2 + n_z^2) \tag{5.1}$$

As we have done before, we define a wave vector $\vec{k}$ by $\vec{k} = \frac{\pi}{L}(n_x, n_y, n_z)$. We use this wave vector to label the single particle orbitals. We also need a spin label, but at this point we assume that there are no magnetic effects. The set of wave vectors is discrete, with step sizes $\Delta k_x = \Delta k_y = \Delta k_z = \frac{\pi}{L}$. Therefore, any sum over orbitals is equal to

$$\sum_{orb} = (2S+1) \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \qquad (5.2)$$

If $h(\epsilon)$ is an arbitrary function of the energy we have

$$\sum_{orb} h(\epsilon_o) = (2S+1) \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} h(\epsilon(n_x,n_y,n_z)) \qquad (5.3)$$

Since the orbital energies do not depend on spin the summation over spins simply gives us a factor $2S+1$.

INTEGRALS ARE EASIER.

Next, we transform the summation variables to $\vec{k}$ and multiply by $1 = \Delta k_x \Delta k_y \Delta k_z (\frac{\pi}{L})^{-3}$. This gives:

$$\sum_{orb} h(\epsilon_o) = (2S+1)(\frac{L}{\pi})^3 \sum_{\vec{k}} h(\epsilon(\vec{k})) \Delta k_x \Delta k_y \Delta k_z \qquad (5.4)$$

If L is large enough the sum can be replaced by an integral and we have (with $V = L^3$):

$$\frac{1}{V} \sum_{orb} h(\epsilon_o) = \frac{2S+1}{\pi^3} \int_{pos} d^3k \; h(\frac{\hbar^2 k^2}{2M}) + \; error \qquad (5.5)$$

where the k-integration runs over all k-vectors with positive components. If we extend the integral over all of k-space, using the fact that $\epsilon(\vec{k})$ is symmetric, we get

$$\frac{1}{V} \sum_{orb} h(\epsilon_o) = \frac{2S+1}{(2\pi)^3} \int d^3k \; h(\frac{\hbar^2 k^2}{2M}) + \; error \qquad (5.6)$$

This gives, of course, the additional factor 2^3. The error is typically of order $\frac{1}{V}$ and hence is unimportant in the limit $V \rightarrow \infty$. Practical calculations always use a finite value for V and in that case the magnitude of the error is determined by the smoothness of the function h on a scale $\frac{\pi}{L}$! This is an important point to keep in mind.

GRAND PARTITION FUNCTION.

The grand energy for a gas of fermions is

$$\Omega(T,\mu,V) = -(2S+1)k_B T \sum_{n_x,n_y,n_z} \log(\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V)) \qquad (5.7)$$

with

$$\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V) = 1 + e^{\frac{\mu - \epsilon(n_x,n_y,n_z)}{k_B T}} \tag{5.8}$$

We now introduce again $\lambda = e^{\frac{\mu}{k_B T}}$, which is always positive, and if we are allowed to replace the sum by an integral we obtain:

$$\frac{1}{V}\Omega(T,\mu,V) = -(2S+1)(2\pi)^{-3}k_B T \int d^3k \ \log(1 + \lambda e^{-\frac{\hbar^2 k^2}{2M k_B T}}) \tag{5.9}$$

Formulas like these make it clear that we really want to use k_B for the Boltzmann constant, in order to avoid confusion with the norm of the wave vector, k.

FREE PARTICLES HAVE A SIMPLE VOLUME DEPENDENCE!

One remarkable thing happened. The right hand side does not depend on volume anymore! When we had a summation over discrete k-vectors, the volume dependence was included in these vectors. But when we changed to a continuous summation (or integration), this volume dependence disappeared, and the only volume dependence is a simple pre-factor V, a linear behavior!

THOSE MATHEMATICAL DETAILS!

What could go wrong here? There is an error associated with the change from summation to integration, and in general we have:

$$\frac{1}{V}\Omega(T,\mu,V) = \left[-(2S+1)(2\pi)^{-3}k_B T \int d^3k \ \log(1 + \lambda e^{-\frac{\hbar^2 k^2}{2M k_B T}})\right][1 + \mathcal{E}(T,V,\mu)] \tag{5.10}$$

where we know that $\lim_{V\to\infty} \mathcal{E}(T,V,\mu) = 0$. But we would like to be able to differentiate the result with respect to T and μ, and hence we need to require uniform convergence! But that in turn also begs the question if the original series is uniformly convergent! In order to discuss the convergence properties we have to take two steps. First, we have to show that the series converges uniformly. Second, we have to show that the series in the thermodynamic limit converges uniformly to the integral. We will discuss the uniform convergence of the series first. The important thing to notice is that both the integral and the series converge.

For large values of the wave vector the terms in the series behave like:

$$t_{n_x,n_y,n_z} = \log(1 + \lambda e^{-\frac{\epsilon(n_x,n_y,n_z)}{k_B T}}) \approx \lambda e^{-\frac{\epsilon(n_x,n_y,n_z)}{k_B T}} \tag{5.11}$$

which approaches zero exponentially fast. Next we consider a domain in (T,μ,L) space with $0 \leqslant T \leqslant T_m$, $-\infty \leqslant \mu \leqslant \mu_m$, and $0 \leqslant L \leqslant L_m$. It is clear that for large values of the indices we have

$$t_{n_x,n_y,n_z} \leqslant e^{\frac{\mu_m}{k_B T} - \frac{\hbar^2 \pi^2}{2ML_m^2 k_B T_m}(n_x^2 + n_y^2 + n_z^2)} \tag{5.12}$$

which means that the series for arbitrary values of T, μ, and L converges faster than for the maximal values. Hence we can use the maximal values of T, μ, and L to describe the convergence of the series in general, and hence the series converges **uniformly** on $0 \leqslant T \leqslant T_m$, $-\infty \leqslant \mu \leqslant \mu_m$, and $0 \leqslant L \leqslant L_m$.

Next we consider a domain for (T, μ, L) space with $T'_m \leqslant T$, $L'_m \leqslant L$, and $-\infty \leqslant \mu \leqslant \mu_m$. The argument of the exponent in (5.7) is

$$\frac{\epsilon(n_x, n_y, n_z)}{k_B T} = \frac{\hbar^2 \pi^2}{2Mk_B T L^2}(n_x^2 + n_y^2 + n_z^2) = x_{n_x}^2 + y_{n_y}^2 + z_{n_z}^2 \tag{5.13}$$

with $x_{n_x} = n_x \sqrt{\frac{\hbar^2 \pi^2}{2Mk_B T L^2}}$ and similar for the y and z components. The size of the steps therefore approaches zero when the temperature becomes very large or when L becomes very large. Therefore, if we can replace the sum by an integral with a certain error at a certain temperature or length, the error will be smaller at all higher temperatures and lengths.

We now make the following observation. Suppose we choose a tolerance ϵ. Because of the uniform convergence of the series we can find a value N_ϵ independent of T, μ, and L (but with $0 \leqslant T \leqslant T_m$, $-\infty \leqslant \mu \leqslant \mu_m$, and $0 \leqslant L \leqslant L_m$) such that:

$$\left| -(2S+1)k_B T \sum_{n_x=1}^{N} \sum_{n_y=1}^{N} \sum_{n_z=1}^{N} \log(\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V)) - S(T,\mu,V) \right| < \epsilon \tag{5.14}$$

for all values of N with $N > N_\epsilon$. Next, use $I(T, \mu, V)$ for the result of the integral. We can find values of T'_m and L'_m in such a manner that

$$\left| -(2S+1)k_B T \sum_{n_x=1}^{\infty} \sum_{n_y=1}^{\infty} \sum_{n_z=1}^{\infty} \log(\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V)) - I(T,\mu,V) \right| < \epsilon \tag{5.15}$$

for all values of the temperature with $T \geqslant T'_m$ and $L \geqslant L'_m$. The last equation implies that the series converges, which we already knew, but does not require uniform convergence.

We are now able to make the following statement. Consider a domain $-\infty \leqslant \mu \leqslant \mu_m$, $T'_m \leqslant T \leqslant T_m$, and $V'_m \leqslant V \leqslant V_m$. On this domain the series for the grand partition function converges uniformly and the sum is close to the value of the integral. We will choose both T_m and V_m very large. We will also choose T'_m very small. Since temperature and volume are multiplied together (via the term TL^2), this means that we need to consider very large values of the volume only. Hence for any value of μ and T in any interval $-\infty \leqslant \mu \leqslant \mu_m$,

$T'_m \leqslant T \leqslant T_m$ we can choose a range of volumes so that the series converges uniformly and is arbitrarily close to the integral, with an error independent of the values of μ and T. Therefore, we can replace the series by the integral for all such values of μ and T, and if the volume is arbitrarily large, the error is arbitrarily small in a uniform manner. As a result, the thermodynamic limit, which we need to take at the end, will not change the results from what we obtained by using the integral.

There are still three problem areas. What if $\mu \to \infty$? What if $T \to 0$? What if $T \to \infty$? The last case is the easiest. If the temperature is very large, the series is very close to the integral for all volumes that are larger than a certain value and all values of μ in the range $-\infty \leqslant \mu \leqslant \mu_m$. The latter is true, since the values of the second order derivative which determine the error are largest at μ_m, and hence this value can be used for uniform convergence criteria. The limit $T \to 0$ is very important and will be discussed after the next paragraph.

What about $\mu \to \infty$? It turns out that this limit is equal to the limit of infinite density. Although we often do calculations at large densities, this limit is not of physical interest since the behavior of the system will be very different from what we are interested in.

There is, however, one more detail. If we replace the summation by an integral, the error is also proportional to the value of the second order derivative of the function somewhere in the interval. If this second order derivative is bounded, there is no problem. If it can go to infinity, there is a problem. As we will see, for fermions there are no problems. But it is good to address this issue for fermions, because after this we will study bosons where there are problems with replacing the sum by an integral. The resulting errors do show up in physics, and are the cause of Bose-Einstein condensation!

The function in the integrant is

$$\log(1 + \lambda e^{-\frac{\hbar^2 k^2}{2Mk_B T}}) \tag{5.16}$$

and is smoothly decaying. The largest order derivatives are when $k \to 0$ and are inversely proportional to the temperature. So the only problem area is when $T \to 0$. In this limit the series has no convergence properties, but we need to investigate whether the series converges to the value of the integral. In this limit the first few terms are so important, that they determine the behavior of the series.

The argument of the summation drops from a value of $\log(1 + e^{\frac{\mu}{k_B T}})$ at the origin of k-space to 0 at infinity. When T is very small, though, the value of the logarithm near the origin becomes very large, and it seems that we need a very dense mesh in k-space to be able to convert the series to an integral with a small error. Hence one should take the limit $V \to \infty$ before the limit $T \to 0$, which is not the correct procedure. For fermions this turns out not to be necessary. Because the series converges uniformly we are able to interchange summation and limit and find

$$\Omega(0,\mu,V) = -(2S+1)\sum_{n_x,n_y,n_z}\lim_{T\to 0}k_BT\log(\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V)) \qquad (5.17)$$

with

$$\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V) = 1 + e^{\frac{\mu-\epsilon(n_x,n_y,n_z)}{k_BT}} \qquad (5.18)$$

The limit of the argument is easy, since $\mu(0) = \epsilon_F(N) > 0$. Therefore the argument is $\mu - \epsilon(n_x,n_y,n_z)$ if $\mu > \epsilon(n_x,n_y,n_z)$, because in that case we have

$$\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V) \approx e^{\frac{\mu-\epsilon(n_x,n_y,n_z)}{k_BT}} \qquad (5.19)$$

and

$$k_BT\log(\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V)) \approx k_BT\frac{\mu-\epsilon(n_x,n_y,n_z)}{k_BT} \qquad (5.20)$$

The argument is 0 if $\mu < \epsilon(n_x,n_y,n_z)$, because now the function $\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V)$ approaches 1 rapidly. As a result we have

$$\frac{\Omega(0,\mu,V)}{V} = -\frac{2S+1}{(2\pi)^3}\int d^3k(\mu-\epsilon(\vec{k}))\Theta(\mu-\epsilon(\vec{k})) + error \qquad (5.21)$$

with $\Theta(x)$ being the well-known step function: $\Theta(x) = 0$ for $x < 0$ and 1 for $x > 0$. In this case the error also goes to zero for $V \to \infty$, only in a different algebraic manner. This shows that we can use

$$\frac{\Omega(T,\mu,V)}{V} = -\frac{2S+1}{(2\pi)^3}\int d^3k k_BT\log(1 + e^{\frac{\mu-\frac{\hbar^2k^2}{2M}}{k_BT}}) \qquad (5.22)$$

for all values of T on $[0,T_m]$ and also in the limit $T \to 0$ with an error less than a small number which only depends on V and not on T. The limit for small temperatures in the last general form of the integral is the same step function we need for the integral that is equal to the sum for $T = 0$. Therefore, we have shown that in the limit $T \to 0$ the series also converges to an integral, and that this is the same integral we obtain by taking the limit $T \to 0$ of the general form. Hence we can use the general form of the integral for all our calculations and derive low temperature series expansions, for example. This situation will be very different for bosons, where these two forms are not the same, and Bose-Einstein condensation occurs!

WHY BOTHER?

So we went through all this math stuff and found that there was nothing special. Why bother? Why would there be anything special? The answer is very simple and important. Every time a convergence fails, we have a singularity and

non-analytic behavior. The physics of such behavior is a phase transition! Hence if we do not discuss convergence issues properly, we will miss many important phase transitions! For example, if we replace sums by integrals for bosons, we see no Bose-Einstein condensation.

EVALUATING THE INTEGRAL.

Our task is to evaluate:

$$\frac{1}{V}\Omega(T,\mu,V) = -(2S+1)(2\pi)^{-3}k_BT \int d^3k \ \log(1 + \lambda e^{-\frac{\hbar^2 k^2}{2Mk_BT}}) \qquad (5.23)$$

The integral is changed into a simple form by a coordinate transformation. Define

$$\vec{x} = \left(\frac{\hbar^2}{2Mk_BT}\right)^{\frac{1}{2}} \vec{k} \qquad (5.24)$$

and we obtain

$$\frac{1}{V}\Omega(T,\mu,V) = -(2S+1)(2\pi)^{-3}k_BT \left(\frac{\hbar^2}{2Mk_BT}\right)^{-\frac{3}{2}} \int d^3x \ \log(1 + \lambda e^{-x^2})$$
$$(5.25)$$

After integrating over the angular variables we find

$$\frac{1}{V}\Omega(T,\mu,V) = -(2S+1)(2\pi)^{-3}k_BT \left(\frac{\hbar^2}{2Mk_BT}\right)^{-\frac{3}{2}} 4\pi \int_0^\infty x^2 dx \ \log(1 + \lambda e^{-x^2})$$
$$(5.26)$$

Next we introduce the thermal wavelength

$$\lambda_T = \left(\frac{2\pi\hbar^2}{Mk_BT}\right)^{\frac{1}{2}} \qquad (5.27)$$

which is approximately the de-Broglie wavelength of a particle with energy k_BT. Again, this combination is introduced for historical reasons to simplify the formulas. The result is a small difference from the purist definition of a de-Broglie wavelength. Note that $n_Q(T) = \lambda_T^{-3}$. At the quantum density a cube of dimension de-Broglie wavelength contains one particle. In other words, the average distance between particles is the de-Broglie wavelength, and it is no surprise that quantum effects will be important!

We also introduce a function $f_{\frac{5}{2}}(\lambda)$ by

$$f_{\frac{5}{2}}(\lambda) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \log(1 + \lambda e^{-x^2}) \qquad (5.28)$$

The notation $\frac{5}{2}$ will become clear in a moment. This function was carefully analyzed, an important occupation before the days of computers when series expansions were badly needed to get numbers. Even now they are important since we can often use simple limit forms of these analytical functions in model calculations. The grand energy is

$$\Omega(T,\mu,V) = - \underbrace{(2S+1)}_{\substack{\text{spin} \\ \text{degeneracy}}} \underbrace{V}_{\substack{\text{simple} \\ \text{volume} \\ \text{dependence}}} \underbrace{k_B T}_{\substack{\text{energy} \\ \text{scale}}} \underbrace{\lambda_T^{-3}}_{\substack{\text{volume} \\ \text{scale}}} \underbrace{f_{\frac{5}{2}}(\lambda)}_{\substack{\text{density} \\ \text{effects}}} \qquad (5.29)$$

The right hand side is independent of volume, depends on μ only via $f_{\frac{5}{2}}(\lambda)$ and on T in three places.

DENSITY DEPENDENCE.

The function $f_{\frac{5}{2}}(\lambda)$ has some simple properties. For large values of x one can expand the logarithm and the integrant is approximately equal to $x^2 \lambda e^{-x^2}$. Therefore, the integral is well behaved at the upper limit of $x \to \infty$ and we are allowed to play all kinds of games with the integrant. If $\lambda < 1$ (remember that $\lambda > 0$) the logarithm can be expanded in a Taylor series for all values of x and since the integral is well behaved it is allowed to interchange summation and integration. Using

$$|z| < 1 \Rightarrow \log(1+z) = \sum_{n=1}^{\infty} \frac{1}{n} z^n (-1)^n \qquad (5.30)$$

and generalizing λ to be able to take on complex values, we have

$$|\lambda| < 1 \Rightarrow \log(1 + \lambda e^{-x^2}) = \sum_{n=1}^{\infty} \frac{1}{n}(-1)^n \lambda^n e^{-nx^2} \qquad (5.31)$$

and since the convergence of the series is uniform we can interchange summation and integration. This leads to

$$|\lambda| < 1 \Rightarrow f_{\frac{5}{2}}(\lambda) = \frac{4}{\sqrt{\pi}} \sum_{n=1}^{\infty} \frac{1}{n}(-1)^n \lambda^n \int_0^{\infty} x^2 dx e^{-nx^2} = \sum_{n=1}^{\infty} \frac{\lambda^n}{n^{\frac{5}{2}}}(-1)^{n+1}$$
$$(5.32)$$

where we have used

$$\int_0^{\infty} y^2 dy e^{-ny^2} = \frac{1}{4}\sqrt{\pi}\frac{1}{n\sqrt{n}} \qquad (5.33)$$

This defines why we used the notation $\frac{5}{2}$. In general, one defines a family of functions $f_\alpha(z)$ by

$$f_\alpha(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^\alpha}(-1)^{n+1} \quad, |z| < 1 \tag{5.34}$$

This defines the function for a disc around the origin. The values for arbitrary values of z are obtained by analytical continuation. For example:

$$f_0(z) = \sum_{n=1}^{\infty} z^n(-1)^{n+1} = 1 - \sum_{n=0}^{\infty} z^n(-1)^n = 1 - \frac{1}{1+z} = \frac{z}{1+z} \tag{5.35}$$

and this defines a function everywhere, except at the pole $z = -1$.

WHY IS THIS IMPORTANT?

One might argue that now we have computers available and we do not need all this analytical detail anymore. Since one use of power series is to be able to evaluate functions over a reasonable domain of arguments, we can simply let computers take over and do the integrations for all values. This is not really correct, however. The analysis above tells us where we have poles and other singularities. These are places where computer calculations will fail! Relying on computers also assumes that there are no errors in the computer calculations, and that is not always true. Therefore, we need to be able to check our computer calculations against the analytical solutions we can obtain via power series and other means. The analytical solutions can tell us that near a limiting point or singularity an energy will behave like $|T - T_0|^{-0.3}$, for example. This is very difficult to extract precisely from a computer calculation. But, of course, computational physics also has its place, because it can extend the analytical solutions to domains of the parameters where we cannot get analytical solutions.

SIMPLE EXAMPLE.

As a simple example, study the differential equation

$$\ddot{x} + x = 0 \tag{5.36}$$

and try a power series of the form $x(t) = \sum_0^\infty c_n t^n$. This gives:

$$\sum_0^\infty c_n n(n - 1)t^{n-2} + \sum_0^\infty c_n t^n = 0 \tag{5.37}$$

or after combining terms with equal powers of t, and defining $c_{-2} = c_{-1} = 0$ we get:

$$\sum_0^\infty (c_{n+2}(n + 2)(n + 1) + c_n)t^n = 0 \tag{5.38}$$

Since this has to be identical to zero, all coefficients of all powers of t have to be zero, and we have $c_{n+2}(n+2)(n+1) + c_n = 0$. We get two sets of solutions. Either we set $c_0 = 1$ and $c_1 = 0$ or do the opposite. A general solution is a linear combination of these two basic solutions. The basic solutions are

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} t^{2k} \tag{5.39}$$

and

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} t^{2k+1} \tag{5.40}$$

which clearly converge for all values of time. Do these functions have special properties? That is very hard to tell from the power series solutions. We know however, that they are equal to $\cos(t)$ and $\sin(t)$ and are periodic. That information comes from a different source of knowledge. We can also solve the differential equation on a computer and plot the solution. Most likely we will see the periodicity right away, but round-off errors will start to change the picture, and perhaps we can see only a certain number of periods. Therefore, can we conclude from the computer calculations that the solutions are periodic? Who knows.

Another example is the differential equation

$$\frac{d^3 x}{dt^3} + x = 0 \tag{5.41}$$

It is easy to construct power series solutions, they are of the form:

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{(3k)!} t^{3k} \tag{5.42}$$

and similar. Are they periodic? What other properties do they have?

FINDING THE CHEMICAL POTENTIAL.

Once we know the grand energy we are able to calculate N and find

$$N(T, \mu, V) = -\left(\frac{\partial \Omega}{\partial \mu}\right)_{T,V} = -\left(\frac{\partial \Omega}{\partial \lambda}\right)_{T,V} \left(\frac{\partial \lambda}{\partial \mu}\right)_{T,V} \tag{5.43}$$

$$N = (2S+1)V k_B T \lambda_T^{-3} \left(\frac{\partial}{\partial \lambda} f_{\frac{5}{2}}(\lambda)\right) \frac{\lambda}{k_B T} \tag{5.44}$$

The derivative for our special function is easy to find for $\lambda < 1$ using the power series.

$$\frac{d}{dz}\sum_{n=1}^{\infty}\frac{z^n}{n^\alpha}(-1)^{n+1}=\sum_{n=1}^{\infty}n\frac{z^{n-1}}{n^\alpha}(-1)^{n+1}=\frac{1}{z}\sum_{n=1}^{\infty}\frac{z^n}{n^{\alpha-1}}(-1)^{n+1} \qquad (5.45)$$

and hence we have for $|z| < 1$:

$$\frac{d}{dz}f_\alpha(z) = \frac{1}{z}f_{\alpha-1}(z) \qquad (5.46)$$

Analytical continuation then leads to the general result that this relation is valid everywhere in the complex plane. As a result, we find:

$$N(T,\mu,V) = \underbrace{(2S+1)}_{\substack{\text{spin}\\\text{degeneracy}}}\underbrace{V}_{\substack{\text{simple}\\\text{volume}\\\text{dependence}}}\underbrace{\lambda_T^{-3}}_{\substack{\text{volume}\\\text{scale}}}\underbrace{f_{\frac{3}{2}}(\lambda)}_{\substack{\text{density}\\\text{effects}}} \qquad (5.47)$$

or, using the density and the quantum density:

$$\frac{n}{n_Q(T)} = (2S+1)f_{\frac{3}{2}}(\lambda) \qquad (5.48)$$

which clearly shows that the effects of the chemical potential only show up through our special function $f_{\frac{3}{2}}(\lambda)$.

We can now use this equation to find $\mu(T,V,N)$. That follows the general procedure we have outlined before. Can we always invert the equation? In thermodynamics we found that $\left(\frac{\partial N}{\partial\mu}\right)_{T,V} > 0$. This allows us to find the chemical potential as a function of N. Does our current result obey the same relation? It should, of course, if we did not make any mistakes. We have

$$\frac{1}{n_Q(T)}\left(\frac{\partial n}{\partial\mu}\right)_{T,V} = \frac{1}{n_Q(T)}\left(\frac{\partial n}{\partial\lambda}\right)_{T,V}\frac{\lambda}{k_BT} = \frac{2S+1}{k_BT}f_{\frac{1}{2}}(\lambda) \qquad (5.49)$$

Is this a positive quantity? It is better to answer this question by looking at the integral forms. We have (using $f_{\alpha-1} = z\frac{d}{dz}f_\alpha$: :

$$f_{\frac{5}{2}}(\lambda) = \frac{4}{\sqrt{\pi}}\int_0^\infty x^2 dx\,\log(1+\lambda e^{-x^2}) \qquad (5.50)$$

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{\pi}}\int_0^\infty x^2 dx\,\frac{\lambda e^{-x^2}}{1+\lambda e^{-x^2}} \qquad (5.51)$$

or

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{\pi}}\int_0^\infty x^2 dx\,\frac{e^{-x^2}}{\lambda^{-1}+e^{-x^2}} \qquad (5.52)$$

The last form already makes it clear that in λ increases the denominator decreases and hence the integral increases. We also find from this expression:

$$f_{\frac{1}{2}}(\lambda) = \frac{4}{\sqrt{\pi}} \frac{1}{\lambda} \int_0^\infty x^2 dx \frac{e^{-x^2}}{\left(\lambda^{-1} + e^{-x^2}\right)^2} \tag{5.53}$$

which is clearly always positive.

If we look at equation(5.48) we see that, since the right hand side is always increasing, there are some simple relations between λ and the temperature T. For $T = 0$ the quantum density goes to zero, and hence the left hand side of equation(5.48) goes to ∞. The right hand side has to go to infinity, too, and this means that λ goes to infinity. If the temperature goes to infinity, the left hand side of the equation goes to zero, which means that λ goes to zero. For values in between, λ decreases when the temperature increases! Hence, **for a fixed value of N**, we have:

$$\lim_{T \to 0} \lambda = \infty \tag{5.54}$$

$$\lim_{T \to \infty} \lambda = 0 \tag{5.55}$$

$$\left(\frac{\partial \lambda}{\partial T}\right)_{V,N} < 0 \tag{5.56}$$

LOW TEMPERATURE EXPANSIONS.

The goal is to find the chemical potential as a function of the number of particles, at a given temperature and volume. We need to use equation (5.48) to do that. We have seen that for a given density low temperatures correspond to values of λ going to infinity. Therefore, we need to investigate the behavior of the function $f_{\frac{3}{2}}$ for large arguments. We start with the integral form, since we are outside the range of the series expansion:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \frac{e^{-x^2}}{\lambda^{-1} + e^{-x^2}} \tag{5.57}$$

which is equal to

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \frac{1}{\lambda^{-1} e^{x^2} + 1} \tag{5.58}$$

Using $y = x^2$ and defining $\lambda = e^\nu$ we get

$$f_{\frac{3}{2}}(\lambda) = \frac{2}{\sqrt{\pi}} \int_0^\infty \frac{1}{e^{y-\nu} + 1} \sqrt{y} dy \tag{5.59}$$

Integration by parts gives via:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \int_0^\infty \frac{1}{e^{y-\nu} + 1} dy^{\frac{3}{2}} \tag{5.60}$$

and using the fact that the integrant is zero at the endpoints of the interval:

$$f_{\frac{3}{2}}(\lambda) = -\frac{4}{\sqrt{3\pi}} \int_0^\infty y^{\frac{3}{2}} d\frac{1}{e^{y-\nu} + 1} \tag{5.61}$$

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \int_0^\infty y^{\frac{3}{2}} \frac{e^{y-\nu}}{\left(e^{y-\nu} + 1\right)^2} dy \tag{5.62}$$

We now define $y = t\nu$ and obtain

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \nu^{\frac{5}{2}} \int_0^\infty t^{\frac{3}{2}} \frac{e^{\nu(t-1)}}{\left(e^{\nu(t-1)} + 1\right)^2} dt \tag{5.63}$$

The form above for $f_{\frac{3}{2}}$ is quite useful. If ν is very large (which means λ very large), the ratio of the exponents on the right hand side is a very sharp function centered at $t = 1$. For example, if t is larger than one, and ν is very large, the exponent is very large, and the ratio of exponents behaves like $e^{\nu(1-t)}$, which is very small. Similarly, if t is less than one, the exponent is very small and the denominator approaches one, while the numerator now is very small.

Since the integrant is sharply peaked at one, this suggests expanding the function $t^{\frac{3}{2}}$ in a Taylor series around one:

$$t^{\frac{3}{2}} = (1 + [t - 1])^{\frac{3}{2}} = \sum_{n=0}^\infty \binom{\frac{3}{2}}{n} (t-1)^n \tag{5.64}$$

This series, however, converges only between zero and two, and hence we cannot use it in the integral as given. However, we can split the integration interval in three parts. The main part is $[\epsilon, 2 - \epsilon]$ where ϵ is a small positive number. The main contribution to the integral comes from this interval. The values at the endpoints are proportional to $e^{-\nu}$ and hence are very small. Therefore:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \nu^{\frac{5}{2}} \int_\epsilon^{2-\epsilon} t^{\frac{3}{2}} \frac{e^{\nu(t-1)}}{\left(e^{\nu(t-1)} + 1\right)^2} dt + \mathcal{O}(\lambda^{-1}) \tag{5.65}$$

The notation $\mathcal{O}(z)$ means terms of order z. But now we can use the series expansion in the integral, and since the series converges uniformly in the range of integration, we can interchange integration and summation:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \nu^{\frac{5}{2}} \sum_{n=0}^\infty \binom{\frac{3}{2}}{n} \int_\epsilon^{2-\epsilon} (t-1)^n \frac{e^{\nu(t-1)}}{\left(e^{\nu(t-1)} + 1\right)^2} dt + \mathcal{O}(\lambda^{-1}) \tag{5.66}$$

or by changing the integration variable:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \nu^{\frac{5}{2}} \sum_{n=0}^\infty \binom{\frac{3}{2}}{n} \int_{-1+\epsilon}^{1-\epsilon} u^n \frac{e^{\nu u}}{\left(e^{\nu u} + 1\right)^2} du + \mathcal{O}(\lambda^{-1}) \tag{5.67}$$

Next, we replace the endpoints in all integrals by $\pm\infty$, which again introduces errors of order $e^{-\nu}$. This seems obvious at first, but we need to take some care since there is a summation in n involved, which easily could give diverging factors. For large values of ν the additional terms in the integral are approximately $\int_{1-\epsilon}^{\infty} u^n e^{-\nu u} du$ and similar for the other part. Since ϵ is very small, we can replace this by zero in this integral. We can now easily show that the integral is proportional to $e^{-\nu} n! \nu_{-n}$. Summing this over n, ignoring the small variation in the binomial factor as a function of n, gives an error term proportional to $e^{-\nu + \frac{1}{\nu}}$. Hence we get:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \nu^{\frac{5}{2}} \sum_{n=0}^{\infty} \binom{\frac{3}{2}}{n} \int_{-\infty}^{\infty} u^n \frac{e^{\nu u}}{(e^{\nu u} + 1)^2} du + \mathcal{O}(\lambda^{-1}) \qquad (5.68)$$

and using $v = \nu u$ we finally get:

$$f_{\frac{3}{2}}(\lambda) = \frac{4}{\sqrt{3\pi}} \nu^{\frac{3}{2}} \sum_{n=0}^{\infty} \nu^{-n} \binom{\frac{3}{2}}{n} \int_{-\infty}^{\infty} v^n \frac{e^v}{(e^v + 1)^2} dv + \mathcal{O}(\lambda^{-1}) \qquad (5.69)$$

The power series is in terms of inverse powers of ν or $\log(\lambda)$, which are the slowest terms to go to zero. For large values of λ these are the only important terms. The ratio of exponents in the integrants is an even function of v, and hence only even powers of n remain in the expansion, because the integrals are zero for odd powers.

Therefore, if $\beta\mu \gg 1$ it follows that

$$f_{\frac{3}{2}}(\lambda) \approx \frac{4}{3\sqrt{\pi}} (\beta\mu)^{\frac{3}{2}} \left(1 + \frac{\pi^2}{8} (\beta\mu)^{-2} \right) \qquad (5.70)$$

Hence in the limit $T \to 0$ we only need to take the first term and we find, with $E_F = \mu(T = 0)$:

$$\frac{n}{n_Q(T)} = (2S + 1) \frac{4}{3\sqrt{\pi}} (\beta E_F)^{\frac{3}{2}} \qquad (5.71)$$

and using the definition of the quantum density:

$$n \left(\frac{2\pi\hbar^2}{M} \right)^{\frac{3}{2}} = (2S + 1) \frac{4}{3\sqrt{\pi}} (E_F)^{\frac{3}{2}} \qquad (5.72)$$

or

$$\left(\frac{3n\sqrt{\pi}}{4(2S + 1)} \right)^{\frac{2}{3}} \frac{2\pi\hbar^2}{M} = E_F \qquad (5.73)$$

which takes the familiar form

$$E_F = \left(\frac{6n\pi^2}{2S + 1} \right)^{\frac{2}{3}} \frac{\hbar^2}{2M} \qquad (5.74)$$

This relates the Fermi energy to the density and this relation is very important in the theory for the conduction electrons in a metal. In that case, of course, $S = \frac{1}{2}$ and the spin factor $2S + 1$ is equal to two.

The correction for low temperatures is obtained by using this $T = 0$ result in the expression for N together with the second term in the series expansion:

$$\frac{n}{n_Q(T)} = (2S + 1)\frac{4}{3\sqrt{\pi}}(\beta\mu)^{\frac{3}{2}}(1 + \frac{\pi^2}{8}(\beta\mu)^{-2}) \qquad (5.75)$$

and by comparing with the $T = 0$ result we obtain:

$$(2S + 1)\frac{4}{3\sqrt{\pi}}(\beta E_F)^{\frac{3}{2}} = (2S + 1)\frac{4}{3\sqrt{\pi}}(\beta\mu)^{\frac{3}{2}}(1 + \frac{\pi^2}{8}(\beta\mu)^{-2}) \qquad (5.76)$$

or

$$(\frac{\mu}{E_F})^{-\frac{3}{2}} = 1 + \frac{\pi^2}{8}(\beta\mu)^{-2} \qquad (5.77)$$

For low temperatures we write $\mu = E_F + \Delta\mu$, with $\Delta\mu$ small, and we can expand the left hand side up to first order:

$$-\frac{3}{2}\frac{\Delta\mu}{E_F} \approx \frac{\pi^2}{8}(\beta\mu)^{-2} \qquad (5.78)$$

On the right hand side we only need the leading order term, and we can replace μ by E_F. This gives

$$\frac{\Delta\mu}{E_F} \approx -\frac{\pi^2}{12}(k_B T)^2 E_F^{-2} \qquad (5.79)$$

or

$$\mu(T, N, V) \approx E_F \left(1 - \frac{\pi^2}{12}\left(\frac{k_B T}{E_F}\right)^2\right) \qquad (5.80)$$

WHAT ARE LOW TEMPERATURES?

The previous equation also gives us a direct measure for the magnitude of the temperature. For a given density $n = \frac{N}{V}$ we can find the Fermi energy from equation (5.74). The corrections depend on the ratio of the thermal energy and the Fermi energy, and hence low temperature means $k_B T \ll E_F$.

HELMHOLTZ FREE ENERGY AL LOW TEMPERATURE.

The Helmholtz free energy is obtained by integrating μ from 0 to N. Since $E_F \propto N^{\frac{2}{3}}$ we can use

$$\int_0^N E_F(N')dN' = \frac{3}{5}NE_F \tag{5.81}$$

$$\int_0^N E_F^{-1}(N')dN' = 3NE_F^{-1} \tag{5.82}$$

and get

$$F(T,N,V) = \int_0^N \mu(T,V,N')dN' \approx \frac{3}{5}NE_F\left(1 - \frac{5\pi^2}{12}\left(\frac{k_BT}{E_F}\right)^2\right) \tag{5.83}$$

OTHER THERMODYNAMIC VARIABLES.

The entropy follows from

$$S(T,N,V) = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \approx \frac{N\pi^2}{2}\frac{k_BT}{E_F}k_B \tag{5.84}$$

and indeed goes to zero if the temperature goes to zero. The internal energy follows from:

$$U(T,N,V) = F + TS \approx \frac{3}{5}NE_F\left(1 + \frac{5\pi^2}{12}\left(\frac{k_BT}{E_F}\right)^2\right) \tag{5.85}$$

At zero temperature we have $U = \frac{3}{5}NE_F$. Hence the average energy per particle is less than the Fermi energy, as expected, since the Fermi energy is the energy of the occupied orbital with the largest energy. The average is not half the Fermi energy, however, since the number of states with a certain energy is increasing with energy. There are more orbitals with an energy larger than half the Fermi energy than there are with an energy less than half the Fermi energy.

The heat capacity at constant volume is:

$$C_V(T,V,N) = \left(\frac{\partial U}{\partial T}\right)_{V,N} \approx Nk_B\frac{\pi^2}{2}\frac{k_BT}{E_F} \tag{5.86}$$

First of all, we notice that the heat capacity goes to zero in a linear fashion, as is required by the third law of thermodynamics. This was not true for the ideal gas, and we clearly have now an improved theory that is valid for low temperature. The ratio of the Fermi-Dirac heat capacity to the ideal heat capacity is:

$$\frac{C_V}{C_V^{ideal}} = \frac{\pi^2}{3}\frac{k_BT}{E_F} \ll 1 \tag{5.87}$$

which seem to indicate that only a small portion of the particles participate in excitation processes. That is indeed true, and will be further explored in solid state physics.

Experiments in solid state physics are able to measure the heat capacity at low temperature. From these measurements on finds a value for the Fermi energy. The density of the electrons can also be measured, from lattice spacing data. Therefore, we can compare the two parts of the equation $E_F = \frac{\hbar^2}{2M}(3n\pi^2)^{\frac{2}{3}}$. In general, we find that the left hand side is not equal to the right hand side. The only available parameter is the mass M of the particles. Apparently, the mass has changed from the free particle mass. This is reasonable, since there are many body interactions. In order to move one electron we need to move others out of the way, too, and hence a certain applied force results in a smaller acceleration. In other words, the mass seems larger. In regular solids this enhancement is a factor between 1 and 10, but in special cases it can be 1000 or more. There are also cases where the enhancement is actually less than one, and lattice effects help to accelerate the electron even more than normal.

Since we also know that $E_F \propto V^{-\frac{2}{3}}$ we find for the pressure

$$p(T, V, N) = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \approx \frac{2}{5}\frac{NE_F}{V}\left(1 + \frac{5\pi^2}{12}\left(\frac{k_B T}{E_F}\right)^2\right) \qquad (5.88)$$

from which we obtain the Gibbs energy:

$$G(T, V, N) = F + pV \approx NE_F\left(1 - \frac{\pi^2}{12}\left(\frac{k_B T}{E_F}\right)^2\right) \qquad (5.89)$$

which shows that also here $G = \mu N$, as expected. Finally, we can also calculate the grand energy:

$$\Omega(T, V, N) = F - \mu N \approx -\frac{2}{5}NE_F\left(1 + \frac{5\pi^2}{12}\left(\frac{k_B T}{E_F}\right)^2\right) \qquad (5.90)$$

which is indeed $-pV$ as expected from the Euler equation in thermodynamics. Now we can go back to the original form for the grand energy derived in this section. Remember that the grand energy had a simple linear dependence on volume? This gives:

$$p = -\left(\frac{\partial \Omega}{\partial V}\right)_{T,\mu} = -\frac{\Omega}{V} \qquad (5.91)$$

indeed. Therefore, the simple volume dependence in the grand energy was required by thermodynamics!

Surprisingly, even at $T = 0$ the pressure of a Fermi gas is greater than zero. This is in contrast to the ideal gas. The origin of this pressure is the Pauli principle which makes that particles of the same spin avoid each other. For example, in a metal we have $n \approx 10^{29}$ m^{-3} and the Fermi energy is a few eV or about 10^{-18} J, leading to a pressure of about 10^5 atm! So why do the conduction electrons stay within a metal? The answer is, of course, the large

Coulomb attraction between the electrons and the ion cores in a solid. A similar large pressure occurs in white dwarf stars. In that case gravity is the force that keeps the particles together.

LARGE TEMPERATURES.

The other limit of interest is $T \to \infty$. In this case we expect, of course, to find the results for an ideal gas. Like we discussed before, since $n = (2S + 1)n_Q(T)f_{\frac{3}{2}}(\lambda)$ the function $f_{\frac{3}{2}}(\lambda)$ has to approach zero and hence $\lambda \to 0$. Therefore this function is dominated by the first term in the power series and we find

$$N(T, \mu, V) \approx (2S + 1)V n_Q(T)\lambda \tag{5.92}$$

Apart from the factor $2S + 1$ this is the result we had before. This last factor is a result of the additional degeneracy we introduced by including spin and reduces the chemical potential. In this high temperature limit we find

$$\Omega(T, \mu, V) \approx -(2S + 1)V k_B T n_Q(T)\lambda \tag{5.93}$$

Together with the formula for N this yields:

$$\Omega(T, \mu, V) \approx -N k_B T \tag{5.94}$$

and using $p = -\frac{\Omega}{V}$:

$$p = \frac{N k_B T}{V} \tag{5.95}$$

Hence the ideal gas law is not influenced by the extra factor $2S+1$. The pressure does not change due to the spin degeneracy, unlike the chemical potential, which is now equal to (found by inverting the equation for N):

$$\mu(T, V, N) = k_B T \log\left(\frac{n}{(2S + 1)n_Q(T)}\right) \tag{5.96}$$

5.2 Bosons in a box.

INTEGRAL FORM.

The discussion in the previous chapter for bosons was again general for all types of bosons, and in order to derive some analytical results we have again to choose a specific model for the orbital energies. Of course we will again take free particles in a box, using $\epsilon_o = \frac{\hbar^2}{2M}k^2$ with $\vec{k} = \frac{\pi}{L}(n_x, n_y, n_z)$. This leads to

$$\Omega(T, \mu, V) = -(2S + 1)k_B T \sum_{n_x, n_y, n_z} \log(\mathfrak{Z}_{n_x, n_y, n_z}(T, \mu, V)) \tag{5.97}$$

with

$$\mathfrak{Z}_{n_x,n_y,n_z}(T,\mu,V) = \left(1 - e^{\frac{\mu - \epsilon(n_x,n_y,n_z)}{k_B T}}\right)^{-1} \tag{5.98}$$

The same trick applies here as we used for fermions, and we replace the series by an integral:

$$\tilde{\Omega}(T,\mu,V) = V k_B T (2S+1) \frac{1}{(2\pi)^3} \int d^3k \log(1 - e^{\frac{\mu - \frac{\hbar^2 k^2}{2M}}{k_B T}}) \tag{5.99}$$

and

$$\Omega(T,\mu,V) = \tilde{\Omega}(T,\mu,V) + \text{error}(T,\mu,V) \tag{5.100}$$

where the error is defined by making this equation exact. One can again show that also for bosons we have $\lim_{V\to\infty} \text{error}(T,\mu,V) = 0$, but the big difference with a system of fermions is that it is not possible to give an upper-bound to the error which is only a function of V. The value of V we need in order to get for example a 1% error becomes infinitely large when $T \to 0$.

Note that $\mu < 0$ for free particles, and hence $0 < \lambda < 1$. Problems also occur in the limit $\mu \to 0$, where the integrant diverges at the lower limit of integration. These can be dealt with. The integral is of the form:

$$\int d^3k \log(\frac{1}{1 - \lambda e^{-\frac{\hbar^2 k^2}{2M k_B T}}}) \tag{5.101}$$

and when k is small the argument of the logarithm is approximately $\log(\frac{2M k_B T}{\hbar^2 k^2})$. Therefore, near the origin we need to evaluate $\int d^3k \log(k)$, which behaves well even though the logarithm diverges, since $k^2 \log(k)$ goes to zero. Hence the function $\tilde{\Omega}$ is well defined for $0 < \lambda \leqslant 1$, or for $\mu \leqslant 0$, and there are no problems in this integral with the limit $\mu \to 0$. As we will see, this limit corresponds to the limit $T \to 0$.

SPECIAL FUNCTIONS FOR BOSONS.

In a manner similar to what we found for fermions, we define a set of special functions by

$$g_\alpha(z) = \sum_{n=1}^{\infty} \frac{z^n}{n^\alpha} \quad , \quad |z| < 1 \tag{5.102}$$

and use analytic continuation to define these functions everywhere. The formula for $\tilde{\Omega}$ is now manipulated in the same way we did for fermions, and we find

$$\tilde{\Omega}(T,\mu,V) = -(2S+1)V k_B T n_Q(T) g_{\frac{5}{2}}(\lambda) \tag{5.103}$$

$$g_{\frac{5}{2}}(\lambda) = -\frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \log(1 - \lambda e^{-x^2}) \qquad (5.104)$$

The limits of the integral cause no problems and it is possible to interchange integration and summation after expanding the logarithm in a Taylor series. This is allowed since we always have $0 < \lambda < 1$. This gives:

$$-\frac{4}{\sqrt{\pi}} \int_0^\infty x^2 dx \log(1 - \lambda e^{-x^2}) = \frac{4}{\sqrt{\pi}} \sum_{n=1}^\infty \frac{1}{n} \lambda^n \int_0^\infty x^2 dx e^{-nx^2} =$$

$$\frac{4}{\sqrt{\pi}} \sum_{n=1}^\infty \frac{1}{n^{\frac{5}{2}}} \lambda^n \int_0^\infty y^2 dy e^{-y^2} \qquad (5.105)$$

which is $g_{\frac{5}{2}}$ indeed.

The number of particles follows from

$$\tilde{N}(T, \mu, V) = -\left(\frac{\partial \tilde{\Omega}}{\partial \mu}\right)_{T,V} = (2S + 1)V n_Q(T) g_{\frac{3}{2}}(\lambda) \qquad (5.106)$$

because we have for the special functions in this case too:

$$\frac{d}{dz} g_\alpha(z) = \frac{1}{z} g_{\alpha-1}(z) \qquad (5.107)$$

The large temperature limit of these equations is the same as for fermions, since the first term in the expansion for $\lambda \to 0$ for $f_\alpha(\lambda)$ and $g_\alpha(\lambda)$ is the same, exactly λ, for both cases. Also, the error term is small for ordinary values of V and we recover the ideal gas laws for both fermions and bosons when the temperature is large or the density is low. The equation for $\tilde{N}$ shows that if $n \ll n_Q(T)$ one needs $g_{\frac{3}{2}}(\lambda) \ll 1$ and this only occurs when $\lambda \ll 1$.

LOW TEMPERATURES.

At very low temperatures we expect that the occupation of the ground state will be large, and hence we need $\mu \to \min(\epsilon)$. In this case we expect problems with replacing the series by an integral. In order to discuss this further we consider the partial sums:

$$\Omega_R(T, \mu, V) = (2S + 1)k_B T \sum_{n_x, n_y, n_z}^{\sqrt{n_x^2 + n_y^2 + n_z^2} < R} \log(1 - \lambda e^{-\frac{\hbar^2 k^2}{2M k_B T}}) \qquad (5.108)$$

with $\vec{k} = \frac{\pi}{L}(n_x, n_y, n_z)$. We see that when $R^2 \frac{\hbar^2 \pi^2}{2ML^2 k_B T} > \log(\lambda)$ the terms in the series become exponentially small, and that convergence goes fast. Like for fermions, we have uniform convergence for any interval $0 < V \leqslant V_m$ and

$0 < T \leqslant T_m$. Also, in case $T \to \infty$ the sum approaches the integral again, and there are no convergence problems for the series.

The next question is if the series always converges to the integral $\tilde{\Omega}$. The integrant in k-space is again steepest for values of k near zero, and we expect problems in the first few terms of the series.

SIMPLE EXAMPLE.

Consider the following series:

$$F(\alpha) = \frac{1}{\sqrt{L}} \sum_{n=0}^{\infty} \frac{e^{-\frac{n}{L}}}{\sqrt{n + \alpha L}} \tag{5.109}$$

Define $x_n = \frac{n}{L}$ and $\Delta x = \frac{1}{L}$. We have:

$$F(\alpha) = \sum_{n=0}^{\infty} \frac{e^{-x_n}}{\sqrt{x_n + \alpha}} \Delta x \tag{5.110}$$

and in the limit $L \to \infty$ this approaches the integral:

$$F(\alpha) = \int dx \frac{e^{-x}}{\sqrt{x + \alpha}} \tag{5.111}$$

If we now take the limit $\alpha \to 0$ we get a well defined result. But we did take the thermodynamic limit $L \to \infty$ first, and that is not correct. If we take the limit $\alpha \to 0$ in the series the very first term, which is equal to $\frac{1}{\sqrt{\alpha L}}$ blows up!

We can find a way out. Write the series in the following form:

$$F(\alpha) = \frac{1}{L\sqrt{\alpha}} + \frac{1}{\sqrt{L}} \sum_{n=1}^{\infty} \frac{e^{-\frac{n}{L}}}{\sqrt{n + \alpha L}} \tag{5.112}$$

Now we can replace the series by an integral for all values of α, since the divergent term is isolated outside the series. Hence we have

$$L \gg 1 \quad \Rightarrow \quad F(\alpha) \approx \frac{1}{L\sqrt{\alpha}} + \int dx \frac{e^{-x}}{\sqrt{x + \alpha}} \tag{5.113}$$

where the first term can be ignored only if $L \gg \frac{1}{\alpha}$! Hence in the limit $\alpha \to 0$ this first term can never be ignored.

THE LOW TEMPERATURE LIMIT FOR BOSONS.

We now return to the question of the magnitude of the error term at low temperature. The easiest way to discuss the nature of the low temperature limit is to consider the expansion for N in terms of orbitals:

$$N(T, \mu, V) = (2S+1) \sum_{n_x, n_y, n_z} \left(e^{\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2) - \mu}{k_B T}} - 1 \right)^{-1} \qquad (5.114)$$

The condition on μ is $\mu \leqslant \frac{\hbar^2 \pi^2}{2M} V^{-\frac{2}{3}} 3$. This guarantees that all terms are positive, and hence each term is smaller than $\frac{N}{2S+1}$. For large values of n_i the terms approach zero exponentially fast, and hence the series converges. That is no surprise, since we know that the result is N. The convergence is uniform for all values of T in $[0, T_m]$ and V in $[0, V_m]$. This is true for all values of μ in the range given above.

It is possible to give an upper-bound to these large terms by

$$\left(e^{\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2) - \mu}{k_B T}} - 1 \right)^{-1} < X e^{-\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2) - \mu}{k_B T_m}} \qquad (5.115)$$

with X of order 1 and for $T < T_m$ if $\sqrt{n_x^2 + n_y^2 + n_z^2} > N(T_m)$. Therefore the series converges uniformly for $T \leqslant T_m$ and $\mu \leqslant 0$, as stated above.

We can now write

$$\lim_{T \to 0} N(T, \mu, V) = (2S+1) \sum_{n_x, n_y, n_z} \lim_{T \to 0} \left(e^{\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2) - \mu}{k_B T}} - 1 \right)^{-1} \quad (5.116)$$

For any value of $\mu \leqslant \frac{\hbar^2 \pi^2}{2M} V^{-\frac{2}{3}} 3$ the limit of each term is zero and hence the limit of N is zero. If the chemical potential is specified, the system contains no particles at zero temperature! This is the case for a system of photons, there is no radiation at zero temperature!

In a gas of bosons, however, we often specify the number of particles $< N >$ and find μ from $< N > = N(T, \mu, V)$. This tells us that μ is a function of T,V, and $< N >$ and that

$$< N > = \lim_{T \to 0} N(T, \mu(T, < N >, V), V) \qquad (5.117)$$

We introduced the notation $< N >$ to distinguish between the actual number of particles in the system and the general function which yield the number of particles when T,μ, and V are specified. Inserting this in the equation for the limits gives:

$$N = (2S+1) \sum_{n_x, n_y, n_z} \lim_{T \to 0} \left(e^{\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2) - \mu(T, N, V)}{k_B T}} - 1 \right)^{-1} \qquad (5.118)$$

where we wrote N again for the number of particles, since the function N did not appear anymore. Because the series converges in a uniform manner, we could interchange the summation and the limit $T \to 0$. For all terms beyond the first we have

$$\frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2) - \mu \geqslant \frac{\hbar^2}{2M}(\frac{\pi}{L})^2(n_x^2 + n_y^2 + n_z^2 - 3) > 0 \qquad (5.119)$$

and hence in the limit $T \to 0$ the exponent goes to infinity and the term goes to zero. As a consequence, we must have

$$N = (2S + 1)\lim_{T \to 0}\left(e^{\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2 3 - \mu(T,N,V)}{k_B T}} - 1\right)^{-1} \qquad (5.120)$$

or

$$\lim_{T \to 0}\frac{\frac{\hbar^2}{2M}(\frac{\pi}{L})^2 3 - \mu(T,N,V)}{k_B T} = \log(1 + \frac{2S+1}{N}) \qquad (5.121)$$

Hence we find the low temperature expansion for the chemical potential and for the absolute activity:

$$\mu(T,N,V) \approx \frac{\hbar^2}{2M}(\frac{\pi}{L})^2 3 - k_B T \log(\frac{N + 2S + 1}{N}) + \mathcal{O}(T^2) \qquad (5.122)$$

WHAT IS A LOW TEMPERATURE?

In the treatment above we call the temperature low if the first term is dominant, or if (comparing the first and the second term):

$$\frac{\hbar^2}{2M}(\frac{\pi}{L})^2 6 - \mu \gg k_B T \qquad (5.123)$$

which is the case (using the limit for μ) if

$$\frac{\hbar^2}{2M}(\frac{\pi}{L})^2 3 \gg k_B T \qquad (5.124)$$

For $L = 1\ m$ and He atoms this gives $k_B T \ll 4 \times 10^{-41}\ J$, or $T \ll 4 \times 10^{-18}\ K$, which is a very low temperature.

But this is not the right answer! We asked the question when are all particles in the ground state. A more important question is when is the number of particles in the ground state comparable to N, say one percent of N. Now we will get a much larger value for the temperature, as we will see a bit later.

LIMITS CANNOT BE INTERCHANGED.

From the formula for the chemical potential we obtain:

$$\lambda(T, N, V) \approx \frac{N}{N + 2S + 1} e^{\frac{\hbar^2}{2Mk_BT}(\frac{\pi}{L})^2 3} \qquad (5.125)$$

The last formula shows that limits cannot be interchanged without penalties:

$$\lim_{T \to 0} \lambda(T, N, V) = \infty \qquad (5.126)$$

$$\lim_{V \to \infty} \lambda(T, N, V) = \frac{N}{N + 2S + 1} \to 1 \qquad (5.127)$$

and hence

$$\lim_{T \to 0} \lim_{V \to \infty} \lambda(T, N, V) \neq \lim_{V \to \infty} \lim_{T \to 0} \lambda(T, N, V) \qquad (5.128)$$

But we need to be careful. Since we required that $\frac{\hbar^2}{2M}(\frac{\pi}{L})^2 3 \gg k_BT$ we cannot take the limit $V \to \infty$ first without violating this condition!

GRAND ENERGY AT LOW TEMPERATURES.

Since we know that the first term in the series for N, and hence in the series for Ω, can be the dominant term and be almost equal to the total sum, we now isolate that term. We could improve on this procedure by isolating more terms, but that will not change the qualitative picture. Hence we write

$$\frac{\Omega(T, \mu, V)}{V} = \frac{(2S + 1)k_BT}{V} \log\left(1 - \lambda e^{-\frac{\epsilon(1,1,1)}{k_BT}}\right) +$$

$$\frac{(2S + 1)k_BT}{V} \sideset{}{'}\sum_{n_x, n_y, n_z} \log(\mathfrak{Z}_{n_x, n_y, n_z}(T, \mu, V)) \qquad (5.129)$$

Now we replace the second term by the integral as before. The lower limit is not strictly zero, but if V is large enough we can replace it by zero. Hence we find

$$\frac{\Omega(T, \mu, V)}{V} = \frac{(2S + 1)k_BT}{V} \log\left(1 - \lambda e^{-\frac{3\hbar^2\pi^2}{2Mk_BTL^2}}\right) -$$

$$(2S + 1)k_BT n_Q(T)g_{\frac{5}{2}}(\lambda) + \text{error}(T, \mu, V) \qquad (5.130)$$

where the error term is due to replacing the sum by the integral. If the temperature is large and the volume is large, the limit of the first term in Ω is zero, the second term dominates, and the error is small compared to the second term. If the temperature is very small, and the volume is large but finite, the error term is large compared to the second term, but small compared to the first term! The error is largest in the transition region between these two cases, and a description could be improved by adding more terms. In the thermodynamic

limit, however, the contribution of these terms will go to zero! For finite volumes these terms can give some contributions, leading to rounding off of sharp curves near the phase transition.

By splitting the range in these two parts, we can show that on each part the convergence is uniform. Hence we have

$$V > V_\epsilon \quad \Rightarrow \quad \left| \frac{\text{error}(T, \mu, V)}{\frac{\Omega}{V}} \right| < \epsilon \tag{5.131}$$

with V_ϵ **independent** of T and μ. Therefore, in all cases the error term is small compared with the value of Ω, assuming a large volume. A value of $1\ m^3$ for the volume certainly satisfies the criteria of being large. Hence for all practical applications we can ignore the error term. Keep in mind, though, that if you do experiments on a small container of helium gas with only 1000 atoms at normal pressure and temperature, you will find deviations from this formula. To analyze such experiments you really have to calculate the sum of the series.

5.3 Bose-Einstein condensation.

The expression for the density $\frac{N}{V}$ is

$$n = \frac{2S+1}{V} f_{BE}(\epsilon_{111}) + \frac{2S+1}{\lambda_T^3} g_{\frac{3}{2}}(\lambda) \tag{5.132}$$

We now take the thermodynamic limit of this expression. For large values of the volume we may replace ϵ_{111} by zero. This gives us

$$n = \frac{2S+1}{V} \frac{\lambda}{1-\lambda} + (2S+1)n_Q(T)g_{\frac{3}{2}}(\lambda) \tag{5.133}$$

It is easy to show that $g_{\frac{3}{2}}(\lambda)$ is a monotonically increasing function of λ, since the derivative is proportional to $g_{\frac{1}{2}}$ which is always positive for $0 \leqslant \lambda \leqslant 1$. As a result, the particle density is an monotonically increasing function of λ. We define $G = g_{\frac{3}{2}}(1) = 2.612 \cdots$.

It is possible to distinguish two regimes. If

$$\frac{n}{n_Q(T)} < (2S+1)G \tag{5.134}$$

it is possible to find a value of $\lambda < 1$ satisfying $n = (2S+1)n_Q(T)g_{\frac{3}{2}}(\lambda)$. In the expression for n we therefore are able to take the thermodynamic limit $V \to \infty$, and the first term in equation (5.133) is zero. This means that in the density regime $n < n_Q(T)$ the ground state orbital does not play an important role, and the gas behaves like a normal gas.

On the other hand, if

$$\frac{n}{n_Q(T)} > (2S+1)G \tag{5.135}$$

this is impossible! For such a combination of density and temperature we need the first term in equation (5.133). For large values of the volume , the value of λ will be close to one, and the value of the second term in equation (5.133) can be replaced by $(2S + 1)G$. In the thermodynamic limit we have

$$\lim_{V \to \infty} \frac{2S + 1}{V} \frac{\lambda}{1 - \lambda} = n - (2S + 1)n_Q(T)G \qquad (5.136)$$

or

$$\lambda = 1 - \frac{2S + 1}{V(n - (2S + 1)n_Q(T)G)} \qquad (5.137)$$

Since in most experiments the temperature is varied and not the density, it is customary to define a critical temperature T_E, the <u>Einstein temperature</u>, by $n = n_Q(T_E)(2S + 1)G$, or

$$T_E = \left(\frac{2\pi\hbar^2}{Mk_B} \right) \left(\frac{n}{(2S + 1)G} \right)^{\frac{2}{3}} \qquad (5.138)$$

The Einstein temperature for Helium is about 3K. The first term in the formula for n, equation (5.133), gives the number of particles in the orbital with the lowest energy. If the temperature is higher than the Einstein temperature, there are essentially no particles in this orbital. If the temperature is below the Einstein temperature, the number of particles in this orbital becomes very large. More precisely, the Einstein temperature is the dividing point between having or not having a **macroscopically large** number of particles in the ground state orbital. In the thermodynamic limit this is the only orbital with a macroscopic number of particles, but for finite volumes other orbitals can be important, too. As we have shown before, for a volume of 1 m^3 we need a temperature less than 10^{-18} K for all particles to be in the ground state.

The relative number of particles in the ground state orbital, $(2S+1)f_{BE}(\epsilon_{111})/N$, for a temperature below the Einstein temperature follows from

$$(2S + 1)\frac{f_{BE}(\epsilon_{111})}{V} = n - (2S + 1)\lambda_T^{-3}G \qquad (5.139)$$

and since the thermal wavelength is proportional to $T^{-\frac{1}{2}}$ we can write this as the fraction of particles in the ground state orbital:

$$(2S + 1)\frac{f_{BE}(\epsilon_{111})}{N} = 1 - \left(\frac{T}{T_E} \right)^{\frac{3}{2}} \qquad (5.140)$$

This fraction is zero for $T > T_E$. At $T = 0$, all particles are in the ground state orbital.

Note that there is some confusion in defining an orbital. If we define an orbital by the quantum numbers n_x, n_y, n_z only, the total number of particles in that orbital is given by $(2S + 1)f_{BE}(\epsilon(n_x, n_y, n_z))$. If we define an orbital by the quantum numbers n plus a quantum number s for spin, the total number of particles in that orbital is $f_{BE}(\epsilon(n_x, n_y, n_z, s))$. But now we have $2S + 1$

degenerate orbitals! Hence at $T = 0$ the number of particles in a ground state orbital $(1, 1, 1, s)$ is $\frac{N}{2S+1}$ and the total number of particles in all ground state orbitals $(1, 1, 1, s)$ together is N. In other words, all particles are in the ground state orbital $(1, 1, 1)$ defined without spin. In a number of books one changes the actual definition of the distribution function by

$$f_{BE}(\epsilon) = \frac{2S + 1}{e^{\frac{\epsilon - \mu}{k_B T}} - 1} \tag{5.141}$$

and similar for fermions. Keep this in mind.

The number of particles in the ground state orbital(s) has a sharp transition according to our formulas. This is because we took the thermodynamic limit. If we keep both terms in the expression for n, however, this transition is smooth for finite volumes. Even for values above the Einstein temperature there are some particles in the ground state orbital. For most experimental situations, however, the differences are small and one is allowed to use the $\frac{3}{2}$ power law formula. Only right at the transition does one see finite size effects. The process in which a macroscopic number of particles go into the ground state orbital is called <u>Bose-Einstein condensation</u>. It is observed experimentally in liquid Helium. Below 2.17 K liquid Helium has no viscosity and a very large thermal conductivity. One can think of these effects as a manifestation of a macroscopic quantum state. Since many particles are in the same orbital, there are no interference effects between their wave functions and they all act in a coherent way.

Note, however, that our calculations are performed for a gas. In a real experiment the temperature is very low and the material is a liquid. Corrections due to particle interactions are needed and the actual calculations are much more difficult. Also, our calculations only show that there is a transition. In order to describe the state with Bose-Einstein condensation we need to take particle interactions into account! Again, these correlation effects are very difficult to incorporate.

5.4   Problems for chapter 5

PROBLEM 1.

An ideal gas consists of H_2 molecules. We need six degrees of freedom to describe such a molecule. Three pertain to the motion of the center of mass of the molecule and these degrees of freedom follow the statistical mechanical description of an ideal gas. There are three internal degrees of freedom, two pertaining to rotations and one to vibrations of each molecule. The values for the energy levels of the rotational degrees of freedom are $k_B T_r j(j + 1)$ and for the vibrational states $k_B T_v(n + \frac{1}{2})$. Quantum effects start to play a role below a temperature T_q, defined by $n_Q(T_q) = n$. The H_2 gas in enclosed in a fixed

volume of 22,4 liters and the number of molecules in this volume is one mole. Assume that $T_r \approx 175\ K$ and $T_v \approx 6,500\ K$.

A() Calculate T_q

(B) Calculate an approximate value for C_v at $T = 50\ K$

(C) Calculate an approximate value for C_v at $T = 4,000\ K$

(D) Calculate an approximate value for C_v at $T = 50,000\ K$

PROBLEM 2.

Using the expansion

$$f_{\frac{3}{2}}(z) = \frac{4}{3\sqrt{\pi}}\left((\log z)^{\frac{3}{2}} + \frac{\pi^2}{8}(\log z)^{-\frac{1}{2}} + \frac{7\pi^4}{640}(\log z)^{-\frac{5}{2}}\cdots\right)$$

for large values of z, calculate the low temperature behavior of $\mu(T)$, $U(T)$, $S(T)$, and $p(T)$ up to fourth order in T.

PROBLEM 3.

At high temperatures we have $\lambda \to 0$ for both bosons and fermions. Use the formula for $\frac{N}{V}$ and expand $f_{\frac{3}{2}}$ and $g_{\frac{3}{2}}$ up to second order. Note that $\frac{n}{n_Q}$ is small!

(A) Find the correction term to $\mu(T)$ due to the second term in that expansion for bosons and fermions.

(B) Calculate the pressure including this second term for bosons and fermions.

PROBLEM 4.

Pauli Paramagnetism.
The energy of non-relativistic electrons in a small magnetic field is given by $\epsilon_{\vec{p},s} = \frac{p^2}{2m} - s\mu_0 B$ where $s = \pm 1$ and μ_0 is the magnetic moment of the electron. Assume $\mu_0 B \ll \epsilon_F$. Note that in this problem we ignore the effect of the magnetic field on the orbit of the electron, that turns out to be OK in first approximation. Evaluate the magnetic susceptibility χ in the following four cases:

(A) For $T = 0$.

(B) For $k_B T \ll \epsilon_F$, one more term beyond (A).

(C) For $T = \infty$. (Note: this one you can get without any detailed calculation).

(D) For $k_B T \gg \epsilon_F$, one more term beyond (C).

PROBLEM 5.

The virial expansion is given by $\frac{p}{kT} = \sum_{j=1}^{\infty} B_j(T) \left(\frac{N}{V}\right)^j$ with $B_1(T) = 1$. Find $B_2(T)$ for non-interacting Fermions in a box.

PROBLEM 6.

The energy of relativistic electrons is given by $\epsilon_{\vec{p},s} = \sqrt{p^2 c^2 + m^2 c^4}$, which is independent of the spin s. These particles are contained in a cubical box, sides of length L, volume V.

(1) Calculate the Fermi energy $\epsilon_F = \mu(T = 0)$ as a function of N and V.

(2) Calculate the internal energy U.

(3) Expand the integral in (2) in a power series, assuming that the density $\frac{N}{V}$ is very low.

(4) Expand the integral in (2) in a power series, assuming that the density $\frac{N}{V}$ is very high.

PROBLEM 7.

Landau diamagnetism. The orbits of an electron in a magnetic field are also quantized. The energy levels of the electron are now given by

$$\epsilon(p_z, j, \alpha, s) = \frac{p_z^2}{2m} + \frac{e\hbar B}{mc}(j + \frac{1}{2})$$

with $p_z = \frac{2\pi}{L}\hbar l$, $l = 0, \pm 1, \pm 2, \cdots$, and $j = 0, 1, \cdots$. The quantum number α counts the degeneracy and α runs from 1 to $g = \frac{eBL^2}{2\pi\hbar c}$. The magnetic field is along the z direction. We have ignored the energy of interaction between the spin and the magnetic field.

(1) Give an expression for the grand partition function $\zeta(T, \mu, V)$.

(2) Calculate the magnetic susceptibility for $T \to \infty$ (no, zero is not an acceptable answer, one more term, please) in a weak magnetic field.

Chapter 6

Density matrix formalism.

6.1 Density operators.

DIFFERENT WAYS TO LOOK AT STATISTICAL MECHANICS.

The theory of statistical mechanics is easily formulated in a quantum mechanical frame-work. This is not the only possibility, however. All original developments were, of course, in terms of classical mechanics. Different questions arise and we will address some of those questions in the next chapter. But even in quantum mechanics there are different ways to obtain information about a system. Using the time-independent Schrödinger equation, like we did, is one way. We can also use matrices, like Heisenberg, or more general, operators.

In statistical mechanics we start with a definition of the entropy. A different formulation of statistical mechanics, as given in this chapter, essentially is based on a different definition of the entropy. Our task is to show that the resulting theories are equivalent. In order to make efficient use of the definition of entropy introduced in this chapter we have to introduce density matrices.

Density matrices are an efficient tool facilitating the formulation of statistical mechanics. A quantum mechanical description of a system requires you to solve for the wave-functions of that system. Pure states are represented by a single wave function, but an arbitrary state could also contain a mixture of wave functions. This is true in statistical mechanics, where each quantum state is possible, and the probability is given by the Boltzmann factor. Density matrices are designed to describe these cases.

Suppose we have solved the eigenvalue system

$$H \left| n \right\rangle = E_n \left| n \right\rangle \tag{6.1}$$

with normalization $\left\langle n' \mid n \right\rangle = \delta_{n'n}$ and closure $\sum_n \left| n \right\rangle \left\langle n \right| = 1$. If the probability for the system of being in state $\left| n \right\rangle$ is p_n, the expectation value of an arbitrary operator A is given by

$$< A >= \sum_n p_n \langle n| A |n\rangle \tag{6.2}$$

The values of p_n are all we know about the state of our system, and they give a complete description of this state. In this case the description of the system is an ensemble, which is a more general way to describe a system than using pure states only. Of course we need $p_n \geqslant 0$ and $\sum_n p_n = 1$. A pure state corresponds to $p_m = 1$ for a particular value of m. We now define the operator ρ by

$$\rho = \sum_n p_n |n\rangle \langle n| \tag{6.3}$$

This operator is called a density operator, and the matrix of its elements in a given representation is called a density matrix. The terms operator and matrix are frequently used for both, however. The operator ρ obeys the following relations:

$$\mathrm{Tr}\ \rho = \sum_n \langle n| \rho |n\rangle = \sum_n p_n = 1 \tag{6.4}$$

$$\rho = \rho^\dagger \tag{6.5}$$

$$\rho^2 \leqslant \rho \tag{6.6}$$

The last equation has to be interpreted in the following manner: we have $\langle x| \rho^2 |x\rangle \leqslant \langle x| \rho |x\rangle$ for all states $|x\rangle$. A consequence of this equation is that ρ is non-negative, $\langle x| \rho |x\rangle \geqslant 0$ for all states $|x\rangle$. A Hermitian operator is defined by $\langle x| \rho^\dagger |y\rangle = (\langle y| \rho |x\rangle)^*$ for all states $|x\rangle$ and $|y\rangle$. The equations (6.4) through (6.6) can also be used as a definition of a density operator. Any operator obeying these requirements is a density operator of the form (6.3) and the probabilities p_n are the eigenvalues of the density operator.

If we know the density matrix for a system, we know everything about this system. This is due to the fact that we can only measure expectation values of observable quantities, and these expectation values are obtained from a density matrix by

$$< A >= \sum_n \langle n| p_n A |n\rangle = \mathrm{Tr}\ (\rho A) \tag{6.7}$$

where we have used the fact that $\rho |n\rangle = p_n |n\rangle$ for the basis set of eigenstates of the density operator. A big advantage of this formulation is that it is basis independent. For example, the trace of a matrix is independent of the choice of basis states. The density matrix can be specified in any basis, or in operator form.

In statistical mechanics the probabilities p_n are equal to the Boltzmann factor, and we have

$$p_n = \frac{1}{\mathcal{Z}} e^{-\frac{E_n}{k_B T}} \tag{6.8}$$

The partition function is

$$\mathcal{Z} = \sum_n e^{-\frac{E_n}{k_B T}} = \sum_n \langle n| e^{-\frac{E_n}{k_B T}} |n\rangle \tag{6.9}$$

Now we use the fact that the states in the sum are eigenstates of the Hamiltonian, and we obtain

$$\mathcal{Z} = \sum_n \langle n| e^{-\frac{H}{k_B T}} |n\rangle = \text{Tr } e^{-\frac{H}{k_B T}} \tag{6.10}$$

At this point we will switch to the more common notation for the inverse temperature, $\beta = \frac{1}{k_B T}$. The density operator is given by

$$\rho = \frac{1}{\mathcal{Z}} \sum_n e^{-\beta E_n} |n\rangle \langle n| \tag{6.11}$$

and using the fact again that we expand in eigenstates of the Hamiltonian:

$$\rho = \frac{1}{\mathcal{Z}} e^{-\beta H} \tag{6.12}$$

These definitions are basis-dependent. The only question is how to define the exponent of an operator, but that turns out to be easy since the power series for exponents always converge uniformly on any finite interval. Hence the exponent of an operator O is defined by $e^O = \sum_{n=0}^{\infty} \frac{1}{n!} O^n$. The mathematical filed of functional analysis, very useful for physicists, gives many more details. Products of operators are well defined. Therefore, the density operator can be specified without knowing the eigenstates of the Hamiltonian. This is definitely an advantage. Note, however, that in many cases the most practical manner to do calculations is by using eigenvalues. But there are certainly other cases where it is better not to do so. Finally, the expectation value of an operator A in this formalism is

$$< A >= \text{Tr } (\rho A) = \frac{\text{Tr } (e^{-\beta H} A)}{\text{Tr } (e^{-\beta H})} \tag{6.13}$$

This is an important form for expectation values. It can be used in diagrammatic techniques in the same manner as Feynman diagrams. The only substitution we need to make is $\beta \rightarrow \imath t$. We replace inverse temperature by imaginary time. Now what does that mean???

Formulas (6.10) and (6.12) contain functions of operators. Elementary functions of an operator are positive integral powers of an operator. For example, $O^2 |x\rangle = O |y\rangle$ with $|y\rangle = O |x\rangle$. Sums and differences of operators are also easily defined. More complicated functions of operators can be defined through power series, just like functions of a complex variable in complex analysis. These power series often have a certain radius of convergence, and expressed in terms

of operators are only valid if the norm of the operator is less than this radius of convergence. But often these functions can be continued for arbitrary operators similar to the analytic continuation procedure in complex analysis. For example, e^O is defined via a power series for all operators O. If we can solve the operator equation $O = e^X$ for the operator X, then we have defined $X = \log(O)$. The logarithm of an operator can be expressed in the form of a power series for $\log(1 + O)$ only if the norm of O is less than one. But $\log(1 + O)$ is defined for all operators O for which $O = e^X$ can be solved; this is a much larger class of operators.

If an operator O is Hermitian, $O = O^\dagger$, the definition of functions of operators can be related to standard functions of real variables. A Hermitian operator has a complete set of eigenfunctions $|n\rangle$ with corresponding eigenvalues λ_n. If $f(x)$ is a well-defined function a real variable x in a certain range, the definition of f can be extended to operators by $O = \sum \lambda_n |n\rangle \langle n| \Rightarrow f(O) = \sum f(\lambda_n) |n\rangle \langle n|$. For example, the logarithm of a Hermitian operator can be defined this way only if none of the eigenvalues of the operator is zero! In other words, $O = e^X$ with O Hermitian can only be solved if none of the eigenvalues of O is zero.

This treatment allows us to check what happens for an operator with $\rho^2 \leqslant \rho$. Suppose we have

$$\rho |n\rangle = r_n |n\rangle \tag{6.14}$$

This gives

$$\langle n| \rho^2 |n\rangle \leqslant \langle n| \rho |n\rangle \Rightarrow r_n^2 \leqslant r_n \tag{6.15}$$

which means that r_n is not negative and at most equal to one, $0 \leqslant r_n \leqslant 1$. Hence for an arbitrary vector:

$$\langle x| \rho^2 |x\rangle = \sum_n \langle x| \rho |n\rangle \langle n| \rho |x\rangle = \sum_n \langle x| r_n |n\rangle \langle n| r_n |x\rangle = \sum_n r_n^2 |\langle n |x\rangle|^2 \tag{6.16}$$

Similarly:

$$\langle x| \rho |x\rangle = \sum_n \langle x| \rho |n\rangle \langle n |x\rangle = \sum_n \langle x| r_n |n\rangle \langle n |x\rangle = \sum_n r_n |\langle n |x\rangle|^2 \tag{6.17}$$

and since $r_n^2 \leqslant r_n$ we have

$$\sum_n r_n^2 |\langle n |x\rangle|^2 \leqslant \sum_n r_n |\langle n |x\rangle|^2 \tag{6.18}$$

or $\langle x| \rho^2 |x\rangle \leqslant \langle x| \rho |x\rangle$ for all possible vectors. Hence the definition for $\rho^2 \leqslant \rho$ is a valid one. Once we define this condition on all eigenstates, it is valid everywhere. This is another example showing that it is sufficient to define properties on eigenstates only.

We can now make the following statement. In general, an arbitrary system is described by a density operator ρ, which is any operator satisfying the conditions (6.4) through (6.6). We also use the term density matrix, but this is basis dependent. If we have an arbitrary basis $|n\rangle$, the density matrix R corresponding to ρ in that basis is defined by $R_{ij} = \langle i| \rho |j\rangle$. In statistical mechanics we make a particular choice for the density matrix. For example, a system at a temperature T and volume V, number of particles N has the density operator given by (6.12). The extensive state variables N and V are used in the definition of the Hamiltonian.

6.2 General ensembles.

The partition function is related to the Helmholtz free energy by $\mathcal{Z} = e^{-\beta F}$. The states $|n\rangle$ in the definitions above all have N particles, and the density operator ρ is defined in the Hilbert space of N-particle states, with volume V. It is also possible to calculate the grand potential Ω in a similar way. The number operator $\mathbb{N}$ gives the number of particles of state $|n\rangle$ via $\mathbb{N}|n\rangle = N_n |n\rangle$. The energy of this state follows from $\mathbb{H}|n\rangle = E_n |n\rangle$.

In this space we still have orthogonality and completeness of the basis. The only difference with the example in the previous section is that our Hilbert space is much larger. We now include all states with arbitrary numbers of particles (even zero!!), but still with volume V. Mathematically these spaces are related by $\mathcal{S}_V = \prod_N \mathcal{S}_{V,N}$.

The grand partition function $\mathfrak{Z}$ is defined by

$$\mathfrak{Z} = \sum_n e^{-\beta(E_n - \mu N_n)} \tag{6.19}$$

where we sum over states $|n\rangle$ with arbitrary numbers of particles. This is often desirable in field-theory, where the number of particles is not fixed and the operators are specified in terms of creation and annihilation operators on this larger space, called Fock-space. The grand partition function is a trace over all states in this general Fock-space:

$$\mathfrak{Z} = \mathrm{Tr}\ e^{-\beta(\mathbb{H} - \mu \mathbb{N})} = e^{-\beta \Omega} \tag{6.20}$$

and the density matrix in this case is $\frac{1}{\mathfrak{Z}} e^{-\beta(\mathbb{H} - \mu \mathbb{N})}$. The operator $\mathbb{H} - \mu \mathbb{N}$ is sometimes called the grand Hamiltonian. In this derivation we have used the fact that we can define energy eigenstates for specific particle numbers, or $[\mathbb{H}, \mathbb{N}] = 0$. That condition is not necessary, however. We can use the partition function in the form (6.20) even if the Hamiltonian does not commute with the particle number operator, like in the theory of superconductivity. The density operator is used to calculate probabilities. The probability of finding a system with temperature T, chemical potential μ, and volume V in a state $|x\rangle$ is given by $\langle x| \frac{1}{\mathfrak{Z}} e^{-\beta(\mathbb{H} - \mu \mathbb{N})} |x\rangle$.

It is also possible to use other Legendre transformations. If $\mathbb{V}$ is the volume operator, one can show that

$$\text{Tr } e^{-\beta(\mathbb{H}+p\mathbb{V})} = e^{-\beta G} \tag{6.21}$$

where G is the Gibbs free energy $G = U - TS + pV$. The trace is a sum over quantum states with arbitrary volumes. This is not necessarily a useful extension, though. Quantum states are often easier to describe at constant volume.

In summary, we have

$$\mathcal{Z}(T,V,N) = \underbrace{\text{Tr}}_{\mathcal{S}_{V,N}} e^{-\beta H} \tag{6.22}$$

where the trace is in the space of all quantum states with volume V and number of particles N. Similarly, we have

$$\mathfrak{Z}(T,\mu,N) = \underbrace{\text{Tr}}_{\mathcal{S}_V} e^{-\beta(\mathbb{H}-\mu\mathbb{N})} \tag{6.23}$$

where the trace is now in the much larger Hilbert space of states with volume V. States do not have to have well defined particle numbers! Independent whether the Hamiltonian commutes with the particle number operator we can always write

$$\mathfrak{Z}(T,\mu,N) = \sum_N \underbrace{\text{Tr}}_{\mathcal{S}_{V,N}} e^{-\beta(\mathbb{H}-\mu\mathbb{N})} \tag{6.24}$$

or

$$\mathfrak{Z}(T,\mu,N) = \sum_N e^{\beta\mu N} \underbrace{\text{Tr}}_{\mathcal{S}_{V,N}} e^{-\beta H} = \sum_N e^{\beta\mu N} \mathcal{Z}(T,V,N) \tag{6.25}$$

as before. This can also be extended. Suppose we can measure the magnetic moment of a system. This magnetic moment M will be an integer times a basic unit for finite particle numbers. Hence we have

$$\mathcal{Z}(T,V,N,M) = \underbrace{\text{Tr}}_{\mathcal{S}_{V,N,M}} e^{-\beta H} \tag{6.26}$$

where the summation is over a smaller space with a specified value of the magnetic moment. If we do calculations as a function of magnetic field we have:

$$\widetilde{\mathcal{Z}}(T,V,N,h) = \underbrace{\text{Tr}}_{\mathcal{S}_{V,N}} e^{-\beta(H-hM)} \tag{6.27}$$

We find again:

$$\widetilde{\mathcal{Z}}(T,V,N,h) = \sum_M e^{\beta hM} \mathcal{Z}(T,V,N,M) \tag{6.28}$$

6.3 Maximum entropy principle.

In Chapter 1 we defined the entropy in terms of the number of states available to a system at constant energy, volume, etc. In Chapter 2 we derived a formula relating the entropy and the probabilities of finding the system in a certain state. One can also turn it around, and use this formula as a definition of the entropy:

$$S = -k_B \text{Tr } \rho \log(\rho) \tag{6.29}$$

It is easy to see that this is the same. If the eigenvalues of ρ are p_n we have

$$S = -k_B \sum_n \langle n| \rho \log(\rho) |n\rangle = -k_B \sum_n p_n \log(p_n) \tag{6.30}$$

where we have used $\log(\rho) |n\rangle = \log(p_n) |n\rangle$. Hence in all cases where we have the eigenstates of the Hamiltonian we get the same answer for the entropy. But equation (6.29) is more general and can be used in all cases where we have no good description of the states of a system in terms of eigenstates of a Hamiltonian. It is a more general foundation of statistical mechanics, and one might argue a better foundation. In all cases we have studied so far both approaches give the same results.

The definition (6.29) is called the information-theoretical definition of the entropy. It has a very useful application for finding the density operator corresponding to a thermodynamical system in a given state. Suppose a system is specified by the values of some state variables. It is then often possible to construct a general form of the density matrix ρ consistent with these state variables. This form will contain other free parameters, though. These free parameters are determined by **maximizing the entropy**! Therefore, one is able to construct the density matrix of a system in a variational way. Once the density matrix is known, all other quantities of interest can be found by simply evaluating expectation values.

Suppose we have a system where we only have states with a given energy, number of particles, and volume. These extensive parameters are sued to define the Hilbert space of possible states of the system, and we maximize the entropy in that space. But suppose we cannot exactly specify one of these variables. For example, the number of particles can vary. In that case we must maximize the entropy in a much larger Hilbert space, containing states with all possible number of particles. But even then we can still measure the *average* number of particles. So we want to find the entropy for a given value of the *average* number of particles. This leads to the use of Lagrange multipliers. In thermodynamics we have made the connection with chemical potentials, and hence we expect that the Lagrange multiplier in this case is related to the chemical potential.

As a first example, we discuss the <u>microcanonical ensemble</u>. Suppose the thermodynamical state of the system is given by exact values of the internal energy U, the volume V, and the number of particles N. There are no variations of these values. The last two conditions constrain the Hilbert space on which ρ

is defined. We restrict the quantum mechanical wave functions to states with a definite volume V and number of particles N. The energy restriction tells us that we only want to consider states with $\langle x | H | x \rangle = U$. The Hilbert space of all possible states $|x\rangle$ obeying these conditions is $\mathcal{S}_{U,V,N}$.

The problem of the microcanonical ensemble is therefore to maximize the entropy (6.29) in this space. hence we want to find the maximal value of the entropy for all operators ρ obeying the conditions (6.4) through (6.6). Limiting our search to Hermitian operators only is easy. The condition (6.6) is an inequality and very hard to incorporate in general during the search. Therefore, this condition we will always have to test at the end. We find all possible maxima of the entropy, and throw away those that correspond to an operator not obeying (6.6). Finally, the equality condition (6.4) is incorporated via a Lagrange multiplier.

The task is now to find the maximum of

$$X(\rho) = -k_B \text{Tr } \rho \log(\rho) + \lambda k_B (\text{Tr } \rho - 1) \tag{6.31}$$

over all Hermitian operators on the space $\mathcal{S}_{U,V,N}$. It turns out to be convenient to extract a factor k_B from the Lagrange multiplier. Based on this euqation, we need some functional derivative equation of the form $\frac{\delta X}{\delta \rho} = 0$, but now using operators. How do we define that? Suppose we change the operator by a small amount, keeping it Hermitian. The we define

$$\Delta X = X(\rho + \Delta \rho) - X(\rho) \tag{6.32}$$

just as we did for functional derivatives. In order to connect with what we did before, assume that we have some basis in the space on which the operators are acting. Define the density matrix by

$$R_{ij} = \langle i | \rho | j \rangle \tag{6.33}$$

and the variation by

$$\Delta R_{ij} = \langle i | \Delta \rho | j \rangle \tag{6.34}$$

In first order we can then write

$$\Delta X = \sum_{i,j} A_{ji} \Delta R_{ij} \tag{6.35}$$

and now we **define** the operator $\frac{\delta X}{\delta \rho}$ by its matrix elements:

$$\langle j | \frac{\delta X}{\delta \rho} | i \rangle = A_{ji} \tag{6.36}$$

This leads to

$$\Delta X = \sum_{i,j} \langle j | \frac{\delta X}{\delta \rho} | i \rangle \langle i | \Delta \rho | j \rangle = \text{Tr } \left(\frac{\delta X}{\delta \rho} \Delta \rho \right) \tag{6.37}$$

and at an extremum this should be zero for all variations $\Delta\rho$, which gives us the condition mentioned before. The changes in the quantity X are related to changes in the operator ρ, and ultimately to changes in the density matrix R_{ij}. Suppose we only change one matrix element, hence only ΔR_{mn} is non-zero for a specific value of m and n. This gives:

$$\Delta X = \langle n| \frac{\delta X}{\delta\rho} |m\rangle \langle m| \Delta\rho |n\rangle \tag{6.38}$$

but we also have in terms of simple partial derivatives

$$\Delta X = \left(\frac{\partial X}{\partial R_{mn}}\right) \langle m| \Delta\rho |n\rangle \tag{6.39}$$

which leads to

$$\langle n| \frac{\delta X}{\delta\rho} |m\rangle = \left(\frac{\partial X}{\partial R_{mn}}\right) \tag{6.40}$$

where we note that the order in which n and m appear is interchanged left and right.

In our particular case we have

$$X = -k_B \sum_{ij} \langle i| \rho |j\rangle \langle j| \log(\rho) |i\rangle + \lambda k_B (\sum_i \langle i| \rho |i\rangle - 1) \tag{6.41}$$

and hence

$$\left(\frac{\partial X}{\partial R_{nm}}\right) = -k_B \langle m| \log(\rho) |n\rangle + \lambda k_B \delta_{nm} - k_B \sum_{ij} \langle i| \rho |j\rangle \left(\frac{\partial \langle j| \log(\rho) |i\rangle}{\partial R_{nm}}\right) \tag{6.42}$$

It is the last term that causes problems in the calculations. This task can be reformulated as follows. Suppose the operator ρ depends on some complex variable x. Calculate $\text{Tr}\,\rho\left(\frac{\partial\rho}{\partial x}\right)$. This can be done by using the definition of a logarithm. We have

$$\log(\rho) = \tau \Leftarrow e^\tau = \rho \tag{6.43}$$

We can relate derivatives of ρ to derivatives of τ. In general we have

$$\frac{\partial}{\partial x} e^{\tau(x)} = \frac{\partial}{\partial x} \sum_n \frac{1}{n!} \tau^n(x) \tag{6.44}$$

and since the series converges uniformly, this gives

$$\frac{\partial}{\partial x} e^{\tau(x)} = \sum_n \frac{1}{n!} \frac{\partial}{\partial x} \tau^n(x) \tag{6.45}$$

At this point we encounter an important difference between functions and operators. We cannot interchange the order of operators if they do not commute, and hence we have to write something of the form $\left(A^2\right)' = A'A + AA'$! In general, an operator and its derivative do not commute. Therefore:

$$\frac{\partial}{\partial x}\tau^n(x) = \sum_{m=1}^{n}\tau^{m-1}(x)\left(\frac{\partial\tau}{\partial x}\right)\tau^{n-m}(x) \tag{6.46}$$

and

$$\frac{\partial}{\partial x}e^{\tau(x)} = \sum_{n}\frac{1}{n!}\sum_{m=1}^{n}\tau^{m-1}(x)\left(\frac{\partial\tau}{\partial x}\right)\tau^{n-m}(x) \tag{6.47}$$

and we cannot say much more. But, in our example we need to take the trace of this equation, and we can use $\mathrm{Tr}\,(ABC) = \mathrm{Tr}\,(CAB)$. Therefore:

$$\mathrm{Tr}\,(\frac{\partial}{\partial x}e^{\tau(x)}) = \sum_{n}\frac{1}{n!}\sum_{m=1}^{n}\mathrm{Tr}\,(\tau^{m-1}(x)\left(\frac{\partial\tau}{\partial x}\right)\tau^{n-m}(x)) \tag{6.48}$$

$$\mathrm{Tr}\,(\frac{\partial}{\partial x}e^{\tau(x)}) = \sum_{n}\frac{1}{n!}\sum_{m=1}^{n}\mathrm{Tr}\,(\tau^{n-1}(x)\left(\frac{\partial\tau}{\partial x}\right)) \tag{6.49}$$

$$\mathrm{Tr}\,(\frac{\partial}{\partial x}e^{\tau(x)}) = \sum_{n}\frac{1}{n!}n\mathrm{Tr}\,(\tau^{n-1}(x)\left(\frac{\partial\tau}{\partial x}\right)) \tag{6.50}$$

$$\mathrm{Tr}\,(\frac{\partial}{\partial x}e^{\tau(x)}) = \mathrm{Tr}\,(\sum_{n}\frac{1}{(n-1)!}\tau^{n-1}(x)\left(\frac{\partial\tau}{\partial x}\right)) \tag{6.51}$$

which finally leads to

$$\mathrm{Tr}\,(\frac{\partial}{\partial x}e^{\tau(x)}) = \mathrm{Tr}\,(e^{\tau(x)}\left(\frac{\partial\tau}{\partial x}\right)) \tag{6.52}$$

or in other words

$$\mathrm{Tr}\,(\frac{\partial\rho}{\partial x}) = \mathrm{Tr}\,(\rho\left(\frac{\partial\log(\rho)}{\partial x}\right)) \tag{6.53}$$

and using $x = R_{nm}$:

$$\mathrm{Tr}\,(\rho\left(\frac{\partial\log(\rho)}{\partial R_{nm}}\right)) = \mathrm{Tr}\,(\frac{\partial\rho}{\partial R_{nm}}) = (\frac{\partial\mathrm{Tr}\,\rho}{\partial R_{nm}}) = \delta_{nm} \tag{6.54}$$

Our final result for the partial derivative is therefore

$$\left(\frac{\partial X}{\partial R_{nm}}\right) = -k_B\left\langle m\right|\log(\rho)\left|n\right\rangle + (\lambda-1)k_B\delta_{nm} \tag{6.55}$$

and setting this equal to zero gives

$$k_B \langle m| \log(\rho) |n\rangle = (\lambda - 1)k_B \delta_{nm} \tag{6.56}$$

which leads to the simple solutions

$$\rho = e^{\lambda - 1}\mathbb{E} \tag{6.57}$$

We still need to check that this is a Hermitian matrix with eigenvalues less than one. That is easy, we simply choose λ to be real and at most one.

How do we find the value of λ? By checking the condition Tr $\rho = 1$. This gives

$$e^{1-\lambda} = \text{Tr } \mathbb{E} \tag{6.58}$$

and the last trace is simply the dimension of the space, or the number of states with the given energy, the multiplicity function:

$$e^{1-\lambda} = g(U, V, N) \tag{6.59}$$

Since the multiplicity function is one or larger, λ is one or less. The entropy for this density matrix is

$$S = -k_B \text{Tr } e^{\lambda - 1}(\lambda - 1)\mathbb{E} = -k_B g^{-1}(U, V, N) \log(g^{-1}(U, V, N))\text{Tr } \mathbb{E} \tag{6.60}$$

or

$$S = k_B \log(g(U, V, N)) \tag{6.61}$$

which is exactly what we expect! Hence we retrieved our old definition, and the new procedure is the same as the old one.

The extremum of X we have found is a maximum. In order to test this, we have to calculate the second order derivatives. Hence we need

$$\left(\frac{\partial^2 X}{\partial \rho_{ij} \partial \rho_{nm}} \right) = -k_B \frac{\partial}{\partial \rho_{ij}} \left(\log(\rho) \right)_{mn} \tag{6.62}$$

In order to find the answer, we use again $\rho(x) = e^{\tau(x)}$. As before, we have

$$\frac{\partial}{\partial x} e^{\tau(x)} = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{m=0}^{n-1} \tau^m(x) \left(\frac{\partial \tau}{\partial x} \right) \tau^{n-m-1}(x) \tag{6.63}$$

In general, as we noted before, this cannot be simplified. Taking the trace of this expression allows us to make it easier. But here we need all values. But we only need to calculate the second order derivative at the extremum. Hence we set $\rho = e^{\lambda - 1}\mathbb{E}$, or $\tau = (\lambda - 1)\mathbb{E}$, and we have

$$\frac{\partial}{\partial x} e^{\tau(x)} = \sum_{n=1}^{\infty} \frac{1}{n!} \sum_{m=0}^{n-1} (\lambda - 1)^m(x) \left(\frac{\partial \tau}{\partial x} \right) (\lambda - 1)^{n-m-1}(x) \tag{6.64}$$

Now we can interchange factors and we find **at the extremum** that

$$\frac{\partial}{\partial x}e^{\tau(x)} = \sum_{n=1}^{\infty}\frac{1}{(n-1)!}(\lambda-1)^{n-1}\left(\frac{\partial\tau}{\partial x}\right) = e^{\lambda-1}\left(\frac{\partial\tau}{\partial x}\right) \qquad (6.65)$$

or

$$\left(\frac{\partial\rho}{\partial x}\right) = \rho\left(\frac{\partial\log(\rho)}{\partial x}\right) \qquad (6.66)$$

Using $x = \rho_{ij}$ we get **at the extremum**

$$\left(\frac{\partial\rho}{\partial\rho_{ij}}\right) = \rho\left(\frac{\partial\log(\rho)}{\partial\rho_{ij}}\right) \qquad (6.67)$$

and we can now take the element nm of this operator equation to get

$$\left(\frac{\partial\rho_{mn}}{\partial\rho_{ij}}\right) = \rho\left(\frac{\partial\log(\rho_{mn})}{\partial\rho_{ij}}\right) \qquad (6.68)$$

The left side is easy, and equal to $\delta_{mi}\delta_{nj}$. Hence **at the extremum** we have

$$\left(\frac{\partial^2 X}{\partial\rho_{ij}\partial\rho_{nm}}\right) = -k_B e^{1-\lambda}\delta_{mi}\delta_{nj} \qquad (6.69)$$

Suppose we make an arbitrary small change in the density matrix around the extremum. In second order we have

$$\Delta X = \sum_{ijmn}\left(\frac{\partial^2 X}{\partial\rho_{ij}\partial\rho_{nm}}\right)\Delta\rho_{ij}\Delta\rho_{nm} = -k_B e^{1-\lambda}\sum_{ijmn}\delta_{mi}\delta_{nj}\Delta\rho_{ij}\Delta\rho_{nm} \qquad (6.70)$$

or

$$\Delta X = -k_B e^{1-\lambda}\sum_{ij}\Delta\rho_{ij}\Delta\rho_{ji} \qquad (6.71)$$

Now we use the fact that the density matrix has to remain Hermitian, and hence $\Delta\rho_{ij}^* = \Delta\rho_{ji}$ to get

$$\Delta X = -k_B e^{1-\lambda}\sum_{ij}|\Delta\rho_{ij}|^2 \qquad (6.72)$$

which shows that the extremum is a maximum, indeed.

WHY BOTHER?

The whole approach sketched above seems very complicated, and for the simple microcanonical ensemble (states with given U, V, N) it is indeed. The advantage is that it is easier to generalize for more complex systems. That is where we gain in this formalism. The only thing we have shown above is that

for a closed system we get the same answers, which shows that the procedure is
OK.

HOW TO DEAL WITH DISCRETE EIGENVALUES.

So far we have described our state vectors in a Hilbert space with states
at a given energy, volume, and particle number. Since the Hamiltonian is the
operator that governs all quantum mechanics, volume and particle number are
ingredients of the Hamiltonian. But energy is the outcome! Therefore, it is
much better to work in the more natural Hilbert space of state vectors that
correspond to the Hamiltonian. The density operator in this case is

$$\rho = C\delta(U - H) \tag{6.73}$$

where C is some constant, determined by $1 = C\mathrm{Tr}\ \delta(U - H)$.

This form of the density operator in the microcanonical ensemble implicitly
assumes that we take the thermodynamic limit. For a finite system the energy
eigenvalues are discrete and the δ -function will yield a number of infinitely high
peaks. Therefore we are not able to take the derivative $\left(\frac{\partial S}{\partial U}\right)$ which is needed
to define the temperature!

We can, however, rewrite the delta function. Using a standard property we
find:

$$\rho = \frac{C}{V}\delta(\frac{U}{V} - \frac{H}{V}) \tag{6.74}$$

and in the thermodynamic limit the spacing in the energy density $\frac{U}{V}$ goes to
zero. But that is the wrong order of limits, which can lead to problems!

MATHEMATICAL DETAILS ABOUT DELTA FUNCTION LIMITS.

In order to be more precise mathematically, we need a limiting procedure
for the delta-function. For example, one can choose a Gaussian approximation

$$\rho = C\frac{1}{\epsilon\sqrt{\pi}}e^{-\frac{(U-H)^2}{\epsilon^2}} \tag{6.75}$$

which gives the correct result in the limit $\epsilon \to 0$. In order to evaluate the trace
of $\rho log \rho$ we now have to sum over states with all possible values of the energy.
Clearly we will get a value for the entropy which is defined for all values of U,
and the partial derivative of the entropy with respect to U will give a good value
for the inverse temperature. But now we have two limits to deal with: $N \to \infty$
and $\epsilon \to 0$. The thermodynamic limit $N \to \infty$ has to be taken after all physical
state variables have been set equal to the values needed in the problem. For
example, if needed the limit $T \to 0$ has to be taken first. The limit $\epsilon \to 0$
is not related to a physical state variable. This limiting procedure is a purely
mathematical trick. Therefore, the order of taking limits is first to assign all

intensive state variables their correct value, then to take the thermodynamic limit $N \to \infty$, and finally consider the results as a function of ϵ and take the limit $\epsilon \to 0$. In practice, one should make N large and choose ϵ as small as possible, but large compared to the average spacing of the energy levels. We will discuss this question by looking at a simple example.

A quantum system in one dimension has single-particle eigenstates with energy $n\omega$, with $n = 1, 2, 3, \cdots$. The energy of N particles is given by $(\sum_i^N n_i)\omega$. Hence the entropy is equal to

$$S(U,N) = -k_B \sum_{n_1,n_2,.,n_N} C \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-(\sum_i^N n_i)\omega)^2}{\epsilon^2}} \log\left(C \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-(\sum_i^N n_i)\omega)^2}{\epsilon^2}}\right)$$

(6.76)

We assume that the particles are not identical. Next we set $x_i = n_i \frac{\omega}{N\epsilon}$ and assume that $\epsilon \gg \frac{\omega}{N}$, in agreement with the statement that we can only take the limit $\epsilon \to 0$ after taking the limit $N \to \infty$. The quantity $\frac{\omega}{N}$ is the spacing between the possible energy values *per particle* and we have to choose ϵ much larger than this value. The summations are replaced by integrations and we have

$$S(U,N) = -k_B \left(\frac{N\epsilon}{\omega}\right)^N \int \cdots \int dx_1 \cdots dx_N$$

$$C \frac{1}{\epsilon\sqrt{\pi}} e^{-(\frac{U}{\epsilon} - Nx_1 - \cdots - Nx_N)^2} \left[\log\left(C \frac{1}{\epsilon\sqrt{\pi}}\right) - (\frac{U}{\epsilon} - Nx_1 - \cdots - Nx_N)^2\right]$$

(6.77)

Now we define $x = \sum_i x_i$ and replace the integration variables by $x, x_2, \cdots, x_N$. Since all the variables x_i have to be positive, we know that for a given value of x all other variables are limited by $\sum_{i=2}^N x_i \leqslant x$. Since the integrand only depends on the value of x, the integration over $x_2, \cdots, x_N$ is just a geometrical factor $g(x)$. Therefore, the entropy is equal to

$$S(U,N) = -k_B \left(\frac{\epsilon N}{\omega}\right)^N \int dx g(x)$$

$$C \frac{1}{\epsilon\sqrt{\pi}} e^{-(\frac{U}{\epsilon} - Nx)^2} \left[\log\left(C \frac{1}{\epsilon\sqrt{\pi}}\right) - (\frac{U}{\epsilon} - Nx)^2\right]$$

(6.78)

The coefficient C follows from $\mathrm{Tr}\, \rho = 1$:

$$C^{-1} = \mathrm{Tr}\, \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-H)^2}{\epsilon^2}}$$

(6.79)

or

$$C^{-1} = \sum_{n_1,n_2,.,n_N} \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-(\sum_i^N n_i)\omega)^2}{\epsilon^2}}$$

(6.80)

Replacing the sum by an integral yields:

$$C^{-1} = \left(\frac{N\epsilon}{\omega}\right)^N \int \cdots \int dx_1 \cdots dx_N \frac{1}{\epsilon\sqrt{\pi}} e^{-(\frac{U}{\epsilon} - Nx_1 - \cdots - Nx_N)^2} \tag{6.81}$$

and with $x = \sum_i x_i$ we get

$$C^{-1} = \left(\frac{N\epsilon}{\omega}\right)^N \int dx \; g(x) \; \frac{1}{\epsilon\sqrt{\pi}} e^{-(\frac{U}{\epsilon} - Nx)^2} \tag{6.82}$$

This reduces the entropy to

$$S(U, N) = -k_B \frac{\int dx \; g(x) \; e^{-(\frac{U}{\epsilon} - Nx)^2} \left[\log\left(C\frac{1}{\epsilon\sqrt{\pi}}\right) - (\frac{U}{\epsilon} - Nx)^2\right]}{\int dx \; g(x) \; e^{-(\frac{U}{\epsilon} - Nx)^2}} \tag{6.83}$$

In order to prepare for taking the thermodynamic limit we now express the entropy as a function of the energy per particle u and the number of particles. The thermodynamic limit is taken in such a way that the energy per particle is constant. This gives for the entropy per particle, $s = \frac{S}{N}$:

$$s(u, N) = -k_B \frac{\int dx \; g(x) \; e^{-N^2(\frac{u}{\epsilon} - x)^2} \left[\frac{1}{N}\log\left(C\frac{1}{\epsilon\sqrt{\pi}}\right) - N(\frac{u}{\epsilon} - x)^2\right]}{\int dx \; g(x) \; e^{-N^2(\frac{u}{\epsilon} - x)^2}} \tag{6.84}$$

This expression can be split into two parts, by separating the two parts of the integral in the numerator, and we have $s(u, N) = s_1(u, N) + s_2(u, N)$ with

$$s_1(u, N) = -k_B \frac{1}{N} \log\left(C\frac{1}{\epsilon\sqrt{\pi}}\right) \tag{6.85}$$

and

$$s_2(u, N) = N k_B \frac{\int dx \; g(x) \; e^{-N^2(\frac{u}{\epsilon} - x)^2} \left[(\frac{u}{\epsilon} - x)^2\right]}{\int dx \; g(x) \; e^{-N^2(\frac{u}{\epsilon} - x)^2}} \tag{6.86}$$

The second term can be written in the form:

$$s_2(u, N) = \frac{1}{2} k_B \frac{\partial}{\partial N} \log\left(\int dx \; g(x) \; e^{-N^2(\frac{u}{\epsilon} - x)^2}\right) \tag{6.87}$$

In the thermodynamic limit the exponent is sharply peaked and the integral is proportional to $g(\frac{u}{\epsilon})\frac{1}{N}$. Hence $s_2(u, N)$ is proportional to $\frac{\partial}{\partial N}\log(N)$ or $\frac{1}{N}$, and goes to zero in the thermodynamic limit.

Using (6.82) for C we find

$$s(u, N) = k_B \frac{1}{N} \log\left(\left(\frac{N\epsilon}{\omega}\right)^N \int dx \; g(x) \; e^{-N^2(\frac{u}{\epsilon} - x)^2}\right) \tag{6.88}$$

For large values of N we can replace the exponent by $\frac{\sqrt{\pi}}{N}\delta(\frac{u}{\epsilon} - x)$ and we get

$$s(u, N) = k_B \frac{1}{N} \log \left(\left(\frac{N\epsilon}{\omega} \right)^N \int dx \; g(x) \; \frac{\sqrt{\pi}}{N} \delta(\frac{u}{\epsilon} - x) \right) \tag{6.89}$$

or

$$s(u, N) = k_B \frac{1}{N} \log \left(\left(\frac{N\epsilon}{\omega} \right)^N g(\frac{u}{\epsilon}) \frac{\sqrt{\pi}}{N} \right) \tag{6.90}$$

The geometrical factor g(x) is easily calculated

$$g(x) = \int_0^x dx_2 \int_0^{x-x_2} dx_3 \cdots \int_0^{x-\cdots-x_{N-1}} dx_N = \frac{1}{(N-1)!} x^{N-1} \tag{6.91}$$

Hence the entropy is

$$s(u, N) = k_B \frac{1}{N} \log \left(\left(\frac{N\epsilon}{\omega} \right)^N \frac{1}{(N-1)!} \left(\frac{u}{\epsilon} \right)^{N-1} \frac{\sqrt{\pi}}{N} \right) \tag{6.92}$$

$$s(u, N) = k_B \frac{1}{N} \log \left(\left(\frac{N}{\omega} \right)^N \frac{1}{N!} u^N \epsilon \sqrt{\pi} \frac{1}{u} \right) \tag{6.93}$$

We can now ignore the factor $\epsilon\sqrt{\pi}\frac{1}{u}$ in the logarithm, since that leads to a term proportional to $\frac{1}{N}$, which disappears in the thermodynamic limit.

For large values of N the value of $N!$ is about $N^N e^{-N}$ and hence we find

$$s(u, N) = k_B \frac{1}{N} \log \left(\omega^{-N} e^N u^N \right) = k_B \left(\log(\frac{u}{\omega}) + 1 \right) \tag{6.94}$$

The first conclusion for our model system is that the entropy is equal to $Nk_B \log(\frac{U}{N\omega})$ in the thermodynamic limit. Hence the temperature and energy are related by $U = Nk_BT$. The second conclusion is that in the thermodynamic limit the value of S does not depend on ϵ. Hence we can now take the limit of $\epsilon \to 0$ very easily, and use delta functions. This is what we tried to show. For a finite system the use of delta functions is not justified, such a procedure only makes sense in the thermodynamic limit.

6.4   Equivalence of entropy definitions for canonical ensemble.

In the previous section we used the information-theoretical definition of the entropy to discuss the microcanonical ensemble. The resulting density matrix was exactly what we expected from a discussion in chapter 2. Only those states with energy U are available, and all have the same probability. Since the canonical and grand-canonical ensemble are based on the microcanonical ensemble, the

results for the entropy and density matrix should also be the same in this case, independent of which definition of the entropy we use. In this section we illustrate this equivalence directly. This direct method also shows how to use the maximum entropy principle.

In the canonical ensemble we specify only the average energy U of the system, the volume V and the number of particles N. If other extensive state variables are known, they are also included in this list. The trace in the expressions for the entropy is taken over all states with volume V, number of particles N, and other extensive state variables. These variables are not allowed to fluctuate. The energy is allowed to vary, however, and we sum over states with all possible energies. The additional requirement we have to impose is $U = \mathrm{Tr}\ H\rho$ for the average energy. This requirement is taken into account via a *Lagrange multiplier* β, and we have to maximize

$$X = -k_B \mathrm{Tr}\ \rho \log \rho + \lambda k_B\ (\mathrm{Tr}\ \rho - 1) - \beta k_B\ (\mathrm{Tr}\ \rho H - U) \qquad (6.95)$$

The choice of sign in front of β is conventional. Clearly we see that $\left(\frac{\partial X}{\partial \lambda}\right) = 0 \Rightarrow \mathrm{Tr}\ \rho = 1$ and $\left(\frac{\partial X}{\partial \beta}\right) = 0 \Rightarrow \mathrm{Tr}\ \rho H = U$. Maximization with respect to ρ leads to

$$0 = \left(\frac{\partial X}{\partial \rho_{nm}}\right) = k_B \lambda \delta_{nm} - k_B (\log \rho)_{mn} - k_B \delta_{nm} - \beta k_B H_{mn} \qquad (6.96)$$

where we have used the results derived in the previous section. The addition of one term did not make the derivative much more complicated. This is valid in general. Most constraints we want to add are linear, and give very easy terms in the Lagrange equation!

The minimization leads to

$$k_B (\log \rho)_{mn} = k_B (\lambda - 1)\delta_{nm} - \beta k_B H_{mn} \qquad (6.97)$$

and hence

$$\rho = e^{\lambda - 1} e^{-\beta H} = C e^{-\beta H} \qquad (6.98)$$

The normalization constant C (or λ) follows again from the requirement that $\mathrm{Tr}\ \rho = 1$. The constant β is obtained by demanding $\mathrm{Tr}\ \rho H = U$. This condition can be compared with thermodynamical expressions via (6.97). If we multiply (6.97) by ρ_{nm} and sum over n and m, we obtain

$$\mathrm{Tr}\ \rho \log \rho = (\lambda - 1)\mathrm{Tr}\ \rho - \beta \mathrm{Tr}\ \rho H \qquad (6.99)$$

and using the definition of the entropy and the constraints on the density matrix:

$$S = -k_B(\lambda - 1) + k_B \beta U \qquad (6.100)$$

The constant $\lambda - 1$ is equal to $\log C$. Since $\mathrm{Tr}\ \rho = 1$ we have $C^{-1} = \mathrm{Tr}\ e^{-\beta H}$. The last expression is the partition function $\mathcal{Z}$ at a temperature $\frac{1}{\beta k_B}$. The

logarithm of the partition function is related to the Helmholtz free energy, $F(T) = -k_B T \log \mathcal{Z}$. Therefore $\lambda - 1 = \frac{F(T)}{k_B T}$ at a temperature $T = \frac{1}{\beta k_B}$. In other words we find

$$TS = -k_B T \frac{F(T)}{k_B T} + k_B T \beta U \qquad (6.101)$$

which is the correct definition of the Helmholtz free energy in thermodynamics only if we require that the Lagrange multiplier β is equal to $\frac{1}{k_B T}$. The temperature is in this formalism related to a Lagrange multiplier which constraints the average energy! This is another way of defining the temperature, and we have just shown that this definition is equivalent to the standard thermodynamical definition, since it leads to the same expression for the Helmholtz free energy.

In the grand canonical ensemble we allow N to vary and take the trace over all states with all possible values of the energy and number of particles. The expression to maximize in this case is

$$\mathbb{X} = -k_B \text{Tr } \rho \log \rho + \lambda k_B \left(\text{Tr } \rho - 1 \right) -$$

$$\beta k_B \left(\text{Tr } \rho \mathbb{H} - U \right) + \beta \mu k_B \left(\text{Tr } \rho \mathbb{N} - N \right) \qquad (6.102)$$

The maximum occurs when

$$\rho = C e^{-\beta(\mathbb{H} - \mu \mathbb{N})} \qquad (6.103)$$

The values of β and C are determined as before, but μ follows from the requirement that the average number of particles is equal to N. One can show that μ must be equal to the standard definition of the chemical potential in the same way as we showed that β is equal to $\frac{1}{k_B T}$.

6.5 Problems for chapter 6

PROBLEM 1.

Show that the expression for the density matrix $\rho = C e^{-\beta(\mathbb{H} - \mu \mathbb{N})}$ corresponds to an **extremum** of (6.102). By analyzing the grand potential Ω show that μ is the chemical potential.

PROBLEM 2.

Show that the solution $\rho = C e^{-\beta H}$ for the density matrix for the canonical ensemble, obtained by finding the extremum of (6.95), corresponds to a **maximum** of X.

PROBLEM 3.

A system is described by a state in a two-dimensional Hilbert space. The Hamiltonian is given by

$$H = \begin{pmatrix} \epsilon & \alpha^* \\ \alpha & 2\epsilon \end{pmatrix}$$

Assume that $\epsilon \gg |\alpha|$ and that $\beta\epsilon \ll 1$.

(A) Calculate the partition function up to second order in β.

(B) Calculate the high temperature limit of the heat capacity.

(C) Suppose that $\alpha = \frac{N}{V}$ and that the density is small. Calculate the pressure in the large volume and high temperature limit.

PROBLEM 4.

A quantum mechanical system is described by a Hamiltonian $H = H_0 + \kappa V$, with $[H_0, V] = 0$. κ is a small constant. The Helmholtz free energy is $F_\kappa(T)$. Calculate the change in Helmholtz free energy, $\Delta F = F_\kappa - F_0$ for this system up to second order in $\frac{\kappa}{k_B T}$.

PROBLEM 5.

In a two-dimensional Hilbert space the density operator is given by its matrix elements:

$$\rho = \begin{pmatrix} x & R \\ R^* & 1 - x \end{pmatrix}$$

This form is clearly Hermitian and has trace one. Calculate the entropy as a function of x and R, and find the values of x and R that make the entropy maximal. Note that you still need to check the condition that the matrix is positive! Also, show that it is a maximum!

PROBLEM 6.

A quantum mechanical system is described by a simple Hamiltonian H, which obeys $H^2 = 1$. Evaluate the partition function for this system. Calculate the internal energy for $T \to 0$ and $T \to \infty$.

PROBLEM 7.

A system is described by a density operator ρ. In this problem, the eigenvalues of this operator are either 0 or 1. The number of particles in the system is N and the volume of the system is V.

(A) How many eigenvalues of ρ are equal to 1?

(B) What is the entropy of the system?

(C) Assume that the particles are independent. If a single particle is described by a density operator ρ_i, how do you construct ρ from the ρ_i?

Chapter 7

Classical statistical mechanics.

7.1 Relation between quantum and classical mechanics.

CHOICE OF BASIS.

Based on the knowledge of quantum statistical mechanics, it is straightforward to derive the expressions valid for a classical description of a many particle system. A classical state of a system is defined by the values of the positions $\vec{r}_i$ and momenta $\vec{p}_i$. These vectors are combined in a single vector $\vec{X}$ in a 6N-dimensional phase-space. Every value of $\vec{X}$ corresponds to a certain state of the system. The classical Hamiltonian $H(\vec{X})$ is constructed from the quantum mechanical one by replacing momentum operators by numbers $\vec{p}_i$. A good basis for the expressions for the density matrix is therefore the set of all vectors $|\vec{X}>$.

In quantum mechanics it is impossible to specify both the position and the momentum of a particle at the same time, since the position and momentum operators do not commute. The commutators are proportional to $\hbar$ and one way of thinking about the classical limit of quantum mechanics is to assume that $\hbar \to 0$. In this sense the position and momentum operators do commute in the classical limit and we are allowed to use their simultaneous eigenvalues as labels characterizing the state of a system.

WIGNER DISTRIBUTION FUNCTION.

In a more formal approach we use so-called Wigner distribution functions, defined by

$$W(\vec{k}, \vec{r}) = \int d^3x e^{i\vec{k}\vec{x}} < \vec{r} + \frac{1}{2}\vec{x}|\rho|\vec{r} - \frac{1}{2}\vec{x} > \qquad (7.1)$$

for a system with only one particle. In the formula above ρ is the density matrix describing this system. This formula represents a partial Fourier transform. If we want to transform the complete density matrix we need two exponentials and two three dimensional integrations, because we need to transform the information contained via $\vec{r}$ and $\vec{r'}$ in $< \vec{r'}|\rho|\vec{r} >$ both. But now we only transform half of the space, which still leaves us with a function of two three dimensional vectors.

It is easy to show that

$$\frac{1}{(2\pi)^3} \int d^3k W(\vec{k}, \vec{r}) = < \vec{r}|\rho|\vec{r} > \qquad (7.2)$$

and

$$\frac{1}{(2\pi)^3} \int d^3r W(\vec{k}, \vec{r}) = < \vec{k}|\rho|\vec{k} > \qquad (7.3)$$

These are the probability functions for finding the particle at position $\vec{r}$ or with momentum $\vec{k}$. In order to derive the last expression we have used

$$|\vec{x} > = (2\pi)^{-\frac{3}{2}} \int d^3k e^{i\vec{k}\vec{x}}|\vec{k} > \qquad (7.4)$$

to get

$$W(\vec{k}, \vec{r}) = \int d^3x e^{i\vec{k}\vec{x}} \frac{1}{(2\pi)^3} \int d^3k' \int d^3k" e^{i[\vec{k}"(\vec{r}-\frac{1}{2}\vec{x})-\vec{k}'(\vec{r}+\frac{1}{2}\vec{x})]} < \vec{k}'|\rho|\vec{k}" > =$$

$$\int d^3k' \int d^3k" \delta(\vec{k} - \frac{1}{2}[\vec{k}' + \vec{k}"])e^{i[\vec{k}"\vec{r}-\vec{k}'\vec{r}]} < \vec{k}'|\rho|\vec{k}" > \qquad (7.5)$$

Integration over $\vec{r}$ gives an extra delta-function $\delta(\vec{k}' - \vec{k}")$ and the stated result follows.

AVERAGES.

Suppose $O(\vec{k}, \vec{r})$ is a classical operator in phase space and if $\mathcal{O}$ is the quantum-mechanical generalization of this classical operator. In a number of cases it is possible to show that the quantum-mechanical expectation value of the operator $< \mathcal{O} >$ is given by

$$< \mathcal{O} > = \int d^3k \int d^3x O(\vec{k}, \vec{x})W(\vec{k}, \vec{x}) \qquad (7.6)$$

which relates classical values and quantum mechanical values of O and $\mathcal{O}$. This procedure is not always possible, though, and our discussion only serves the

purpose to illustrate that a connection between quantum mechanical density operators and classical distribution functions can be made. In general, the expressions one needs for the distribution function W are more complicated than the form given by Wigner.

Classical density matrix.

In order to give the formulas for classical statistical mechanics we consider the canonical ensemble first. Since the Hamiltonian in the classical limit is diagonal in $\vec{X}$, we find that the density matrix is diagonal too, and that

$$< \vec{X}|\rho|\vec{X} >= \rho(\vec{X}) = Ce^{-\beta H(\vec{X})} \tag{7.7}$$

In many ways this form of ρ can be directly compared with the function W introduced in the previous paragraph.

Volume per state.

The normalization constant C follows again from the requirement that $\mathrm{Tr}\rho = 1$. This trace implies a sum over all possible states. In phase space, this is replaced by an integral. We do have to multiply this integral by a volume factor, though. The magnitude of this volume factor follows from the old quantum mechanical quantization condition

$$\oint pdq = nh \tag{7.8}$$

where p is the generalized momentum corresponding to the generalized coordinate q. This contour integral is an integral multiple of Planck's constant h. In two-dimensional phase space $[p, q]$ the area inside a quantum orbit is nh, hence the area between two quantum orbits is h itself. In other words, the area per orbit is h. This value is consistent with the Heisenberg uncertainty relation. If we construct a wave-package from plane-waves, we find that the position is specified within Δx and the corresponding momentum coordinate within Δp. Also, $\Delta x \Delta p > \alpha \hbar$ for a value of α which is about one. If we represent the states of a system by a grid of points in the p-x plane the distance between the points in the x-direction has to be about Δx. If grid points are closer together than Δx they could not be distinguished and if they are further apart they do not cover all space. Similarly, the distance of the grid points in p-space should be Δp and hence the volume per point in one dimension should be $\Delta x \Delta p$, or about $\alpha \hbar$. The formula above shows that we have to take $\alpha = 2\pi$.

Classical integrals.

If the N particles in our system are **not identical**, we therefore have the rule

$$\sum_{\vec{X}} \Rightarrow \frac{1}{h^{3N}} \int d\vec{X} \qquad (7.9)$$

If the particles are **identical**, however, any permutation of the coordinates gives the same state, and we have to replace the sum by

$$\sum_{\vec{X}} \Rightarrow \frac{1}{N!h^{3N}} \int d\vec{X} \qquad (7.10)$$

As a result, the classical partition function is equal to

$$\mathcal{Z} = \frac{1}{N!h^{3N}} \int d\vec{X} e^{-\beta H(\vec{X})} \qquad (7.11)$$

and the density matrix is

$$\rho(\vec{X}) = \frac{1}{\mathcal{Z}} e^{-\beta H(\vec{X})} \qquad (7.12)$$

7.2 Classical formulation of statistical mechanical properties.

ENTROPY.

The entropy follows from the information-theoretical definition 6.29 and we find

$$S(T, V, N) = -\frac{k_B}{\mathcal{Z}N!h^{3N}} \int d\vec{X} e^{-\beta H(\vec{X})} \log\left(\frac{1}{\mathcal{Z}} e^{-\beta H(\vec{X})}\right) =$$

$$\frac{k_B}{\mathcal{Z}N!h^{3N}} \int d\vec{X} e^{-\beta H(\vec{X})} \left(\log(\mathcal{Z}) + \beta H(\vec{X})\right) = k_B \log \mathcal{Z} + k_B \beta U \qquad (7.13)$$

Note that the volume V is included in this formula, it simply limits the integrations over the coordinates $\vec{x}_i$. As expected, our equation leads to $F = U - TS$ and our definitions are equivalent to the old definitions of the entropy. This was already shown explicitly in the previous chapter for the canonical ensemble.

DENSITY OF STATES.

The micro-canonical ensemble can be treated in a similar way. In this case the density matrix is equal to

$$\rho(\vec{X}) = C\delta(U - H(\vec{X})) \qquad (7.14)$$

where the normalization constant C follows from $\text{Tr}\rho = 1$, or

$$C^{-1} = \mathcal{D}(U, V, N) = \frac{1}{N!h^{3N}} \int d\vec{X} \delta(U - H(\vec{X})) \qquad (7.15)$$

Note that in classical mechanics there are no problems with discrete levels, the set of possible values for $H(\vec{X})$ is continuous. If the particles are not identical, the factor N! should be omitted. The function $\mathcal{D}(U)$ is called the density of states. This density of states also relates the micro-canonical and the canonical ensemble via

$$\mathcal{Z} = \frac{1}{N!h^{3N}} \int d\vec{X} e^{-\beta H(\vec{X})} = \frac{1}{N!h^{3N}} \int d\vec{X} \int du \delta(u - H(\vec{X})) e^{-\beta u} \qquad (7.16)$$

which leads to

$$\mathcal{Z}(T, V, N) = \int du \mathcal{D}(u, V, N) e^{-\beta u} \qquad (7.17)$$

The partition function is the Laplace transform of the density of states. In many cases it is easier to evaluate $\mathcal{Z}$ and then this relation can be inverted to obtain the density of states $\mathcal{D}$. One important technical detail is that the energy integral is bounded at the lower end. There is a ground state energy, and a minimum energy level. We can always shift our energy scale so that the minimum energy is zero. In that case we can write

$$\mathcal{D}(U, V, N) = \frac{1}{2\pi\imath} \int_{c-\imath\infty}^{c+\imath\infty} dt e^{st} \mathcal{Z}(t, V, N) \qquad (7.18)$$

GRAND PARTITION FUNCTION.

The grand canonical ensemble follows from the density matrix defined in the much larger space of states with all possible numbers of particles. The grand partition function now involves a sum over all possible combinations of particles and is equal to

$$\mathfrak{Z}(T, \mu, V) = \sum_N \frac{1}{N!h^{3N}} \int d\vec{X} e^{-\beta(H(\vec{X}, N) - \mu N)} \qquad (7.19)$$

This leads to

$$\mathfrak{Z}(T, \mu, V) = \sum_N e^{\frac{\mu N}{k_B T}} \mathcal{Z}(T, N, V) \qquad (7.20)$$

which is exactly the same relation as we derived in before. This shows that our classical formalism is consistent with the quantum statistical approach. In deriving classical statistical mechanics from its quantum analogue, the factors N! and h^{3N} come in naturally. It is also possible to postulate classical statistical mechanics directly, but then these factors have to be introduced in some ad-hoc fashion.

7.3 Ergodic theorem.

CONNECTION WITH EXPERIMENT.

Most experiments involve systems that involve many degrees of freedom. In all those cases we are only interested in average quantities. All average quantities follow from the quantum mechanical formula, and can be expressed as classical averages if we change sums to integrals:

$$< A > = \frac{1}{N! h^{3N}} \int d\vec{X} \rho(\vec{X}) A(\vec{X}) \tag{7.21}$$

The factor $N!$ is needed for indistinguishable particles, and should be omitted for distinguishable particles.

The previous formula does, of course, not describe the way we measure quantities! In a measurement we often consider time averaged quantities or space averaged quantities. In the first case, we start the system at an initial state $\vec{X}_{in}$ at $t = 0$ and use Hamilton's equations $\left(\frac{\partial \vec{p}_i}{\partial t}\right) = -\left(\frac{\partial H}{\partial \vec{x}_i}\right)$ and $\left(\frac{\partial \vec{x}_i}{\partial t}\right) = +\left(\frac{\partial H}{\partial \vec{p}_i}\right)$ to calculate $\vec{X}(t)$. We then evaluate

$$< A >_\tau (t) = \frac{1}{\tau} \int_t^{t+\tau} A(\vec{X}(t')) dt' \tag{7.22}$$

where $\vec{X}(t)$ is the state of the system at time t. The value of this average depends on the measurement time τ and on the initial state. Birkhoff showed that in general $\lim_{\tau \to \infty} < A >_\tau (t)$ is independent of the initial conditions. So if we take τ large enough it does not matter where we start. Note that in cases where we have external potentials that vary with time we cannot make τ too large, it has to be small compared to the time scale of the external potential. In different words, the effects of external potentials that vary too rapidly are not described by thermodynamic equilibrium processes anymore.

EQUIVALENCY OF AVERAGES.

Now we ask the more important question. Under what conditions is this measured quantity equal to the thermodynamical ensemble average? In the quantum-mechanical case we argued that all accessible quantum states are equally probable if the total energy U of the system is fixed. We needed to invoke some random external small perturbation to allow the system to switch quantum states. In the classical case we need a similar argument.

The time evolution of a classical system is determined by Hamilton's equations:

$$\frac{\partial \vec{p}_i}{\partial t} = -\frac{\partial H}{\partial \vec{x}_i} \quad , \quad \frac{\partial \vec{x}_i}{\partial t} = \frac{\partial H}{\partial \vec{p}_i} \tag{7.23}$$

The trajectory $\vec{X}(t)$ is well-described. For an isolated system the energy is conserved, and $\vec{X}(t)$ is confined on the surface in phase space given by $H(\vec{X}) = U$. In the micro-canonical ensemble $\rho(\vec{X}) = C\delta(H(\vec{X}) - U)$ and **all** states with $H(\vec{X}) = U$ are included in the average. In the time average, only a small subset of these states is sampled. The requirement that both averages are the same leads to the requirement that in the time τ, which has to be short compared with the time scale of changes in $< A > (t)$, the trajectory $\vec{X}(t)$ in phase space samples all parts of the accessible phase space equally well. If this is true, and the time-average is equal to the ensemble average, the system is called **ergodic**. Most systems with many degrees of freedom seem to obey this ergodic hypothesis.

If a system is ergodic the orbit $\vec{X}(t)$ should approach any point in the accessible region in phase space arbitrarily close. Originally it was thought that this statement could be related to Poincaré's theorem. This theorem states that *a system having a finite energy and a finite volume will, after a sufficiently long time, return to an arbitrary small neighborhood of almost any given initial state.* The time involved, however, is prohibitively long. It is on the order of e^N, where N is the number of degrees of freedom.

CHAOS.

A more useful approach combines the classical trajectories with small external perturbations, like in the quantum-mechanical treatment. An orbit $\vec{X}(t)$ is called **chaotic** if an arbitrary small change in the initial conditions in general leads to a very large change in the state of the system at a later time t. Classical statistical mechanics therefore requires that most of the orbits of the system are chaotic. The time τ still has to be large in order to obtain a good sampling of phase space. For meta-stable systems like glass the time τ can be longer than a human life-span. This means that the ergodic theorem in itself is not sufficient for the time-average and ensemble-average to be the same in a real experiment.

If we want to measure properties of a system, we have to connect the system to some outside instrument. If we want to specify any intensive state variable of a system, we have to connect it to a reservoir. In all these cases there will be small perturbations on the system and the description given in the previous paragraph is valid. What happens in a truly isolated system though? The answer to this question is impossible to derive from experiments, since any experiment involves a perturbation. Hence questions pertaining to a statistical mechanical description of a truly isolated system in terms of ensembles are purely academic. Nevertheless scientists have tried to answer this question and two opposing points of view have been defended. Landau claims that it is not possible to give such a description and that you need outside perturbations, while Khinchin tells us that it is possible to give such a description.

7.4   What is chaos?

The nature of chaos is best understood by considering some simple examples. First, we investigate a one-dimensional harmonic oscillator. The Hamiltonian is given by

$$H(x,p) = \frac{1}{2}(p^2 + x^2) \qquad (7.24)$$

with solutions

$$x(t) = x_0 \cos(t) + p_0 \sin(t) \quad , \quad p(t) = -x_0 \sin(t) + p_0 \cos(t) \qquad (7.25)$$

Phase space is two-dimensional, and the surface of states with energy U is a circle with radius $\sqrt{2U}$. Each orbit $(x(t), p(t))$ samples all states on this available surface. These orbits are not chaotic, since a small change in the initial conditions gives only a small change in the state at arbitrary times t. Also, the time needed to sample all of the available states is the period of the oscillator, and this is a long time for the system. Any average over a complete period essentially throws away all information about details of the system! In a case like this one would like to measure over a much shorter time. Hence a one-dimensional harmonic oscillator is an exceptional case, *it is too simple!*

A two-dimensional harmonic oscillator is described by the Hamiltonian

$$H(x, y, p_x, p_y) = \frac{1}{2}(p_x^2 + p_y^2 + x^2 + y^2) \qquad (7.26)$$

and the solutions for the orbits are

$$x(t) = x_0 \cos(t) + p_{x0} \sin(t) \quad , \quad y(t) = y_0 \cos(t) + p_{y0} \sin(t)$$

$$p_x(t) = -x_0 \sin(t) + p_{x0} \cos(t) \quad , \quad p_y(t) = -x_0 \sin(t) + p_{y0} \cos(t) \qquad (7.27)$$

In this case phase space is four-dimensional, and the surface of constant energy is a three-dimensional hyper-sphere.

In order to present results for systems with many degrees of freedom we often use a simple representation. An arbitrary orbit is represented by its cross-sections with a number of two-dimensional cuts through the constant energy

surface. For the two-dimensional oscillator at a given energy U only three vari-
ables are independent. Suppose that y is the dependent variable. In order to
get a two-dimensional representation of the orbit, we have to eliminate one ad-
ditional variable. We will choose x and consider all states of the system with
$x = 0$. Therefore we work in the plane (p_x, p_y). The energy U is equal to
$\frac{1}{2}(p_x^2 + p_y^2) + \frac{1}{2}y^2$ and hence the available part of phase space in this plane is
given by

$$p_x^2 + p_y^2 \leqslant 2U \qquad (7.28)$$

A chaotic orbit would sample all values of the momentum within this circle.
The trajectories of the harmonic oscillator are very regular, however. If we
assume that $x_0 = 1$ and $p_{x0} = 0$, the condition $x = 0$ occurs at time $t = 2\pi n$
for $n = 0, 1, 2, \cdots$ At these times $p_x = 0$ and $p_y = \pm p_{y0}$; hence only two points
are sampled!

Harmonic systems are too simple. There are three constants of motion.
$J_x = x^2 + p_x^2$ and $J_y = y^2 + p_y^2$ are conserved, and the phase angle between the
x and y part of the orbit is also constant.

LARGE SYSTEMS.

This conclusion can be generalized to a system of N three-dimensional har-
monic oscillators. All orbits are non-chaotic, and such a system cannot be
described by statistical mechanical methods! If N is very large, however, and
the initial conditions of the oscillators are random, such a system does give the
correct statistical mechanical results if we measure variables as spatial averages!
Randomness has to be brought in, however. Either the initial states are random,
or each oscillator is perturbed randomly and continuously.

This trick only works when the oscillators are identical, and we essentially
map all coordinates into a six-dimensional phase space for a single harmonic
oscillator. Measurements which go beyond this simple projection of phase space
still have to be discussed carefully. Also, there are problems with correlations.
If there are no random perturbations the relative motion of any two oscillators is
completely determined by the initial conditions and hence cannot be described
by a statistical mechanical ensemble average.

ADDITIONAL INTERACTIONS.

The harmonic oscillator is easy since it only contains second order terms.
One has to include higher order terms to see interesting effects. A standard
example is the Hénon-Heiles potential

$$V(x, y) = \frac{1}{2}x^2 + \frac{1}{2}y^2 + x^2 y - \frac{1}{3}y^3 \qquad (7.29)$$

Near the origin particles are still bound, because the third order terms are small.
In the y direction, however, the potential goes to minus infinity, and a quantum

particle would always tunnel out. A classical particle will remain bound, as long as its energy is smaller than the minimum barrier. The y direction gives us this minimum, and for $x = 0$ we have $V(0, y) = \frac{1}{2}y^2 - \frac{1}{3}y^3$, which has a maximum at $y = 1$ with value $\frac{1}{6}$. Hence we have to restrict the energy U to values less then one-sixth and larger than zero. The values for x and y corresponding to $V(x, y) = \frac{1}{6}$ are given by

$$(2y + 1)(3x^2 - (y - 1)^2) = 0 \tag{7.30}$$

which corresponds to three straight lines. A contour map of V(x,y) is not hard to sketch.
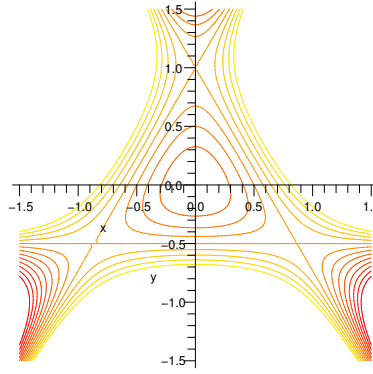

Figure 7.1: **Hénon-Heiles potential.**

The part of phase space with $x = 0$ available in the (p_x, p_y) plane is given by a circle of radius $\sqrt{2U}$. The equations of motion governing this system are easily derived:

$$\frac{\partial x}{\partial t} = p_x \quad , \quad \frac{\partial y}{\partial t} = p_y$$

$$\frac{\partial p_x}{\partial t} = -x - 2xy \quad , \quad \frac{\partial p_y}{\partial t} = -y - x^2 + y^2 \tag{7.31}$$

and they can be integrated starting from any value of (x, y, p_x, p_y) yielding and energy U and a bound orbit. The results are very interesting. If U is very small, all orbits behave more or less like normal harmonic oscillator orbits. If U is close to $\frac{1}{6}$, almost all states correspond to a single chaotic orbit. For values of the energy in between part of the states correspond to regular orbits, and part of the states to a chaotic orbit. This chaotic orbit is the feature we need in statistical mechanics. When a system contains many degrees of freedom and higher order interactions, chaotic orbits become more pronounced at more values of the energy. As a result, large systems can be described by statistical mechanical tools.

This example is studied in detail in a course on non-linear dynamics. An important question, studied actively in that field, is the following. What are the conditions for the Hamiltonian so that most of phase space for most energies is covered by a chaotic orbital? In our words, can we construct a non-trivial many body Hamiltonian for which statistical mechanics does not work? Note that every Hamiltonian will have regular orbits, but the ergodic theorem is valid only if these orbits have a measure zero.

7.5 Ideal gas in classical statistical mechanics.

The classical Hamiltonian for a system of N independent particles of mass m is given by

$$H = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m} \tag{7.32}$$

Collisions are needed.

Such a system is clearly not ergodic. The solutions for the trajectories are $\vec{x}_i(t) = \vec{x}_i(0) + \vec{v}_i t$ and sample only a small fraction of phase space. We have to perform a spatial average in order to get some meaningful answers. In reality, of course, there are collisions. These collisions are elastic and are assumed to be instantaneous and they randomly redistribute the energy between all particles. In this case, we can perform a time average to get a meaningful answer. The latter procedure is more satisfying from an experimental point of view, and allows us to compare experimental results with ensemble averages derived from the Hamiltonian above.

If we assume that the atoms are hard spheres, instantaneous collisions are the only possibility. But the interactions between atoms are more than that, there are long range tails in the potential. This means that when the density of the gas is too high, the interaction potential plays a role. Hence the ideal gas is a first approximation of a description of a gas at low density, where we can ignore the interaction terms.

Density of states.

We will calculate the entropy of the ideal gas in the micro-canonical ensemble. We replace the delta functions by Gaussian exponentials, and use the limit $\epsilon \to 0$. We have to use this ϵ-procedure in order to avoid problems with the product of a delta-function and its logarithm. The entropy follows from

$$S(U,V,N) = -k_B \frac{1}{N!h^{3N}} \int d\vec{X} C \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-H)^2}{\epsilon^2}} \log\left(C \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-H)^2}{\epsilon^2}}\right) \tag{7.33}$$

Similar to what we derived before, it is possible to show that the last, exponential, factor in the logarithm does not give a contribution to the integral. We expand the logarithm to get:

$$S(U,V,N) = -k_B \frac{1}{N!h^{3N}} \int d\vec{X} C \frac{1}{\epsilon\sqrt{\pi}} e^{-\frac{(U-H)^2}{\epsilon^2}} \left[\log(C) - \log(\epsilon\sqrt{\pi}) - \frac{(U-H)^2}{\epsilon^2} \right]$$
$$(7.34)$$

The third term in the integral is very small when $|U - H| \gg \epsilon$ due to the exponential. It is also very small when $|U - H| \ll \epsilon$ because of the quadratic term. Hence it only contributes when $|U - H| \approx \epsilon$. In that case the integrant is proportional to $\frac{1}{\epsilon}$. But the volume in phase space that corresponds to this value is proportional to ϵ^{6N-1} and hence the total contribution vanishes when ϵ goes to zero.

The second factor in the logarithm gives a contribution $k_B \log(\epsilon\sqrt{\pi})$ to the entropy, and this factor vanishes in the thermodynamic limit. Hence we find

$$S(U,V,N) = -k_B \log(C) = k_B \log(\mathcal{D}(U,V,N)) \qquad (7.35)$$

with

$$\mathcal{D}(U,V,N) = C^{-1} = \frac{1}{N!h^{3N}\epsilon\sqrt{\pi}} \int d\vec{X} e^{-\frac{(U-H)^2}{\epsilon^2}} \qquad (7.36)$$

Here we can again take the limit $\epsilon \to 0$ to recover the expression we derived before. But note that we have used the thermodynamic limit, and we did take it before the limit $\epsilon \to 0$. That is technically not correct. This was a topic of much discussion in the early days of the development of statistical mechanics.

For an ideal gas the density of states $\mathcal{D}$ follows from

$$\mathcal{D}(U,V,N) = \frac{1}{N!h^{3N}} \int dp_1^3 \cdots dp_N^3 \delta(U - \sum_{i=1}^{N} \frac{\vec{p_i}^2}{2m}) \int dr_1^3 \cdots dr_N^3 \qquad (7.37)$$

The spatial integral is easy and gives a factor V^N. The integral over the momenta is equal to the surface area of a 3N-dimensional hyper-sphere of radius $\sqrt{2mU}$. This leads to

$$\mathcal{D}(U,V,N) = \frac{V^N}{N!h^{3N}} \frac{2\pi^{\frac{3N}{2}} \left(\sqrt{2mU}\right)^{3N-1}}{\Gamma(\frac{3N}{2})} \qquad (7.38)$$

If N is even, $\Gamma(\frac{3N}{2}) = (\frac{3N}{2}-1)!$ and if N is odd, $\Gamma(\frac{3N}{2}) = \sqrt{\pi}2^{-\frac{3N-1}{2}}(3N-2)!!$ In both cases we can approximate $\log(\Gamma(\frac{3N}{2}))$ by $(\frac{3N-1}{2})\log(\frac{3N}{2}) - \frac{3N}{2}$.

This assumes that N is very large (thermodynamic limit) and that we can replace $N-1$ by N. The entropy in this limit is (again ignoring terms not linear in N)

$$S(U,V,N) = k_B \log(\frac{V^N}{N!h^{3N}} \left(\sqrt{2\pi mU}\right)^{3N}) - \frac{3}{2}Nk_B \log(\frac{3}{2}N) + \frac{3}{2}k_B N =$$

$$k_B N \log(\frac{V(2\pi mU)^{\frac{3}{2}}}{Nh^3}) + Nk_B - Nk_B \log\left((\frac{3}{2}N)^{\frac{3}{2}}\right) + \frac{3}{2}Nk_B =$$

$$Nk_B \log\left(\frac{V}{N}(\frac{4\pi mU}{3Nh^2})^{\frac{3}{2}}\right) + \frac{5}{2}Nk_B \qquad (7.39)$$

This is in agreement with what we derived before. Using $\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_{V,N}$ we find $\frac{1}{T} = \frac{3}{2}\frac{Nk_B}{U}$ as expected. Replacing U by $\frac{3}{2}Nk_BT$ in the formula for the entropy yields the old Sackur-tetrode formula. Hence the ideal gas is the described by the same formulas in classical statistical mechanics as in quantum statistical mechanics. In quantum statistical mechanics we needed to take the high temperature (or low density) limit to arrive at the Sackur-tetrode formula. In this limit quantum effects were not important, and it comes as no surprise that we recover the same formula in the classical limit!

7.6 Normal systems.

In the derivations for the entropy in the micro-canonical ensemble we have often made use of the thermodynamic limit. This means that if we take the limit $N \rightarrow \infty$ with $\frac{U}{N} = \mathcal{U}$ and $\frac{V}{N} = \mathcal{V}$ constant. For the ideal gas we see immediately that in this limit $\frac{S}{N}$ also approaches a constant value $\mathcal{S}$. Systems for which this is true are called **normal**. The ideal gas is a normal system, and all other systems we have discussed and will discuss are normal. If a system is not normal, standard statistical mechanics does not apply. It is easy to show that all systems with short-range interactions are normal. An abnormal system therefore has to have long-range interactions between the particles. A simple example of a abnormal system is a gravitational system with a uniform mass density μ inside a sphere of radius R. The potential energy of such a system is

$$V_{pot} = \int_{r,r'<R} d^3r d^3r' \frac{\mu^2}{|\vec{r} - \vec{r'}|} \qquad (7.40)$$

For large values of R this scales with $R^5 \propto V^{\frac{5}{3}}$. Hence in the limit $N \rightarrow \infty$ with $\frac{V}{N}$ constant the potential energy does not approach a constant, but diverges.

7.7 Quadratic variables.

The kinetic energy for an ideal gas has a very simple form. This is related to the fact that the momentum appears as a second power in the Hamiltonian. A

very general statement for such variables is easy to prove in classical statistical mechanics. Suppose we have one component c of the 6N-dimensional vector $\vec{X}$ in phase space which appears in the Hamiltonian in a quadratic way:

$$H(\vec{X}) = H' + \alpha c^2 \tag{7.41}$$

where H' does not depend on c. The partition function then has the form

$$\mathcal{Z} = \mathcal{Z}' \int_{-\infty}^{\infty} dc e^{-\frac{\alpha c^2}{k_B T}} = \mathcal{Z}' \sqrt{\frac{\pi k_B T}{\alpha}} \tag{7.42}$$

where $\mathcal{Z}'$ does not contain any reference to the component c anymore. The internal energy U is related to $\mathcal{Z}$ by $U = -\frac{\partial}{\partial \beta} \log(\mathcal{Z})$ and hence we find

$$U = U' + \frac{1}{2} k_B T \tag{7.43}$$

Hence the energy associated with a quadratic variable is simply $\frac{1}{2} k_B T$.

DIATOMIC GAS.

For an ideal mono-atomic gas we have kinetic energy only, and since this involves $3N$ components, the internal energy of an ideal gas is $\frac{3}{2} N k_B T$. For an ideal diatomic gas the situation is more complicated. At relatively low temperatures only the motion of the center of mass plays a role and the internal energy is again $\frac{3}{2} N k_B T$. At intermediate temperatures the molecules will rotate freely, adding two quadratic variables to the Hamiltonian, e.g. the two rotational momenta. Hence the energy is $\frac{5}{2} N k_B T$. At high temperatures the atoms are able to vibrate. This now adds two quadratic coordinates to the Hamiltonian, one momentum variable and one variable representing the distance between the atoms. Hence the energy is $\frac{7}{2} N k_B T$. At very high temperatures the molecules will start to break apart and we lose the vibrational bond energy associated with the position coordinate. The two atoms together are now described by six momentum coordinates and the energy becomes $3 N k_B T$.

7.8 Effects of the potential energy.

The ideal gas is an exceptional system, since it contains no potential energy term in the Hamiltonian. As a result $\mathcal{D}$ is proportional to V^N and the only volume dependence of the entropy is a term $N k_B \log(V)$. Because $\left(\frac{\partial S}{\partial V}\right)_{U,N} = \frac{p}{T}$ we find $pV = N k_B T$. Hence the ideal gas law holds for any system for which the Hamiltonian does not depend on the generalized coordinates, but only on the momenta! A second case occurs when we assume that each molecule has a finite volume b. The integration over each spatial coordinate then excludes a volume $(N-1)b$ and $\mathcal{D}$ is proportional to $(V - Nb)^N$ for large values of N, where $(N-1)b \approx Nb$. This gives rise to the factor $V - Nb$ in the van der Waals equation of state.

In a more general case, the Hamiltonian for the classical particles contains a potential energy U, and we have

$$H = \sum_{i=1}^{N} \frac{p_i^2}{2m} + U(\vec{r}_1, \vec{r}_2, \cdots, \vec{r}_N) \tag{7.44}$$

In many cases the potential energy obeys a given scaling relation. For example, assume that

$$U(\lambda\vec{r}_1, \cdots, \lambda\vec{r}_N) = \lambda^\gamma U(\vec{r}_1, \cdots, \vec{r}_N) \tag{7.45}$$

If the interactions are pure Coulomb interactions, then $\gamma = -1$. The partition function in this case is

$$\mathcal{Z} = \frac{1}{N!h^{3N}} \left(\int dp\, e^{-\frac{p^2}{2mk_BT}} \right)^{3N} \int \cdots \int d^3r_1 \cdots d^3r_N e^{-\frac{U(\vec{r}_1,\ldots,\vec{r}_N)}{k_BT}} \tag{7.46}$$

With $\beta = \frac{1}{k_BT}$ this is equal to

$$\mathcal{Z} = \frac{1}{N!} \left(\frac{\sqrt{2mk_BT\pi}}{h} \right)^{3N} \int_V d^3r_1 \cdots \int_V d^3r_N e^{-U(\beta^{\frac{1}{\gamma}}\vec{r}_1,\ldots,\beta^{\frac{1}{\gamma}}\vec{r}_N)} =$$

$$\frac{1}{N!} \left(\frac{\sqrt{2mk_BT\pi}}{h} \right)^{3N} \beta^{-\frac{3N}{\gamma}} \int_{V\beta^{\frac{3}{\gamma}}} d^3x_1 \cdots \int_{V\beta^{\frac{3}{\gamma}}} d^3x_N e^{-U(\vec{x}_1,\cdots,\vec{x}_N)} \tag{7.47}$$

In other words, in terms of V and T we have

$$\mathcal{Z}(T,V,N) = \frac{1}{N!} \left(\frac{\sqrt{2mk_B\pi}}{h} \right)^{3N} k_B^{\frac{3N}{\gamma}} T^{3N(\frac{1}{2}+\frac{1}{\gamma})} g(N, VT^{-\frac{3}{\gamma}}) \tag{7.48}$$

where g(N,x) is the result of the integration over all spatial coordinates. The Helmholtz free energy $F = -k_BT \log \mathcal{Z}$ has therefore only one term which depends on volume:

$$F(T,V,N) = k_BT \log(N!) - 3Nk_BT \log \left(\frac{\sqrt{2mk_B\pi}}{h} \right) - \frac{3Nk_BT}{\gamma} \log(k_B) -$$

$$3Nk_BT(\frac{1}{2} + \frac{1}{\gamma}) \log(T) - k_BT \log(g(N, VT^{-\frac{3}{\gamma}})) \tag{7.49}$$

The pressure follows from $- \left(\frac{\partial F}{\partial V} \right)_{N,T}$. The equation of state takes the general form

$$pT^{-1+\frac{3}{\gamma}} = f(N, VT^{-\frac{3}{\gamma}}) \tag{7.50}$$

where $f(N, x)$ is a function related to $g(N, x)$.

If there are only Coulomb interactions, $\gamma = -1$, and we find $p = T^4 f(N, VT^3)$. In general we can write the equation in the form

$$\frac{pV}{T} = VT^{-\frac{3}{\gamma}} f(N, VT^{-\frac{3}{\gamma}}) \tag{7.51}$$

The value of γ can therefore be determined from an experiment in which we vary p and T in such a way that pV/T is constant. That implies V^γ/T^3 is constant and a plot of V versus T gives γ. This description is valid for a gas of Na^+ and Cl^- ions at low density. The forces between ions can only be approximated by simple Coulomb forces when the ions are far apart. At small distances one has to include the effects of induced dipoles. Also, for a gas of inert atoms we expect to have van der Waals interactions, in which case we have $\gamma = -6$.

7.9 Problems for chapter 7

PROBLEM 1.

In the presence of a magnetic induction $\vec{B}$ the classical Hamiltonian representing the energy of the system, including the interaction energy with the external field, follows from the Hamiltonian $H(\vec{p}_1, \cdots, \vec{p}_N, \vec{r}_1, \cdots, \vec{r}_N)$ without a field by minimal substitution. The momenta $\vec{p}_i$ are replaced by $\vec{p}_i - \frac{e}{c}\vec{A}(\vec{r}_i)$ where $\vec{A}$ is the vector potential defined by $\vec{B} = \vec{\nabla} \times \vec{A}$.

A. Calculate the free energy $\mathcal{G}(T, \vec{B}, N, V) = U - TS - \vec{M} \cdot \vec{B}$

B. Prove van Leeuwen's theorem that diamagnetism does not exist in classical mechanics by showing that $\vec{M}(\vec{B}) = \vec{0}$ for all values of the induction $\vec{B}$.

PROBLEM 2.

Consider a classical system of identical rod-like molecules which have a permanent electric dipole. The direction of each molecule is given by the angular coordinates θ_i, ϕ_i. The Hamiltonian is given by

$$H(\vec{p}_i, \vec{r}_i, p_{\theta i}, p_{\phi i}, \theta_i, \phi_i) = \sum_{i=1}^{N} \left(\frac{\vec{p}_i^2}{2m} + \frac{1}{2}I[\dot{\theta}_i^2 + \dot{\phi}_i^2 \sin^2(\theta_i)] - dE\cos(\theta_i) \right)$$

where I is the moment of inertia of the molecule and d the electric dipole moment. The electric field $\vec{E}$ is applied along the z-axis.

A. Calculate the free energy $\mathfrak{G}(T, \vec{E}, N, V) = U - TS - \vec{P} \cdot \vec{E}$.

B. Calculate the polarization $\vec{P}(T, \vec{E}, N, V)$.

C. Show that the dielectric constant ϵ in the high temperature limit is given by $\epsilon = 1 + \frac{4\pi N d^2}{3VkT}$.

PROBLEM 3.

The Hamiltonian for a classical system of two identical particles is given by

$$H(\vec{r}_1, \vec{r}_2, \vec{p}_1, \vec{p}_2) = \frac{1}{2m}(\vec{p}_1^2 + \vec{p}_2^2) + \frac{1}{2}m\omega^2(\vec{r}_1 - \vec{r}_2)^2$$

Evaluate the Helmholtz free energy and the internal energy of this system as a function of temperature.

PROBLEM 4.

The positions of the atoms in a one-dimensional solid are given by $r_n = na + x_n$, $n = 1 \cdots N$. Neighboring atoms are connected by springs, and the Hamiltonian is

$$H(x_1, \cdots, x_N, p_1, \cdots, p_N) = \frac{1}{2m} \sum_{i=1}^{N} p_i^2 + \frac{K}{2} \sum_{i=2}^{N} (x_{i-1} - x_i)^2$$

Evaluate the internal energy and heat capacity of this solid.

PROBLEM 5.

Consider a system of N identical particles in one dimension. The positions x_n of the particles are connected via $|x_n - x_{n+1}| = a$. This is a model of a polymer chain with $N - 1$ links. The total length of the chain is $L = |x_1 - x_N|$. The potential energy of the n-th particle is mgx_n. Calculate the coefficient of linear thermal expansion for this system, and show that it is negative, like for a rubber band. Ignore kinetic energy terms in the energy.

PROBLEM 6.

The Hamiltonian for a classical single molecule consisting of two atoms is given by:

$$H(\vec{r}_1, \vec{r}_2, \vec{p}_1, \vec{p}_2) = \frac{1}{2m}(\vec{p}_1^2 + \vec{p}_2^2) + \frac{K}{2}(|\vec{r}_1 - \vec{r}_2| - d)^2$$

Evaluate the Helmholtz free energy and the internal energy of a system of N independent molecules of this type as a function of temperature.

PROBLEM 7.

The virial for a classical system is defined by $\mathcal{V} = \sum_i \vec{r}_i \frac{d\vec{p}_i}{dt}$. Show that the thermal average of $\mathcal{V}$ is equal to $-3Nk_BT$.

PROBLEM 8.

Consider the quantity $L(t) = \sum_i \vec{r}_i(t)\vec{p}_i(t)$. Show that the thermal average of $L(t)$ is independent of time.

Chapter 8

Mean Field Theory: critical temperature.

8.1   Introduction.

WHEN ARE SYSTEMS INTERESTING?

In the previous chapters we have developed the principles of statistical mechanics. Most of the applications were to simple models. The only important system we discussed in detail was a gas of independent particles. The model we discussed for that system is very important, since a gas is almost always close to ideal. Quantum effects can be explained in this model in a qualitative way. The theoretical model of an ideal gas is also a very good starting point for the treatment of a real gas, as we will see in a next chapter.

In this chapter we discuss systems of real interest. By that we mean systems that show a phase transition from one state to another. Real systems, of course, are quite complex, so we need to find a simple model that applies to realistic systems, yet can be approximately solved by analytical means. Therefore, we focus on a theoretical model which applies to many effects encountered in solids.

The most important characteristic of a solid is the fact that the atomic positions are fixed, if we ignore the small fluctuations due to phonons. With each atomic position we can associate other degrees of freedom, like spin or atom-type. Solid state physics is concerned with the procedures of obtaining information pertaining to these degrees of freedom, experimentally as well as theoretically.

Theoretical models of internal degrees of freedom coupled to fixed lattice sites are called lattice models. These models play an important role in many areas of physics and can be generalized to cases where particles are moving and are not restricted to specific lattice sites. Here we will only discuss the most elementary member of the class of lattice models, the Ising model. We will use

this model to show many of the techniques employed in statistical mechanical calculations of solids. In this sense the Ising model is a standard workhorse or toy-problem. More complicated and realistic lattice models can be solved along similar lines, although the calculations are more complicated and are often only feasible numerically.

BASIS FOR THE ISING MODEL.

The Ising model is a simple description of a magnetic solid. We assume that each atom labelled with i has a total angular moment $\vec{S}_i$. This total angular moment is a combination of the total spin and the total orbital angular momentum of all the electrons. In a real system, both the direction and magnitude of $\vec{S}_i$ are a function of time. The magnitude fluctuates around an average value, which could be zero. The directions are able to change. For example, in iron above the Curie temperature the individual atoms still have a local spin associated with them, but the orientation of these spins is random. Only below the Curie temperature do we find an ordered, ferromagnetic state.

A simple model for the internal energy of a collection of atomic spin-moments is the Heisenberg model (the traditional semantics uses spin moment here, but it can easily include orbital effects):

$$\mathcal{H} = -\sum_{i<j} \mathcal{J}_{ij}^H \vec{\mathcal{S}}_i \bullet \vec{\mathcal{S}}_j \tag{8.1}$$

There are only interactions between pairs of spins and the form of the interaction is a simple *exchange* type. The minus sign in front of the whole expression is a standard convention. There are N particles, and we keep N fixed. In expression 8.1 each bond between two atoms is counted once and hence the values of i and j are restricted by $i < j$. The total spin-moment $\vec{S}$ is given by

$$\vec{\mathcal{S}} = \sum_i \vec{\mathcal{S}}_i \tag{8.2}$$

and can be changed by applying an external field $\vec{H}$. Changing the volume means changing the interatomic distances and this affects the coefficients $\mathcal{J}_{ij}^H$. Note that both the Hamiltonian H and the spins $\vec{S}_i$ are quantum-mechanical operators. One possible limit is to make these into classical numbers, and we can use the same expression 8.1 in a classical theory.

We will assume that the volume is fixed and do not calculate the pressure. Hence the only extensive state variables are the entropy S and the total spin $\vec{S}$; the corresponding intensive variables are the temperature T and the applied magnetic field $\vec{H}$. N does show up in all formulas, however, and the Gibbs-Duhem relation is $U = TS + \mu N + \vec{H}\vec{M}$, where $\vec{M}$ is the magnetization density associated with $\vec{S}$.

APPROXIMATIONS.

In a first approximation the strength of the exchange interaction depends on the distance between lattice sites only. We denote the lattice sites by $\vec{R}_i$ and have

$$\mathcal{H} = -\sum_{i<j} J(|\vec{R}_i - \vec{R}_j|)\vec{\mathcal{S}}_i \bullet \vec{\mathcal{S}}_j \tag{8.3}$$

A complete basis for the Hilbert space in which we diagonalize this Hamiltonian is

$$|S^1, \cdots, S^N, S_z^1, \cdots, S_z^N> \tag{8.4}$$

Unfortunately, these are not eigenstates of of the Hamiltonian, since the operators $\vec{\mathcal{S}}_i$ are able to lower and raise the z quantum number. We have

$$\vec{\mathcal{S}}_i \bullet \vec{\mathcal{S}}_j = \mathcal{S}_i^+ \mathcal{S}_j^- + \mathcal{S}_i^- \mathcal{S}_j^+ + \mathcal{S}_{iz}\mathcal{S}_{jz} \tag{8.5}$$

Because we do not know what to do with the first two terms, we ignore them. The Hamiltonian is now

$$\mathcal{H} = -\sum_{i<j} J(|\vec{R}_i - \vec{R}_j|)\mathcal{S}_{iz}\mathcal{S}_{jz} \tag{8.6}$$

and the states in 8.4 are now eigenstates of this Hamiltonian. This system is still too complicated, though. Next, we assume that all spins are equal to one-half, $S^i = \frac{1}{2}$. This means that the z-component can take two values only, and we write for the values $\frac{1}{2}\sigma_i\hbar$, which implies that $\sigma_i = \pm 1$.

In the previous paragraphs we stated that even the Heisenberg model is hard to analyze. It can be simplified, however, for a system for which we know that a phase transition to an ordered magnetic state exists. Another way of looking a this is the following. Assume that the external magnetic field $\vec{H}$ is always applied along a given direction. This direction is used to measure the spin-components. Further, we ignore changes in the magnitude of the spin and only allow the direction to change. Quantum-mechanically, the only degree of freedom for each atomic site is the projection of the spin on the direction of the external field. Therefore the eigenvalues of this projection are

$$S_i^{par} = \frac{1}{2}\hbar\sigma_i \tag{8.7}$$

with $\sigma_i = \pm 1$. Hence our quantum states are completely described by a set of numbers $\{\sigma_1, \cdots, \sigma_N\}$ and as we have seen before there are 2^N possible states for this system. The eigenstates are denoted by $|\sigma_1, \cdots, \sigma_N>$. The original energy expression contains the product of spin-components perpendicular to the quantization axis, but leaving out these terms, like we did above, corresponds to assuming that they average out to zero. Therefore, the energy eigenvalue of a given quantum state $|\sigma_1, \cdots, \sigma_N>$ is

$$E\{\sigma_1, \cdots, \sigma_N\} = -\sum_{i<j} J(|\vec{R}_i - \vec{R}_j|)\frac{\hbar^2}{4}\sigma_i\sigma_j \tag{8.8}$$

Even this form is hard to analyze. Fortunately, in most cases the interactions in solids are screened and as a result they have a short range. We therefore take interactions into account between nearest neighbors only on the lattice and assume that the interactions are the same between all nearest neighbor pairs. A sum over nearest neighbors only is indicated by $< ij >$ under the summation sign. We also include the factor of four and Planck's constant in the interaction strength J and find for our final model for the energy eigenvalues

$$E\left\{\sigma_1, \cdots, \sigma_N\right\} = -J \sum_{<ij>} \sigma_i \sigma_j \tag{8.9}$$

The sum over nearest neighbors again counts each bond only once. For example, for a simple mono-atomic structure $\sum_{<ij>} 1 = \frac{1}{2}Nq$ where q is the number of nearest neighbors of each site. Although geometrical effects are no longer included via the exchange interaction strength, they are important due to the lattice summations. The state labels do not depend on geometry.

INCLUDING A MAGNETIC FIELD.

Very often we need the presence of a magnetic field. The energy term associated with an external field is simplified by assuming that each total spin couples to the external field $\vec{H}$ according to

$$\mathcal{H}_{int}(\vec{\mathcal{S}}_i) = -\vec{H} \bullet \gamma_G \vec{\mathcal{S}}_i \tag{8.10}$$

This represents the fact that the magnetic moment is proportional to the orbital moment, according to $\vec{M}_i = \gamma_G \vec{S}_i$. The components of $\vec{S}$ parallel to $\vec{H}$ are the only ones that play a role. The factors γ_G and $\frac{1}{2}\hbar$ are now included in the strength of the field h, leading to

$$E_{int} = -\vec{H} \bullet \vec{M} = -h \sum_i \sigma_i \tag{8.11}$$

Hence we define the total magnetic moment or total spin (equivalent in this case, since we redefined the units of magnetic field by including the appropriate factors)

$$M\left\{\sigma_1, \cdots, \sigma_N\right\} = \sum_i \sigma_i \tag{8.12}$$

THERMODYNAMIC LIMIT.

So far we have not specified the actual positions of the atoms. The atomic models are valid for all values of N, but a comparison with thermodynamics is only possible in the thermodynamic limit. If we want to study bulk properties, we assume that $N \rightarrow \infty$ and calculate the energy and magnetic moment per particle. In that limit all atomic positions become equivalent if the structure

is a monoatomic periodic lattice. We are also able to study surfaces in such a limit by assuming that the atomic positions only fill half of space. Even in the limit $N \to \infty$ this still leaves us with a surface. Due to the reduced symmetry atoms at the surface do not have the same environment as atoms far inside, and calculations are in general harder to perform.

8.2   Basic Mean Field theory.

The easiest way to obtain an approximate solution for the thermodynamics of almost all models is called mean field theory. In this section we discuss the basic mean field theory in the context of the Ising model, and from this it will be clear how to generalize this approximation for arbitrary physical models. In later sections we extend the idea of mean field theory to obtain better approximations.

First, we need to cast the energy function 8.9) in the form of an operator. Assume that the quantum state in which atom i has spin σ_i is given by $|\sigma_1, \cdots, \sigma_N >$. The Hamiltonian operator representing the energy of the system and acting on the space of all possible quantum states $|\sigma_1, \cdots, \sigma_N >$ is

$$\mathcal{H} = \sum_{\sigma_1, \cdots, \sigma_N} |\sigma_1, \cdots, \sigma_N > E\{\sigma_1, \cdots, \sigma_N\} < \sigma_1, \cdots, \sigma_N| \qquad (8.13)$$

The operator for the magnetic moment is

$$\mathcal{M} = \sum_{\sigma_1, \cdots, \sigma_N} |\sigma_1, \cdots, \sigma_N > M\{\sigma_1, \cdots, \sigma_N\} < \sigma_1, \cdots, \sigma_N| \qquad (8.14)$$

At this point we have to make a choice whether to sum over quantum states with a given value of the total moment or to introduce a magnetic field and sum over all states. The latter is much easier. The change is similar to what we did when we replaced the canonical (fixed N) ensemble by the grand canonical (fixed μ) ensemble. We will evaluate probabilities for a given temperature and magnetic field. The thermodynamic energy function we want to get is the magnetic Gibbs free energy

$$\mathfrak{G}(T, h, N) = U - TS - hM \qquad (8.15)$$

where M is the thermodynamic average of the magnetization density $\mathcal{M}$. The partition function to calculate is

$$\mathfrak{Z}(T, h, N) = \text{Tr } e^{-\beta(\mathcal{H} - h\mathcal{M})} \qquad (8.16)$$

The trace is best evaluated in the basis of eigenstates, and we have

$$\mathfrak{Z}(T, h, N) = \sum_{\sigma_1, \cdots, \sigma_N} < \sigma_1, \cdots, \sigma_N | e^{-\beta(\mathcal{H} - h\mathcal{M})} | \sigma_1, \cdots, \sigma_N > \qquad (8.17)$$

$$3(T, h, N) = \sum_{\sigma_1, \cdots, \sigma_N} e^{\beta(J\sum_{<ij>} \sigma_i\sigma_j + h\sum_i \sigma_i)} \qquad (8.18)$$

The magnetic Gibbs energy $\mathfrak{G}$ follows from the partition function according to

$$\mathfrak{G}(T, h, N) = -k_B T \log\left(3(T, h, N)\right) \qquad (8.19)$$

The expectation value of the Hamiltonian 8.9 for a given quantum state is a quadratic function in the variables σ_i. This is the source of all calculational problems. In mean field theory we approximate this Hamiltonian by a form that is *linear* in the spin-variables σ_i.

The spin-variables σ_i are pure numbers, and it will be useful to connect them with operators representing them. We have

$$\mathcal{S}_{zi}|\sigma_1, \cdots, \sigma_N> = \sigma_i|\sigma_1, \cdots, \sigma_N> \qquad (8.20)$$

and

$$\mathcal{H} = -J \sum_{<ij>} \mathcal{S}_{iz}\mathcal{S}_{jz} \qquad (8.21)$$

Note that these are not quite the usual spin operators, we have taken a factor $\frac{1}{2}\hbar$ out. The average value of the spin per site is given by

$$m(T, h, N) = \frac{1}{N} \sum_{i=1}^{N} \langle \mathcal{S}_i \rangle_{T,h} \qquad (8.22)$$

where we have defined the thermodynamical average by

$$\langle \mathcal{S}_i \rangle_{T,h} = \frac{1}{3(T, h, N)} \text{Tr } \mathcal{S}_i e^{-\beta(\mathcal{H} - h\mathcal{M})} \qquad (8.23)$$

If it is obvious which type of ensemble we are using, we will drop the labels on the average. In the remainder of this section it is understood that averages are at a given value of the temperature T and magnetic field h. We now consider an infinite solid and ignore surface effects. Also we assume that there is only one type of atomic site. The spin averages are in that case independent of the atomic position i and we have

$$m(T, h, N) = \langle \mathcal{S}_i \rangle \qquad (8.24)$$

We now rewrite the energy of a given quantum state using the average of the spin variables and deviations from average. We find

$$\mathcal{H} = -J \sum_{<ij>} (\mathcal{S}_{iz} - m)(\mathcal{S}_{jz} - m) - J \sum_{<ij>} \mathcal{S}_{iz}m - J \sum_{<ij>} m\mathcal{S}_{jz} + J \sum_{<ij>} m^2 \quad (8.25)$$

This can be rewritten by using the number of nearest neighbors q for each site and the total number N of sites.

$$\mathcal{H} = -J \sum_{<ij>} (\mathcal{S}_{iz} - m)(\mathcal{S}_{jz} - m) - Jm\frac{1}{2}q \sum_i \mathcal{S}_{iz} - Jm\frac{1}{2}q \sum_j \mathcal{S}_{jz} + Jm^2\frac{1}{2}Nq$$
(8.26)

where the factors $\frac{1}{2}$ occur because we are counting nearest neighbor bonds. For each site we include only half of the neighbors, in agreement with the requirement $i < j$. The two terms in the middle are now combined, and we end with:

$$\mathcal{H} = -J \sum_{<ij>} (\mathcal{S}_{iz} - m)(\mathcal{S}_{jz} - m) - Jmq \sum_i \mathcal{S}_{iz} + Jm^2\frac{1}{2}Nq \qquad (8.27)$$

The internal energy is the thermodynamical average of the Hamiltonian. Hence we find

$$U = \langle \mathcal{H} \rangle = -J \sum_{<ij>} \langle (\mathcal{S}_{iz} - m)(\mathcal{S}_{jz} - m) \rangle - Jmq \langle \sum_i \mathcal{S}_{iz} \rangle + Jm^2\frac{1}{2}Nq \quad (8.28)$$

$$U = -J \sum_{<ij>} \langle (\mathcal{S}_{iz} - m)(\mathcal{S}_{jz} - m) \rangle - Jm^2\frac{1}{2}Nq \qquad (8.29)$$

The second term on the right hand side is the energy of a system where each spin has its average value. The first term is a correction to this simple expression related to fluctuations in the spin variables. In mean field theory we ignore these fluctuations. We assume that the fluctuations on different sites are **independent**, they are **uncorrelated**, and write

$$\langle (\mathcal{S}_{iz} - m)(\mathcal{S}_{jz} - m) \rangle = \langle (\mathcal{S}_{iz} - m) \rangle \langle (\mathcal{S}_{jz} - m) \rangle = 0 \qquad (8.30)$$

We approximate the expression for the Hamiltonian operator by ignoring all terms containing products of differences of the form $\mathcal{S}_i - m$. Hence the mean-field Hamiltonian is

$$\mathcal{H}^{mf} = -Jmq \sum_i \mathcal{S}_{iz} + Jm^2\frac{1}{2}Nq \qquad (8.31)$$

The name mean field (or average field) is derived from the physical interpretation of this Hamiltonian. The energy of a spin at a given site i is determined only by the average of the spins on the neighboring sites. Another way of obtaining the mean-field is the following. We would like to write the Hamiltonian 8.21 in a linearized form:

$$\mathcal{H}^{lin} = -h_{eff} \sum_i \mathcal{S}_i + H_0 \qquad (8.32)$$

where the functions h_{eff} and H_0 are independent of the eigenvalues of the operators $\mathcal{S}$. We need a criterion to find h_{eff}. A simple physical picture of the state of a spin $\mathcal{S}_i$ shows all neighboring spins having their average value m, and hence $h_{eff} = mqJ$ as we found in 8.31. The term H_0 is determined by requiring that the energy is correct if all spins have the value m, in other words it removes the effects of double counting. The internal energy in linearized approximation is

$$\langle \mathcal{H}^{lin} \rangle = -mqJ \langle \sum_i \mathcal{S}_i \rangle + H_0 \tag{8.33}$$

$$\langle \mathcal{H}^{lin} \rangle = -m^2 qJN + H_0 \tag{8.34}$$

and this needs to be equal to $-\frac{1}{2}NqJm^2$, which gives H_0 as before.

One final remark. The basic form of mean field theory ignores correlations between fluctuations on different sites. But we also know that fluctuations are directly related to response functions. Therefore, we expect that response functions in mean field theory will not be extremely accurate. Especially, we will see errors in the results for critical exponents.

8.3 Mean Field results.

In order to proceed further we have to calculate the average spin m. We use the mean-field Hamiltonian to do that and we find

$$m(T,h) = \frac{\text{Tr } \mathcal{S}_i e^{-\beta(\mathcal{H}^{mf} - h\mathcal{M})}}{\mathfrak{Z}^{mf}(T,h,N)} \tag{8.35}$$

with

$$\mathfrak{Z}^{mf}(T,h,N) = \text{Tr } e^{-\beta(\mathcal{H}^{mf} - h\mathcal{M})} \tag{8.36}$$

But this cannot be done directly. The formula for m contains $\mathcal{H}^{mf}$ and hence it depends on m. Therefore, m cannot be chosen arbitrarily, but has to satisfy the two equations above. This is an implicit equation for m.

The partition function is easy to calculate in this approximation. We use the basis of eigenstates and we find

$$e^{\frac{1}{2}\beta NqJm^2} \mathfrak{Z}^{mf}(T,h,N) = \sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} e^{\beta(h+mqJ)\sum_{i=1}^N \sigma_i} =$$

$$\sum_{\sigma_1 = \pm 1} \cdots \sum_{\sigma_N = \pm 1} \prod_i e^{\beta(h+mqJ)\sigma_i} =$$

$$\prod_i \left[\sum_{\sigma_i = \pm 1} e^{\beta(h+mqJ)\sigma_i} \right] =$$

$$\left[\sum_{\sigma=\pm1} e^{\beta(h+mqJ)\sigma}\right]^N = 2^N \cosh^N(\beta(h+mqJ)) \qquad (8.37)$$

In a similar way we find

$$e^{\frac{1}{2}\beta NqJm^2}\mathrm{Tr}\ \mathcal{S}_k e^{-\beta(\mathcal{H}^{mf}-h\mathcal{M})} =$$

$$\sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_N=\pm1}\sigma_k e^{\beta(h+mqJ)\sum_{i=1}^N \sigma_i} =$$

$$\sum_{\sigma_1=\pm1}\cdots\sum_{\sigma_N=\pm1}\sigma_k \prod_i e^{\beta(h+mqJ)\sigma_i} =$$

$$\sum_{\sigma_k=\pm1}\sigma_k e^{\beta(h+mqJ)\sigma_k}\prod_{i\neq k}\left[\sum_{\sigma_i=\pm1} e^{\beta(h+mqJ)\sigma_i}\right] =$$

$$\sum_{\sigma_k=\pm1}\sigma_k e^{\beta(h+mqJ)}\left[\sum_{\sigma=\pm1} e^{\beta(h+mqJ)\sigma}\right]^{N-1} =$$

$$2^N \cosh^{N-1}(\beta(h+mqJ))\sinh(\beta(h+mqJ)) \qquad (8.38)$$

This expression is independent of k indeed, as required. Combining the last two formulas we find that $m(T,h)$, the average of the spin variable, has to obey

$$m = \tanh\left(\beta(mqJ + h)\right) = \tanh\left(\beta qJ\ \left(m + \frac{h}{qJ}\right)\right) \qquad (8.39)$$

The next question we have to answer is what are the solutions for m? We see that there are two ways to modify the nature of the solutions. We can either change βqJ, which changes the slope of the hyperbolic tangent at the origin, of change $\frac{h}{qJ}$, which shifts the whole curve. The magnitude of the coupling constant J or the number of neighbors q together do only set the scales for the temperature and the magnetic field. If we express the temperature in unit of $\frac{qJ}{k_B}$ and the magnetic field in units of qJ then the solutions for m do not depend directly on q and J anymore. Hence we define $h^* = \frac{h}{qJ}$ and $\beta^* = \beta qJ$. The resulting equation we have to solve is

$$m = \tanh\left(\beta^*\ (m + h^*)\right) \qquad (8.40)$$

This has the form $m = f(m)$ and a standard technique to solve such an equation is to draw a graph. We look at the graph of $y = f(m)$ and check where it intersects the straight line $y = m$. In figure 8.1 we show the results for $\beta^* = \frac{1}{2}$ and $h^* = 2$.

In this case we see that there is one solution. In general, there will always be at least on solution, since the hyperbolic tangent stays between minus and plus one, and the line $y = m$ has to intersect it. If one draws a picture of the left
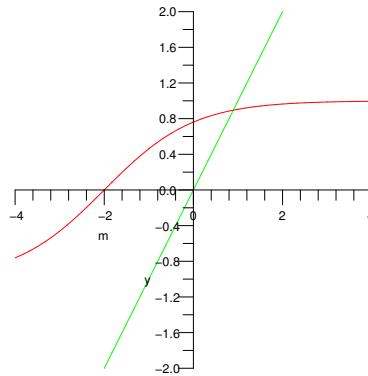
Figure 8.1: $\beta^* = \frac{1}{2}$, $h^* = 2$

and right hand sides of this equation, it is easy to see that one either has one or three solutions. In figure 8.2 we show the results for $\beta^* = 2$ and $h^* = 0.1$, and now we see three solutions.
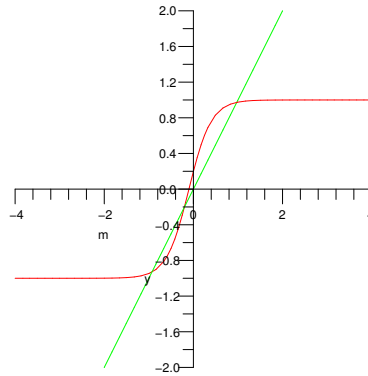
Figure 8.2: $\beta^* = 2$, $h^* = 0.1$

When there is only one solution, we have found the function $m(\beta^*, h^*)$ and are able to calculate the internal energy and the partition function from. In case there are three solutions, we need to compare the magnetic Gibbs energy for the three solutions and find the lowest energy solution. Note that when $|h^*| > 1$ there is only one solution, the model system is always magnetized. In a strong magnetic field the spins are always aligned in a ferromagnetic manner.

Spontaneous magnetic order.

An interesting case is related to the possibility of having spontaneous magnetization. When there is no external field we need to solve

$$m = \tanh\left(\beta^* \, m\right) \tag{8.41}$$

If we compare the slope of the hyperbolic tangent at the origin with the number one, the slope of the left hand side, we see that for $\beta^* > 1$ there are three solutions for m. When $\beta^* \leqslant 1$, there is only one solution. Obviously, $m = 0$ is always a solution. When β^* is only slightly larger than one, the solutions of the previous equation can be found by expanding the hyperbolic tangent, since in this case all solutions have values for m which are very small. Therefore we find

$$m \approx \beta^* m - \frac{1}{3}(\beta^* m)^3 \tag{8.42}$$

This leads to $m = 0$, as expected, or

$$m^2 \approx \frac{3(\beta^* - 1)}{\beta^{*3}} \approx 3(\beta^* - 1) \tag{8.43}$$

If there is a non-zero solution which is very small, we need β^* to be slightly larger than one. Since such a small non-zero solution only occurs very close to T_c, we find immediately that $\beta_c^* = 1$. Hence for $T < T_c$, where T_c is defined by $\beta_c^* = 1$ or $k_B T_c = qJ$, we have a second solution that vanishes at T_c like

$$m \propto \sqrt{3(\beta^* - 1)} = \sqrt{3(\frac{T_c}{T} - 1)} = \sqrt{\frac{3}{T}(T_c - T)} \tag{8.44}$$

which gives a critical exponent β (do not confuse this β with $k_B T$) of $\frac{1}{2}$, just as we argued in thermodynamics for mean field theory!

In order to find the thermodynamically stable state we have to compare the free energies. From the result for the partition function we obtain

$$\mathfrak{G}(T, h = 0, N) = -Nk_B T \log(2\cosh(\beta^* m)) + \frac{1}{2}NqJm^2 \tag{8.45}$$

Since m is close to zero for T near T_c the right hand side can be expanded like

$$\mathfrak{G}(T, 0, N) \approx -Nk_B T \log(2) - Nk_B T \log(1 + \frac{1}{2}(\beta^* m)^2) + \frac{1}{2}NqJm^2 \approx$$

$$-Nk_B T \log 2 - Nk_B T \frac{1}{2}(\beta^* m)^2 + \frac{1}{2}NqJm^2 =$$

$$-Nk_B T \log 2 - \frac{NqJ}{2}\left[k_B T \beta^2 qJ - 1\right]m^2$$

$$-Nk_B T \log 2 - \frac{NqJ}{2k_B T}\left[qJ - k_B T\right]m^2$$

$$-Nk_BT\log 2 - \frac{NqJ}{2T}\left[T_C - T\right]m^2 \tag{8.46}$$

Because $T < T_c$ the coefficient in front of the m^2 term is negative, and lowers the energy. This shows that the solution with $m \neq 0$ has the lowest energy. The problem is also symmetric, for $h = 0$ there are two possible values for m with opposite sign.

Note that for $T > T_c$ and hence $m = 0$ we have

$$\mathfrak{G}(T > T_c, h = 0, N) = -Nk_BT\log(2) \tag{8.47}$$

This result makes sense. The internal energy for zero moment is zero in the mean field approximation. Also, the magnetization is zero, and hence from $\mathfrak{G} = U - TS - HM$ we have $S = Nk_B\log(2)$, which tells us that the number of states available to the system is equal to 2^N. All states are available, which is what we would expect.

8.4   Density-matrix approach (Bragg-Williams approximation.

A second way to obtain the solutions for the Ising model is via the maximum entropy principle. This approach might seem quite tedious at first, and it is indeed, but it is also much easier to improve upon in a systematic fashion. This principle states that we have to maximize

$$S = -k_B\text{Tr}\ [\rho\log(\rho)] \tag{8.48}$$

over all density matrices ρ consistent with our model. Similar to what we found in a previous chapter for the canonical ensemble, with constraints $\text{Tr}\rho = 1$, $\text{Tr}\rho(\mathcal{H}) = U$, and $\text{Tr}\rho(\mathcal{M}) = M$, the density matrix maximizing this expression for the entropy is given by

$$\rho = \frac{1}{\text{Tr}\ \left[e^{-\beta(\mathcal{H}-h\mathcal{M})}\right]}e^{-\beta(\mathcal{H}-h\mathcal{M})} \tag{8.49}$$

This gives the exact solution, and the temperature and field have to be set in such a manner as to give U and M. The trace of the operator in 8.49 is still hard to calculate, however.

Maximum (and minimum) principles are very powerful. Instead of varying the entropy over all possible density matrices, we select a subset and maximize over this subset only. This gives us an approximate density matrix and the quality of the approximation is determined by our ability to find a good subset. Many calculations in physics in general are based on minimum or maximum principles.

The constraints we need are incorporated via Lagrange multipliers, and the function we need to maximize is

$$-k_B \text{Tr} \ [\rho \log(\rho)] - \beta k_B \ [\text{Tr}(\rho \mathcal{H}) - U] + \beta k_B h \ [\text{Tr}(\rho \mathcal{M}) - M] + k_B \lambda \ [\text{Tr}\rho - 1] \tag{8.50}$$

with $\text{Tr} \ \rho = 1$, $\rho = \rho^\dagger$, and $\rho^2 \leqslant \rho$, the last inequality in an operator sense. But we can rewrite the previous expression in the form:

$$-k_B \text{Tr} \ [\rho \log(\rho)] - \beta k_B \text{Tr}(\rho \mathcal{H}) + \beta k_B h \text{Tr}(\rho \mathcal{M}) + k_B \lambda \ [\text{Tr}\rho - 1] + k_B \beta U - k_B \beta h M \tag{8.51}$$

The operator for the Gibbs-like free energy for a magnetic problem is given by

$$\mathcal{G} = \mathcal{H} - T\mathcal{S}_{en} - h\mathcal{M} = \text{Tr}(\rho \mathcal{H}) + k_B T \text{Tr} \ [\rho \log(\rho)] - h \text{Tr}(\rho \mathcal{M}) \tag{8.52}$$

This means that we have to maximize the following expression for ρ:

$$-\frac{1}{T}\mathcal{G} + k_B \lambda \ [\text{Tr}\rho - 1] + k_B \beta U - k_B \beta h M \tag{8.53}$$

and then choose h and T to give the correct values of U and M. In other words we have to minimize $\mathcal{G}$!!! This change is equivalent to what we saw in thermodynamics, where we derived minimum energy principles from the maximum entropy principle.

Therefore, what we plan to do is to use a smaller set of density matrices and minimize the operator form of the Gibbs energy within this set to get an upper bound of the true thermodynamical Gibbs energy.

In our case, we will use the following approximation of the density matrix, as defined via the matrix elements

$$< \sigma_1, \cdots, \sigma_N |\rho| \sigma_1', \cdots, \sigma_N' > = \rho_1(\sigma_1, \sigma_1') \rho_2(\sigma_2, \sigma_2') \ \cdots \ \rho_N(\sigma_N, \sigma_N') \tag{8.54}$$

The functions $\rho_i(\sigma, \sigma')$ represent two by two matrices. Essentially, we decompose the density matrix as a direct product of N two by two matrices:

$$\rho = \rho_1 \bigotimes \rho_2 \bigotimes \cdots \bigotimes \rho_N \tag{8.55}$$

Writing the density matrix as a direct product in this way is equivalent to ignoring correlations between different sites. The correlation is not completely gone, of course, since the energy still connects neighboring sites.

Why does this form ignore correlations between different sites? If we use this form to calculate the thermodynamic average of a quantity which only depends on a single spin variable σ_k we find

$$\langle f(\sigma_k) \rangle = \left\{ \sum_{\sigma_k = \pm 1} f(\sigma_k) \rho_k(\sigma_k, \sigma_k) \right\} \prod_{i \neq k} \left\{ \sum_{\sigma_i = \pm 1} \rho_i(\sigma_i, \sigma_i) \right\} \tag{8.56}$$

and since

$$1 = \text{Tr } \rho = \prod_i \left\{ \sum_{\sigma_i = \pm 1} \rho_i(\sigma_i, \sigma_i) \right\} \tag{8.57}$$

it follows that

$$\langle f(\sigma_k) \rangle = \frac{\sum_{\sigma_k = \pm 1} f(\sigma_k) \rho_k(\sigma_k, \sigma_k)}{\sum_{\sigma_k = \pm 1} \rho_k(\sigma_k, \sigma_k)} \tag{8.58}$$

This result is independent of the state of the spins on all other sites. There are no correlations between the sites for expectation values.

We still have to minimize the Gibbs energy and that procedure will determine the form of ρ_k. Hence the nature of the system will only enter via the thermodynamic average for the energy and entropy, and hence the density matrix in this uncorrelated model depends only on these average quantities. We expect therefore that this will give the same results as the mean field or average field approximation. At this level the previous formulation of mean field theory is often easier to apply, but the present formulation is an important starting point for improving the model by including correlations in the density matrix.

For a general system the local density matrices ρ_i depend on the atomic site. For an infinite solid without surfaces all sites are equivalent and all functions ρ_i are identical to a simple function $\widetilde{\rho}$. The matrix elements $\widetilde{\rho}(\sigma, \sigma')$ have to obey certain requirements, though.

First of all, the density matrix has trace one, $\text{Tr } \rho = 1$, which gives

$$\sum_{\sigma_1, \cdots, \sigma_N} \widetilde{\rho}(\sigma_1, \sigma_1) \cdots \widetilde{\rho}(\sigma_N, \sigma_N) = 1 \tag{8.59}$$

or $[\text{Tr } \widetilde{\rho}]^N = 1$. The density matrix is Hermitian and if we consider one off diagonal element only we have

$$< \sigma_1, \cdots, \sigma'_k, \cdots, \sigma_N | \rho | \sigma_1, \cdots, \sigma_k, \cdots, \sigma_N > =$$

$$(< \sigma_1, \cdots, \sigma_k, \cdots, \sigma_N | \rho | \sigma_1, \cdots, \sigma'_k, \cdots, \sigma_N >)^* \tag{8.60}$$

We write the density matrix as a direct product again, and sum over all values σ_i with $i \neq k$. This gives

$$[\text{Tr } \widetilde{\rho}]^{N-1} < \sigma'_k | \widetilde{\rho} | \sigma_k > = \left([\text{Tr } \widetilde{\rho}]^{N-1} < \sigma_k | \widetilde{\rho} | \sigma'_k > \right)^* \tag{8.61}$$

Now we use $[\text{Tr } \widetilde{\rho}]^N = 1$ and get

$$\frac{< \sigma'_k | \widetilde{\rho} | \sigma_k >}{\text{Tr } \widetilde{\rho}} = \left(\frac{< \sigma_k | \widetilde{\rho} | \sigma'_k >}{\text{Tr } \widetilde{\rho}} \right)^* \tag{8.62}$$

The last equation means that the operator $\frac{\widetilde{\rho}}{\text{Tr } \widetilde{\rho}}$ is Hermitian with trace one, and therefore has real eigenvalues. Call the eigenvalues λ_1 and λ_2 and the

eigenvectors $e_1(\sigma)$ and $e_2(\sigma)$. We have $\lambda_1 + \lambda_2 = 1$. The eigen equations are now:

$$\widetilde{\rho}\vec{e}_m = \lambda_j m \left[\mathrm{Tr}\ \widetilde{\rho}\right] \vec{e}_m \tag{8.63}$$

The last result is interesting. It implies that the trace of the operator is a simple factor, which does not alter the eigenstates, and does not affect the ratio of probabilities for spin up and down. Also, the resulting factor in the complete density matrix ρ is $\left[\mathrm{Tr}\ \widetilde{\rho}\right]^N$. First of all, this trace is equal to one, and hence the trace of $\widetilde{\rho}$ has to be a phase factor. Second, the phase factor is not important, since it does not affect the total density matrix. Therefore, we can take $\mathrm{Tr}\ \widetilde{\rho} = 1$.

We also require that the total density matrix ρ is positive definite, meaning that

$$< \Psi|\rho|\Psi > \geqslant < \Psi|\rho^2|\Psi > \tag{8.64}$$

for all states $|\Psi>$. Suppose the components of $|\Psi>$ are given by

$$< \sigma_1, \cdots, \sigma_N|\Psi > = \prod_i e_{n(i)}(\sigma_i) \tag{8.65}$$

where the function $n(i)$ gives either one or two. We have

$$< \Psi|\rho|\Psi > = \lambda_1^P \lambda_2^{N-P} \left[\mathrm{Tr}\ \widetilde{\rho}\right]^N = \lambda_1^P \lambda_2^{N-P} \tag{8.66}$$

where P is the number of times $n(i)$ is equal to one, and we used the fact that $\mathrm{Tr}\ \rho = 1$. Similarly, we have

$$< \Psi|\rho^2|\Psi > = \lambda_1^{2P} \lambda_2^{2N-2P} \left[\mathrm{Tr}\ \widetilde{\rho}\right]^{2N} = \lambda_1^{2P} \lambda_2^{2N-2P} \tag{8.67}$$

As a result we have $\lambda_1^P \lambda_2^{N-P} \geqslant \lambda_1^{2P} \lambda_2^{2N-2P} \geqslant 0$ for all values of P. Taking $P = N$ gives $\lambda_1^N \geqslant \lambda_1^{2N}$ which means (with λ_1 being real) that $\lambda_N \leqslant 1$. This implies $\lambda_1 \leqslant 1$, and since $\lambda_1 + \lambda_2 = 1$ we have $\lambda_2 \geqslant 0$. Similarly, $\lambda_2 \leqslant 1$ and $\lambda_1 \geqslant 0$. Therefore, $\widetilde{\rho}$ is positive definite with trace one.

The most general form of a matrix $\widetilde{\rho}$ is

$$\widetilde{\rho} = \begin{pmatrix} \frac{1}{2}(1+m) & a^* \\ a & \frac{1}{2}(1-m) \end{pmatrix} \tag{8.68}$$

Here m is again the average value of the spin variable, $m = \sum_\sigma \sigma\widetilde{\rho}(\sigma,\sigma)$. Also, the trace of this matrix is one. The number a is complex. Therefore, we are left with three free parameters only! The number is three since a is complex and a combination of two independent real numbers. Note that at this moment m is still a parameter, but it will be the average magnetic moment **after we completed the minimization**.

We started from the following minimization:

$$\mathfrak{G}(h, T) = \min_\rho \mathrm{Tr}\ \left[\rho\mathcal{H} - \rho h\mathcal{M} + k_B T\rho \log(\rho)\right] \tag{8.69}$$

but now restrict ourselves to density matrices $\rho = (\widetilde{\rho})^N$, where the power means a direct product and where we use the three parameters a and m as defined before. This leads to

$$\mathfrak{G}(h,T) \leqslant \min_{m,a} \text{Tr} \; \left[(\widetilde{\rho})^N \mathcal{H} - (\widetilde{\rho})^N h\mathcal{M} + k_B T (\widetilde{\rho})^N \log((\widetilde{\rho})^N) \right] \qquad (8.70)$$

and this procedure gives an upper bound to the energy. Note that this does not give any bounds on derivatives. But we know that derivatives are monotonous, and hence we cannot have large oscillations in them. Therefore, derivatives have errors similar to the errors in the free energy!

The energy $U - Mh$ is calculated from $\text{Tr} \; (\mathcal{H} - h\mathcal{M})\rho$, which becomes

$$U - Mh = \sum_{\sigma_1} \cdots \sum_{\sigma_N} \widetilde{\rho}(\sigma_1,\sigma_1) \cdots \widetilde{\rho}(\sigma_N,\sigma_N) \left[-J \sum_{<ij>} \sigma_i\sigma_j - h \sum_i \sigma_i \right] \quad (8.71)$$

A given variable σ_j occurs only in a linear fashion in this summation, and hence the sums over the spin-variables either give a factor one or m, because $\sum_\sigma \widetilde{\rho}(\sigma,\sigma) = 1$ and $\sum_\sigma \sigma\widetilde{\rho}(\sigma,\sigma) = m$. This leads to

$$U - Mh = -\frac{1}{2}JNqm^2 - hNm \qquad (8.72)$$

as expected. This expression is independent of the values of the complex number a.

The harder task is to evaluate the entropy, because it has the logarithm of a matrix in the expression.

The entropy can be obtained from

$$S = -k_B \sum_{\sigma_1,\cdots,\sigma_N} \sum_{\sigma'_1,\cdots,\sigma'_N} < \sigma_1,\cdots,\sigma_N|\rho|\sigma'_1,\cdots,\sigma'_N >< \sigma'_1,\cdots,\sigma'_N| \log \rho|\sigma_1,\cdots,\sigma_N >$$

$$(8.73)$$

So how does one calculate the logarithm of a matrix? The final result is not too surprising, it looks like $\log(xy) = \log(x) + \log(y)$, but here we work with matrices in stead of numbers, so we have to be a little careful. In our case we define N additional matrices related to $\widetilde{\rho}$ by

$$< \sigma_1,\cdots,\sigma_N|R_n|\sigma'_1,\cdots,\sigma'_N >= \delta_{\sigma_1,\sigma'_1}\delta_{\sigma_2,\sigma'_2} \cdots \widetilde{\rho}(\sigma_n,\sigma_n) \cdots \qquad (8.74)$$

which is diagonal except for the n-th spin variable. Hence $\rho = R_1 R_2 \cdots R_N$ where all the matrices are of rank 2^N. These matrices commute, since the off diagonal elements occur on different blocks that are not connected. So we have $[R_i, R_j] = 0$.

The exponent of a matrix is defined via a power-series as usual, $e^A = \sum_n \frac{1}{n!}A^n$, and the logarithm is the inverse of the exponential, $e^A = B \;\; \Rightarrow$

$A = \log(B)$. We first look at the logarithm of R_n, $T_n = \log(R_n)$. The operators R_n have a simple form, they are Hermitian, and the eigenvectors are simple. The only states that can be mixed are states with different values of σ_n. But vectors that have different values of σ_i with $i \neq n$ are not mixed, and all eigenvectors of R_n have values for σ_i either $+1$ or -1. Since the eigenvectors of T_n are identical to the eigenvectors of R_n, the same is true for T_n. This implies that we can write

$$< \sigma_1, \cdots, \sigma_N | T_n | \sigma'_1, \cdots, \sigma'_N > = \delta_{\sigma_1, \sigma'_1} \delta_{\sigma_2, \sigma'_2} \cdots \widetilde{\tau}(\sigma_n, \sigma_n) \cdots \quad (8.75)$$

The only part of the space that is mixed corresponds to σ_n, but since the eigenvectors of both operators are teh same, we need to have

$$\widetilde{\tau} = \log(\widetilde{\rho}) \quad (8.76)$$

Finally, because $[R_i, R_j] = 0$ we have $\log(R_i R_j) = \log(R_i) \log(R_j)$. This gives us that $\log(\rho) = \sum_i \log(T_i)$. Hence

$$< \sigma_1, \cdots, \sigma_N | \log(\rho) | \sigma'_1, \cdots, \sigma'_N > = \sum_{n=1}^{N} \delta_{\sigma_1, \sigma'_1} \delta_{\sigma_2, \sigma'_2} \cdots \langle \log(\widetilde{\rho}) \rangle (\sigma_n, \sigma_n) \cdots$$
$$(8.77)$$

and

$$\mathrm{Tr}\, \rho \log(\rho) = \sum_{n=1}^{N} \sum_{\sigma_1} < \sigma_1 | \widetilde{\rho} | \sigma_1 > \sum_{\sigma_2} < \sigma_2 | \widetilde{\rho} | \sigma_2 > \cdots \sum_{\sigma_n, \sigma'_n} < \sigma_n | \widetilde{\rho} | \sigma'_n >< \sigma'_n | \log(\widetilde{\rho}) | \sigma_n > \cdots$$
$$(8.78)$$

The trace of the operators is equal to one, so all those sums in the equation are replaced by one. This leaves

$$\mathrm{Tr}\, \rho \log(\rho) = \sum_{n=1}^{N} \sum_{\sigma_n, \sigma'_n} < \sigma_n | \widetilde{\rho} | \sigma'_n >< \sigma'_n | \log(\widetilde{\rho}) | \sigma_n > = N \mathrm{Tr}\, \widetilde{\rho} \log(\widetilde{\rho}) \quad (8.79)$$

and we have

$$S = -k_B N \mathrm{Tr}\, \widetilde{\rho} \log(\widetilde{\rho}) \quad (8.80)$$

as expected, since all sites are equivalent and contributions from different sites are not mixed. Nevertheless, it was important to follow the mathematical derivation to make sure that our intuitive expectation was indeed correct.

Next we need to minimize the free energy

$$\mathfrak{G}(h, T) \leqslant \min_{m, a} \left[-\frac{1}{2} J N q m^2 - h N m + N k_B T \mathrm{Tr}\, \widetilde{\rho} \log(\widetilde{\rho}) \right] \quad (8.81)$$

Since the first parts are independent of a we can write this as

$$\mathfrak{G}(h,T) \leqslant \min_m \left[-\frac{1}{2} JNqm^2 - hNm + Nk_BT \min_a \{\text{Tr } \widetilde{\rho} \log(\widetilde{\rho})\} \right] \qquad (8.82)$$

and we find the value of a from

$$\min_a \{\text{Tr } \widetilde{\rho} \log(\widetilde{\rho})\} \qquad (8.83)$$

This number a is now determined by requiring

$$\frac{d}{da} \text{Tr } \widetilde{\rho} \log(\widetilde{\rho}) = 0 \qquad (8.84)$$

This expression is similar to the one we discussed in the chapter on operator methods, and here, too, we can write

$$\frac{d}{da} \text{Tr } \widetilde{\rho} \log(\widetilde{\rho}) = \text{Tr } \frac{d\widetilde{\rho}}{da} \log(\widetilde{\rho}) + \text{Tr } \frac{d\widetilde{\rho}}{da} \qquad (8.85)$$

Since Tr $\widetilde{\rho}$ has to remain equal to one, the second term is zero. The first term is $[\log \widetilde{\rho}]_{12}$. Similarly, we find that the derivative with respect to a^* is $[\log \widetilde{\rho}]_{21}$. In order to vary a and a^* we can either vary the real and imaginary part of a independently, but also can vary a and a^* independently. Both procedures give

$$[\log \widetilde{\rho}]_{12} = [\log \widetilde{\rho}]_{21} = 0 \qquad (8.86)$$

If $\log \widetilde{\rho}$ is diagonal, $\widetilde{\rho}$ must be diagonal and hence $a = 0$. This follows from $A = e^{\log(A)} = \sum \frac{1}{n!} [\log(A)]^n$, and because $\log A$ is diagonal, all powers of $\log A$ are diagonal. Therefore we find that

$$\widetilde{\rho} = \left[\begin{array}{cc} \frac{1}{2}(1+m) & 0 \\ 0 & \frac{1}{2}(1-m) \end{array} \right] \qquad (8.87)$$

Note that $\widetilde{\rho}$ is Hermitian with trace one, and that the condition of being positive definite requires $|m| \leqslant 1$, which is also expected.

There is one more detail, however. We found an extremum for the expression Tr $\widetilde{\rho} \log(\widetilde{\rho})$, but is it a minimum? Since there is only one extremum, independent of the value of m, we can look at one other easy case and check if the results are smaller or larger. Let us consider a real, positive, and small, and $m = 0$.

$$\widetilde{\rho} = \left[\begin{array}{cc} \frac{1}{2} & a \\ a & \frac{1}{2} \end{array} \right] \qquad (8.88)$$

Perturbation theory tells us that the eigenvalues of this matrix are $\frac{1}{2} \pm a$. Hence the trace is now equal to $T(a) = (\frac{1}{2}+a)\log(\frac{1}{2}+a) + (\frac{1}{2}-a)\log(\frac{1}{2}-a)$. If we take the derivative $\frac{d}{da}$ of this expression we find $\frac{dT}{da} = \log(\frac{1}{2}+a) - \log(\frac{1}{2}-a)$. This expression is indeed zero for $a = 0$, as needed, and is also positive for $a > 0$, as can be seen after combining the logarithms to $\frac{dT}{da} = \log(\frac{1+2a}{1-2a})$. Hence in this

particular direction our extremum is a minimum, and since we only have one extremum, it is a minimum no matter how we make changes.

Finally, we find that the entropy is

$$S = -Nk_B \left(\frac{1+m}{2} \log(\frac{1+m}{2}) + \frac{1-m}{2} \log(\frac{1-m}{2}) \right) \tag{8.89}$$

At this point we still have m as an unknown parameter. The thermodynamic variables which are specified are h, T, and N. The minimization needed to get to the free energy is now

$$\mathfrak{G}(h,T) \leqslant \min_m \left[-\frac{1}{2}JNqm^2 - hNm + Nk_BT \left(\frac{1+m}{2} \log(\frac{1+m}{2}) + \frac{1-m}{2} \log(\frac{1-m}{2}) \right) \right] \tag{8.90}$$

where m is restricted between -1 and $+1$. The value of m which minimizes this expression is then used to construct $\widetilde{\rho}$ and hence ρ. Once we have the density matrix ρ, we are able to calculate all thermodynamic quantities of interest.

The expression above looks like Landau theory, and, of course, it is. Landau theory has indeed a very good microscopic basis. We can use any parametrization of the density matrix, and the Gibbs like free energy has to be the minimum as a function of these parameters. The general form of the energy then follows, and for a magnetic model takes the form above.

The slope of $\mathfrak{G}(T,h,N;m)$ as a function of m is given by

$$\frac{\partial \mathfrak{G}}{\partial m}(T,h,N;m) = -NqJm - Nh + \frac{1}{2}Nk_BT \log \frac{1+m}{1-m} \tag{8.91}$$

At $m = -1$ the slope is $-\infty$ and at $m = +1$ the slope is $+\infty$. Hence $\mathfrak{G}(T,h,N;m)$ always has at least one minimum as a function of m for a value of m between -1 and $+1$. The value of m is found by equating the derivative of $\mathfrak{G}$ to zero, which yields

$$\frac{mqJ + h}{k_BT} = \log \sqrt{\frac{1+m}{1-m}} \tag{8.92}$$

It is not hard to show that this is equivalent to $m = tanh\beta(qJm + h)$. If we have $y = \log \sqrt{\frac{1+m}{1-m}}$ we have $(1-m)e^{2y} = 1+m$, and $e^{2y} - 1 = m(e^{2y} + 1)$, or $e^y - e^{-y} = m(e^y + e^{-y})$.

If there is only one solution, this has to be the minimum value of $\mathfrak{G}$ according to our observation of the values of the derivatives at the end-points. Since the derivative of $\mathfrak{G}$ with respect to m is a continuous function of m we are also able to deduce that there has to be an odd number of extrema between -1 and $+1$ due to the values of the slopes at the end-points. They have to be ordered minimum..maximum..minimum etc. Of course, one can always check this by calculating the second derivative

$$\frac{\partial^2 \mathfrak{G}}{\partial m^2} = -NqJ + \frac{Nk_BT}{1-m^2} \tag{8.93}$$

The results of this section are identical to the mean-field results. In this section we derived the mean-field results via the density matrix and approximated the density matrix by ignoring all correlations. This is another way of understanding how mean-field theory ignores correlations.

The form of the function $\mathfrak{G}$ is very similar to forms we used in our treatment of Landau theory in thermodynamics. We can show this in more detail by expanding the function in powers of m for values of m near zero. Using

$$(1+m)\log(1+m) = (1+m)\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}m^k = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}m^k + \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k}m^{k+1}$$
(8.94)

and the fact that we have to add the same expression with negative m shows that even exponents are the only ones to survive. Hence

$$(1+m)\log(1+m)+(1-m)\log(1-m) = 2\sum_{k=2,even}^{\infty} \frac{(-1)^{k+1}}{k}m^k + 2\sum_{k=2,even}^{\infty} \frac{(-1)^{k}}{k-1}m^k$$
(8.95)

or

$$(1+m)\log(1+m)+(1-m)\log(1-m) = 2\sum_{k=2,even}^{\infty} \frac{1}{k(k-1)}m^k$$
(8.96)

This is an even function of m as needed, and we have

$$\mathfrak{G}(h,T,N;m) = -\frac{1}{2}JNqm^2 - hNm + Nk_BT\left(-\log(2) + \frac{1}{2}m^2 + \frac{1}{12}m^4\right)$$
(8.97)

or

$$\frac{1}{N}\mathfrak{G}(h,T,N;m) = [-hm] + [-Nk_BT\log(2)] + \frac{1}{2}[k_BT - Jq]m^2 + \frac{1}{4}[\frac{k_BT}{3}]m^4$$
(8.98)

This is exactly the model for a second order phase transition that we discussed in Thermodynamics. Hence the Ising model in the mean field approximation shows a second order phase transition at a temperature T_c given by $k_BT_c = qJ$. The critical exponent β is therefore $\frac{1}{2}$ (Note that β is in statistical mechanics used for $\frac{1}{k_BT}$, but also denotes a critical exponent. Do not confuse these two!). The relation between the macroscopic Landau theory and statistical mechanics is very important, it shows how the parameters in a thermodynamical model can be derived from a statistical mechanical theory on a microscopic level.

In the previous sections we have derived the same result in two different ways, and therefore have shown that these two approximations are equivalent. In **Mean Field Theory** we replace the quadratic coupling terms in the Hamiltonian by the product of an average and a variable. This is equivalent to coupling each spin to the average field due to the neighbors. In the **Bragg-Williams approximation** we replace the density matrix by a form that does not contain coupling between neighboring sites. As a result, in the Hamiltonian again the average field of the neighbors plays a role only.

8.5   Critical temperature in different dimensions.

In the mean-field solution of the Ising model the critical temperature does not explicitly depend on the dimensionality of the problem. Indirectly it does, of course, because the possible values of q depend on the number of dimensions. For a mono-atomic, periodic lattice in one dimension q has to be two, in two dimensions q can be two, four, or six, and in three dimensions q can be two, four, six, eight, or twelve. Hence the higher the number of dimensions, the larger the critical temperature can be. For closed packed systems the number of neighbors is maximal, and the critical temperature would increase according to the pattern 2:6:12 going from one to three dimensions. This is also true in general. When the number of neighbors increases, the total strength of the interactions driving an ordered state increases and the ordered state can persist up to a higher temperature.

The mean-field solution of the Ising model is only an approximation. In general, the errors in the mean-field solutions become larger when the number of dimensions decreases. In one dimension, for example, there are no phase transitions in the Ising model! What causes the error in mean-field theory? In essence, we did something dangerous by interchanging two limits. We first assumed that $N \to \infty$ in order to deduce that $\langle \mathcal{S}_i \rangle$ is the same for all positions i. Or that $\widetilde{\rho}_i$ is independent of i. Then we used this common value to solve the equations for finite values of N. The results made of this procedure made sense in the thermodynamic limit, since m did not depend on N. The error is actually in the first step. We have to solve the whole problem at finite N and calculate $m_i(N, T)$. In this expression we should take the limit $N \to \infty$. The results will be different in that case. We cannot interchange these limits!

There is an old argument, due to Landau (who else?), that the Ising chain does not show a phase transition. Consider a finite chain of N sites, numbered one through N sequentially. Assume that $h = 0$. There are two completely ordered states, either all spins are up or down. Next we consider all states which include a so-called Bloch wall. The first m spins have one direction, while the remaining $N - m$ spins have the opposite direction. The number of these states is $2(N-1)$. The energy of a completely ordered chain is $E_0 = -J(N-1)$. The energy of a Bloch wall state is different only due to the contribution of the two spin variables at opposite sites of the Bloch wall. Therefore the energy of a Bloch wall state is $-J(N-2) + J = E_0 + 2J$. At a fixed temperature and at

$h = 0$ the difference in Gibbs free energy, $\Delta\mathcal{G}$, is therefore:

$$\Delta\mathcal{G} = 2J - T\Delta S = 2J - k_B T \left(\log 2(N-1) - \log(2)\right) = 2J - k_B T \log(N-1) \tag{8.99}$$

For every value of the temperature not equal to zero, this free energy difference becomes negative in the thermodynamic limit $N \to \infty$. Hence at a non-zero temperature the completely ordered state is not stable in the thermodynamic limit and the Ising chain should not show a phase transition in the traditional sense, even though at $T = 0$ the system is stable in the completely ordered state. The temperature $T = 0$ is a singularity, but for a real phase transition we require that it must be possible to define values of the temperature below T_c.

We will now look further into the physical explanation of this phenomenon. Consider a finite chain with all spins up at non-zero temperature. Each spin is able to fluctuate, and spins that are not at the end have to pay an energy $4J$ because they break two bonds by switching. The probability for this to happen is $e^{-4J\beta}$. It is far more likely for a spin at the end to switch, since the energy cost is only $2J$ and hence the probability is $e^{-2J\beta}$.

How long does it take for such a defect to be introduced? Due to thermal effects (in real systems most likely related to phonons, but any small outside perturbation will do) the spin tries to change. We can represent these effects by an *attempt frequency* A. The time scale typical for the introduction of such defects, t_{intro}, then has to obey:

$$1 = A e^{-2J\beta} t_{intro} \tag{8.100}$$

and we say that every t_{intro} seconds a new defect is introduced.

What is the lifetime of such a defect? The easiest way is to think about it in terms of the motion of a Bloch wall. Such motion does NOT cost energy, since the energy is directly associated with the two sites defining the Bloch wall, and no matter where it is the energy is the same $2J$. Of course, at the end points the Bloch wall can disappear, with a gain of energy $2J$.

The motion of a Bloch wall can be represented by a random walk, since at each stage the probability of moving to the left is equal to the probability of moving to the right. We define t_{hop} to be the average time between hops. This is again a thermal effect, and the hopping time is the inverse of a different attempt frequency, $A' t_{hop} = 1$, where there is no Boltzmann factor. We need some outside mechanism to cause these hops. They are most likely coupled to the same heat bath which drives the spin flips, but the coupling strength is different because it is a different mechanism. Since the motion of the Bloch wall is a random walk, the average distance covered in lattice spacings in a time t is $\sqrt{\frac{t}{t_{hop}}}$. The time it takes for a Bloch wall to move through the chain is therefore found by setting this expression equal to N, and we have

$$t_{life} = t_{hop} N^2 \tag{8.101}$$

We can now distinguish two cases. If $t_{life} \gg t_{intro}$ we find at each moment many Bloch walls in the system. This means that the system looks like it is not ordered. On the other hand, if $t_{life} \ll t_{intro}$ most of the time there are no Bloch walls in the system, and the chain is fluctuating back and forth between two completely ordered states, one with all spins up and one with all spins down.

In both cases we have for the time average of a spin

$$\langle \sigma_i \rangle = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau \sigma_i(t) dt \tag{8.102}$$

which is zero in both case. Note that we always need $\tau \gg t_{life}, t_{intro}$.

What are values we expect for realistic systems. If we assume that attempts are caused by phonons,the attempt frequencies for both creation and hopping are of order 10^{12} s^{-1}. This means that

$$\frac{t_{intro}}{t_{life}} \approx \frac{A^{-1}e^{2J\beta}}{A^{-1}N^2} \tag{8.103}$$

and with $N = 10^8$ as typical for a 1 cm chain and $J = 1$ meV as a typical spin energy, we have $t_{intro} \approx t_{life}$ for $T \approx 1$ K. But this temperature depends on the size of the system. At high temperatures the shorter time is the introduction time. We always see many domain walls. This is the normal state of the system. If we go to lower temperatures this picture should remain the same, but due to the finite length of the chain we see a switch in behavior. Below a certain temperature the chain at a given time is always uniformly magnetized, but over time it switches. The temperature at which this change of behavior takes place depends on the length of the chain. This is an example of a **finite size effect**. For large systems finite size effects take place at very low temperatures, see example numbers above. For small samples this can be at much higher temperatures!

In the previous discussion we have three limits to worry about. In order for the averages to be the same, $\langle \sigma \rangle_{time} = \langle \sigma \rangle_{ensemble}$, we need $\tau \to \infty$. We are interested in low temperatures, or the limit $T \to 0$. Finally, we have the thermodynamic limit $N \to \infty$. Since we need $\tau \gg t_{life} = t_{hop}N^2$ we see that we need to take the limit $\tau \to \infty$ before the thermodynamic limit. Also, because $\tau \gg t_{intro} = A^{-1}e^{2J\beta}$ we need to take the limit $\tau \to \infty$ before the limit $T \to 0$. Form our discussion it is clear that $\lim_{\tau \to \infty} \langle \sigma \rangle_\tau = 0$. Hence we have $\lim_{T \to 0} \langle \sigma \rangle_\infty = 0$ and this remains zero in the TD limit. The one dimensional Ising chain is never ordered!

Calculations at $T = 0$ do not make sense. At $T = 0$ we have $t_{intro} = \infty$ and we always have $\tau < t_{intro}$. That violates the ergodic theorem. So, even though people do calculations at $T = 0$, one has to argue carefully if it makes sense. That is what we did for the system of Fermions, for example.

After taking the limit $\tau \to \infty$, what does the system look like? If we now take $T \to 0$ we find a system that is most of the time completely ordered, but once in a while flips. The net magnetization is zero. Therefore, the real $T = 0$ ground state of a finite system is a system that is ordered most of the time,

but sometimes flips. A bar magnet will switch poles once in a while (but don't wait for it to happen, folks). If we would take the thermodynamic limit first, the ground state of the system would always be disordered! This is the wrong picture of the ground state with zero magnetization. On the other hand, if we take a system at a small non-zero temperature, the thermodynamic limit then implies that the equilibrium state of the system is a disordered state. Now the spatial and temporal averages of variables are the same. But at $T = 0$ they are not, and the point $T = 0$ is a singularity for the one dimensional Ising chain.

Realistic systems have finite values of N and non-zero values of T. Keep this in mind, and watch out for sample size dependent behavior near phase transitions. Near a phase transition the correlation length becomes larger than the system size, just as in our example of the one dimensional Ising chain. Near a phase transition we have to look carefully at the thermodynamic limit and scale out finite size effects, using appropriate scaling laws.

The same argument can be applied in two dimensions. Assume that the spins are arranged on a square of $L \times L$ sites. Consider all Bloch wall states which divide the sample into two parts, each of which is connected and contains at least one surface spin. The average length of the Bloch wall L_W is proportional to L, $L_W = cL$. The way we defined the Bloch wall does not allow for fractal dimensions, because we have discrete building blocks, and finite steps along the wall.

The number of possible walls depends on the number of neighbors of each site. In first approximation it will depend on L in the form $2Lb^{L_W}$. In this case b is the average number of choices we have for a wall to continue at a given point. For example, on a square lattice we expect b to be between two and three. At each site the next point in the Bloch wall can either go right, straight, or left. This would mean three choices for each link in the wall. Since we have to avoid wall crossings, the actual value will be lower. The pre-factor $2L$ is due to the fact that we have L possible starting points at a given side, and we can start at two sides. The increase in energy due to the wall will be $L_W 2J$, because each element of the wall breaks a bond. Hence we have

$$\Delta \mathcal{G} = 2JL_W - k_B T \log(2[2Lb^{L_W}]) + k_B T \log(2) \qquad (8.104)$$

where we included a factor two because we can go spin up to down or down to up. This gives

$$\Delta \mathcal{G} = 2JL_W - k_B T \log(b^{L_W}) - k_B T \log(2[2L]) + k_B T \log(2) \qquad (8.105)$$

In the limit $L \to \infty$ the last two terms are not important and we see that

$$\Delta \mathcal{G} = L_W(2J - k_B T \log(b)) \qquad (8.106)$$

Therefore the ordered state is stable against the thermal introduction of Bloch walls if $k_B T \log(b) < 2J$. This gives an estimate of

$$k_B T_c = \frac{2J}{\log(b)} \tag{8.107}$$

The mean field results for a square lattice is $k_B T_c^{mf} = 4J$; if $b > 1.6$ the Landau estimate of T_c is below the mean-field result. The exact result for a two-dimensional square Ising lattice is $k_B T_c = 2.27J$, or $b \approx 2.4$. This is a very reasonable value for b considering our discussion on the relation between the parameter b and choices of continuing a Bloch wall. The largest value of b we can have is three, which would give $k_B T_c = 1.8J$, and this gives a lower limit on calculations for T_c.

8.6 Bethe approximation.

The next question is: how can we improve the mean-field results. In the mean field approximation we do not have any information on the collective behavior of two neighboring spins. We need that information, and want to find ways to obtain that information. One possibility is via the density matrices. If we introduce some correlation in the approximate density matrix, we will obtain better results. Mathematically, this is a rigorous procedure, and can be performed easily on a computer.

A simple model of introducing correlations is the Bethe approximation, which does not introduce density matrices explicitly. The basic idea of the Bethe approximation is the following. Consider a cluster of atoms. The interactions between spins in this cluster are taken into account exactly, but the interactions with the environment are treated in a mean field manner.

The basic philosophy of the mean field approximation was the following. The energy at a given site depends on the effects of the neighboring sites through some averaged additional magnetic field:

$$E_{mf}(\sigma_0) = -(h + h')\sigma_0 + f(h') \tag{8.108}$$

where h is the regular magnetic field, and h' is the additional field. The term $f(h')$ has to be introduced to avoid double counting of bonds. At the central site all bonds to neighbors are counted, and if we multiply by N that means that we would count the effect of each bond twice. We determine h' by requiring translational symmetry, $\langle \sigma_0 \rangle = m$, which leads to $m = \tanh(\beta(h + h'))$. We need an additional condition to relate h' and m and make the obvious choice $h' = qJm$.

How can we improve the mean field result? If we want to describe bonds with neighbors exactly, we need to consider a cluster of atoms. All sites in the cluster are treated exactly. There is still an outside of the cluster, and that outside will be represented by some effective field. In this way we do describe correlations between spins in the cluster, and obtain more information that we had before. But since the outside is still represented by an average field, we expect that some elements of the mean field theory might still survive. We do

expect that if we take the limit of an infinite cluster we might be able to get the exact solution, as long as we do take limits in the appropriate order.

In our case we have short range (nearest neighbor) interactions only. So we would expect rapid convergence with cluster size. But that is not necessarily true, since near a phase transition the correlation length becomes infinitely large. The cluster we consider here consists of a central atom and its q neighbors. This is the approach taken by Bethe. Because of symmetry we assume that all neighboring sites are equivalent. We label the central spin 0 and the neighbors 1,2,..,q. The energy of this cluster is

$$E_c(\sigma_0, \cdots, \sigma_q) = -J\sigma_0 \sum_{i=1}^{q} \sigma_i - h \sum_{i=0}^{q} \sigma_i - h' \sum_{i=1}^{q} \sigma_i + f(h') \qquad (8.109)$$

The last term represents the interactions with the environment. We assume that the environment is in some average thermodynamic state and that the effects of the environment on the outer spins are noticeable through an effective magnetic field h' acting on the surface of the cluster. The average field due to the outside acts only on the atoms that are in contact with the outside. Our result can be generalized to larger and non-symmetric clusters, which makes the mathematics much more complicated. For example, it would introduce effective fields that are site specific. In our case we use the same effective field h' on all neighbors, since the cluster is symmetric.

For this cluster we now have to impose translational symmetry, and we need $\langle \sigma_i \rangle = \langle \sigma_0 \rangle$. This gives an equation which determines the value of h'. We did already make one choice here, and we assumed that the effective field on the central atom is zero. That makes sense. If we make the cluster larger, however, we have to introduce an effective field for each shell around the center, and we get an equation equation magnetic moments for each shell, and can again determine all fields. Only the central effective field can be set equal to zero! An individual atom in the nearest neighbor shell is not at a symmetric position with respect to the surface of the cluster, and hence not equivalent to the central cell. It therefore needs an effective field.

Next, we need to calculate the partition function. In order to do so, we drop the term $f(h')$ in the energy, since it will give a factor in the partition function that does not depend on the values of σ and hence does not influence the spin averages. We do need it at the end again, however, to calculate the energy. The partition function for the cluster is

$$\mathcal{Z}_c = \sum_{\{\sigma_0, \cdots, \sigma_q\}} e^{-\beta E_c(\sigma_0, \cdots, \sigma_q)} \qquad (8.110)$$

The derivation is a bit tedious. First we introduce the energy formula and separate the sum on the central spin from the others.

$$\mathcal{Z}_c = \sum_{\{\sigma_1, \cdots, \sigma_q\}} \sum_{\sigma_0} e^{\beta J \sigma_0 \sum_{i=1}^{q} \sigma_i} e^{\beta h \sum_{i=0}^{q} \sigma_i} e^{\beta h' \sum_{i=1}^{q} \sigma_i} \qquad (8.111)$$

Next, we perform the sum over the central spin

$$\mathcal{Z}_c = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \left[e^{\beta J \sum_{i=1}^q \sigma_i} e^{\beta h} e^{\beta(h+h') \sum_{i=1}^q \sigma_i} + e^{-\beta J \sum_{i=1}^q \sigma_i} e^{-\beta h} e^{\beta(h+h') \sum_{i=1}^q \sigma_i} \right] \tag{8.112}$$

$$\mathcal{Z}_c = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \left[e^{\beta h} \prod_{i=1}^q e^{\beta(J+h+h')\sigma_i} + e^{-\beta h} \prod_{i=1}^q e^{\beta(-J+h+h')\sigma_i} \right] \tag{8.113}$$

The sums and products can now be interchanged, since $\sum_{x_1,x_2,\cdots} f(x_1)f(x_2)\cdots = [\sum_x f(x)]^N$. This leads to

$$\mathcal{Z}_c = e^{\beta h}[2\cosh(\beta(J+h+h'))]^q + e^{-\beta h}[2\cosh(\beta(-J+h+h'))]^q \tag{8.114}$$

The spin averages are calculated from the expressions

$$\langle \sigma_0 \rangle = \frac{1}{\mathcal{Z}_c} \sum_{\{\sigma_0,\cdots,\sigma_q\}} \sigma_0 e^{-\beta E_c(\sigma_0,\cdots,\sigma_q)} = \frac{S_0}{\mathcal{Z}_c} \tag{8.115}$$

and

$$\langle \sigma_j \rangle = \frac{1}{\mathcal{Z}_c} \sum_{\{\sigma_0,\cdots,\sigma_q\}} \sigma_j e^{-\beta E_c(\sigma_0,\cdots,\sigma_q)} = \frac{S_j}{\mathcal{Z}_c} \tag{8.116}$$

The additional sums are not that hard to evaluate once we have done the partition function.

$$S_0 = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \sum_{\sigma_0} \sigma_0 e^{\beta J \sigma_0 \sum_{i=1}^q \sigma_i} e^{\beta h \sum_{i=0}^q \sigma_i} e^{\beta h' \sum_{i=1}^q \sigma_i} \tag{8.117}$$

$$S_0 = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \left[e^{\beta J \sum_{i=1}^q \sigma_i} e^{\beta h} e^{\beta(h+h') \sum_{i=1}^q \sigma_i} - e^{-\beta J \sum_{i=1}^q \sigma_i} e^{-\beta h} e^{\beta(h+h') \sum_{i=1}^q \sigma_i} \right] \tag{8.118}$$

$$S_0 = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \left[e^{\beta h} \prod_{i=1}^q e^{\beta(J+h+h')\sigma_i} - e^{-\beta h} \prod_{i=1}^q e^{\beta(-J+h+h')\sigma_i} \right] \tag{8.119}$$

$$S_0 = e^{\beta h}[2\cosh(\beta(J+h+h'))]^q - e^{-\beta h}[2\cosh(\beta(-J+h+h'))]^q \tag{8.120}$$

and

$$S_j = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \sum_{\sigma_0} \sigma_j e^{\beta J \sigma_0 \sum_{i=1}^q \sigma_i} e^{\beta h \sum_{i=0}^q \sigma_i} e^{\beta h' \sum_{i=1}^q \sigma_i} \tag{8.121}$$

$$S_j = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \left[\sigma_j e^{\beta J \sum_{i=1}^q \sigma_i} e^{\beta h} e^{\beta(h+h') \sum_{i=1}^q \sigma_i} + \sigma_j e^{-\beta J \sum_{i=1}^q \sigma_i} e^{-\beta h} e^{\beta(h+h') \sum_{i=1}^q \sigma_i} \right] \tag{8.122}$$

$$S_j = \sum_{\{\sigma_1,\cdots,\sigma_q\}} \left[e^{\beta h} \sigma_j \prod_{i=1}^q e^{\beta(J+h+h')\sigma_i} + e^{-\beta h} \sigma_j \prod_{i=1}^q e^{\beta(-J+h+h')\sigma_i} \right] \tag{8.123}$$

$$S_j = e^{\beta h} [2\cosh(\beta(J+h+h'))]^{q-1} [2\sinh(\beta(J+h+h'))] +$$

$$e^{-\beta h} [2\cosh(\beta(-J+h+h'))]^{q-1} [2\sinh(\beta(-J+h+h'))] \tag{8.124}$$

The last expression is independent of j, as expected.

The value for h' is determined by requiring that the average spin is the same everywhere, or

$$m = \langle \sigma_0 \rangle = \langle \sigma_j \rangle \tag{8.125}$$

which leads to $S_0 = S_j$ or

$$e^{\beta h} [2\cosh(\beta(J+h+h'))]^q - e^{-\beta h} [2\cosh(\beta(-J+h+h'))]^q =$$

$$e^{\beta h} [2\cosh(\beta(J+h+h'))]^{q-1} [2\sinh(\beta(J+h+h'))] +$$

$$e^{-\beta h} [2\cosh(\beta(-J+h+h'))]^{q-1} [2\sinh(\beta(-J+h+h'))] \tag{8.126}$$

or

$$e^{\beta h} [\cosh(\beta(J+h+h'))]^{q-1} \left[\cosh(\beta(J+h+h')) - \sinh(\beta(J+h+h'))\right] =$$

$$e^{-\beta h} [\cosh(\beta(-J+h+h'))]^{q-1} \left[\cosh(\beta(J-h-h')) - \sinh(\beta(J-h-h'))\right] \tag{8.127}$$

which leads to

$$e^{\beta h} [\cosh(\beta(J+h+h'))]^{q-1} \left[e^{-\beta(J+h+h')}\right] =$$

$$e^{-\beta h}[\cosh(\beta(-J+h+h'))]^{q-1}\left[e^{-\beta(J-h-h')}\right] \qquad (8.128)$$

or

$$\frac{\cosh(\beta(J+h+h'))}{\cosh(\beta(J-h-h'))} = e^{\frac{2}{q-1}\beta h'} \qquad (8.129)$$

The self-consistency equations can be solved numerically. The interesting question, as usual, pertains to spontaneous magnetic order. Hence we now focus on the case $h = 0$. A trivial solution is $h' = 0$. This gives $S_0 = 0$ and hence $m = 0$, and corresponds to a non-magnetic state.

In order to see if there are other solutions one can consider both sides of 8.129 as a function of h'. The right hand side is a simple exponentially increasing function, which has a positive curvature. The left hand side is also an increasing function, as can be seen from

$$\frac{\partial}{\partial h'}\frac{\cosh(\beta(J+h'))}{\cosh(\beta(J-h'))} =$$

$$\beta\frac{\sinh(\beta(J+h'))\cosh(\beta(J-h'))+\cosh(\beta(J+h'))\sinh(\beta(J-h'))}{\cosh^2(\beta(J-h'))} =$$

$$\beta\frac{\sinh(2\beta J)}{\cosh^2(\beta(J-h'))} \qquad (8.130)$$

This is positive indeed. Since the hyperbolic cosine is minimal at zero, we see that this derivative is maximal for $h' = J$. Hence the left hand side of 8.129 is one for $h' = 0$, increases with a positive curvature until $h' = J$, and then increases with a negative curvature until it approaches $e^{2\beta J}$ in the limit $h' \to \infty$. Plots for two cases are shown in the following figures.
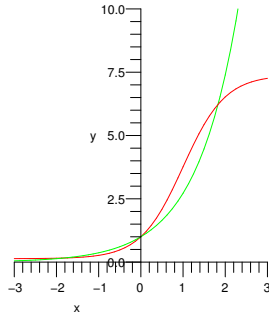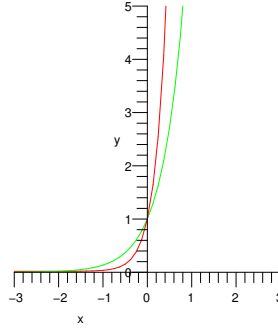

Figure 8.3: $\beta = 1$, $J = 1$, $q = 3$

Figure 8.4: $\beta = 2$, $J = 1$, $q = 3$

If the slope of the left hand side of 8.129 at $h' = 0$ is larger than the slope of the exponent, there has to be at least one intersection of the two curves for some positive value of h'. Hence if

$$\beta \frac{\sinh(2\beta J)}{\cosh^2(\beta J)} > \frac{2\beta}{q - 1} \tag{8.131}$$

or

$$\tanh(\beta J) > \frac{1}{q - 1} \tag{8.132}$$

there is a solution of 8.129 with a value of $h' > 0$. What happens if the slope of the left hand side is smaller? One has to look at the curvatures. It turns out that there are no solutions in that case. The condition for more solutions will always happen if $q > 2$, since for small values of the temperature the hyperbolic tangent approaches the value one. For $q = 2$, however, there is no solution. Hence the one dimensional Ising chain in the Bethe approximation does not show a phase transition. But the point $T = 0$ is still special, since in the limit the equation is obeyed.

The critical temperature in the Bethe approximation is therefore given by

$$k_B T_c = \frac{J}{\coth^{-1}(q - 1)} \tag{8.133}$$

which can be compared to the mean field result of $k_B T_c = qJ$. For a two-dimensional square lattice $q = 4$, and we find $k_B T_c = 2.885J$. This is much closer to the exact answer $2.269J$ than the mean field result $4J$. The Bethe approximation greatly improves the estimates of the critical temperature.

We can solve $\coth(\beta J) = q - 1$ using

$$\frac{e^{\beta J} + e^{-\beta J}}{e^{\beta J} - e^{-\beta J}} = q - 1 \tag{8.134}$$

$$e^{\beta J}(2 - q) = e^{-\beta J}(-q) \tag{8.135}$$

$$2\beta J = \log(\frac{q}{q - 2}) \tag{8.136}$$

which leads to

$$k_B T_c = J \frac{2}{\log(\frac{q}{q-2})} \tag{8.137}$$

In the limit $q \gg 1$ we have $\log(\frac{q}{q-2}) \approx \log(1 + \frac{2}{q}) \approx \frac{2}{q}$ and hence $k_B T_c \approx qJ$ like in mean field. The mean field results is best when the number of neighbors is large, which is certainly true in high dimensional systems. As we saw in Thermodynamics, mean field theory is exact for high dimensions.

The final, important, question is the following. When there is a solution with a non-zero value of h', does this solution have the lowest energy? Now we need to address the question of finding $f(h')$. The energy of the cluster is given by

$$\langle E_c \rangle = -Jq\langle \sigma_0 \sigma_j \rangle - h(q+1)m - h'm + f(h') \tag{8.138}$$

and the correlation function $\langle \sigma_0 \sigma_j \rangle$ follows from

$$\langle \sigma_0 \sigma_j \rangle = \frac{1}{\mathcal{Z}_c} \sum_{\{\sigma_0, \cdots, \sigma_q\}} \sigma_0 \sigma_j e^{-\beta E_c(\sigma_0, \cdots, \sigma_q)} = \frac{S_{0j}}{\mathcal{Z}_c} \tag{8.139}$$

Following the calculation for S_j we find that

$$S_{0j} = e^{\beta h}[2\cosh(\beta(J + h + h'))]^{q-1}[2\sinh(\beta(J + h + h'))] -$$

$$e^{-\beta h}[2\cosh(\beta(-J + h + h'))]^{q-1}[2\sinh(\beta(-J + h + h'))] \tag{8.140}$$

where the only difference with S_j is the minus sign in front of the second term. So here we see another advantage of the Bethe approximation, we do have an estimate for the correlation function.

The value of $f(h')$ can be determined by considering a situation where all spins are equal to m and are uncorrelated, in which case we know the energy. This is too tedious, though. But if we think in terms of Landau theory and consider h' to be a parameter in the energy expression, we can deduce that with three solutions for h' the order has to be minimum-maximum-minimum, and hence the $h' = 0$ solution, which is always the middle one, has to be a maximum.

8.7 Problems for chapter 8

PROBLEM 1.

A model of a binary alloy can be constructed similar to an Ising model. Assume that the possible atomic sites i are fixed and that the actual atom on site i is either type A or type B. The effects of the chemical bond extend only to nearest neighbors. The energy of an AA-bond is ϵ_{AA}, of an AB-bond ϵ_{AB}, and of a BB-bond ϵ_{BB}. The parameter $\epsilon = \frac{1}{2}\epsilon_{AA} + \frac{1}{2}\epsilon_{BB} - \epsilon_{AB}$ is useful. The concentration of A atoms is $c_A = \frac{N_A/}{N}$ and the concentration of B atoms is $c_B = 1 - c_A$. Introduce variables σ_i related to the number of A atoms n_{Ai} on site i by $n_{Ai} = \frac{1}{2}(1 + \sigma_i)$. Obviously, $\sigma_i = \pm 1$ and $n_{Bi} = \frac{1}{2}(1 - \sigma_i)$.

A. Calculate the energy of the binary alloy in a state $\{\sigma_1, \cdots, \sigma_N\}$.

B. Define variables J and h in such a way that this expression looks like the Ising model.

C. If $J > 0$ one finds a critical temperature T_c. What happens below T_c in this case?

D. Suppose $J < 0$. What is the structure of the alloy at low temperatures?

PROBLEM 2.

Consider a one-dimensional Ising model. Assume that $J < 0$. Introduce new spin-variables τ_i related to the σ_i variables by $\tau_i = (-1)^i \sigma_i$.

A. Calculate T_c for this system.

B. What is happening below T_c?

PROBLEM 3.

Consider the Heisenberg model for spin $\frac{1}{2}$ particles. The Hamiltonian is given by

$$\mathcal{H} = -J \sum_{<ij>} \vec{\mathcal{S}}_i \bullet \vec{\mathcal{S}}_j$$

The spin operators $\vec{\mathcal{S}}$ are the standard Pauli matrices $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & -\imath \\ \imath & 0 \end{pmatrix}$, and $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$. The state of each individual spin is a two-spinor $\begin{pmatrix} \gamma \\ \mu \end{pmatrix}$. The state of the whole system is now given by the direct product of N spinors $|i>$, in the form $|1> |2> \cdots |N>$. Assume that the density matrix ρ can be written as a direct product of 2×2 matrices $\widetilde{\rho}$ like we did for the Ising model. Use the same parametrization for $\widetilde{\rho}$ as for the Ising model.

A. Calculate the internal energy as a function of a and m. What is the difference compared with the result for the Ising model?

B. Find three equations relating the matrix elements $\log(\widetilde{\rho})_{ij}$ and a and m by minimizing the free energy with respect to a and m.

C. A fourth equation can be found using $\log \det A = \mathrm{Tr} \ \log A$ for arbitrary matrices. Use this together with the results of B to find an equation for a and m of the form $\log \widetilde{\rho} =$ matrix containing a and m.

D. Show that $a = 0$ corresponds to the thermodynamic equilibrium state and that the results for the Heisenberg and Ising model in this approximation are the same.

PROBLEM 4.

Probably the simplest cluster one can imagine in a square lattice is a square of four atomic sites. Treat the interactions between these four sites exactly but treat the interactions with the remainder of the lattice in a mean-field approach.

A. How many inequivalent sites are there within this cluster?

B. The effective field on each site is found by assuming that the spin on all neighboring sites that are not inside the cluster is equal to m, the average magnetization. What is the cluster Hamiltonian?

C. What is the self-consistency condition?

D. Calculate T_c. How does this compare with the mean-field value $k_B T_c = 4J$ and the cluster value $k_B T_c = 2.885J$?

PROBLEM 5.

A more complicated cluster in a square lattice is a square of nine sites.

A. How many inequivalent sites are there in this cluster?

B. We need to impose conditions of the form $< \sigma_i >=< \sigma_j >$. How many constraints does this give?

C. What is the cluster Hamiltonian? How many effective fields are there?

D. Indicate how you would try to solve this cluster problem.

PROBLEM 6.

Consider a linear chain with alternating spin one-half and spin one atoms. The state of this system is given by $\{s_1, s_2, \cdots\}$ where $s_i = \pm 1$ if i is odd and $s_i = -2, 0, 2$ for i even. The energy of such a state is given by

$$E\{s_1, s_2, \cdots\} = -J \sum_i s_i s_{i+1} - h \sum_i s_i$$

The values of the average magnetization on the two types of sites are given by $m_{1/2}$ and m_1. Generalize the mean field approach to construct two coupled equations for these quantities and calculate T_c for this system.

PROBLEM 7.

The positions of the sites on a lattice are given by so-called Bravais lattice vectors $\vec{R}_i$. The Ising model can be generalized to include interactions between all spins. Typically, the strength of the interaction depends only on the distance between the spins:

$$E\left\{\sigma_i\right\} = -\frac{1}{2}\sum_{i\neq j} J(|\vec{R}_i - \vec{R}_j|)\sigma_i\sigma_j - h\sum_i \sigma_i$$

Calculate T_c in the mean field approximation.

PROBLEM 8.

Consider the following generalization of the Ising model. The value of the spin parameter S_i on lattice site i can be ± 1 or 0. The energy of the configuration $\{S_i\}$ is

$$E\left\{S_i\right\} = -J\sum_{<ij>} S_iS_j - h\sum_i S_i$$

Use the density operator approach. Assume that the fluctuations in the spins are independent, hence

$$< S_1, S_2, \cdots |\rho| S_1', S_2', \cdots >=< S_1|\widetilde{\rho}|S_1' >< S_2|\widetilde{\rho}|S_2' > \cdots$$

Derive a self-consistency equation for the average moment M on each site. Show that T_c is proportional to qJ.

PROBLEM 9.

Consider the following generalization of the Ising model. The value of the spin parameter S_i on lattice site i can be ± 1 or 0. The energy of the configuration $\{S_i\}$ is

$$E\left\{S_i\right\} = -J\sum_{<ij>} S_iS_j - h\sum_i S_i$$

Using the mean field approximation, derive a self-consistency equation for the average moment M on each site.

PROBLEM 10.

Consider a two-dimensional triangular lattice (q=6). A cluster in this lattice is formed by a triangle of three sites. The interactions between the atoms in this cluster are treated exactly. The effect of the rest of the lattice is treated in the mean field approach. Evaluate T_c in this case. Compare your result to the results of the mean-field and Bethe cluster approach.

Chapter 9

General methods: critical exponents.

9.1 Introduction.

In this chapter we will consider a variety of methods that can give us approximations in statistical mechanics. In the previous chapter we looked at the mean field approximation and its cousins. We found that mean field gives a reasonable description if there is a phase transition. It also leads to a decent estimate of the transition temperature, except for the one dimensional Ising chain. Cluster approximations improve the estimates of the transition temperature. All in all these methods are very useful to describe phase transitions and help us understand different models. Where they fail, however, is in providing values for critical exponents. They always give the same values, no matter what the model is, due to the mean field nature. This is easy to understand. In a mean field model we approximate effects further away by an average. When correlation lengths become large, a cluster will always sample that average value and essentially see a mean field. So on this chapter we focus in more detail on correlation functions and critical exponents.

There are two possible goals of calculations in statistical mechanics. One useful result is to be able to find certain thermal averages of quantities. But the real goal is to find the appropriate free energy, since everything follows from there. That second goal is more difficult. For example, in the Bethe approximation we needed to include a difficult term $f(h')$ in the energy. For the calculation of averages like $\langle \sigma_0 \sigma_j \rangle$ that term dropped out. But without that term we cannot evaluate free energies.

9.2   Integration over the coupling constant.

Since exact answers are hard to get for most realistic models, it is useful to know a variety of tricks to obtain good approximations. The procedure of integration over the coupling constant is one of those tricks. We will not investigate the relations between this method and others, but simply remark that this method is another way of obtaining a correlated density matrix. We use the Ising model again for illustration. We will break up the Hamiltonian into two parts:

$$\mathcal{V} = - \sum_{\{\sigma_1, \cdots, \sigma_N\}} |\sigma_1, \cdots, \sigma_N > \sum_{<ij>} \sigma_i \sigma_j < \sigma_1, \cdots, \sigma_N| \tag{9.1}$$

$$\mathcal{H}_0 = -h \sum_{\{\sigma_1, \cdots, \sigma_N\}} |\sigma_1, \cdots, \sigma_N > \sum_i \sigma_i < \sigma_1, \cdots, \sigma_N| \tag{9.2}$$

and study the general Hamiltonian

$$\mathcal{H} = \mathcal{H}_0 + \lambda \mathcal{V} \tag{9.3}$$

In this division the term $\mathcal{H}_0$ contains all the easy terms. In other systems it would include the kinetic energy in many cases. The important point is that the operator $\mathcal{H}_0$ can be diagonalized, we can find eigenvalues and eigenvectors. The operator $\mathcal{V}$ contains all the hard terms, the ones that we do not know how to deal with, and have approximated before. We are obviously interested in the case $\lambda = J$. We will, however, treat λ as a variable parameter. The division is made in such a way that the problem is simple for $\lambda = 0$. Again, we always make sure that for $\mathcal{H}_0$ the solutions are known.

The free energy as a function of λ is

$$\mathcal{G}(\lambda) = -k_B T \log \left(\text{Tr } e^{-\beta \mathcal{H}_0 - \beta \lambda \mathcal{V}} \right) \tag{9.4}$$

and the derivative with respect to λ is

$$\frac{d\mathcal{G}}{d\lambda} = -\frac{k_B T}{\text{Tr } e^{-\beta \mathcal{H}_0 - \beta \lambda \mathcal{V}}} \text{Tr } \frac{d}{d\lambda} e^{-\beta \mathcal{H}_0 - \beta \lambda \mathcal{V}} \tag{9.5}$$

One always has to be careful when taking derivatives of operators. In this case $\mathcal{H}_0$ and $\mathcal{V}$ commute and the exponent of the sum is the product of the single exponents.

$$[\mathcal{H}_0, \mathcal{V}] = 0 \quad \Rightarrow \quad e^{-\beta \mathcal{H}_0 - \beta \lambda \mathcal{V}} = e^{-\beta \mathcal{H}_0} e^{-\beta \lambda \mathcal{V}} \tag{9.6}$$

The derivative with respect to λ affects the second term only, and we have

$$\frac{d}{d\lambda} e^{-\beta \lambda \mathcal{V}} = \frac{d}{d\lambda} \sum_{n=0}^{\infty} \frac{1}{n!} (-\beta \lambda)^n \mathcal{V}^n = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} \lambda^{n-1} (-\beta)^n \mathcal{V}^n = -\beta \mathcal{V} e^{-\beta \lambda \mathcal{V}}$$

$$\tag{9.7}$$

which leads to

$$\frac{d\mathcal{G}}{d\lambda} = \frac{\text{Tr } \mathcal{V}e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}}}{\text{Tr } e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}}} \tag{9.8}$$

It is useful to digress here and show that this result is true in general. We have

$$\frac{d}{d\lambda}e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}} = \sum_{n=0}^{\infty} \frac{1}{n!}\frac{d}{d\lambda}\left[-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}\right]^n \tag{9.9}$$

Because operators do not commute we need to keep track of the order of the operators and derivatives:

$$\frac{d}{d\lambda}e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}} = \sum_{n=1}^{\infty} \frac{1}{n!}\sum_{m=1}^{n}\left[-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}\right]^{m-1}\left[-\beta\mathcal{V}\right]\left[-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}\right]^{n-m} \tag{9.10}$$

The derivative of the very first term is zero, so the sum starts at $n = 1$. Next we realize that we need the trace of this expression, and that the trace obeys $\text{Tr}(ABC) = \text{Tr}(CAB)$. This gives

$$\text{Tr } \frac{d}{d\lambda}e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}} = \sum_{n=1}^{\infty} \frac{1}{n!}\sum_{m=1}^{n}\text{Tr }\left[-\beta\mathcal{V}\right]\left[-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}\right]^{n-1} \tag{9.11}$$

Now we can do the sum on m, it simply gives a factor n and find

$$\text{Tr } \frac{d}{d\lambda}e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!}\text{Tr }\left[-\beta\mathcal{V}\right]\left[-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}\right]^{n-1} \tag{9.12}$$

Finally, we reassemble the sum and obtain

$$\text{Tr } \frac{d}{d\lambda}e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}} = \text{Tr }\left[-\beta\mathcal{V}\right]e^{-\beta\mathcal{H}_0-\beta\lambda\mathcal{V}} \tag{9.13}$$

leading to the required result. Integrating over the coupling constant λ then gives the following formula for the free energy

$$\mathcal{G}(\lambda) = \mathcal{G}(0) + \int_0^J d\lambda \frac{d\mathcal{G}}{d\lambda} = \mathcal{G}(0) + \int_0^J \frac{d\lambda}{\lambda}\langle\lambda\mathcal{V}\rangle_\lambda \tag{9.14}$$

where it is standard practice to put a factor λ inside the ensemble average. Remember that $\mathcal{G}(0)$ is a known quantity.

The formula above tells us that once we know $\langle\mathcal{V}\rangle_\lambda$ we can calculate the free energy. That is not surprising, because once we have the pair correlation function and hence the potential we essentially have solved the problem. The trick is not how to get the pair correlation function exactly, but how to approximate it. Using an approximation to the pair correlation function will give us in general an improved free energy.

As an example we will consider the result obtained in the Bethe approximation for the one-dimensional Ising model without an external field ($h = 0$). The effective field h' is a function of λ and is determined by

$$\frac{\cosh \beta(\lambda + h')}{\cosh \beta(\lambda - h')} = e^{h'2\beta} \tag{9.15}$$

But we already know that there are no phase transitions, independent of the value of λ, and hence $h' = 0$. In the Bethe approximation the only interaction term pertains to the bonds between the central site and its neighbors, and hence the interaction energy of a cluster is

$$\langle \lambda \mathcal{V} \rangle_\lambda = \langle -\lambda \sigma_0 \sum_{i=1}^{2} \sigma_i \rangle_\lambda \tag{9.16}$$

Now we use the results from the previous chapter to calculate this average. In order to find the potential energy we calculate the cluster energy from the partition function. The cluster partition function was determined in the previous chapter, and in this case is given by

$$\mathfrak{Z}_c = 8 \cosh^2(\beta \lambda) \tag{9.17}$$

and the cluster energy E_c follows from

$$E_c = -\frac{\partial}{\partial \beta} \log \mathfrak{Z}_c = -2\lambda \tanh \beta \lambda \tag{9.18}$$

Since $h = 0$ we have $\mathcal{H}_0 = 0$, and all the internal energy is due to the interaction term. The internal energy of the whole system of N particles is

$$U = \langle \lambda \mathcal{V} \rangle = \frac{1}{2} N E_c \tag{9.19}$$

The factor one-half is needed to avoid double counting of all bonds. A cluster contains two bonds! As a result we find that

$$\mathcal{G}(J) = \mathcal{G}(0) - N \int_0^J d\lambda \tanh \beta \lambda = \mathcal{G}(0) - N k_B T \log \cosh \beta J \tag{9.20}$$

Next, we determine the free energy at $\lambda = 0$. There is only an entropy term, since the internal energy is zero. Without a magnetic field and without interactions all configurations are possible, thus $S = N k_B \log 2$, and we find after combining the two logarithmic terms:

$$\mathcal{G}(J) = -N k_B T \log(2 \cosh \beta J) \tag{9.21}$$

The average of the interaction energy was obtained in the Bethe approximation. It turns out, however, that the calculated free energy is exact! The reason for that is the simplicity of the system. There is no phase transition, and the correlation function in the cluster approximation is actually correct.

Note that in this case we could also have obtained the same result by calculating the entropy from the specific heat

$$S = \int_0^T \frac{dT'}{T'} \frac{dU}{dT'} = N k_B J \int_\beta^\infty \frac{\beta' d\beta'}{\cosh^2 \beta' J} \qquad (9.22)$$

This integral is harder to evaluate, however. But it can be done, especially since we already know the answer!

We can also analyze the Ising model in some more detail. The energy of a configuration is

$$E\{\sigma_1, \sigma_2, \cdots\} = -J \sum_{<ij>} \sigma_i \sigma_j - h \sum_i \sigma_i \qquad (9.23)$$

The interaction term is large, of course, and the division we made before is not optimal. We would like to subtract some average value of the spins and write

$$E\{\sigma_1, \sigma_2, \cdots\} = -J \sum_{<ij>} (\sigma_i - \mu)(\sigma_j - \mu) - (J\mu q + h) \sum_i \sigma_i + J\mu^2 \frac{1}{2} Nq \quad (9.24)$$

and we define this for a variable coupling constant via

$$E_\lambda\{\sigma_1, \sigma_2, \cdots\} = -\lambda \sum_{<ij>} (\sigma_i - \mu)(\sigma_j - \mu) - (J\mu q + h) \sum_i \sigma_i + J\mu^2 \frac{1}{2} Nq \quad (9.25)$$

It would be natural to define μ as the average magnetization. But that value depends on the value of λ, which makes the λ dependence of the Hamiltonian quite complicated. Hence we need to choose μ as some kind of average magnetization, and an appropriate value of the coupling constant.

The reference system has energy eigenvalues given by

$$E_0\{\sigma_1, \sigma_2, \cdots\} = -(J\mu q + h) \sum_i \sigma_i + J\mu^2 \frac{1}{2} Nq \qquad (9.26)$$

and the partition function is

$$\mathfrak{Z}_0(T, h, N) = e^{-\beta J\mu^2 \frac{1}{2} Nq} \sum_{\sigma_1, \sigma_2, \cdots} e^{\beta(J\mu q + h) \sum_i \sigma_i} \qquad (9.27)$$

which gives

$$\mathfrak{Z}_0(T, h, N) = e^{-\beta J\mu^2 \frac{1}{2} Nq} [2 \cosh \beta(J\mu q + h)]^N \qquad (9.28)$$

and the free energy is

$$\mathcal{G}_0(T, h, N) = J\mu^2 \frac{1}{2} Nq - N k_B T \log(2 \cosh \beta(J\mu q + h)) \qquad (9.29)$$

The interaction term is now

$$\langle \mathcal{V} \rangle_\lambda = -\langle \sum_{<ij>} (\sigma_i - \mu)(\sigma_j - \mu) \rangle_\lambda = -\frac{1}{2} Nq \langle (\sigma_i - \mu)(\sigma_j - \mu) \rangle_\lambda \qquad (9.30)$$

We now introduce the average magnetization m_λ and write

$$\langle \mathcal{V} \rangle_\lambda = -\frac{1}{2} Nq \langle (\sigma_i - m_\lambda)(\sigma_j - m_\lambda) \rangle_\lambda - \frac{1}{2} Nq \langle 2(m_\lambda - \mu)\sigma_i \rangle_\lambda + \frac{1}{2} Nq(m_\lambda^2 - \mu^2)$$
$$(9.31)$$

Next we approximate the correlations between the fluctuations around the average by zero and obtain

$$\langle \mathcal{V} \rangle_\lambda \approx -\frac{1}{2} Nq \langle 2(m_\lambda - \mu)\sigma_i \rangle_\lambda + \frac{1}{2} Nq(m_\lambda^2 - \mu^2) \qquad (9.32)$$

or

$$\langle \mathcal{V} \rangle_\lambda \approx -Nq(m_\lambda - \mu)m_\lambda + \frac{1}{2} Nq(m_\lambda^2 - \mu^2) \qquad (9.33)$$

or

$$\langle \mathcal{V} \rangle_\lambda \approx -\frac{1}{2} Nq(m_\lambda - \mu)^2 \qquad (9.34)$$

This gives

$$\left(\frac{\partial \mathcal{G}}{\partial \lambda} \right)_{T,h,N} = -\frac{1}{2} Nq(m_\lambda - \mu)^2 \qquad (9.35)$$

The magnetic moment is related to the Gibbs like free energy via

$$Nm_\lambda = -\left(\frac{\partial \mathcal{G}}{\partial h} \right)_{T,N} \qquad (9.36)$$

and hence

$$N \left(\frac{\partial m}{\partial \lambda} \right)_{h,T} = -\left(\frac{\partial^2 \mathcal{G}}{\partial h \partial \lambda} \right)_{T,N} = Nq(m - \mu) \left(\left(\frac{\partial m}{\partial h} \right)_{\lambda,T} - \left(\frac{\partial \mu}{\partial h} \right)_T \right) \quad (9.37)$$

where we have $m(h, \lambda, T)$ and $\mu(h, T)$. This equation shows that if $m_\lambda = \mu$ the magnetization will not change as a function of λ. Also, because for large fields the magnetization approaches one, we see that if $m > \mu$ one needs $\frac{dm}{dh} < \frac{d\mu}{dh}$ and hence the right hand side of the partial differential equation is zero. This means that for large values of λ the solutions for m will approach μ. Now suppose that we use the mean field solution for the magnetization with $\lambda = J$ for μ. Then we conclude that the true magnetization of the system will be smaller than the mean field value, because we integrate to J only and not to infinity. Of course, this uses an approximation for the correlation function, so the conclusion does not have to be general.

9.3 Critical exponents.

Most of the interesting behavior of the phase transitions occurs near the critical temperature. We have shown that the mean field results are equivalent to the Landau theory discussed in the thermodynamics, and hence near T_c we have the magnetization proportional to the square root of $T_c - T$, and hence the critical exponent β is $\frac{1}{2}$ in mean field theory. In the Bethe approximation we have to find the value of h' for temperatures just below T_c. It turns out that in the Bethe approximation the phase transition is also second order, and hence near T_c the value of h' will be small. Hence we can expand the equation for h' in terms of h' and we find

$$\frac{\cosh \beta(J + h')}{\cosh \beta(J - h')} \approx 1 + h' 2\beta \tanh \beta J + (h')^2 2\beta^2 \tanh^2 \beta J + \mathcal{O}(h')^3 \quad (9.38)$$

$$e^{\frac{2\beta h'}{q-1}} \approx 1 + h' \frac{2\beta}{q-1} + (h')^2 \frac{2\beta^2}{(q-1)^2} + \mathcal{O}(h')^3 \quad (9.39)$$

The condition for a phase transition is the equality of the derivatives at $h' = 0$. This gave our self-consistency condition. If this condition is satisfied, the second derivatives are also the same! It is not hard to show that the third order derivatives are really different. Therefore, when h' is small, it is determined by

$$h' 2\beta \tanh \beta J + (h')^2 2\beta^2 \tanh^2 \beta J + a(h')^3 = h' \frac{2\beta}{q-1} + (h')^2 \frac{2\beta^2}{(q-1)^2} + b(h')^3$$
$$(9.40)$$

where a and b are different numbers. At the critical temperature $h' = 0$ is a triple solution. This has to be the case, since at a temperature just below the critical temperature the solutions are $h' = 0$ and $h' = \pm\epsilon$, where ϵ is a small number. These three solutions merge together at the critical temperature. At a temperature slightly below the critical temperature we can write

$$0 = h' \left((T - T_c)(c + dh') + (a - b)(h')^2 \right) \quad (9.41)$$

where c and d are constants. This shows that near T_c the effective field h' is proportional to $\sqrt{T_c - T}$. The equations in the previous chapter can be used to show that for small values of h the magnetization m is proportional to h', and hence we find that the critical exponent β is equal to $\frac{1}{2}$ in the Bethe approximation, just as it is in mean field theory.

The susceptibility χ is defined by

$$\chi(h, T) = \left(\frac{\partial m}{\partial h} \right)_T \quad (9.42)$$

and can be calculated easily in the mean field approximation from

$$m = \tanh \beta(qJm + h) \quad (9.43)$$

Differentiating both sides and using $h = 0$ gives

$$\chi(h = 0, T) = \frac{\beta}{\cosh^2 \beta J q m} \left\{ qJ\chi(h = 0, T) + 1 \right\} \tag{9.44}$$

If T is larger than T_c we have $m = 0$ and we find that near T_c

$$\chi(0, T) \approx \frac{1}{k_B(T - T_c)} \tag{9.45}$$

If, on the other hand, T is less than T_c, the magnetization m is non-zero. Near T_c the value of m is small, however, and we can expand the square of the hyperbolic cosine

$$\cosh^2 \beta q J m \approx 1 + (\beta q J)^2 \alpha^2 (T_c - T) \tag{9.46}$$

where we used that near the critical temperature $m \approx \alpha \sqrt{T_c - T}$. The value of α follows from the results in the previous chapter and we find $\alpha^2 = \frac{3}{T_c}$. From 9.44 we get

$$\left\{ \cosh^2 \beta q J m - \beta q J \right\} \chi(h = 0, T) = \beta \tag{9.47}$$

Since $\beta q J = \frac{T_c - T}{T} \approx \frac{1}{T_c}(T_c - T)$ near T_c we find

$$\chi(0, T) \approx \frac{1}{2k_B(T_c - T)} \tag{9.48}$$

Hence near the critical temperature we find in general that

$$\chi(0, T) = A_{\pm} |T - T_c|^{-\gamma} \tag{9.49}$$

where the value of the critical exponent $\gamma = 1$, just like we found in mean field theory in chapter four.

The calculation in the Bethe approximation is harder. We have from the previous chapter:

$$m = \frac{\cosh \beta(J + h' + h) - \cosh \beta(J - h' - h)}{\cosh \beta(J + h' + h) + \cosh \beta(J - h' - h)} \tag{9.50}$$

with self-consistency equation 8.129:

$$\frac{\cosh(\beta(J + h + h'))}{\cosh(\beta(J - h - h'))} = e^{\frac{2}{q-1}\beta h'} \tag{9.51}$$

The first of these two equations gives

$$\beta^{-1} \left(\frac{\partial m}{\partial h} \right)_T =$$

$$[(\sinh \beta(J + h' + h) + \sinh \beta(J - h' - h))(\cosh \beta(J + h' + h) + \cosh \beta(J - h' - h)) -$$

$$(\cosh \beta(J + h' + h) - \cosh \beta(J - h' - h))(\sinh \beta(J + h' + h) - \sinh \beta(J - h' - h))]$$

$$\frac{1}{(\cosh \beta(J + h' + h) + \cosh \beta(J - h' - h))^2} \left(\left(\frac{\partial h'}{\partial h} \right)_T + 1 \right) \tag{9.52}$$

Above T_c we have for $h = 0$ that also $h' = 0$ and hence

$$\left(\frac{\partial m}{\partial h} \right)_T = \tanh(\beta J) \left(\left(\frac{\partial h'}{\partial h} \right)_T + 1 \right) \tag{9.53}$$

On the other hand, from the self-consistency equation we find

$$\beta \frac{\sinh(\beta(J + h + h')) \cosh(\beta(J - h - h')) + \cosh(\beta(J + h + h')) \sinh(\beta(J - h - h'))}{\cosh^2(\beta(J - h - h'))}$$

$$\left(\left(\frac{\partial h'}{\partial h} \right)_T + 1 \right) = \frac{2}{q - 1} \beta e^{\frac{2}{q-1}\beta h'} \left(\frac{\partial h'}{\partial h} \right)_T \tag{9.54}$$

which for $h = 0$ and $h' = 0$ leads to

$$\tanh(\beta J) \left(\left(\frac{\partial h'}{\partial h} \right)_T + 1 \right) = \frac{1}{q - 1} \left(\frac{\partial h'}{\partial h} \right)_T \tag{9.55}$$

or

$$\left(\frac{\partial h'}{\partial h} \right)_T = \frac{\tanh(\beta J)}{\frac{1}{q-1} - \tanh(\beta J)} \tag{9.56}$$

which shows that $\left(\frac{\partial h'}{\partial h} \right)_T$ diverges like $(T - T^c)^{-1}$ and hence that the susceptibility diverges in that manner, again giving $\gamma = 1$. One can also calculate the behavior at the other side of the critical temperature, but that is even more messy.

The exact solutions for the Ising model show that there is no phase transition in one dimension, that $\gamma = \frac{7}{4}$ in two dimensions and that $\gamma \approx 1.25$ (and perhaps equal to $\frac{5}{4}$?) in three dimensions. The higher the number of dimensions, the closer the value is to the mean field result and the result of the Bethe approximation.

We could continue calculating all other critical exponents for the Ising model in both the mean-field and the Bethe approximation. Since we have shown that the mean field approximation is equivalent to a second order Landau theory, we already know the answers. The answers for the Bethe approximation will be the same! This is due to the following. Near the critical temperature the only important length scale is the correlation length ξ, which diverges. If the correlation length is smaller than the radius of the cluster used in the Bethe

approximation, the answers in the Bethe approximation and in mean-field theory will be different. If, on the other hand, the correlation length ξ is much larger than the size of the cluster, the size of the cluster is not important and the problem is essentially solved in a mean-field approach. Since ξ diverges near T_c, the Bethe approximation is equivalent to a mean field theory in this limit. Therefore, the Bethe approximation will give the same critical exponents as mean-field theory. In order to make a real improvement, we need to take into account correlations of infinite range.

We can improve the analysis, however. We will use the susceptibility as an example, but the analysis woks for all critical exponents. Near the critical temperature we have $\chi \approx A_\pm |T - T_c|^\gamma$, plus higher order terms. This means that near the critical temperature we have $\log(\chi) = \approx \log(A_\pm) + \gamma \log |T - T_c|$. The second term will dominate if we are close enough to the critical temperature. Hence we have

$$\frac{\log(\chi)}{\log |T - T_c|} \quad \to \gamma \tag{9.57}$$

near the critical temperature. Hence if we plot the ratio on the left as a function of temperature, we should be able to extract the value of γ.

Very simply stated, so what can be wrong. First of all, we assume that we know the value of the critical temperature. That might be true or not. What happens if we are wrong? If we do not use the exact value of the critical temperature a plot of the ratio of the logarithms will show two peaks. The ratio will go to zero when the denominator goes to infinity, or when the temperature reaches the erroneous value of the critical temperature. The ratio will go to infinity when the enumerator diverges, at the true value of the critical temperature. For example, see figure 9.1 where we have used $\chi = \frac{1}{|t-1|}(1 + 0.1|t - 1|)$ and moved T_c by a small amount to $T_c = 1.01$. We can use this behavior to find the critical temperature. If we have sufficient data near the phase transition we try different values of T_c and see when the peak and zero merge.

But what is near the critical temperature? One could guess that the correction in χ is a factor of the form $1 + B_\pm |T - T_c|$. This gives an extra term in the denominator equal to the logarithm of this expression, and near the critical temperature we now find

$$\frac{\log(\chi)}{\log |T - T_c|} \quad \to \gamma + B_\pm \frac{|T - T_c|}{\log |T - T_c|} \tag{9.58}$$

The second term goes to zero, indeed. But how fast? We do not know. See figure 9.2 for a typical case. We used the same form as before, but now with the correct value of the critical temperature. In addition, the previous argument assumes that the behavior near the critical temperature is given by powers only. There are theoretical results where near the critical temperature we have logarithmic factors, $\chi \propto |T - T_c|^\gamma \log |T - T_c|$. Because the logarithm goes to infinity slower than any inverse power, such a factor does not change the value of the critical exponent, but it certainly changes the behavior near the phase transition. We
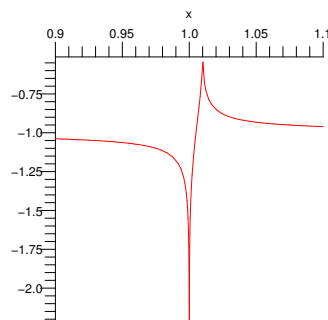
Figure 9.1: **Ideal case to find critical exponent, with wrong guess of critical temperature.**
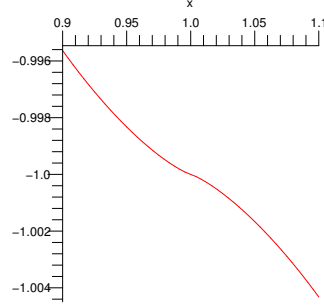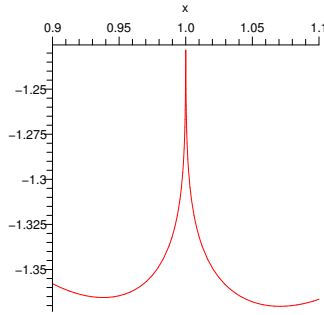
need to get very close to the phase transition to get rid of such terms, see figure 9.3.

Unfortunately, reality is even more complicated. The best way to discuss this is to think about the correlation length ξ, the length over which spins act coherently. This length diverges near the critical temperature. Using the correlation length we can define the range of temperatures needed to find the values of the critical exponents. If the range of the interactions in the material is R_{int} we need temperatures close enough to T_c so that $\xi(T - T_c) \gg R_{int}$. But in real life the correlation length cannot increase beyond the size of the sample. Suppose the sample has similar dimensions in all directions, given by a number R_s. If we have $\xi(T - T_c) > R_s$ the sample will act as a large molecule with a given susceptibility, which will remain the same. There is no divergence compensating the denominator, and the ratio of the logarithms goes to zero, see figure 9.4.

This shows that we can still find the critical exponent for an infinite sample by connecting both sides of the curve, but the precision does suffer. Also, if the sample dimensions are not the same in all directions we will see a transition from three to two to one to zero dimensional behavior!

Finally, we can use this to analyze cluster data. If the size of the cluster is larger than the effective range of the interactions the ratio of the logarithms will be close to the correct value, but if the correlation length becomes larger than the cluster size we will get the mean field results. See figure 9.5, where we changed the model so the exact results for γ is two.

The previous discussion shows the importance of the thermodynamic limit. For a finite sample the behavior right at the critical temperature is always normal or analytic. Nothing diverges. But close to the critical temperature we see the characteristics of an infinite sample. The larger the sample, the closer we can get to the critical temperature an still see the behavior of an infinite

Figure 9.2:   **Ideal case to find critical exponent.**

Figure 9.3:   **Non-analytic case to find critical exponent.**

sample. Hence the procedure to find the critical exponents needs to consider critical behavior close to the phase transition as a function of sample size. If we replace the outside of the sample by something else, the values right at the critical temperature change. For example, if we replace the outside by an effective medium we obtain mean field results right at T_c. But away from the critical temperature we get the characteristics of the sample, and the outside world does not play a role. For a finite sample there are no singularities at the phase transition, only smooth changes in behavior. If these changes become abrupt for in the thermodynamic limit, then we have a singularity and a phase transition.
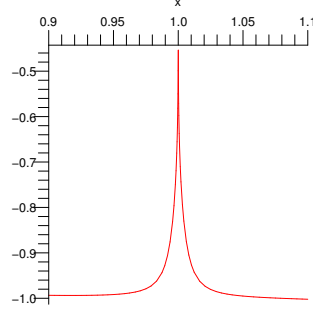
Figure 9.4:   **Finite sample case to find critical exponent.**
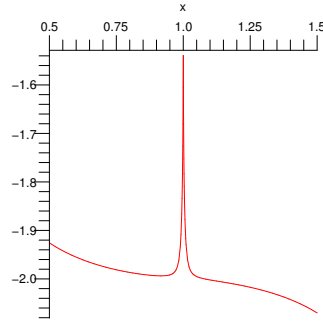
Figure 9.5:   **Cluster results to find critical exponent.**

9.4   Relation between susceptibility and fluctuations.

As we have remarked a few times, response functions are related to fluctuations. This can easily be shown for the susceptibility. By definition we have

$$\chi = \frac{\partial}{\partial h} \frac{\text{Tr } \mathcal{S}_0 e^{-\beta(\mathcal{H}-h\mathcal{M})}}{\text{Tr } e^{-\beta(\mathcal{H}-h\mathcal{M})}} \tag{9.59}$$

This derivative can be performed explicitly, since $\mathcal{H}$ and $\mathcal{M} = \sum_i \mathcal{S}_i$ commute. We find

$$\chi = \frac{\text{Tr } \mathcal{S}_0 \beta \mathcal{M} e^{-\beta(\mathcal{H}-h\mathcal{M})}}{\text{Tr } e^{-\beta(\mathcal{H}-h\mathcal{M})}} - \frac{\left[\text{Tr } \mathcal{S}_0 e^{-\beta(\mathcal{H}-h\mathcal{M})}\right]\left[\text{Tr}\beta\mathcal{M}\ e^{-\beta(\mathcal{H}-h\mathcal{M})}\right]}{\left(\text{Tr } e^{-\beta(\mathcal{H}-h\mathcal{M})}\right)^2} \quad (9.60)$$

or

$$\chi = \beta \sum_i \left\{ \langle \mathcal{S}_0 \mathcal{S}_i \rangle - \langle \mathcal{S}_0 \rangle \langle \mathcal{S}_i \rangle \right\} \quad (9.61)$$

This can be transformed to

$$\chi = \beta \sum_i \langle \mathcal{S}_0 - m \rangle \langle \mathcal{S}_i - m \rangle \quad (9.62)$$

which clearly shows that χ is directly related to the fluctuations in the spin variables. Here we used $\langle \mathcal{S}_0 \rangle = \langle \mathcal{S}_i \rangle = m$.

If fluctuations are uncorrelated we have $< AB >=< A >< B >$ which would seem to imply that for uncorrelated fluctuations $\chi = 0$. That is not correct, however. Fluctuations on the same site are always correlated, and in that case we would get $\chi =< (\mathcal{S}_0 - m)^2 >$, which is always positive indeed.

It is customary to define a spin-correlation function Γ_i by

$$\Gamma_i(T) = \langle \mathcal{S}_0 \mathcal{S}_i \rangle_T - \langle \mathcal{S}_0 \rangle_T \langle \mathcal{S}_i \rangle_T \quad (9.63)$$

If the spins at site 0 and site i are uncorrelated, $\Gamma_i = 0$. This spin-correlation function can also be expressed in terms of real coordinates $\vec{r}$, as long as we understand that the value $\Gamma(\vec{r}, T)$ is actually an average over a number of atomic sites in a volume ΔV which is small compared to the total volume of the crystal. In that case we find

$$\chi = \beta \int d^3 r \Gamma(\vec{r}, T) \quad (9.64)$$

If for large values of r the spin-correlation function is proportional to r^{-p} the integral is well-defined only if $p > 3$. Therefore, at T_c where the spin-correlation function diverges we need to have $p \leqslant 3$. At temperatures away from the critical temperature the spin correlation function does not diverge. If the correlation function would always be a power law, $\Gamma \propto r^{-p(T)}$, this would imply $p = 3$ at the critical temperature, since $p(T) > 3$ away from the critical point. That is not correct, however, because we also have an exponential factor, which disappears at the critical temperature. In general we write $\Gamma \propto r^{-p} e^{-\alpha(T)r}$ and we use $\alpha(T) > 0$ away from the critical point and hence $\alpha(T_c) = 0$. Since these are the only two possibilities and since we know from experiment that the first one is wrong, we find that correlations have to die out exponentially. The length scale corresponding to this exponential decay is called the correlation length, and at the critical point the correlation length diverges and becomes infinite.

If we have no correlations we find $\chi = \beta(< \mathcal{S}_0^2 > -m^2) = \beta(1 - m^2)$. Here we know $< \mathcal{S}_0^2 >= 0$ because the value of the spin variable is ± 1 and hence

the value of the square is always one. But the susceptibility diverges even in mean-field theory! This means that the spin-correlation in mean field theory for large distances is non-zero. There seems to be a contradiction here, but that is not true. The only assumption we make in mean field theory is that $\Gamma_i = 0$ for nearest-neighbors! For further neighbors we do not make any assumptions. We are not able to calculate the correlation function at larger distances. Similar, in the Bethe approximation we can only calculate the spin correlation function inside the cluster. In this case the condition involving h' is not readily expressed in terms of the spin-correlation function.

9.5 Exact solution for the Ising chain.

The Ising model in one dimension can be solved exactly. The same statement is true in two dimensions, but the solution is much more difficult. We will not do that here. First consider a chain of N atoms without an external field. The energy of a configuration $\{\sigma_1, \cdots, \sigma_N\}$ is given by

$$H(\sigma_1, \cdots, \sigma_N) = -J \sum_{i=1}^{N-1} \sigma_i \sigma_{i+1} \qquad (9.65)$$

The partition function is

$$\mathcal{Z}(T, N) = \sum_{\sigma_1, \cdots, \sigma_N} e^{\beta J \sum_i \sigma_i \sigma_{i+1}} \qquad (9.66)$$

which can be calculated by starting at the end via

$$\mathcal{Z}(T, N) = \sum_{\sigma_1} \sum_{\sigma_2} e^{\beta J \sigma_1 \sigma_2} \cdots \sum_{\sigma_N} e^{\beta J \sigma_{N-1} \sigma_N} \qquad (9.67)$$

Since the last factor is the only place where σ_N plays a role, the sum over σ_N can be performed giving

$$\sum_{\sigma_N} e^{\beta J \sigma_{N-1} \sigma_N} = 2 \cosh(\beta J) \qquad (9.68)$$

Note that σ_{N-1} drops out since it only takes the values ± 1 and since the hyperbolic cosine is symmetric! Therefore the partition function is

$$\mathcal{Z}(T, N) = 2 \cosh(\beta J) \sum_{\sigma_1} \sum_{\sigma_2} e^{\beta J \sigma_1 \sigma_2} \cdots \sum_{\sigma_{N-1}} e^{\beta J \sigma_{N-1} \sigma_N} \qquad (9.69)$$

This process is now repeated and leads to

$$\mathcal{Z}(T, N) = 2^{N-1} \cosh^{N-1}(\beta J) \sum_{\sigma_1} = 2^N \cosh^{N-1}(\beta J) \qquad (9.70)$$

and the free energy at $h = 0$ is

$$\mathcal{G}(T, N, h = 0) = -Nk_BT \log(2 \cosh(\beta J) + k_BT \log \cosh(\beta J) \tag{9.71}$$

In the thermodynamic limit the second term can be ignored, and the answer is the same as we found by integrating over the coupling constant.

The previous paragraph gave the results for the Ising chain without a magnetic field. What happens when a magnetic field is present? In that case the calculation is more complicated if we use the same boundary conditions. The main problem is that the term with the product of neighboring spins has fewer terms than the term connecting to the magnetic field. It is hard treat both terms at the same time.

It is possible, however, to change the boundary conditions, since we always want to take the limit $N \to \infty$. Therefore, we assume **periodic boundary conditions** by connecting spin N back with spin 1. Hence in the calculations we take $\sigma_0 = \sigma_N$ and $\sigma_1 = \sigma_{N+1}$. The energy in the extra bond is small compared to the total energy if N is large, and disappears in the thermodynamic limit. The energy of a configuration $\{\sigma_1, \cdots, \sigma_N\}$ is

$$H(\sigma_1, \cdots, \sigma_N) = -J \sum_{i=1}^{N} \sigma_i \sigma_{i+1} - \frac{h}{2} \sum_{i=1}^{N} [\sigma_i + \sigma_{i+1}] \tag{9.72}$$

where we made the magnetic contribution symmetric, which is possible because we are using periodic boundary conditions. This is written with a single summation in the form

$$H(\sigma_1, \cdots, \sigma_N) = -\sum_{i=1}^{N} f(\sigma_i, \sigma_{i+1}) \tag{9.73}$$

where we defined $f(\sigma, \sigma') = J\sigma\sigma' + \frac{h}{2}(\sigma + \sigma')$. The partition function is

$$\mathcal{Z}(T, N) = \sum_{\sigma_1, \cdots, \sigma_N} e^{\beta \sum_i f(\sigma_i, \sigma_{i+1})} = \sum_{\sigma_1, \cdots, \sigma_N} \prod_i e^{\beta f(\sigma_i, \sigma_{i+1})} \tag{9.74}$$

This looks like the product of a set of two by two matrices. Hence we define the matrix $\mathfrak{T}$ by

$$\mathfrak{T}(\sigma, \sigma') = e^{\beta J f(\sigma, \sigma')} \tag{9.75}$$

This matrix $\mathfrak{T}$ is often called a **transfer matrix**. It is a two by two matrix looking like

$$\mathfrak{T} = \begin{pmatrix} e^{\beta(J+h)} & e^{-\beta J} \\ e^{-\beta J} & e^{\beta(J-h)} \end{pmatrix} \tag{9.76}$$

The partition function in terms of $\mathfrak{T}$ is

$$3(T, h, N) = \sum_{\sigma_1, \cdots, \sigma_N} \mathfrak{T}(\sigma_1, \sigma_2)\mathfrak{T}(\sigma_2, \sigma_3)\cdots \mathfrak{T}(\sigma_N, \sigma_1) \qquad (9.77)$$

or simply

$$3(T, h, N) = \text{Tr }\ \mathfrak{T}^N \qquad (9.78)$$

This looks like a very simple expression, and it is. There are many situation in physics where we can write some complicated expression as the trace or determinant of a product of simple matrices. Our case is particularly easy, since all matrices are the same. Making progress in the calculation of matrix expressions is easiest if we know the eigenvalues and eigenvectors. We need to solve

$$\mathfrak{T}|e >= t|e > \qquad (9.79)$$

In our case the matrix is real and symmetric, and we know that there are two eigenvectors with real eigenvalues. Finding the eigenvalues of a real and symmetric two by two matrix can always be done. We need to solve $\det(\mathfrak{T} - t\mathfrak{E}) = 0$, where $\mathfrak{E}$ is the identity matrix. This gives

$$(e^{\beta(J+h)} - t)(e^{\beta(J-h)} - t) - e^{-2\beta J} = 0 \qquad (9.80)$$

or

$$t^2 - t[e^{\beta(J+h)} + e^{\beta(J-h)}] + e^{2\beta J} - e^{-2\beta J} = 0 \qquad (9.81)$$

which has solutions

$$t_{\pm} = \frac{1}{2}\left(e^{\beta(J+h)} + e^{\beta(J-h)} \pm \sqrt{[e^{\beta(J+h)} + e^{\beta(J-h)}]^2 - 4[e^{2\beta J} - e^{-2\beta J}]}\right)$$
$$(9.82)$$

$$t_{\pm} = e^{\beta J}\cosh(\beta h) \pm \frac{1}{2}\sqrt{e^{2\beta(J+h)} + e^{2\beta(J-h)} - 2e^{2\beta J} + 4e^{-2\beta J}} \qquad (9.83)$$

$$t_{\pm} = e^{\beta J}\cosh(\beta h) \pm \frac{1}{2}\sqrt{[e^{\beta(J+h)} - e^{\beta(J-h)}]^2 + 4e^{-2\beta J}} \qquad (9.84)$$

which leads to the final expression

$$t_{\pm} = e^{\beta J}\cosh(\beta h) \pm \sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}} \qquad (9.85)$$

There are two real solutions as expected for a real and symmetric matrix. Note that we have $t_+ > t_-$. Also, we have $t_+ + t_- = \text{Tr }\mathfrak{T} = e^{\beta(J+h)} + e^{\beta(J-h)} > 0$ and $t_+ t_- = \det(\mathfrak{T}) = e^{2\beta J} - e^{-2\beta J} > 0$. Therefore both eigenvalues have to be positive. The corresponding eigenvectors are $|e_{\pm} >$ and the partition function in terms of these eigenvectors is

$$\mathfrak{Z}(T, h, N) = <e_+|\mathfrak{T}^N|e_+> + <e_-|\mathfrak{T}^N|e_-> \tag{9.86}$$

Therefore the partition function in terms of the eigenvalues is

$$\mathfrak{Z}(T, h, N) = t_+^N + t_-^N \tag{9.87}$$

and the magnetic free energy is

$$\mathfrak{G}(T, h, N) = -k_B T \log(t_+^N + t_-^N) \tag{9.88}$$

Now we use $t_+ > t_-$ and rewrite this in the form

$$\mathfrak{G}(T, h, N) = -k_B T \log(t_+^N [1 + \left(\frac{t_-}{t_+}\right)^N]) \tag{9.89}$$

$$\mathfrak{G}(T, h, N) = -N k_B T \log(t_+) - k_B T \log(1 + \left(\frac{t_-}{t_+}\right)^N) \tag{9.90}$$

In the thermodynamic limit $N \to \infty$ the second term becomes zero, because $|\frac{t_-}{t_+}| < 1$ and we find

$$\mathfrak{G}(T, h, N) = -N k_B T \log\left(e^{\beta J} \cosh(\beta h) + \sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}}\right) \tag{9.91}$$

In the limit $h = 0$ we recover our previous result. The magnetization per particle m follows from

$$m = -\frac{1}{N} \left(\frac{\partial \mathfrak{G}}{\partial h}\right)_{T,N} =$$

$$k_B T \frac{e^{\beta J} \beta \sinh(\beta h) + \frac{1}{2}[e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}]^{-\frac{1}{2}}[e^{2\beta J} 2 \sinh(\beta h) \cosh(\beta h)\beta]}{e^{\beta J} \cosh(\beta h) + \sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}}} \tag{9.92}$$

or

$$m = \frac{e^{\beta J} \sinh(\beta h)}{\sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}}}$$

$$\frac{\sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}} + e^{\beta J} \cosh(\beta h)}{e^{\beta J} \cosh(\beta h) + \sqrt{e^{2\beta J} \sinh^2(\beta h) + e^{-2\beta J}}} \tag{9.93}$$

with the final result

$$m = \frac{\sinh(\beta h)}{\sqrt{\sinh^2(\beta h) + e^{-4\beta J}}} \tag{9.94}$$

This formula shows clearly that there is no phase transition in the one-dimensional Ising chain. In the limit $h \to 0$ the magnetization m is zero as long as $T \neq 0$. Next we ask the question how large a field do we need to get a large magnetization. The condition for a value of m close to one is

$$\sinh^2(\beta h) \gg e^{-4\beta J} \tag{9.95}$$

Note that this is never satisfied in the limit $h \to 0$, because in that limit the right hand side goes to zero very rapidly. If we define a field h_o by $\sinh^2(\beta h_o) = e^{-4\beta J}$, then this field represents the field at which the chain becomes magnetic. In the limit $T \to 0$ the exponent will be very small, and therefore the hyperbolic sine is very small. As a result we find

$$T \to 0 \quad h_o \approx k_B T e^{-2\beta J} \tag{9.96}$$

Although the Ising chain is non-magnetic at low temperatures, we only need a very small magnetic field to change its state into a magnetic one. It goes to zero exponentially fast. We see the susceptibility at zero field is very large near $T = 0$, which is the situation close to a phase transition. We can evaluate $\chi(h, T, N) = \left(\frac{\partial m}{\partial h}\right)_{T,N}$ and find

$$\chi(h, T, N) = \frac{\begin{array}{c} e^{\beta J}\beta\cosh(\beta h)\sqrt{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}} - \\ e^{\beta J}\sinh(\beta h)\frac{1}{2}[e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}]^{-\frac{1}{2}}e^{2\beta J}2\sinh(\beta h)\cosh(\beta h)\beta \end{array}}{e^{2\beta J}\sinh^2(\beta h) + e^{-2\beta J}} \tag{9.97}$$

At $h = 0$ this is equal to

$$\chi(h = 0, T, N) = \frac{e^{\beta J}\beta\sqrt{e^{-2\beta J}}}{e^{-2\beta J}} = \beta e^{2\beta J} \tag{9.98}$$

which indeed diverges for $T \to 0$.

The solution for m is singular at $T = 0$, and the effects of this singularity are felt at small temperatures. It is not surprising that one might interpret the data as showing a phase transition. For example, the following figure shows m versus T and h with $4J = 1$:

This figure seems to indicate that at low temperatures there is an abrupt change from negative to positive magnetization as a function of h, with a critical temperature of about 0.5, which is indeed equal to $2J$. But now look at it with h plotted on a logarithmic scale. Now we see that we always go through an area with zero magnetization.

For a real phase transition we have a positive critical temperature, below which we see in the $m(T, h)$ plots an abrupt change. The key word is below the critical temperature. In the one dimensional Ising chain this point is pushed to $T = 0$, and we cannot get below anymore. We still see the effects of the singularity at $T = 0$, though. The ordered state at $h = 0$ might live long enough that for all practical purposes the system behaves like a system with a
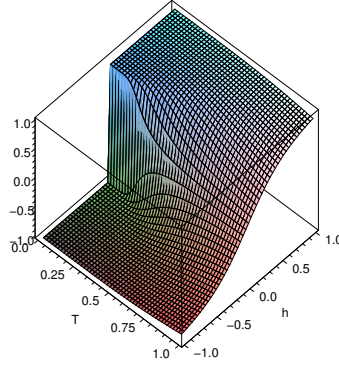
Figure 9.6:   **Magnetization of the one dimensional Ising chain.**

phase transition. This requires switching times much longer than the time of the experiment.

Phase transitions are related to singular behavior of the free energy. The previous example shows that the singularity can be at zero temperature. If one extends the temperature range to negative values, the singular point will be at negative values for systems without phase transitions. It is interesting to study how the position of the singularity depends on parameters in the model hamiltonian, so one can predict when phase transitions occur.

9.6   Spin-correlation function for the Ising chain.

The spin correlation function Γ_i is an important quantity used to study the effects of the singular behavior at $T = 0$. Note that semantically a critical temperature can never be zero, since it is not possible to go below the critical temperature in that case. In this section we calculate the spin-correlation function for the Ising chain without an external field, that is for $h = 0$. We use periodic boundary conditions and assume that the temperature is non-zero.

The spin correlation function is related to the pair distribution function g_i, which is defined by

$$g_i = \langle \mathcal{S}_0 \mathcal{S}_i \rangle = \langle \sigma_0 \sigma_i \rangle \tag{9.99}$$

This function contains the information that we need to discuss how values of the spin on one site relate to values of the spin on another site. For quantities that are not correlated we have $< AB > = < A > < B >$ and hence if the values on the different sites are not correlated we have $g_i = \langle \sigma_0 \sigma_i \rangle = \langle \sigma_0 \rangle \langle \sigma_i \rangle = m^2$.

The spin correlation function measures the correlation between fluctuations from average between different sites. We have
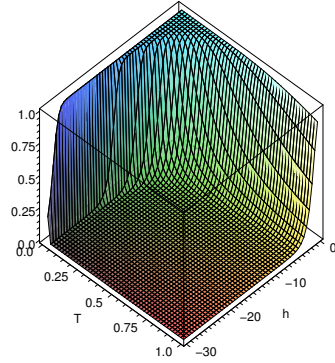
Figure 9.7: **Magnetization of the one dimensional Ising chain, as a function of log(h).**

$$\Gamma_i = \langle (\sigma_0 - m)(\sigma_i - m) \rangle = g_i - m^2 \tag{9.100}$$

This is often a more useful quantity to study. In a true mean field theory all fluctuations on different sites would be uncorrelated, and we would have $\Gamma_i = \delta_{i0}(1 - m^2)$. That would imply $\chi = \beta\Gamma_0$ and this quantity does not diverge at T_c. That is wrong, we looked at that before, so the mean field theory makes another approximation.

For the average energy we have

$$\langle H \rangle = -JNqm^2 - hN - JNq\Gamma_{nn} \tag{9.101}$$

and mean field theory is obtained by requiring that the spin correlations between neighboring sites are zero. We only need fluctuations on neighboring sites to be uncorrelated. Further correlations can be non zero, and will have to be non-zero because the susceptibility is diverging!

To solve the real problem we need the value of Γ_{nn}. We can find this function by considering the exact behavior of a pair of atoms. But this pair is emerged in the rest of the system. We now need to ask the question how large the value of the spin on of of the sites is if the connections are all made to averages. That requires the knowledge of both the nearest neighbor and the next nearest neighbor spin correlation function. It is possible to build up a system of equations where the equation for spin correlation functions at a given distance requires knowledge of spin correlation functions one distance further apart. We need to know all spin correlation functions to solve this system, or we need to have a good approximation for how spin correlation functions at large distances decay.

Here we consider the one dimensional Ising chain again. We set the external field equal to zero. Since without an external field there is no magnetization, the spin-correlation function is given by

$$\Gamma_j = \frac{1}{\mathfrak{Z}(T, h = 0, N)} \sum_{\sigma_0, \cdots, \sigma_{N-1}} \sigma_0 \sigma_j e^{\beta J \sum_{i=0}^{N-1} \sigma_i \sigma_{i+1}} \tag{9.102}$$

where we use periodic boundary conditions and have $\sigma_0 = \sigma_N$.

In order to evaluate this expression we use a simple trick. First we make the problem more complicated by assuming that all bonds can be different. This means that the coupling constant between spins i and $i+1$ is J_i. We treat these coupling constants as independent parameters. In the end we need, of course, $J_i = J$. We define a function $\mathcal{S}$ by

$$\mathcal{S}(J_0, .., J_{N-1}) = \sum_{\sigma_0, \cdots, \sigma_{N-1}} e^{\beta \sum J_i \sigma_i \sigma_{i+1}} \tag{9.103}$$

This is similar to a partition function, and we have that $\mathfrak{Z} = \mathcal{S}(J, J, J, \cdots)$.

Next we take partial derivatives with respect to J_0 to J_{j-1}. This gives us

$$\frac{\partial}{\partial J_0} \frac{\partial}{\partial J_1} \cdots \frac{\partial}{\partial J_{j-1}} \mathcal{S} = \sum_{\sigma_0, \cdots, \sigma_{N-1}} e^{\beta \sum J_i \sigma_i \sigma_{i+1}} [\beta \sigma_0 \sigma_1][\beta \sigma_1 \sigma_2] \cdots [\beta \sigma_{j-1} \sigma_j] \tag{9.104}$$

Since $\sigma_i^2 = 1$ this simplifies to

$$\frac{\partial}{\partial J_0} \frac{\partial}{\partial J_1} \cdots \frac{\partial}{\partial J_{j-1}} \mathcal{S} = \beta^j \sum_{\sigma_0, \cdots, \sigma_{N-1}} e^{\beta \sum J_i \sigma_i \sigma_{i+1}} \sigma_0 \sigma_j \tag{9.105}$$

This leads to

$$g_j = \frac{1}{\beta^j \mathcal{S}} \frac{\partial}{\partial J_0} \frac{\partial}{\partial J_1} \cdots \frac{\partial}{\partial J_{j-1}} \mathcal{S} \bigg|_{J_i = J} \tag{9.106}$$

Therefore we would like to evaluate $\mathcal{S}$. This is done using transfer matrices like we did in the previous section. We define

$$\mathfrak{T}^{(i)} = \begin{pmatrix} e^{\beta J_i} & e^{-\beta J_i} \\ e^{-\beta J_i} & e^{\beta J_i} \end{pmatrix} \tag{9.107}$$

or $\mathfrak{T}^{(i)}_{\sigma, \sigma'} = e^{\beta J_i \sigma \sigma'}$. The definition of $\mathcal{S}$ has a sum in the exponent and we write this exponent of a sum as the product of exponents in the form

$$\mathcal{S} = \sum_{\sigma_0, \cdots, \sigma_{N-1}} e^{\beta J_0 \sigma_0 \sigma_1} e^{\beta J_1 \sigma_1 \sigma_2} \cdots e^{\beta \sigma_{N-1} \sigma_{N-1} \sigma_0} \tag{9.108}$$

From this we see that

$$\mathcal{S} = \text{Tr} \ \mathfrak{T}^{(0)} \mathfrak{T}^{(1)} \cdots \mathfrak{T}^{(N-1)} \tag{9.109}$$

If we need to calculate the trace or determinant of a matrix, it is always very useful to know the eigenvalues and eigenvectors. The eigenvectors of the matrices $\mathfrak{T}^{(i)}$ are easy, they are $e_\pm = \frac{1}{\sqrt{2}}(1, \pm 1)$ independent of the index i.

The eigenvalues do depend on the index i, but they are easy to find, they are $\lambda_\pm^{(i)} = e^{\beta J_i} \pm e^{-\beta J_i}$. We can now evaluate $\mathcal{S}(J, J, \cdots)$ and find

$$\mathcal{S}(J, J, \cdots) =$$

$$< e_+|\mathfrak{T}^{(0)}\mathfrak{T}^{(1)}\cdots\mathfrak{T}^{(N-1)}|e_+ >\Big|_{J_i=J} + < e_-|\mathfrak{T}^{(0)}\mathfrak{T}^{(1)}\cdots\mathfrak{T}^{(N-1)}|e_- >\Big|_{J_i=J} = \lambda_+^N + \lambda_-^N \tag{9.110}$$

where we have defined $\lambda_\pm = e^{\beta J} \pm e^{-\beta J}$.

Next we discuss the derivative with respect to J_i. Because the factor J_i occurs in one place only, we see

$$\frac{\partial}{\partial J_i}\mathcal{S} = \text{Tr } \mathfrak{T}^{(0)}\mathfrak{T}^{(1)}\cdots\frac{\partial\mathfrak{T}^{(i)}}{\partial J_i}\cdots\mathfrak{T}^{(N-1)} \tag{9.111}$$

The derivative of the matrix is easy, and we define

$$\mathfrak{U}^{(i)} = \frac{\partial\mathfrak{T}^{(i)}}{\partial J_i} \tag{9.112}$$

and find

$$\mathfrak{U}^{(i)} = \beta \begin{pmatrix} e^{\beta J_i} & -e^{-\beta J_i} \\ -e^{-\beta J_i} & e^{\beta J_i} \end{pmatrix} \tag{9.113}$$

The eigenvectors of this matrix are again $e_\pm$ independent of the index i, but the eigenvalues are now interchanged, e_+ goes with $\beta\lambda_-^{(i)}$ and vice versa.

We are now in the position to calculate the spin correlation function. We have

$$\frac{\partial}{\partial J_0}\frac{\partial}{\partial J_1}\cdots\frac{\partial}{\partial J_{j-1}}\mathcal{S} = \text{Tr } \mathfrak{U}^{(0)}\cdots\mathfrak{U}^{(j-1)}\mathfrak{T}^{(j)}\cdots\mathfrak{T}^{(N-1)} \tag{9.114}$$

which is equal to

$$< e_+|\mathfrak{U}^{(0)}\cdots\mathfrak{U}^{(j-1)}\mathfrak{T}^{(j)}\cdots\mathfrak{T}^{(N-1)}|e_+ > + < e_-|\mathfrak{U}^{(0)}\cdots\mathfrak{U}^{(j-1)}\mathfrak{T}^{(j)}\cdots\mathfrak{T}^{(N-1)}|e_- > \tag{9.115}$$

If we now calculate this expression at $J_i = J$ we find

$$\beta^j \lambda_-^j \lambda_+^{N-j} + \beta^j \lambda_+^j \lambda_-^{N-j} \tag{9.116}$$

Since the average magnetization is zero, the spin correlation functions is equal to the pair distribution function, we we get

$$\Gamma_j = g_j = \frac{\lambda_-^j \lambda_+^{N-j} + \lambda_+^j \lambda_-^{N-j}}{\lambda_+^N + \lambda_-^N} \tag{9.117}$$

Because of the periodic boundary conditions this result is symmetric under the change $j \leftrightarrow N - j$. Also, we have $\Gamma_N = 1$, which is to be expected because site N is equal to site 0, and $< \sigma_0 \sigma_N >=< \sigma_0^2 >= 1$.

The expression is simplified when we divide the numerator and denominator by the eigenvalue λ_+ to the power N. Note that $\lambda_+ > \lambda_- > 0$. The ratio of the two eigenvalues is simple, we have

$$\frac{\lambda_-}{\lambda_+} = \frac{e^{\beta J} - e^{-\beta J}}{e^{\beta J} + e^{-\beta J}} = \tanh(\beta J) \tag{9.118}$$

This gives

$$\Gamma_j = \frac{\tanh^j(\beta J) + \tanh^{N-j}(\beta J)}{1 + \tanh^N(\beta J)} \tag{9.119}$$

This expression is complicated, but we have to remember that we always need to invoke the thermodynamic limit after all calculations have been done. That is the case here, we have our final result. If the temperature is non-zero, $\tanh(\beta J) < 1$ and in the limit $N \gg j$ only the lowest order terms survives. Note that we have given explicit meaning to the idea of the thermodynamic limit, we require that N is large compared to j! In that case we can write

$$\Gamma_j = \tanh^j(\beta J) \tag{9.120}$$

This formula shows that as a function of the distance j the spin correlation function decays as a power law. Hence it can be related to a correlation length via

$$\Gamma_j = e^{-\frac{j}{\xi}} \tag{9.121}$$

where the correlation length ξ is found by

$$\xi = -\frac{j}{\log(\Gamma_j)} = -[\log \tanh(\beta J)]^{-1} \tag{9.122}$$

As expected, we find that the spin correlation decays exponentially. If Γ_j is close to one, spins tend to point in the same direction, and the magnetization is strongly correlated. The spin correlation length measures the distance over which spins are correlated and tend to point in the same direction. Note that this common direction can fluctuate!

The correlation length increases with decreasing temperature. We see that in the limit $T \to 0$ the hyperbolic tangent approaches one and the logarithm gives zero. We can find the behavior by

$$\tanh(\beta J) = \frac{1 - e^{-2\beta J}}{1 + e^{-2\beta J}} \approx 1 - 2e^{-2\beta J} \tag{9.123}$$

and the logarithm of one plus a small term is approximated by the small term. That gives

$$\xi \approx \frac{1}{2}e^{2\beta J} \qquad (9.124)$$

which diverges faster that any power law. Normally we would expect that $\xi \propto |T - T_c|^{-\nu}$, but that is not the case here.

Before we draw any conclusions we need to check one thing. The thermodynamic limit should always be taken last. So we need to take the limit to zero temperature first! We should not take the limit $T \to 0$ in the expression above. As usual, the limit $T \to 0$ has to be taken before the limit $N \to \infty$. Before the thermodynamic limit we have

$$\xi = -j \left[\log \left(\frac{\tanh^j(\beta J) + \tanh^{N-j}(\beta J)}{1 + \tanh^N(\beta J)} \right) \right]^{-1} \qquad (9.125)$$

Using the exact expression (no approximation) $\tanh(\beta J) = 1 - x$, with $|x| \ll 1$ we have

$$\xi = -j \left[\log \left(\frac{(1-x)^j + (1-x)^{N-j}}{1 + (1-x)^N} \right) \right]^{-1} \qquad (9.126)$$

If we take linear terms in the powers only, we see that

$$\frac{(1-x)^j + (1-x)^{N-j}}{1 + (1-x)^N} \approx \frac{(1-jx) + (1 - (N-j)x)}{1 + (1 - Nx)} = 1 \qquad (9.127)$$

Therefore, we need second order terms. We have

$$\frac{(1-x)^j + (1-x)^{N-j}}{1 + (1-x)^N} \approx \frac{2 - Nx + \frac{j(j-1)}{2}x^2 + \frac{(N-j)(N-j-1)}{2}x^2}{2 - Nx + \frac{N(N-1)}{2}x^2} \qquad (9.128)$$

This is equal to

$$\frac{1 + \frac{1}{2-Nx}\left(\frac{j(j-1)}{2}x^2 + \frac{(N-j)(N-j-1)}{2}x^2 \right)}{1 + \frac{1}{2-Nx}\frac{N(N-1)}{2}x^2} \qquad (9.129)$$

which can be approximated by

$$(1 + \frac{1}{2 - Nx}\left(\frac{j(j-1)}{2}x^2 + \frac{(N-j)(N-j-1)}{2}x^2 \right))(1 - \frac{1}{2 - Nx}\frac{N(N-1)}{2}x^2) \qquad (9.130)$$

which is approximated by

$$1 + \frac{1}{2 - Nx}\left(\frac{j(j-1)}{2}x^2 + \frac{(N-j)(N-j-1)}{2}x^2 \right) - \frac{1}{2 - Nx}\frac{N(N-1)}{2}x^2 \qquad (9.131)$$

which gives up to second order

$$1 + \left(\frac{j(j-1)}{4} + \frac{(N-j)(N-j-1)}{4} - \frac{N(N-1)}{4} \right) x^2 \tag{9.132}$$

or

$$1 + \frac{1}{4} \left(j(j-1) - jN - N(j+1) + j(j+1) + N \right) x^2 \tag{9.133}$$

or

$$1 + \frac{1}{4} \left(2j^2 - 2jN \right) x^2 \tag{9.134}$$

We now have for the correlation length

$$\xi \approx -j \left[\log \left(1 + \frac{1}{2} j(j-N) x^2 \right) \right]^{-1} \tag{9.135}$$

and after approximating the logarithm

$$\xi \approx \frac{2}{(N-j)x^2} \tag{9.136}$$

But now we are able to approximate the value of x by $-2e^{-2\beta J}$ and obtain

$$\xi \approx \frac{e^{4\beta J}}{2(N-j)} \tag{9.137}$$

This is a different result, indeed. In the limit of zero temperature the hyperbolic tangent is very close to one, and all terms in the pair correlation function play a role. But the divergence is still exponential in stead of a power law, so that qualitative conclusion did not change. This is another consequence of the fact that at zero temperature there is not really a phase transition.

The previous results have some interesting consequences. First, suppose that the correlation length is equal to the length of the chain. The whole chain will act in a correlated fashion. At a given time it is very likely that all the spins point in the same direction and one would be tempted to call the chain magnetic. There will be times, however, where part of the chain is magnetized in one direction, and part in the other. Such a disturbance has a higher energy and will disappear. Nevertheless, this disturbance might switch the magnetization of the whole chain. An average over an infinite time therefore gives a non-magnetic state! If we make the chain longer, we have to go to a lower temperature for the correlation length to be equal to the chain-length. Fluctuations become more unlikely, and it will take longer for the chain to switch its magnetization. From a fundamental point of view, the chain is still non-magnetic. From a practical point of view, the chain is magnetic in a meta-stable state. Therefore, we could define a transition temperature T^* by $\xi(T^*) = N$. But phase transitions are only defined in the thermodynamic limit, and we have $\lim_{N \to \infty} T^* = 0$. But it does tell us when finite size effects play a role. For a given N we can find T^*

and we expect to observe things that are dependent on the sample size below that temperature.

Finally, we can check the relation between the spin correlation function and the susceptibility. We found before that $\chi = \beta \sum \Gamma_i$ and this gives

$$\beta \sum_i \Gamma_i = \beta \sum_i \tanh^i(\beta J) = \frac{\beta}{1 - \tanh(\beta J)} \tag{9.138}$$

and for small temperatures this is about $\frac{1}{2}\beta e^{2\beta J}$. That is off by a factor of two. Again, this is not surprising. We used the result obtained after taking the thermodynamic limit to calculate the result. We need to take the whole expression as a function of N, though, and evaluate

$$\sum_{j=0}^{N-1} \frac{\tanh^j(\beta J) + \tanh^{N-j}(\beta J)}{1 + \tanh^N(\beta J)} \tag{9.139}$$

This is equal to

$$\frac{1}{1 + \tanh^N(\beta J)} \left(\sum_{j=0}^{N-1} \tanh^j(\beta J) + \tanh^N(\beta J) \sum_{j=0}^{N-1} \tanh^{-j}(\beta J) \right) \tag{9.140}$$

or

$$\frac{1}{1 + \tanh^N(\beta J)} \left(\frac{1 - \tanh^N(\beta J)}{1 - \tanh(\beta J)} + \tanh^N(\beta J) \frac{1 - \tanh^{-N}(\beta J)}{1 - \tanh^{-1}(\beta J)} \right) \tag{9.141}$$

or

$$\frac{1}{1 + \tanh^N(\beta J)} \left(\frac{1 - \tanh^N(\beta J)}{1 - \tanh(\beta J)} + \tanh(\beta J) \frac{\tanh^N(\beta J) - 1}{\tanh(\beta J) - 1} \right) \tag{9.142}$$

or

$$\frac{1 - \tanh^N(\beta J)}{1 + \tanh^N(\beta J)} \frac{1 + \tanh(\beta J)}{1 - \tanh(\beta J)} \tag{9.143}$$

and now the results are consistent. If we take the thermodynamic limit in the last equation we get an extra factor in the denominator, which is equal to two at low temperature, and which is needed to get the results we derived before. Also, if we first take the limit to zero temperature the result becomes N, which is correct because all atoms now act in the same way and the magnetization of one atom as a function of h follows Curie's law with $\chi = N\beta$.

The observations in the last section might seem a bit esoteric, but they are important. Too often approximated results are used to obtain erroneous answers. One needs to know when approximations can be made.

9.7   Renormalization group theory.

A very powerful way of treating critical phenomena is via renormalization group theory. The basic idea behind this theory is very simple. Near a critical point the dominant length scale for any system is the correlation length ξ. Exactly at the critical temperature this correlation length diverges and somehow a picture of a given state of a system should be independent of any length scale. In other words, if we paint spin-up atoms red and spin-down atoms blue, in a magnetic system the pattern we observe should be independent of any length scale. If we look at the pattern of spins with atomic resolution we will see the individual red and blue dots. If we diminish the resolution we will start to see average colors, obtained by mixing the effects of a few dots. If we decrease the resolution even more, the average color will depend on the effects of more dots in a larger region. Imagine a whole series of pictures taken this way, with decreasing resolution. The pictures with atomic resolution will look somewhat different, since we see the individual dots. All the other pictures should look very similar, however, since the divergence of the correlation length has taken away our measure of length. There is no way to distinguish the pictures by their pattern and deduce the magnification from the observed pattern. Notice that in real life this fails since samples are of finite dimensions and at some point we will see the boundaries of the samples. Theorists do not bother with these trivial details.

The previous paragraph sketches the concepts of renormalization group theory. Essentially we apply scaling theory but now start at the microscopic level. We build the macroscopic theory up from the ground, by averaging over larger and larger blocks. The next step is to formulate this idea in mathematical terms. We need to define some kind of transformation which corresponds to decreasing the magnetization. The easiest way is to define a procedure that tells us how to average over all variables in a certain volume. The one-dimensional Ising model will again serve as an example. The atomic sites are labelled with an index i. Suppose we want to decrease the magnification by an integral factor p. Hence we define a new index j which also takes all integer values and label the atoms in group j by $i = pj + k$, with $k = 1, 2, \cdots, p$. The spin in cell j can be defined in several ways. For example, we could assign it the average value or the value of the left-most element of the cell. In the end, our result should be independent of this detail, and we choose the procedure that is easiest to treat.

This procedure defines a transformation R_p in the set of states that can be characterized by $\{\cdots, \sigma_{-1}, \sigma_0, \sigma_1, \cdots\}$. It is not a one-to-one mapping and hence the word group is a misnomer, the inverse of R_p does not exist. This transformation is also applied to the Hamiltonian and the requirement is that at a critical point the Hamiltonian is invariant under the operations of the renormalization group.

These ideas are best illustrated by using the one-dimensional Ising ring as an example, even though this model does not show a phase transition. It is the easiest one for actual calculations, and it will serve to get the ideas across. The partition function is given by

$$\mathfrak{Z}(T, N, h) = \sum_{\{\sigma_1, \cdots, \sigma_N\}} e^{\beta J \sum_{i=1}^N \sigma_i \sigma_{i+1} + \beta h \sum_{i=1}^N \sigma_i} \tag{9.144}$$

The transformation we have in mind is a decrease of the magnification by a factor of two. Hence we plan to combine spins $2j$ and $2j + 1$ into a single average spin with value $\sigma'_j = \sigma_{2j}$. We also combine βJ into one symbol $\tilde{J}$ and similarly write $\tilde{h} = \beta h$. For simplicity we assume that N is even. Of course, in the end we need to take the limit $N \rightarrow \infty$, and the fact that N is even or odd does not play a role. The parameters in the partition function are therefore $\tilde{J}$, $\tilde{h}$, and N. Our goal is to write the partition function in the form

$$\mathfrak{Z}(N, \tilde{J}, \tilde{h}) = \sum_{\sigma'_1, \cdots, \sigma'_{\frac{N}{2}}} e^{E'(\sigma'_1, \cdots, \sigma'_{\frac{N}{2}})} \tag{9.145}$$

This can be accomplished very easily by separating out the odd and even values of i in the original formula of the partition function. Using the periodic boundary conditions to give the original Hamiltonian a symmetrical form we find

$$\mathfrak{Z}(N, \tilde{J}, \tilde{h}) = \sum_{\sigma_2, \sigma_4, \cdots} \sum_{\sigma_1, \sigma_3, \cdots} e^{\beta J \sum_{i=1}^N \sigma_i \sigma_{i+1} + \beta h \sum_{i=1}^N \sigma_i} \tag{9.146}$$

We can think about this in terms bonds between spins. If we consider the spins with even indices to be the "master" spin, the spins with odd indices represent bonds between these master spins. Summation over all possible odd spin states is equivalent to a summation over all possible bonds between master spins.

Next, we rewrite the summation in the exponent as a summation over odd spin indices only. That is easy, and we have

$$\tilde{J} \sum_{i=1}^N \sigma_i \sigma_{i+1} + \tilde{h} \sum_{i=1}^N \sigma_i = \sum_{i \text{ odd}} \left(\tilde{J} \sigma_i (\sigma_{i-1} + \sigma_{i+1}) + \tilde{h}(\sigma_i + \frac{1}{2}(\sigma_{i-1} + \sigma_{i+1})) \right) \tag{9.147}$$

which gives

$$\mathfrak{Z}(N, \tilde{J}, \tilde{h}) = \sum_{\sigma_2, \sigma_4, \cdots} \sum_{\sigma_1, \sigma_3, \cdots} \prod_{i \text{ odd}} e^{\tilde{J} \sigma_i (\sigma_{i-1} + \sigma_{i+1}) + \tilde{h}(\sigma_i + \frac{1}{2}(\sigma_{i-1} + \sigma_{i+1}))} \tag{9.148}$$

Now we perform the summation over the variables with odd indices. Each such sum occurs only in one factor of the product of exponents, and hence these sums can be done term by term. We arrive at

$$\mathfrak{Z}(N, \tilde{J}, \tilde{h}) = \sum_{\sigma_2, \sigma_4, \cdots} \prod_{i \text{ odd}} \sum_{\sigma} e^{\tilde{J} \sigma (\sigma_{i-1} + \sigma_{i+1}) + \tilde{h}(\sigma + \frac{1}{2}(\sigma_{i-1} + \sigma_{i+1}))} \tag{9.149}$$

or

$$3(N, \tilde{J}, \tilde{h}) = \sum_{\sigma_2, \sigma_4, \cdots} \prod_{i \text{ odd}} e^{\tilde{h}(\frac{1}{2}(\sigma_{i-1}+\sigma_{i+1}))} 2 \cosh(\tilde{J}(\sigma_{i-1} + \sigma_{i+1}) + \tilde{h}) \quad (9.150)$$

This looks like something for a partition function where only even spins play a role. On the right hand side we have exponential like things. Can we write this in a form

$$3(N, \tilde{J}, \tilde{h}) = \sum_{\sigma_2, \sigma_4, \cdots} e^{E(\sigma_2, \sigma_4, \cdots)} \quad (9.151)$$

where the form of the energy is similar to the original energy? Clearly, the value of the coupling constant will be different, and also the zero of energy can be shifted. Therefore we try

$$E(\sigma_2, \sigma_4, \cdots) = \sum_{i \text{ odd}} \left(2g + \tilde{J}' \sigma_{i-1} \sigma_{i+1} + \frac{1}{2} \tilde{h}'(\sigma_{i-1} + \sigma_{i+1}) \right) \quad (9.152)$$

The question is, is this possible, and if yes, what are the functions $g(\tilde{J}, \tilde{h})$, $\tilde{J}'(\tilde{J}, \tilde{h})$, and $\tilde{h}'(\tilde{J}, \tilde{h})$.

The answer whether it is possible can be given easily. Because we wrote the forms symmetric for interchanges of left and right, we need to consider only three combinations of neighboring even spins, $(+, +)$, $(+, -) = (-, +)$, and $(-, -)$. That gives us three equations, from which we can find three independent functions!

The equations to solve are

$$(+, +) \; : \; e^{\tilde{h}} 2 \cosh(2\tilde{J} + \tilde{h}) = e^{2g + \tilde{J}' + \tilde{h}'} \quad (9.153)$$

$$(+, -) \; : \; 2 \cosh(\tilde{h}) = e^{2g - \tilde{J}'} \quad (9.154)$$

$$(-, -) \; : \; e^{-\tilde{h}} 2 \cosh(-2\tilde{J} + \tilde{h}) = e^{2g + \tilde{J}' - \tilde{h}'} \quad (9.155)$$

The road to solving these is a bit tricky, but here is one approach. Multiply the first and the third, which gives

$$4 \cosh(2\tilde{J} + \tilde{h}) \cosh(-2\tilde{J} + \tilde{h}) = e^{4g + 2\tilde{J}'} \quad (9.156)$$

Multiply by the square of the second equation and we get

$$16 \cosh(2\tilde{J} + \tilde{h}) \cosh(-2\tilde{J} + \tilde{h}) \cosh^2(\tilde{h}) = e^{8g} \quad (9.157)$$

which gives us g. We can also divide the product of the first and third by the square of the second, and now we find

$$\frac{\cosh(2\tilde{J} + \tilde{h}) \cosh(-2\tilde{J} + \tilde{h})}{\cosh^2(\tilde{h})} = e^{4\tilde{J}'} \quad (9.158)$$

which gives us $\tilde{J}'$. Finally, we divide the first and third equation to get

$$e^{2\tilde{h}} \frac{\cosh(2\tilde{J} + \tilde{h})}{\cosh(-2\tilde{J} + \tilde{h})} = e^{2\tilde{h}'} \tag{9.159}$$

which gives us $\tilde{h}'$. The solutions are

$$\tilde{J}'(\tilde{J}, \tilde{h}) = \frac{1}{4} \log \left(\frac{\cosh(2\tilde{J} + \tilde{h})\cosh(-2\tilde{J} + \tilde{h})}{\cosh^2(\tilde{h})} \right) \tag{9.160}$$

$$\tilde{h}'(\tilde{J}, \tilde{h}) = \tilde{h} + \frac{1}{2} \log \left(\frac{\cosh(2\tilde{J} + \tilde{h})}{\cosh(-2\tilde{J} + \tilde{h})} \right) \tag{9.161}$$

$$g(\tilde{J}, \tilde{h}) = \frac{1}{16} \log \left(16 \cosh(2\tilde{J} + \tilde{h}) \cosh(-2\tilde{J} + \tilde{h}) \cosh^2(\tilde{h}) \right) \tag{9.162}$$

We are now able to relate the original partition function to a new partition function describing half the particles on twice the length scale. Because of our redefining the variables, the partition function depends on the variables N, $\tilde{J}$, and $\tilde{h}$. This is equivalent to the original set N, T, h. We have redefined the temperature scale by using βJ and redefined the field by using βh. Hence the partition function is a function $\mathfrak{Z}(N, \tilde{J}, \tilde{h})$. We have expressed this in the form of a partition function with new coupling constants and fields at half the number of particles, We needed to add an energy shift. This translates to

$$\mathfrak{Z}(N, \tilde{J}, \tilde{h}) = e^{Ng(\tilde{J}, \tilde{h})} \mathfrak{Z}(\frac{1}{2}N, \tilde{J}', \tilde{h}') \tag{9.163}$$

This equation relates the properties of the original system to the properties of a system with double the length scale, and renormalized interaction strengths. Because the partition functions are directly related, thermodynamic properties are similar! Our goal is now to repeat this process over and over again. We have

$$\mathfrak{Z}(\frac{1}{2}N, \tilde{J}', \tilde{h}') = e^{\frac{1}{2}Ng(\tilde{J}', \tilde{h}')} \mathfrak{Z}(\frac{1}{4}N, \tilde{J}'', \tilde{h}'') \tag{9.164}$$

Note that the functional forms $g(x, y)$, $\tilde{J}''(x, y)$, and $\tilde{h}''(x, y)$ are the same as before!

When we apply this procedure repeatedly we therefore find

$$\mathfrak{Z}(N, \tilde{J}, \tilde{h}) = e^{Ng(\tilde{J}, \tilde{h})} e^{\frac{1}{2}Ng(\tilde{J}', \tilde{h}')} \cdots \mathfrak{Z}(0, \tilde{J}^\infty, \tilde{h}^\infty) \tag{9.165}$$

The free energy follows from $\log(\mathfrak{Z}) = -\beta \mathfrak{G}$. Therefore we have

$$-\beta \mathfrak{G} = N \left(g(\tilde{J}, \tilde{h}) + \frac{1}{2} g(\tilde{J}', \tilde{h}') + \left[\frac{1}{2} \right]^2 g(\tilde{J}'', \tilde{h}'') + \cdots \right) \tag{9.166}$$

if the series converges.

At this point we have established a general procedure relating the partition function of a system to a partition function containing the combined effects of the spins in a block of two. The coupling constants $\tilde{J}$ and $\tilde{h}$ changed values, however. They had to be renormalized in order for the expressions to be valid. Hence the new system is different from the old one, since it is not described by the same Hamiltonian. It is possible that in certain cases both $\tilde{J}' = \tilde{J}$ and $\tilde{h}' = \tilde{h}$. In that case the old and new system do represent the same physics, no matter how many times we apply the demagnification operation. Critical points therefore correspond to fixed points of the renormalization formulas. One has to keep in mind that a fixed point does not necessarily correspond to a critical point; there are more fixed points than critical points.

The one-dimensional Ising model depends only on two coupling constants. These represent the interaction energy between the spins and the energy of a spin in an external field. Both constants are scaled with respect to $k_B T$. It is in general always possible to scale the constants with respect to the temperature, since the partition function always combines β and H in a product. In a general model, one has a number of coupling constants, and a search for critical points corresponds to a search for fixed points in a many-dimensional space. The easiest example of such a search is again for the one-dimensional Ising model, this time without an external field. Hence $\tilde{h} = 0$ and 9.161 shows that $\tilde{h}' = 0$ too. In every step of the renormalization procedure the coupling constant for the external field remains zero. The renormalization equation 9.160 is now very simple:

$$\tilde{J}'(\tilde{J}, 0) = \frac{1}{2} \log \cosh(2\tilde{J}) \qquad (9.167)$$

Since $\cosh(x) \leqslant e^x$ we see that $\log \cosh(2\tilde{J}) \leqslant \log e^{2\tilde{J}} = 2\tilde{J}$ and hence $0 \leqslant \tilde{J}' \leqslant \tilde{J}$. This also implies $g(\tilde{J}', 0) \leqslant g(\tilde{J}, 0)$, and hence the series 9.166 for the free energy converges and is bounded by $\sum_k \left(\frac{1}{2}\right)^k g(\tilde{J}, 0) = 2g(\tilde{J}, 0)$.

Suppose we start with a coupling constant $\tilde{J} = \tilde{J}^{(0)}$. Each iteration adds a prime to the value according to the equation 9.160 and after k iterations we have the value $\tilde{J} = \tilde{J}^{(k)}$. If we keep going, we arrive at $\lim_{k \to \infty} \tilde{J}^{(k)} = \tilde{J}^{(\infty)}$. Because of the limiting conditions we have $0 \leqslant \tilde{J}^{(\infty)} \leqslant \tilde{J}$.

Since the function g is also decreasing, we can get a lower bound on the free energy by using the infinite value, and we find

$$2Ng(\tilde{J}^{(\infty)}, 0) \leqslant -\beta \mathfrak{G} \leqslant 2Ng(\tilde{J}^{(0)}, 0) \qquad (9.168)$$

It is clear that the value of $\tilde{J}^{(\infty)}$ depends on the initial value of $\tilde{J}$ and hence on the temperature T. The possible values of $\tilde{J}^{(\infty)}$ can be found by taking the limit in 9.160 on both sides. That leads to

$$\tilde{J}^{(\infty)} = \frac{1}{2} \log \left(\cosh(2\tilde{J}^{(\infty)}) \right) \qquad (9.169)$$

This equation has solutions $\tilde{J}^{(\infty)} = 0$ and $\tilde{J}^{(\infty)} = \infty$. These solutions are called the **fixed points** of the equation. If we start with one of these values, the scaling transformations will not change the values. Hence they represent physical situations that are length independent.

In the case $\tilde{J}^{(\infty)} = 0$ we find that the free energy is given by

$$-\beta\mathfrak{G} = 2N\frac{1}{4}\log(4) = \log(2^N) \tag{9.170}$$

The right hand side is equal to the entropy of a completely disordered chain divided by the Boltzmann constant. Therefore we have $\mathfrak{G} = -TS$, as expected for large temperatures or zero coupling constant.

In the case $\tilde{J}^{(\infty)} = \infty$ we find that the free energy is given by

$$-\beta\mathfrak{G} = 2N\frac{1}{4}\log(2e^{2\tilde{J}^{(\infty)}}) \;\Rightarrow\; \mathfrak{G} \approx -Nk_BT\tilde{J} = -NJ \tag{9.171}$$

The right hand side is equal to the energy of a completely ordered chain, which happens if the temperature is zero of the coupling constant is infinity.

If we start with an arbitrary value of $\tilde{J}$ the next value will be smaller and so on. So we will end at the fixed point zero. The only exception is when we start exactly at infinity, then we stay at infinity. Hence we can say the following. If we start at a fixed point, we remain at that point. If we make a small deviation from the fixed point for our starting value, we always end up at zero. Therefore, $\tilde{J}^{(\infty)} = 0$ is a **stable** fixed point and $\tilde{J}^{(\infty)} = \infty$ is an **unstable** fixed point.

The stable fixed point $\tilde{J} = 0$ corresponds to $T = \infty$. At an infinite temperature any system is completely disordered and the states of the individual spins are completely random and uncorrelated. If we average random numbers the results will remain random and a system at an infinite temperature will look the same for any magnification. This fixed point is therefore a trivial fixed point and is expected to occur for any system. The same is true for the second fixed point in our simple model, which corresponds to $T = 0$. At zero temperature all spins are ordered and again the system looks the same under any magnification. It is again a fixed point which will always show up. It does not correspond to a critical point since it does not divide the temperature range in two parts, there are no negative temperatures. Close to zero temperature the effects of this fixed point are noticeable, however, and show up in a divergence of the correlation length to infinity at $T = 0$. Note that at infinite temperature the correlation length becomes zero.

What we do in renormalization group theory is to replace one spin by the effective value of a block of spins. That changes the length scales of our problem, and in general changes our observations. That is not true, however, at the critical point. In that case the correlation length is infinity and repeated transformations will give the same results. Therefore , the critical point will show up as a fixed point in the scaling equations. The other case where changing length scales does not affect the physics is when the correlation length is zero. Hence we always have a fixed point corresponding to zero correlation length or infinite temperature.

Note that we cannot go any further in renormalization than to have one spin left. If at this point values are converged, fine, but if not, we see finite size effects. Only in the thermodynamic limit do we get a sharp transition temperature independent of the initial conditions. For a finite sample we always find a small range of fixed points.

The application of renormalization theory to a one-dimensional problem is often straightforward because a one-dimensional space has a very simple topology. In more dimensions one has to deal with other complications. For example, we could apply the same set of ideas to a two-dimensional Ising model. The sites are numbered (i, j) and the easiest geometry is that of a square lattice where the actual lattice points are given by $\vec{R} = a(i\hat{x} + j\hat{y})$. We could try to sum first over all sites with $i + j$ odd. The remaining sites with $i + j$ even form again a square lattice, but with lattice constant $a\sqrt{2}$. Suppose we start with a model with only nearest-neighbor interactions. It is easy to see that after one step in this renormalization procedure we have a system with nearest and next-nearest neighbor interactions! As a result we have to consider the Ising model with all interactions included and in the end find fixed points which correspond to nearest neighbor interactions only. Real life is not always as easy as the current section seemed to suggest. But if we do it correctly, we will find that the two-dimensional Ising model has a non-trivial fixed point even for $h = 0$. The value $\tilde{J}_{fp}$ is related to T_c by $\tilde{J}_{fp} = \frac{J}{k_B T_c}$.

If renormalization theory would only give critical temperatures, its value would be quite small. The most important aspect of this theory, however, is that it also yields critical exponents. Suppose that $\tilde{J}^*$ and $\tilde{h}^*$ are the values of the coupling parameters at a critical point. The critical exponents are related to the behavior of the renormalization group equations near the critical point. Assume that we are close to the critical point and that we have $\tilde{J} = \tilde{J}^* + \delta\tilde{J}$, $\tilde{h} = \tilde{h}^* + \delta\tilde{h}$ with $\delta\tilde{J}$ and $\delta\tilde{h}$ small. In first approximation we use a linearized form, valid for small deviations, and we have

$$\delta\tilde{J}' = \left(\frac{\partial\tilde{J}'}{\partial\tilde{J}}\right)\delta\tilde{J} + \left(\frac{\partial\tilde{J}'}{\partial\tilde{h}}\right)\delta\tilde{h} \tag{9.172}$$

and

$$\delta\tilde{h}' = \left(\frac{\partial\tilde{h}'}{\partial\tilde{J}}\right)\delta\tilde{J} + \left(\frac{\partial\tilde{h}'}{\partial\tilde{h}}\right)\delta\tilde{h} \tag{9.173}$$

where the partial derivatives are calculated at the fixed point.

The deviations in the coupling constants from the fixed point are combined in a two-vector $\vec{d}$. If we have more coupling constants, we have a larger dimensional space, so it is easy to generalize to include a larger number of coupling constants. Hence near a critical point we have

$$\vec{d}' = \mathcal{M}\vec{d} \tag{9.174}$$

where $\mathcal{M}$ is the Jacobian matrix of the renormalization group transformation. The eigenvalues of this matrix determine the stability of the critical point. If an eigenvalue has an absolute value less than one, the direction of the corresponding eigenvector corresponds to a direction in which the fixed point is stable. The deviations will become smaller in this direction. Eigenvalues larger than one correspond to unstable directions. A stable fixed point has all eigenvalues less than one.

For the one-dimensional Ising chain we have

$$\left(\frac{\partial \tilde{J}'}{\partial \tilde{J}}\right) = \frac{1}{2}\frac{\sinh(2\tilde{J}+\tilde{h})}{\cosh(2\tilde{J}+\tilde{h})} + \frac{1}{2}\frac{\sinh(2\tilde{J}-\tilde{h})}{\cosh(2\tilde{J}-\tilde{h})} \tag{9.175}$$

$$\left(\frac{\partial \tilde{J}'}{\partial \tilde{h}}\right) = \frac{1}{4}\frac{\sinh(2\tilde{J}+\tilde{h})}{\cosh(2\tilde{J}+\tilde{h})} - \frac{1}{4}\frac{\sinh(2\tilde{J}-\tilde{h})}{\cosh(2\tilde{J}-\tilde{h})} - \frac{1}{2}\frac{\sinh(\tilde{h})}{\cosh(\tilde{h})} \tag{9.176}$$

$$\left(\frac{\partial \tilde{h}'}{\partial \tilde{J}}\right) = \frac{\sinh(2\tilde{J}+\tilde{h})}{\cosh(2\tilde{J}+\tilde{h})} - \frac{\sinh(2\tilde{J}-\tilde{h})}{\cosh(2\tilde{J}-\tilde{h})} \tag{9.177}$$

$$\left(\frac{\partial \tilde{h}'}{\partial \tilde{h}}\right) = 1 + \frac{1}{2}\frac{\sinh(2\tilde{J}+\tilde{h})}{\cosh(2\tilde{J}+\tilde{h})} + \frac{1}{2}\frac{\sinh(2\tilde{J}-\tilde{h})}{\cosh(2\tilde{J}-\tilde{h})} \tag{9.178}$$

and the matrix $\mathcal{M}$ for the one-dimensional Ising model at the point $(\tilde{J},0)$ is simple

$$\begin{pmatrix} \tanh(2\tilde{J}) & 0 \\ 0 & 1+\tanh(2\tilde{J}) \end{pmatrix} \tag{9.179}$$

At the critical point $(0,0)$ the form is simply

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \tag{9.180}$$

This means that along the $\tilde{J}$ direction the critical point $(0,0)$ is very stable; first order deviations disappear and only higher order terms remain. In the $\tilde{h}$ direction the deviations remain constant.

Suppose that we have found a fixed point $\tilde{J}^*$ for a more-dimensional Ising model at $h = 0$. Since the problem is symmetric around $h = 0$ the direction corresponding to $\tilde{J}$ must be the direction of an eigenvector of the matrix $\mathcal{M}$. Suppose the corresponding eigenvalue is λ. For small deviations at constant field $h = 0$ we can therefore write

$$\delta\tilde{J}' = \lambda\delta\tilde{J} \tag{9.181}$$

Now assume that we constructed this small deviation in $\tilde{J}$ by changing the temperature slightly. Hence we assume that $\delta\tilde{J} = \alpha(T - T_c)$ where α is a

constant. The deviation $\delta\tilde{J}'$ corresponds to a different temperature T' according to $\delta\tilde{J} = \alpha(T' - T_c)$ and we have

$$(T' - T_c) = \lambda(T - T_c) \tag{9.182}$$

On the other hand the correlation length depends on the temperature via $\xi(T) = c(T - T_c)^{-\nu}$. By construction the primed system corresponds to a re-scaled system with correlation length $\frac{\xi}{f}$. The scaling factor f is equal to two in our example above. Hence we also have

$$(T' - T_c)^{-\nu} = (T - T_c)^{-\nu}\frac{1}{f} \tag{9.183}$$

combining these two results gives

$$(T - T_c)^{-\nu}\lambda^{-\nu} = (T - T_c)^{-\nu}\frac{1}{f} \tag{9.184}$$

or $\lambda^\nu = f$. Hence if we calculate the matrix $\mathcal{M}$ **at the fixed point** the eigenvalues are related to critical exponents. We could also vary h to get similar results. The one dimensional Ising model is again too simple. The fixed points are $\tilde{J}^* = 0$ or ∞ with a value of $\lambda = 0$ or 1. If λ approaches zero from above the value of ν approaches zero. This means that the correlation length is independent of the temperature. In other words, the correlation length is not important and can be taken to be zero for all practical purposes, as expected. On the other hand, for the other fixed point λ approaches one from above, and this means that ν goes to infinity. The correlation length increases faster than a power law, which is what we already found from the exact results.

9.8   Problems for chapter 9

PROBLEM 1.

Consider the one-dimensional Ising model in a cluster approximation. A cluster contains a central atom 0 and the neighboring sites $j = \pm 1$.

A. Calculate $<\sigma_0\sigma_j>$ in this approach.

B. Assume that there is an additional potential in this model of the form $V = \lambda\sum_{j=\pm 1}\sigma_0\sigma_j$. Show that we indeed have $\left(\frac{\partial G}{\partial \lambda}\right) = <V>_\lambda$.

PROBLEM 2.

Suppose the earth magnetic field is due to a linear chain of atoms with spin one-half polarized along the chain. Take $J = 1\ eV$ and $N = 10^{16}$. Use the results for the one-dimensional Ising model to estimate the time between spontaneous reversals of the earth magnetic field.

PROBLEM 3.

Evaluate $\langle \sigma_0 \sigma_i \rangle$ in the Bethe cluster approximation for $h = 0$. Use this expression to calculate the free energy for $T > T_c$. Derive the value of the critical exponent α.

PROBLEM 4.

Consider a one-dimensional chain of spin one particles. Use the transfer matrix method to obtain the free energy of this system.

Appendix A

Solutions to selected problems.

A.1   Solutions for chapter 1.

PROBLEM 4.

There are several ways to attack this problem. Use a computer, or use the formula in the notes:

$$N! = N^N e^{-N} \sqrt{2\pi N} e^{-\frac{1}{12}\frac{1}{N}+R(N)}$$

where $R(N)$ is of order N^{-2}.
The values we use in this problem are

  (I) $N! \approx N^N$

 (II) $N! \approx N^N e^N$

(III) $N! \approx N^N e^{-N} \sqrt{N}$

(IV) $N! \approx N^N e^{-N} \sqrt{2\pi N}$

This gives for the relative error $\epsilon = |1 - \frac{approximation}{N!}|$:

  (I) $\epsilon_1 \approx |1 - e^N \frac{1}{\sqrt{2\pi N}} e^{+\frac{1}{12}\frac{1}{N}-R(N)}|$

 (II) $\epsilon_2 \approx |1 - \frac{1}{\sqrt{2\pi N}} e^{+\frac{1}{12}\frac{1}{N}-R(N)}|$

(III) $\epsilon_3 \approx |1 - \frac{1}{\sqrt{2\pi}} e^{+\frac{1}{12}\frac{1}{N}-R(N)}|$

(IV) $\epsilon_4 \approx |1 - e^{+\frac{1}{12}\frac{1}{N}-R(N)}|$

Therefore in the limit $N \to \infty$ we have

- (I) $\displaystyle\lim_{N\to\infty} \epsilon_1 = \infty$

- (II) $\displaystyle\lim_{N\to\infty} \epsilon_1 = 1$

- (III) $\displaystyle\lim_{N\to\infty} \epsilon_1 = 1 - \frac{1}{\sqrt{2\pi}}$

- (IV) $\displaystyle\lim_{N\to\infty} \epsilon_1 = 0$

and only in the last case will the relative error become small. For a 1% error we need $e^{+\frac{1}{12}\frac{1}{N} - R(N)} = 0.99$ or approximately (using $e^x \approx 1 + x$) we need $\frac{1}{12}\frac{1}{N} = 0.01$ which will be true if $N > 8$.

PROBLEM 5.

This is a problem which shows how probabilities change when you add information. This one is called the principle of restricted choice. Bridge players should be familiar with it!

Label the door you chose with A, the others with B and C. When you make your first choice, your chance of finding the car is of course $\frac{1}{3}$. Hence there are three equally probable scenarios.

- (1) $A = \text{car}, B = 0, C = 0$ Probability $\frac{1}{3}$

- (2) $A = 0, B = \text{car}, C = 0$ Probability $\frac{1}{3}$

- (3) $A = 0, B = 0, C = \text{car}$ Probability $\frac{1}{3}$

Next, the game show host points at a door. Now there are six scenarios:

- (1a) $A = \text{car}, B = 0, C = 0$ Host chooses B Probability $\frac{1}{6}$

- (1b) $A = \text{car}, B = 0, C = 0$ Host chooses C Probability $\frac{1}{6}$

- (2a) $A = 0, B = \text{car}, C = 0$ Host chooses B Probability $\frac{1}{6}$

- (2b) $A = 0, B = \text{car}, C = 0$ Host chooses C Probability $\frac{1}{6}$

- (3a) $A = 0, B = 0, C = \text{car}$ Host chooses B Probability $\frac{1}{6}$

- (3b) $A = 0, B = 0, C = \text{car}$ Host chooses C Probability $\frac{1}{6}$

which are all equally probable. But now we introduce the fact that the host cannot open the door behind which the car is hidden, which eliminates sequences 2a and 3b. In scenario 2 where the car is behind door B, the host has to choose door C! **His choice is restricted!** Hence the four possible sequences of events are:

(1a) $A = \text{car}, B = 0, C = 0$ Host chooses B Probability $\frac{1}{6}$

(1b) $A = \text{car}, B = 0, C = 0$ Host chooses C Probability $\frac{1}{6}$

(2b) $A = 0, B = \text{car}, C = 0$ Host chooses C Probability $\frac{1}{3}$

(3a) $A = 0, B = 0, C = \text{car}$ Host chooses B Probability $\frac{1}{3}$

which are not all equally probable anymore. If we now label the door which the host chose by B, we only have sequences 1a and 3a and we have

(1a) $A = \text{car}, B = 0, C = 0$ Host chooses B Probability $\frac{1}{3}$

(3a) $A = 0, B = 0, C = \text{car}$ Host chooses B Probability $\frac{2}{3}$

and therefore the probability that the car is behind the other door is twice as large as the probability of the car being behind the original door. Another way of saying this is that when the car is behind the original door A the host will choose door B only half the time and door C the other half. If the car is behind door C he has to choose door B. This reduces the probability in case 1 by one-half.

PROBLEM 6.

Define the number of atoms with spin $s_i = s$ to be N_s. Therefore:

$$\sum_{s=-S}^{S} N_s = N$$

$$\sum_{s=-S}^{S} sN_s = M$$

and the number of ways to find the set of values $N_{-S}, N_{-S+1}, \cdots, N_{S-1}, N_S$ is

$$\frac{N!}{N_{-S}!N_{-S+1}!\cdots N_{S-1}!N_S!}$$

which leads to

$$g(N, M) = \sum_{N_{-S}, N_{-S+1}, \cdots, N_{S-1}, N_S} \frac{N!}{N_{-S}!N_{-S+1}!\cdots N_{S-1}!N_S!}\delta_{\sum_{s=-S}^{S} N_s, N}\delta_{\sum_{s=-S}^{S} sN_s, M}$$

where we used the Kronecker delta to limit the sums to the right cases.
We now approximate the factorials, keeping only terms that depend on N in the exponent, $N! \approx N^N e^{-N}$, and get

$$g(N, M) = \sum_{N_{-S}, N_{-S+1}, \cdots, N_{S-1}, N_S} \left(\frac{N}{N_{-S}} \right)^{N_{-S}} \cdots \left(\frac{N}{N_S} \right)^{N_S} \delta_{\sum_{s=-S}^S N_s, N} \delta_{\sum_{s=-S}^S s N_s, M}$$

In the limit $N \to \infty$ the log of the multiplicity function is again approximated by the log of the largest term, and hence we need to find the maximum of

$$T(N_{-S}, N_{-S+1}, \cdots, N_{S-1}, N_S) = \sum_{s=-S}^S N_s(\log(N) - \log(N_s))$$

with the conditions $\sum_{s=-S}^S N_s = N$ and $\sum_{s=-S}^S s N_s = M$. This can be done by introducing two Lagrange multiplyers, and we need to minimize:

$$U(N_{-S}, N_{-S+1}, \cdots, N_{S-1}, N_S, \alpha, \beta) =$$

$$\sum_{s=-S}^S N_s(\log(N) - \log(N_s)) + \alpha(\sum_{s=-S}^S N_s - N) + \beta(\sum_{s=-S}^S s N_s - M)$$

Taking the derivatives with respect to the variables N_s and equating these to zero gives:

$$(\log(N) - \log(N_s)) - 1 + \alpha + \beta s = 0$$

or

$$N_s = N e^{-1+\alpha+\beta s}$$

with α and β determined from

$$N = \sum_{s=-S}^S N_s = N \sum_{s=-S}^S e^{-1+\alpha+\beta s}$$

$$M = xN = \sum_{s=-S}^S N_s = N \sum_{s=-S}^S s e^{-1+\alpha+\beta s}$$

Therefore, the value of T at this point is:

$$T_{max} = \sum_{s=-S}^S N_s(1 - \alpha - \beta s) = (1 - \alpha)N - \beta x N$$

which is proportional to N indeed. In order to find how this depends on x we need to solve the equations for α and β, which are (after dividing by N):

$$1 = e^{-1+\alpha} \sum_{s=-S}^S e^{+\beta s}$$

$$x = e^{-1+\alpha} \sum_{s=-S}^{S} s e^{+\beta s}$$

The derivatives with respect to x of the first equation is :

$$0 = e^{-1+\alpha} \frac{d\alpha}{dx} \sum_{s=-S}^{S} e^{+\beta s} + e^{-1+\alpha} \sum_{s=-S}^{S} s e^{+\beta s} \frac{d\beta}{dx}$$

which gives

$$0 = \frac{d\alpha}{dx} + x \frac{d\beta}{dx}$$

Also we then get

$$\frac{dT_{max}}{dx} = -\frac{d\alpha}{dx} N - \beta N - x \frac{d\beta}{dx} N = -\beta N$$

which shows that this term is maximal for $\beta = 0$ and this gives immediately $x = 0$. When β is small the equations for α and β give

$$1 = e^{-1+\alpha} \sum_{s=-S}^{S} (1 + \beta s) = e^{-1+\alpha}(2S+1)$$

$$x = e^{-1+\alpha} \sum_{s=-S}^{S} s(1 + \beta s) = \frac{1}{2S+1} \beta \sum_{s=-S}^{S} s^2$$

which shows that α is constant and $\beta \propto x$, and hence

$$T_{max} = \log(2S+1)N - cx^2 N$$

with $c = \frac{2S+1}{\sum_{s=-S}^{S} s^2}$, which is again a Gaussian distribution for the multiplicity function:

$$g(N, x) \approx g(N, 0)e^{-cNx^2}$$

PROBLEM 7.

Twelve physics graduate students go to the bookstore to buy textbooks. Eight students buy a copy of Jackson, six students buy a copy of Liboff, and two students buy no books at all. What is the probability that a student who bought a copy of Jackson also bought a copy of Liboff? What is the probability that a student who bought a copy of Liboff also bought a copy of Jackson?

We have two students who buy no book, two students who only buy Liboff, four students who only buy Jackson, and four students who buy both books. Therefore

the probability that a student who bought a copy of Jackson also bought a copy
of Liboff is four out of eight, or 50%. *Similarly, the probability that a student*
who bought a copy of Liboff also bought a copy of Jackson is four out of six, or
66.67%.

PROBLEM 8.

For an ideal gas we have $U = \frac{3}{2}Nk_BT$, where N is the number of particles. Use
the relation between the entropy $S(U, N)$ and the multiplicity function $g(U, N)$
to determine how $g(U, N)$ depends on U.

From

$$\frac{1}{T} = \left(\frac{\partial S}{\partial U}\right)_N$$

we find

$$\left(\frac{\partial S}{\partial U}\right)_N = \frac{3}{2}Nk_B\frac{1}{U}$$

which gives

$$S = \frac{3}{2}Nk_B\log(U) + C(N)$$

and with

$$g(U, N) = e^{\frac{S}{k_B}}$$

we get

$$g(U, N) = C'(N)U^{\frac{3}{2}N}$$

Does this make sense? The kinetic energy of one particle is $\frac{\hbar^2 k^2}{2m}$ and hence the
total kinetic energy is given by:

$$E_k = \frac{\hbar^2}{2m}\sum_{i=1}^{N}\sum_{j=x}^{z}k_{ij}^2$$

A surface of constant energy U is a hypersphere in 3N dimensions, with radius
R given by $U = \frac{\hbar^2 R^2}{2m}$. The number of states with a given energy is proportional
to the surface area of that sphere, which is proportional to R^{3N-1}. For large N
we ignore the -1 and hence the area is proportional to $U^{\frac{3}{2}N}$.

PROBLEM 9.

The energy eigenstates of a harmonic oscillator are $\epsilon_n = \hbar\omega(n + \frac{1}{2})$ for $n = 0, 1, 2, \cdots$ Consider a system of N such oscillators. The total energy of this system in the state $\{n_1, n_2, \cdots, n_N\}$ is

$$U = \sum_{i=1}^{N} \epsilon_{n_i} = (M + \frac{1}{2}N)\hbar\omega$$

where we have defined

$$M = \sum_{i=1}^{N} n_i$$

Calculate the multiplicity function $g(M, N)$. Hint: relate $g(M, N)$ to $g(M, N + 1)$ and use the identity

$$\sum_{k=0}^{m} \binom{n+k}{n} = \binom{n+1+m}{n+1}$$

Show that for large integers $g(M, N)$ is a narrow Gaussian distribution in $x = \frac{M}{N}$.

It is easy to show that

$$g(M, N) = \sum_{m=0}^{M} g(m, n)g(M - m, N - n)$$

because the states of parts of the system are independent. We also know that

$$g(M, 1) = 1$$

because for one particle there is only one way of obtaining M units of energy. Therefore

$$g(M, N + 1) = \sum_{m=0}^{M} g(m, N)$$

Comparing with the hint we see that

$$g(M, N) = \binom{n + M}{n}$$

for some value of n linearly related to N, or $n = N + n_0$. Because $g(M, 1) = 1$ and $\binom{0+M}{0} = 1$ we find that $n = N - 1$ and hence

$$g(M, N) = \binom{N - 1 + M}{N - 1}$$

Next we take the logarithms:

$$\log g(M, N) = \log(N - 1 + M)! - \log(N - 1)! - \log M!$$

and using Stirling

$$\log N! = N \log(N) - N + \frac{1}{2} \log(2\pi N)$$

we arrive at

$$\log g(M, N) = (N - 1 + M) \log(N - 1 + M) - (N - 1) \log(N - 1) - M \log(M)$$

$$-\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(\frac{N - 1 + M}{(N - 1)M})$$

Insert $M = xN$ and replace $N - 1$ by N:

$$\log g(M, N) = N(1 + x) \log(N(1 + x)) - N \log N - xN \log(xN)$$

$$-\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(\frac{1 + x}{xN})$$

$$\log g(M, N) = N(1 + x) \log(1 + x) - xN \log(x)$$

$$-\frac{1}{2} \log(2\pi) + \frac{1}{2} \log(\frac{1 + x}{xN})$$

The maximum occurs if $\frac{\partial}{\partial x} g = 0$ which gives:

$$N \log(1 + x) - N \log(x) + \frac{1}{2} \left(\frac{1}{1 + x} - \frac{1}{x} \right) = 0$$

Consider the function:

$$f(x) = N \log(x) + \frac{1}{2} \frac{1}{x}$$

and we need to solve

$$f(x + 1) = f(x)$$

The slope of the function $f(x)$ is given by:

$$\frac{df}{dx} = \frac{N}{x} - \frac{1}{2} \frac{1}{x^2}$$

which is positive for $x > \frac{1}{2N}$. Because $g = f(x + 1) - f(x)$ this shows that g does not have a maximum. The largest value is at $x = \infty$, which makes sense if you think about the physics. But my apologies for making the last part of the problem misleading.

A.2 Solutions for chapter 2.

PROBLEM 4.

Assume that the protons are labelled by i, with i running from 1 to N, and that the spin of the proton is $s_i\mu$. The energy of such a state is $\sum_i(-s_i\mu H)$. The partition function is

$$Z(T) = \sum_{s_1,\cdots,s_N} e^{\frac{1}{kT}\sum_i s_i\mu H}$$

which is

$$Z(T) = (Z_1(T))^N$$

$$Z_1(T) = e^{-\frac{\mu H}{kT}} + e^{+\frac{\mu H}{kT}}$$

The internal energy of the sample is $U = -\sum_i s_i\mu H$, which is proportional to the difference in population. Therefore the power absorbed is proportional to U. We also have:

$$U = kT^2\frac{\partial}{\partial T}\log(Z(T))$$

and hence

$$U = kT^2 N\frac{e^{+\frac{\mu H}{kT}} - e^{-\frac{\mu H}{kT}}}{e^{+\frac{\mu H}{kT}} + e^{-\frac{\mu H}{kT}}}\left(\frac{-\mu H}{kT^2}\right)$$

When $\mu H \ll kT$ we replace the exponents by $e^x = 1 + x$ and get

$$U \approx -N\mu H\frac{+\frac{\mu H}{kT} - (-\frac{\mu H}{kT})}{2} = -N\frac{\mu^2 H^2}{kT}$$

and the power is inversely proportional to T.

PROBLEM 5.

(a) At low temperature $\bar{E} \approx N\epsilon_1$, all particles in ground state. At high temperature all states are equally probable, and $\bar{E} \approx \frac{N}{2}(\epsilon_1 + \epsilon_2)$. In thermodynamics we have seen that $\left(\frac{\partial U}{\partial T}\right)_V \to 0$ for $T \to 0$, because the entropy becomes zero. Therefore the curve $\bar{E}(T)$ starts with horizontal slope at $T = 0$ bends upwards, bends down again and approaches the high temperature limit in an asymptotic fashion. We expect the change from low to high temperature behavior to take place near a temperature T_0 given by $\epsilon_2 - \epsilon_1 = kT_0$, because now we have enough thermal energy to make the transition from state 1 to state 2.

(b) Because $C_V = \left(\frac{\partial U}{\partial T}\right)_V$ we start at zero for low temperature, go through a maximum near T_0 and approach zero again for high temperature.

(c) Like in the previous problem we can write the partition function in the following manner, by defining a variable $s_i = \pm 1$ for each particle by $\epsilon_i = \frac{1}{2}(\epsilon_1 + \epsilon_2) + s_i \frac{1}{2}(\epsilon_2 - \epsilon_1) = A + Bs_i$. This gives

$$Z(T) = \sum_{s_1,\cdots,s_N} e^{-\frac{1}{kT}\sum_i (A + Bs_i)}$$

which is

$$Z(T) = e^{-\frac{NA}{kT}} \left(Z_1(T)\right)^N$$

$$Z_1(T) = e^{\frac{B}{kT}} + e^{-\frac{B}{kT}}$$

Using

$$U = kT^2 \frac{\partial}{\partial T} \log(Z(T))$$

we get

$$U = kT^2 \frac{\partial}{\partial T} \left(-\frac{NA}{kT} + N\log(Z_1)\right)$$

or

$$U = NA + NkT^2 \frac{e^{\frac{B}{kT}} - e^{-\frac{B}{kT}}}{e^{\frac{B}{kT}} + e^{-\frac{B}{kT}}} \left(\frac{-B}{kT^2}\right)$$

which leads to

$$U = N(A - Bf(T))$$

or

$$U = N(\frac{1}{2}(\epsilon_1 + \epsilon_2) - \frac{1}{2}(\epsilon_2 - \epsilon_1)f(T))$$

with

$$f(T) = \frac{e^{\frac{B}{kT}} - e^{-\frac{B}{kT}}}{e^{\frac{B}{kT}} + e^{-\frac{B}{kT}}}$$

This gives the correct limits indeed. $f(0) = 1$ which gives $U(0) = N\epsilon_1$. $f(\infty) = 0$ which gives $U(\infty) = \frac{1}{2}(\epsilon_1 + \epsilon_2)$.

Also

$$C_V = -\frac{N}{2}(\epsilon_2 - \epsilon_1)f'(T)$$

and using

$$\frac{d}{dx}\frac{e^x - e^{-x}}{e^x + e^{-x}} = 1 - \frac{(e^x - e^{-x})^2}{(e^x + e^{-x})^2} = \frac{4}{(e^x + e^{-x})^2}$$

we get

$$C_V = \frac{4NB^2}{kT^2}\frac{1}{(e^{\frac{B}{kT}} + e^{-\frac{B}{kT}})^2}$$

For T to zero this approaches zero exponentially, for T to infinity this goes to zero inverse quadratically, and the approaching zero confirms what we found in (b).

We can write $C_V = Nkg(\frac{B}{kT})$ with

$$g(x) = \frac{4x^2}{(e^x + e^{-x})^2}$$

$$g'(x) = \frac{8x}{(e^x + e^{-x})^2} - \frac{4x^2(e^x - e^{-x})}{(e^x + e^{-x})^3} = g(x)\left(\frac{2}{x} - \frac{e^x - e^{-x}}{e^x + e^{-x}}\right)$$

which is zero if

$$0 = \frac{2}{x} - \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

which happens when x is somewhat larger than 2, or $kT_0 \approx \frac{1}{4}(\epsilon_2 - \epsilon_1)$, more or less as in (b).

PROBLEM 6.

Each positive ion can be in one of four states. If the central atom is at the origin, the x coordinates of these states are $-\frac{1}{2}a$ for i=1,2 and $+\frac{1}{2}a$ for i=3,4. The component of the dipole along the x direction is $-\frac{1}{2}ea$ and $+\frac{1}{2}ea$ respectively, and the energies are $+\frac{1}{2}eaE$ and $-\frac{1}{2}eaE$. The state of atom j is therefore given by a number s_j which runs from 1 to 4, and the partition function is

$$Z(T) = \sum_{s_1, \cdots, s_N} e^{-\frac{1}{kT}\sum_j \epsilon(s_j)}$$

Again, the energies are independent and we get

$$Z(T) = (Z_1(T))^N$$

$$Z_1(T) = 2\left(e^{+\frac{eaE}{2kT}} + e^{-\frac{eaE}{2kT}}\right)$$

The probability of finding a state with quantum numbers $s_1, \cdots, s_N$ is therefore:

$$Prob(s_1, \cdots, s_N) = \frac{1}{Z(T)}e^{-\frac{1}{kT}\sum_j \epsilon(s_j)}$$

and the average dipole moment is

$$< P >= \sum_{s_1,\cdots,s_N} \sum_j P(s_j) Prob(s_1,\cdots,s_N)$$

which simplifies to

$$< P >= N \frac{(\frac{1}{2}ea)e^{+\frac{eaE}{2kT}} + (-\frac{1}{2}ea)e^{-\frac{eaE}{2kT}}}{e^{+\frac{eaE}{2kT}} + e^{-\frac{eaE}{2kT}}}$$

or

$$< P >= \frac{1}{2}Nae\tanh(\frac{eaE}{2kT})$$

which reduces for small fields, or $eaE \ll 2kT$ to

$$< P >\approx NE\frac{a^2e^2}{4kT}$$

PROBLEM 7.

The probability of finding a system in a state s is P_s. In this case the entropy of the system is $S = -k_B \sum_s P_s \log(P_s)$. Assume that the system consists of two independent subsystems and that the state s is the combination of subsystem 1 being in state s_1 and system 2 being in state s_2. Use the formula for the entropy given above to show that $S = S_1 + S_2$.

The state of the system is given by the states of the subsystems, $s = (s_1, s_2)$. Because the systems are independent we have $P_s = P_{s_1}P_{s_2}$ and hence:

$$S = -k_B \sum_s P_s \log(P_s) = -k_B \sum_{s_1,s_2} P_{s_1} P_{s_2} \log(P_{s_1} P_{s_2})$$

$$S = -k_B \sum_{s_1} \sum_{s_2} P_{s_1} P_{s_2} \left(\log(P_{s_1}) + \log(P_{s_1})\right)$$

With $\sum P_{s_i} = 1$ this gives

$$S = -k_B \sum_{s_1} P_{s_1} \log(P_{s_1}) + \sum_{s_2} P_{s_2} \log(P_{s_1}) = S_1 + S_2$$

PROBLEM 8.

The energy eigenvalues for a single particle in a given potential are ϵ_n. These energy values are independent of how many other particles are present. Show that the partition function $\mathcal{Z}(T,N)$ obeys the relation $\mathcal{Z}(T,N) = (\mathcal{Z}(T,1))^N$.

For one particle we have:

$$\mathcal{Z}(T,1) = \sum_{n_1} e^{-\beta \epsilon_{n_1}}$$

When we have N particles they are in states $n_1, n_2, \cdots, n_N$ *and we have*

$$\mathcal{Z}(T,N) = \sum_{n_1} \sum_{n_2} \cdots \sum_{n_N} e^{-\beta\left(\epsilon_{n_1}+\epsilon_{n_2}+\cdots+\epsilon_{n_N}\right)}$$

Which is equal to :

$$\mathcal{Z}(T,N) = \sum_{n_1} e^{-\beta \epsilon_{n_1}} \sum_{n_2} e^{-\beta \epsilon_{n_2}} \cdots \sum_{n_N} e^{-\beta \epsilon_{n_N}} = \left(\mathcal{Z}(T,1)\right)^N$$

PROBLEM 9.

The energy eigenvalues of a system are given by 0 and $\epsilon + n\Delta$ for $n = 0, 1, 2, \cdots$. We have both $\epsilon > 0$ and $\Delta > 0$. Calculate the partition function for this system. Calculate the internal energy and the heat capacity. Plot the heat capacity as a function of temperature for $0 < k_B T < \epsilon$ for (a) $\Delta \gg \epsilon$, (b) $\Delta = \epsilon$, and (c) $\Delta \ll \epsilon$.

$$\mathcal{Z} = e^{\beta 0} + \sum_{n=0}^{\infty} e^{-\beta(\epsilon+n\Delta)} = 1 + e^{-\beta\epsilon} \sum_{n} e^{-n\beta\Delta}$$

The sum can be done using $\sum_n r^n = \frac{1}{1-r}$ *and gives:*

$$\mathcal{Z} = 1 + \frac{e^{-\beta\epsilon}}{1 - e^{-\beta\Delta}} = \frac{1 + e^{-\beta\epsilon} - e^{-\beta\Delta}}{1 - e^{-\beta\Delta}}$$

$$U = k_B T^2 \frac{\partial}{\partial T} \log(\mathcal{Z}) = -\frac{\partial}{\partial \beta} \log(\mathcal{Z})$$

$$U = \frac{\epsilon e^{-\beta\epsilon} - \Delta e^{-\beta\Delta}}{1 + e^{-\beta\epsilon} - e^{-\beta\Delta}} + \frac{\Delta e^{-\beta\Delta}}{1 - e^{-\beta\Delta}}$$

$$U = \epsilon \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon} - e^{-\beta\Delta}} + \Delta \left(\frac{e^{-\beta\Delta}}{1 - e^{-\beta\Delta}} - \frac{e^{-\beta\Delta}}{1 + e^{-\beta\epsilon} - e^{-\beta\Delta}} \right)$$

We have three cases:

$$\Delta \gg \epsilon \Rightarrow U \approx \epsilon \frac{e^{-\beta\epsilon}}{1 + e^{-\beta\epsilon}} = \epsilon \frac{1}{e^{\beta\epsilon} + 1}$$

$$\Delta = \epsilon \Rightarrow U = \epsilon \frac{e^{-\beta\epsilon}}{1 - e^{-\beta\epsilon}} = \epsilon \frac{1}{e^{\beta\epsilon} - 1}$$

$$\Delta \ll \epsilon \Rightarrow U \approx \epsilon + k_B T (\text{for } k_B T \gg \Delta)$$

Finally, we have $C_V = \left(\frac{\partial U}{\partial T}\right) = -k_B\beta^2\left(\frac{\partial U}{\partial \beta}\right)$ *and hence in these cases we find:*

$$\Delta \gg \epsilon \Rightarrow C_V \approx k_B(\beta\epsilon)^2\frac{e^{\beta\epsilon}}{\left(e^{\beta\epsilon} + 1\right)^2}$$

$$\Delta = \epsilon \Rightarrow C_V = k_B(\beta\epsilon)^2\frac{e^{\beta\epsilon}}{\left(e^{\beta\epsilon} - 1\right)^2}$$

$$\Delta \ll \epsilon \Rightarrow C_V \approx k_B(\text{for } k_BT \gg \Delta)$$

A.3   Solutions for chapter 3

PROBLEM 3.

A system contains an ideal gas of atoms with spin $\frac{1}{2}$ in a magnetic field $B(\vec{r})$. The concentration of the spin up (down) particles is $n_\uparrow(\vec{r})$ ($n_\downarrow(\vec{r})$). The temperature is T.

(A) Evaluate the **total** chemical potential for the spin up and down particles.

The magnetic moment of an atom with spin $\vec{S}$ is $\vec{m} = \gamma\mu_B\vec{S}$. Assuming that the quantization axis is along the B-field, and that the energy per atom is $U_{mag} = -\vec{m} \cdot \vec{B}$, we get:

$$\mu_\uparrow(\vec{r}) = k_BT\log(\frac{n_\uparrow(\vec{r})}{n_Q(T)}) - \frac{1}{2}\gamma\mu_B B(\vec{r})$$

$$\mu_\downarrow(\vec{r}) = k_BT\log(\frac{n_\downarrow(\vec{r})}{n_Q(T)}) + \frac{1}{2}\gamma\mu_B B(\vec{r})$$

(B) These two chemical potentials have to be the same and independent of $\vec{r}$. Explain why.

If either chemical potential were position dependent, there would be no equilibrium and a flow of particles would result. Hence both the spin-up and spin-down chemical potentials are independent of position. If these two potentials were not the same, the spins of the atoms would flip, and again that would be a non equilibrium situation.

(C) Calculate the magnetic moment of this gas as a function of position.

$$\mu = k_BT\log(\frac{n_\uparrow(\vec{r})}{n_Q(T)}) - \frac{1}{2}\gamma\mu_B B(\vec{r})$$

gives

$$n_\uparrow(\vec{r}) = n_Q(T)e^{\frac{1}{k_BT}(\mu + \frac{1}{2}\gamma\mu_B B(\vec{r}))}$$

and

$$\mu = k_BT\log(\frac{n_\downarrow(\vec{r})}{n_Q(T)}) + \frac{1}{2}\gamma\mu_B B(\vec{r})$$

gives

$$n_\downarrow(\vec{r}) = n_Q(T)e^{\frac{1}{k_BT}(\mu - \frac{1}{2}\gamma\mu_B B(\vec{r}))}$$

Therefore, the magnetization density is:

$$m(\vec{r}) = \gamma\mu_B\frac{1}{2}(n_\uparrow(\vec{r}) - n_\downarrow(\vec{r})) = \gamma\mu_B\frac{1}{2}n_Q(T)e^{\frac{\mu}{k_BT}}\left(e^{\frac{\gamma\mu_B B(\vec{r})}{2k_BT}} - e^{-\frac{\gamma\mu_B B(\vec{r})}{2k_BT}}\right)$$

or

$$m(\vec{r}) = \gamma\mu_B n_Q(T)e^{\frac{\mu}{k_BT}}\sinh(\frac{\gamma\mu_B B(\vec{r})}{2k_BT})$$

(D) Show that the concentration of magnetic particles is high in regions with a large magnetic field.

$$n(\vec{r}) = n_\uparrow(\vec{r}) + n_\downarrow(\vec{r}) = n_Q(T)e^{\frac{\mu}{k_BT}}\left(e^{\frac{\gamma\mu_B B(\vec{r})}{2k_BT}} + e^{-\frac{\gamma\mu_B B(\vec{r})}{2k_BT}}\right)$$

or

$$n(\vec{r}) = 2n_Q(T)e^{\frac{\mu}{k_BT}}\cosh(\frac{\gamma\mu_B B(\vec{r})}{2k_BT})$$

which is clearly largest when B is largest.

P.S. Also note that when $\gamma\mu_B B(\vec{r}) \ll 2k_BT$ we have:

$$n(\vec{r}) \approx 2n_Q(T)e^{\frac{\mu}{k_BT}} = n_0$$

which is constant and defines the chemical potential. The magnetization density is:

$$m(\vec{r}) \approx \gamma\mu_B n_Q(T)e^{\frac{\mu}{k_BT}}\frac{\gamma\mu_B B(\vec{r})}{2k_BT} = n_0\frac{\gamma^2\mu_B^2}{4k_BT}B(\vec{r})$$

which is the famous Curie-Weiss law.

PROBLEM 4.

The state of a many body system is characterized by two quantum numbers, n and m. The possible values of the quantum number n are $0, 1, 2, \cdots, \infty$, while the values of m are in the range $0, 1, \cdots, n$. The energy of the system in the state (n, m) is $n\hbar\omega$ and the number of particles is m. Evaluate the grand partition function for this system.

$$\mathfrak{Z} = \sum_{n=0}^{\infty} \sum_{m=0}^{n} e^{-\beta(n\hbar\omega - \mu m)}$$

$$\mathfrak{Z} = \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega} \sum_{m=0}^{n} e^{m\beta\mu}$$

$$\mathfrak{Z} = \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega} \frac{1 - e^{(n+1)\beta\mu}}{1 - e^{\beta\mu}}$$

$$\mathfrak{Z} = \frac{1}{1 - e^{\beta\mu}} \sum_{n=0}^{\infty} e^{-n\beta\hbar\omega} - \frac{e^{\beta\mu}}{1 - e^{\beta\mu}} \sum_{n=0}^{\infty} e^{-n\beta(\hbar\omega - \mu)}$$

$$\mathfrak{Z} = \frac{1}{1 - e^{\beta\mu}} \frac{1}{1 - e^{-\beta\hbar\omega}} - \frac{e^{\beta\mu}}{1 - e^{\beta\mu}} \frac{1}{1 - e^{\beta(\mu - \hbar\omega)}}$$

or

$$\mathfrak{Z} = \frac{1}{1 - e^{\beta\mu}} \left(\frac{1}{1 - e^{-\beta\hbar\omega}} - \frac{1}{e^{-\beta\mu} - e^{-\beta\hbar\omega}} \right)$$

PROBLEM 5.

An ideal gas of atoms with mass m is contained in a cylinder that spins around with angular frequency ω. The system is in equilibrium. The distance to the axis of the cylinder is r. The radius of the cylinder is R. Calculate the density of the gas as a function of r.

Reasoning similar to problem 1 .

$$\mu = k_B T \log(\frac{n(\vec{r})}{n_Q(T)}) + \frac{1}{2}m\omega^2 r^2$$

$$n(\vec{r}) = n_Q(T)e^{\beta\mu}e^{-\frac{m\omega^2}{2k_B T}r^2}$$

PROBLEM 6.

Extremely relativistic particles obey the relation $E(\vec{k}) = \hbar c|\vec{k}|$. Assume we have a gas of these identical particles at low density, or $n \ll n_Q(T)$.

(A) Calculate the partition function $\mathcal{Z}_1(T, V)$ for N=1.

(B) Calculate $\mathcal{Z}(T, V, N)$.

(C) Calculate $p(T, V, N)$, $S(T, V, N)$, and $\mu(T, V, N)$.

Ignore spin

$$\mathcal{Z}_1(T, V) = \sum_{n_x, n_y, n_z} e^{-\frac{\hbar c \pi}{k_B T L} \sqrt{n_x^2 + n_y^2 + n_z^2}}$$

Transform to integral

$$\vec{x} = \frac{\hbar c \pi}{k_B T L}(n_x, n_y, n_z)$$

$$\mathcal{Z}_1(T, V) = \left(\frac{\hbar c \pi}{k_B T L}\right)^{-3} \frac{1}{8} \int d^3 x e^{-x}$$

where we extended the range of integration to all values of $\vec{x}$, and not only one octant.

$$\mathcal{Z}_1(T, V) = \left(\frac{k_B T L}{\hbar c 2 \pi}\right)^3 \int_0^\infty 4\pi x^2 dx e^{-x}$$

$$\mathcal{Z}_1(T, V) = \left(\frac{k_B T L}{\hbar c 2 \pi}\right)^3 8\pi = V \left(\frac{k_B T}{\hbar c 2 \pi}\right)^3 8\pi$$

This is allowed at low density, when the points in x-space are close together

$$\mathcal{Z}(T, V, N) = \frac{1}{N!} V^N \left(\frac{k_B T}{\hbar c 2 \pi}\right)^{3N} (8\pi)^N$$

$$F = -k_B T \log(\mathcal{Z}) = -N k_B T \log(8\pi \frac{V}{N} \left(\frac{k_B T}{\hbar c 2 \pi}\right)^3) - N k_B T$$

Where we used Stirling

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

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} = N k_B \log(8\pi \frac{V}{N} \left(\frac{k_B T}{\hbar c 2 \pi}\right)^3) + 4 N k_B$$

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} = -k_B T \log(8\pi \frac{V}{N} \left(\frac{k_B T}{\hbar c 2 \pi}\right)^3)$$

Check:

$$U = k_B T^2 \frac{\partial}{\partial T} \log(\mathcal{Z}) = 3N k_B T$$

$$TS - pV + \mu N = 3N k_B T$$

PROBLEM 7.

$$\mu = k_B T \log(\frac{n}{n_Q(T)})$$

We need to include **both inside and outside** the bottle!

$$\Delta F = (\mu_{in} - \mu_{out})\Delta N$$

when ΔN molecules move from outside to inside the bottle. This leads to

$$\Delta F = k_B T \log(\frac{n_{in}}{n_{out}})\Delta N$$

This is only negative for $\Delta N > 0$ when $n_{in} < n_{out}$, and hence when the outside density is higher.

PROBLEM 8.

This is Planck's law for the distribution of photons. Assume that the volume dependence is in the frequency ω, but that the frequency does not depend on the number of photons present. This gives:

$$N(T, V, \mu) = \left[e^{\frac{\hbar \omega(V)}{k_B T}} - 1 \right]^{-1}$$

which is independent of μ! The grand potential is therefore

$$\Omega(T, V, \mu) = \Omega(T, V, 0) - \int_0^\mu N d\mu' = \Omega(T, V, 0) - N\mu$$

and therefore the Helmholtz free energy is

$$F(T, V, N) = \Omega(T, V, \mu(N)) + N\mu(N) = \Omega(T, V, 0)$$

which does not depend on N!!
Hence

$$\mu = \left(\frac{\partial F}{\partial N} \right)_{T,V} = 0$$

which means that it does not cost energy to create and destroy photons, which is, of course, what should happen in a cavity, where photons are always created and destroyed.

PROBLEM 9

$$\mathcal{Z}(T,\mu) = 1 + e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} + e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}$$

$$N(T,\mu) = \frac{1}{\mathcal{Z}}\left(e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} + 2e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}\right)$$

$$\left(\frac{\partial N}{\partial T}\right)_\mu = -\frac{1}{\mathcal{Z}^2}\left(\frac{\partial \mathcal{Z}}{\partial T}\right)_\mu\left(e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} + 2e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}\right) +$$

$$\frac{1}{\mathcal{Z}}\left((\epsilon_1-\mu)e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + (\epsilon_2-\mu)e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} + 2(\epsilon_1+\epsilon_2+I-2\mu)e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}\right)\frac{1}{k_BT^2}$$

$$\mathcal{Z}k_BT^2\left(\frac{\partial N}{\partial T}\right)_\mu =$$

$$-N(T,\mu)\left((\epsilon_1-\mu)e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + (\epsilon_2-\mu)e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} + (\epsilon_1+\epsilon_2+I-2\mu)e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}\right) +$$

$$\left((\epsilon_1-\mu)e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + (\epsilon_2-\mu)e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} + 2(\epsilon_1+\epsilon_2+I-2\mu)e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}\right) =$$

$$(1-N(T,\mu))(\epsilon_1-\mu)e^{-\frac{1}{k_BT}(\epsilon_1-\mu)} + (1-N(T,\mu))(\epsilon_2-\mu)e^{-\frac{1}{k_BT}(\epsilon_2-\mu)} +$$

$$(2-N(T,\mu))(\epsilon_1+\epsilon_2+I-2\mu)e^{-\frac{1}{k_BT}(\epsilon_1+\epsilon_2+I-2\mu)}$$

This can be negative if $N(T,\mu) > 1$, which requires that the lowest energy state is $\epsilon_1 + \epsilon_2 + I$, or $I < -\epsilon_1 - \epsilon_2$. In that case the binding energy between the two particles $(-I)$ is so strong that particles are most likely to be found two at a time. In that case $N(T=0,\mu) = 2$ and this value can only become smaller with temperature.

A.4 Solutions for chapter 4.

PROBLEM 4.

The Maxwell distribution function f_M is given by $f_M(\epsilon;T,\mu) = e^{\frac{1}{k_BT}(\mu-\epsilon)}$. Show that

$$S(T, \mu, V) = Nk_B - \sum_o f_M(\epsilon_o; T, \mu) \log(f_M(\epsilon_o; T, \mu))$$

where the sum is over orbitals.

There are several ways to do this. For example:

$$< n >_o = k_B T \left(\frac{\partial \log 3_o}{\partial \mu} \right)_{T,V}$$

$$< n >_o = f_M(\epsilon_o; T, \mu) = e^{\frac{1}{k_B T}(\mu - \epsilon_o)}$$

$$\left(\frac{\partial \log 3_o}{\partial \mu} \right)_{T,V} = \frac{1}{k_B T} e^{\frac{1}{k_B T}(\mu - \epsilon_o)}$$

$$\log 3_o = e^{\frac{1}{k_B T}(\mu - \epsilon_o)} = f_M(\epsilon_o; T, \mu)$$

$$\Omega = -k_B T \sum_o \log(3_o) = -k_B T \sum_o f_M(\epsilon_o; T, \mu)$$

$$S = -\left(\frac{\partial \Omega}{\partial T} \right)_{V,\mu} = k_B \sum_o f_M(\epsilon_o; T, \mu) + k_B T \sum_o \frac{\partial}{\partial T} f_M(\epsilon_o; T, \mu)$$

$$\sum_o f_M(\epsilon_o; T, \mu) = N$$

$$\frac{\partial}{\partial T} f_M(\epsilon_o; T, \mu) = -\frac{1}{k_B T^2}(\mu - \epsilon_o) e^{\frac{1}{k_B T}(\mu - \epsilon_o)} = -\frac{1}{T} \log(f_M(\epsilon_o; T, \mu)) f_M(\epsilon_o; T, \mu)$$

which gives the result we need

PROBLEM 5.

Consider a system of independent particles. The number of orbitals with energy between E and $E + dE$ is given by $N(E)dE$. The function $N(E)$ is called the density of states. One measures the expectation value of a certain operator $\mathcal{O}$. For a particle in an orbital o the value of the operator depends only on the energy of the orbital, or $\mathcal{O}_o = O(\epsilon_o)$. Show that in the thermodynamic limit the ensemble average of the operator is given by

$$< \mathcal{O} > = \int_{-\infty}^{\infty} O(E) N(E) f(E; T, \mu) dE$$

where $f(E; T, \mu)$ is the distribution function for these particles.

$$< \mathcal{O} > (T, \mu, V) = \sum_o \mathcal{O}_o f(\epsilon_o; T, \mu)$$

$$< \mathcal{O} > (T, \mu, V) = \sum_o O(\epsilon_o) f(\epsilon_o; T, \mu)$$

Define an energy mesh by $E_i = idE$ where dE will be very small

$$< \mathcal{O} > (T, \mu, V) = \sum_i \sum_{o, E_i < \epsilon_o < E_{i+1}} O(\epsilon_o) f(\epsilon_o; T, \mu)$$

$$< \mathcal{O} > (T, \mu, V) = \sum_i \sum_{o, E_i < \epsilon_o < E_{i+1}} O(E_i) f(E_i; T, \mu)$$

$$< \mathcal{O} > (T, \mu, V) = \sum_i N(E_i) dE O(E_i) f(E_i; T, \mu)$$

which leads to the integral

PROBLEM 6.

The orbital energies of a system of Fermions are given by $\epsilon_i = i\Delta$, with $\Delta > 0$ and $i = 1, 2, 3, \cdots, \infty$. These energies are non-degenerate. If the system has N particles, show that the low temperature limit of the chemical potential gives $\epsilon_F = (N + \frac{1}{2})\Delta$.

$$N = \sum_i \frac{1}{e^{\frac{1}{k_B T}(i\Delta - \mu)} + 1}$$

At $T = 0$ the lowest N states are occupied:

$$N = \sum_{i=1}^{N} 1$$

Taking the difference:

$$0 = \sum_{i=1}^{N} \left(\frac{1}{e^{\frac{1}{k_B T}(i\Delta - \mu)} + 1} - 1 \right) + \sum_{i=N+1}^{\infty} \left(\frac{1}{e^{\frac{1}{k_B T}(i\Delta - \mu)} + 1} \right)$$

$$\sum_{i=1}^{N} \left(1 - \frac{1}{e^{\frac{1}{k_B T}(i\Delta - \mu)} + 1} \right) = \sum_{i=N+1}^{\infty} \left(\frac{1}{e^{\frac{1}{k_B T}(i\Delta - \mu)} + 1} \right)$$

$$\sum_{i=1}^{N} \left(\frac{1}{e^{-\frac{1}{k_B T}(i\Delta - \mu)} + 1} \right) = \sum_{i=N+1}^{\infty} \left(\frac{1}{e^{\frac{1}{k_B T}(i\Delta - \mu)} + 1} \right)$$

for small T the terms closest to i=N are the most important:

$$\left(\frac{1}{e^{-\frac{1}{k_BT}(N\Delta-\mu)}+1}\right)\sum_{i=1}^{N}\left(\frac{e^{-\frac{1}{k_BT}(N\Delta-\mu)}+1}{e^{-\frac{1}{k_BT}(i\Delta-\mu)}+1}\right)=$$

$$\left(\frac{1}{e^{\frac{1}{k_BT}((N+1)\Delta-\mu)}+1}\right)\sum_{i=N+1}^{\infty}\left(\frac{e^{\frac{1}{k_BT}((N+1)\Delta-\mu)}}{e^{\frac{1}{k_BT}(i\Delta-\mu)}+1}\right)$$

When the temperature is small, the sums go to one and we have:

$$\frac{1}{e^{-\frac{1}{k_BT}(N\Delta-\mu)}+1}=\frac{1}{e^{\frac{1}{k_BT}((N+1)\Delta-\mu)}+1}$$

$$e^{-\frac{1}{k_BT}(N\Delta-\mu)}+1=e^{\frac{1}{k_BT}((N+1)\Delta-\mu)}+1$$

$$-\frac{1}{k_BT}(N\Delta-\mu)=\frac{1}{k_BT}((N+1)\Delta-\mu)$$

$$-(N\Delta-\mu)=((N+1)\Delta-\mu)$$

which gives the desired result, since at $T=0$ $\mu=\epsilon_F$. In words, for low temperature the amount on electrons leaving the state $i=N$ is equal to the amount gained in $i=N+1$, and $f_{FD}(\mu+x;T,\mu)+f_{FD}(\mu-x;T,\mu)=1$

PROBLEM 7.

The entropy for a system of independent Fermions is given by

$$S=-k_B\sum_{o}\left(f_{FD}\log(f_{FD})+(1-f_{FD})\log(1-f_{FD})\right)$$

Calculate $\lim_{T\to0}f_{FD}(\epsilon,T,\mu)$ for $\epsilon<\mu$,$\epsilon=\mu$, and $\epsilon>\mu$.

$$f_{FD}(\epsilon,T,\mu)=\frac{1}{e^{\beta(\epsilon-\mu)}+1}$$

$$f_{FD}(\mu,T,\mu)=\frac{1}{2}$$

Therefore, the limits are 1, $\frac{1}{2}$, and 0, respectively.
The number of orbitals with energy ϵ_o equal to the Fermi energy ϵ_F is M. Calculate the entropy at $T=0$ in this case.

$$S=-k_BM\left(\frac{1}{2}\log(\frac{1}{2})+\frac{1}{2}\log(\frac{1}{2})\right)=k_B\log(2^M)$$

Explain your answer in terms of a multiplicity function. Pay close attention to the issue of dependent and independent variables.

The number of ways to occupy M orbitals is 2^M, which included all states ranging from no particles at the Fermi energy to all orbitals at the Fermi energy filled. This is correct, since the formula gives the entropy as a function of μ, and not at a given number of particles!

A.5 Solutions for chapter 5.

PROBLEM 4.

Pauli Paramagnetism.
The energy of non-relativistic electrons in a small magnetic field is given by $\epsilon_{\vec{p},s} = \frac{p^2}{2m} - s\mu_0 B$ where $s = \pm 1$ and μ_0 is the magnetic moment of the electron. Assume $\mu_0 B \ll \epsilon_F$. Note that in this problem we ignore the effect of the magnetic field on the orbit of the electron, that turns out to be OK in first approximation. Evaluate the magnetic susceptibility χ in the following four cases:

(A) For $T = 0$.

(B) For $k_B T \ll \epsilon_F$, one more term beyond (A).

(C) For $T = \infty$. (Note: this one you can get without any detailed calculation).

(D) For $k_B T \gg \epsilon_F$, one more term beyond (C).

$$N = V\hbar^{-3} \sum_s \int d^3 p f_{FD}(\epsilon_{\vec{p},s})$$

$$M = V\hbar^{-3}\mu_0 \sum_s \int d^3 p s f_{FD}(\epsilon_{\vec{p},s})$$

We use the first integral to find $\mu(N,V,T)$ and then the second to find $M(N,V,T)$. For low temperature we write the integrals in the following way:

$$I(T,\mu) = V\hbar^{-3} \int d^3 p f_{FD}(\frac{p^2}{2m},T,\mu)$$

Since the argument of the Fermi-Dirac distribution is given by $\mu - \epsilon = \mu - \frac{p^2}{2m} + s\mu_0 B$, we can rewrite the expressions for N and M as:

$$N = I(T,\mu + \mu_0 B) + I(T,\mu - \mu_0 B)$$

$$M = \mu_0(I(T,\mu + \mu_0 B) - I(T,\mu - \mu_0 B))$$

Using the fact that B is very small we find up to first order:

$$N = 2I(T, \mu)$$

$$M = 2\mu_0^2 B \left(\frac{\partial I}{\partial \mu}\right)_T (T, \mu)$$

$$\chi = 2\mu_0^2 \left(\frac{\partial I}{\partial \mu}\right)_T (T, \mu)$$

$$\chi = N\mu_0^2 \left(\frac{\partial \log I}{\partial \mu}\right)_T (T, \mu)$$

At $T = 0$ the integral is easy, since $\mu(T = 0) = \epsilon_F$:

$$I(T = 0, \epsilon_F) = V\hbar^{-3}\frac{4\pi}{3}p_F^3 = V\hbar^{-3}\frac{4\pi}{3}(2m\epsilon_F)^{\frac{3}{2}}$$

with $\epsilon_F = \frac{p_F^2}{2m}$. Hence

$$\chi(T = 0) = N\mu_0^2\frac{3}{2}\frac{1}{\epsilon_F}$$

For small T we use $\mu = \epsilon_F - \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F}$ which was obtained from:

$$N = 2I(T, \epsilon_F - \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F})$$

Therefore, we have at low temperature:

$$I(T, \epsilon_F - \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F}) = I(T = 0, \epsilon_F) = V\hbar^{-3}\frac{4\pi}{3}(2m\epsilon_F)^{\frac{3}{2}}$$

and taking the derivative with ϵ_F:

$$\left(\frac{\partial I}{\partial \mu}\right)_T (T, \epsilon_F - \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F})(1 + \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F^2}) = V\hbar^{-3}\frac{4\pi}{3}(2m\epsilon_F)^{\frac{3}{2}}\frac{3}{2}\frac{1}{\epsilon_F}$$

which gives

$$\chi(T)(1 + \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F^2}) = \chi(T = 0)$$

or

$$\chi(T) = \chi(T = 0)(1 - \frac{\pi^2}{12}\frac{k_B^2 T^2}{\epsilon_F^2})$$

Large T, looks like Maxwell-Boltzmann

$$N = V\hbar^{-3}\sum_s \int d^3p\, e^{\frac{1}{k_B T}(\mu - \frac{p^2}{2m} + s\mu_0 B)}$$

$$M = V\hbar^{-3}\mu_0 \sum_s \int d^3p\, s e^{\frac{1}{k_BT}(\mu - \frac{p^2}{2m} + s\mu_0 B)}$$

$$\frac{M}{N} = \mu_0 \frac{\sinh(\frac{\mu_0 B}{k_BT})}{\cosh(\frac{\mu_0 B}{k_BT})}$$

For $T \to \infty$ this is zero. That makes sense, all states are equally probable, independent of B, and hence there is no M as a function of B. Therefore $\chi = 0$. The next order term is:

$$\frac{M}{N} = \mu_0 \frac{\mu_0 B}{k_BT}$$

$$\chi = \frac{N\mu_0^2}{k_BT}$$

PROBLEM 5.

The virial expansion is given by $\frac{p}{kT} = \sum_{j=1}^{\infty} B_j(T) \left(\frac{N}{V}\right)^j$ with $B_1(T) = 1$. Find $B_2(T)$ for non-interacting Fermions in a box.

Using

$$\Omega = -2Vk_BT\lambda_T^{-3}f_{\frac{5}{2}}(\lambda)$$

$$N = 2V\lambda_T^{-3}f_{\frac{3}{2}}(\lambda)$$

and the fact that low density corresponds to small λ we get:

$$\Omega \approx -2Vk_BT\lambda_T^{-3}(\lambda - \lambda^2 2^{-\frac{5}{2}})$$

$$N \approx 2V\lambda_T^{-3}(\lambda - \lambda^2 2^{-\frac{3}{2}})$$

The pressure is still simple:

$$p = -\frac{\Omega}{V} \approx 2k_BT\lambda_T^{-3}(\lambda - \lambda^2 2^{-\frac{5}{2}})$$

We need to solve for λ as a function of N. Using the fact that the density $n = \frac{N}{V}$ is small:

$$n \approx 2\lambda_T^{-3}\lambda(1 - \lambda 2^{-\frac{3}{2}})$$

$$n(1 + \lambda 2^{-\frac{3}{2}}) \approx 2\lambda_T^{-3}\lambda$$

$$n(1 + n\lambda_T^3 2^{-\frac{5}{2}}) \approx 2\lambda_T^{-3}\lambda$$

$$\lambda \approx n\frac{1}{2}\lambda_T^3(1 + n\lambda_T^3 2^{-\frac{5}{2}})$$

Inserting in the pressure, and only keeping terms to second order:

$$p \approx 2k_B T\lambda_T^{-3}\lambda(1 - \lambda 2^{-\frac{5}{2}})$$

$$p \approx 2k_B T\lambda_T^{-3} n\frac{1}{2}\lambda_T^3(1 + n\lambda_T^3 2^{-\frac{5}{2}})(1 - n\frac{1}{2}\lambda_T^3 2^{-\frac{5}{2}})$$

$$p \approx k_B Tn(1 + n\lambda_T^3 2^{-\frac{5}{2}} - n\frac{1}{2}\lambda_T^3 2^{-\frac{5}{2}})$$

$$B_2(T) = \lambda_T^3 2^{-\frac{7}{2}}$$

This is positive, increasing the pressure, due to the fact that fermions do not like to be in the same place.

SOLUTION PROBLEM 6.

The chemical potential follows from the equation for N:

$$N = \sum_o f_{FD}(\epsilon_o; T, \mu)$$

which is in this case

$$N = \frac{2V}{(2\pi)^3}\int d^3k f_{FD}(\sqrt{\hbar^2 c^2 k^2 + m^2 c^4}; T, \mu)$$

and there are no convergence problems, like in the notes. This gives

$$N = \frac{2V}{(2\pi)^3}\int d^3k \frac{1}{e^{\beta(\sqrt{\hbar^2 c^2 k^2 + m^2 c^4} - \mu)} + 1}$$

with $\beta = \frac{1}{k_B T}$. At $T = 0$ this reduces to

$$N = \frac{2V}{(2\pi)^3}\int_{k<k_F} d^3k$$

and

$$\epsilon_F = \sqrt{\hbar^2 c^2 k_F^2 + m^2 c^4}$$

The equation for N gives

$$N = \frac{2V}{(2\pi)^3}\frac{4\pi}{3}k_F^3$$

or

$$k_F = \left(3\pi^2 \frac{N}{V}\right)^{\frac{1}{3}}$$

The internal energy at $T = 0$ follows from a similar expression:

$$U(T=0) = \frac{2V}{(2\pi)^3} \int_{k<k_F} d^3k \sqrt{\hbar^2 c^2 k^2 + m^2 c^4}$$

or

$$U(T=0) = \frac{V}{\pi^2} \int_0^{k_F} k^2 dk \sqrt{\hbar^2 c^2 k^2 + m^2 c^4}$$

If the density is very low, we have for all values $\hbar^2 c^2 k^2 \ll m^2 c^4$ and we can write

$$\sqrt{\hbar^2 c^2 k^2 + m^2 c^4} = mc^2 \sqrt{\frac{\hbar^2}{m^2 c^2} k^2 + 1} \approx mc^2 + \frac{\hbar^2 k^2}{2m}$$

which gives

$$U \approx Nmc^2 + \frac{3}{5} N \frac{\hbar^2 k_F^2}{2m}$$

which is the non-relativistic result, with the addition of the rest mass energy. For large densities k_F is very large and we can use $\hbar^2 c^2 k^2 \gg m^2 c^4$ and

$$\sqrt{\hbar^2 c^2 k^2 + m^2 c^4} \approx \hbar c k$$

which leads to the extreme relativistic limit

$$U \approx \frac{V\hbar c}{\pi^2} \int_0^{k_F} k^3 dk = \frac{V\hbar c}{\pi^2} \frac{1}{4} k_F^4 = \frac{3}{4} N\hbar c k_F$$

or

$$U \approx \frac{3}{4} N \epsilon_F$$

where we now have a factor $\frac{3}{4}$ in front.

SOLUTION PROBLEM 7.

The beginning is copied from the previous solution. Since we include the magnetic field in the energy part, we really calculate $\mathcal{G}(T, \mu, V, B) = \Omega - MB$ from the partition function, similar to the notes:

$$\mathcal{G} = -k_B T \sum_o \log(1 + \lambda e^{-\frac{\epsilon_o}{k_B T}})$$

this can be written in the form

$$\mathcal{G} = -k_B T \sum_o \log(1 + \lambda e^{-\frac{1}{k_B T}(\frac{p_z^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))})$$

If we have the free energy, we also have the partition function, of course.

$$\mathcal{G} = -k_B T 2 \frac{eBL^2}{2\pi\hbar c} \sum_{p_z,j} \log(1 + \lambda e^{-\frac{1}{k_B T}(\frac{p_z^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))})$$

$$\mathcal{G} = -k_B T 2 \frac{eBL^2}{2\pi\hbar c} \frac{L}{2\pi\hbar} \int_{-\infty}^{\infty} dp \sum_j \log(1 + \lambda e^{-\frac{1}{k_B T}(\frac{p^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))})$$

or

$$\mathcal{G} = -k_B TV \frac{eB}{2\pi^2\hbar^2 c} \int_{-\infty}^{\infty} dp \sum_j \log(1 + \lambda e^{-\frac{1}{k_B T}(\frac{p^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))})$$

Using $N = -\left(\frac{\partial\mathcal{G}}{\partial\mu}\right)_{T,V,B}$ and $M = -\left(\frac{\partial\mathcal{G}}{\partial B}\right)_{T,\mu,V}$ we find

$$N = V \frac{eB}{2\pi^2\hbar^2 c} \int_{-\infty}^{\infty} dp \sum_j \frac{1}{\lambda^{-1} e^{\frac{1}{k_B T}(\frac{p^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))} + 1}$$

which gives in the limit $\lambda \to 0$ (for high temperature):

$$N \approx \lambda V \frac{eB}{2\pi^2\hbar^2 c} \int_{-\infty}^{\infty} dp \sum_j e^{-\frac{1}{k_B T}(\frac{p^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))}$$

or

$$N \approx \lambda V \frac{eB}{2\pi^2\hbar^2 c} e^{-\frac{e\hbar B}{2mck_B T}} \int_{-\infty}^{\infty} dp e^{-\frac{p^2}{2mk_B T}} \sum_j e^{-j\frac{e\hbar B}{mck_B T}}$$

$$N \approx \lambda V \frac{eB}{2\pi^2\hbar^2 c} e^{-\frac{e\hbar B}{2mck_B T}} \sqrt{2mk_B T} \int_{-\infty}^{\infty} dx e^{-x^2} \frac{1}{1 - e^{-\frac{e\hbar B}{mck_B T}}}$$

Check: If B goes to zero we get

$$N \approx \lambda V \frac{1}{4\pi^2\hbar^3} \left(\sqrt{2mk_B T}\right)^3 \int_{-\infty}^{\infty} dx e^{-x^2}$$

which is equal to the old result.
We can also take the small λ limit in the free energy and get

$$\mathcal{G} = -k_B TV \frac{eB}{2\pi^2\hbar^2 c} \int_{-\infty}^{\infty} dp \sum_j \lambda e^{-\frac{1}{k_B T}(\frac{p^2}{2m} + \frac{e\hbar B}{mc}(j+\frac{1}{2}))}$$

Comparing with the formula for N we see $\mathcal{G} = -Nk_B T$, which should not come as a surprise, since $\mathcal{G} = -pV$ and we are in the limit of an ideal gas!! Therefore:

$$M = k_B T \left(\frac{\partial N}{\partial B} \right)_{T,\mu,V}$$

For small values of B we have

$$N \approx \lambda V \frac{eB}{2\pi^2\hbar^2 c} e^{-\frac{e\hbar B}{2mck_B T}} \sqrt{2\pi mk_B T} \frac{1}{1 - e^{-\frac{e\hbar B}{mck_B T}}}$$

For small values of x we have

$$\frac{e^{-x}}{1 - e^{-2x}} \approx \frac{1-x}{2x - 2x^2 + \frac{4}{3}x^3} = \frac{1}{2x}\frac{1-x}{1-x+\frac{2}{3}x^2} \approx \frac{1}{2x}(1-x)(1+x+\frac{1}{3}x^2) \approx \frac{1}{2x}(1-\frac{1}{6}(2x)^2)$$

and hence

$$N \approx \lambda V \frac{eB}{2\pi^2\hbar^2 c} \sqrt{2\pi mk_B T}\frac{mck_B T}{e\hbar B}(1 - \frac{1}{6}\left(\frac{e\hbar B}{mck_B T}\right)^2)$$

$$N \approx \lambda V \frac{mk_B T}{2\pi^2\hbar^3} \sqrt{2\pi mk_B T}(1 - \frac{1}{6}\left(\frac{e\hbar B}{mck_B T}\right)^2)$$

Hence

$$M = k_B T \lambda V \frac{mk_B T}{2\pi^2\hbar^3} \sqrt{2\pi mk_B T}\frac{B}{3}\left(\frac{e\hbar}{mck_B T}\right)^2$$

$$\chi = k_B T \lambda V \frac{mk_B T}{2\pi^2\hbar^3} \sqrt{2\pi mk_B T}\frac{1}{3}\left(\frac{e\hbar}{mck_B T}\right)^2$$

and using the expression for N at $T = 0$:

$$\chi = k_B T N \frac{1}{3}\left(\frac{e\hbar}{mck_B T}\right)^2 = \frac{N}{3k_B T}\left(\frac{e\hbar}{mc}\right)^2$$

A.6 Solutions for chapter 6.

PROBLEM 4.

A quantum mechanical system is described by a Hamiltonian $H = H_0 + \kappa V$, with $[H_0, V] = 0$. κ is a small constant. The Helmholtz free energy is $F_\kappa(T)$. Calculate the change in Helmholtz free energy, $\Delta F = F_\kappa - F_0$ for this system up to second order in $\frac{\kappa}{k_B T}$.

$$\mathcal{Z} = Tre^{-\beta H} = Tre^{-\beta H_0 - \beta \kappa V} = Tre^{-\beta H_0}e^{-\beta \kappa V}$$

since the matrices commute. The value of κ is small, hence:

$$\mathcal{Z} = Tre^{-\beta H_0}\left(1 - \beta\kappa V + \frac{1}{2}\beta^2\kappa^2 V^2\right)$$

Using the definition of thermodynamic average we have

$$< X > Tre^{-\beta H_0} = TrXe^{-\beta H_0}$$

and hence

$$\mathcal{Z} = Tre^{-\beta H_0}\left(1 - \beta\kappa < V > +\frac{1}{2}\beta^2\kappa^2 < V^2 >\right)$$

Using $\log(\mathcal{Z}) = -\beta F$ we find

$$-\beta F_\kappa = -\beta F_0 + \log(1 - \beta\kappa < V > +\frac{1}{2}\beta^2\kappa^2 < V^2 >)$$

Expanding the log in second order gives

$$-\beta F_\kappa = -\beta F_0 - \beta\kappa < V > +\frac{1}{2}\beta^2\kappa^2 < V^2 > -\frac{1}{2}\beta^2\kappa^2 < V >^2$$

or

$$F_\kappa - F_0 = \kappa < V > -\frac{1}{2}\beta\kappa^2\left(< V^2 > - < V >^2\right)$$

PROBLEM 5.

In a two-dimensional Hilbert space the density operator is given by its matrix elements:

$$\rho = \begin{pmatrix} x & R \\ R^* & 1 - x \end{pmatrix}$$

This form is clearly Hermitian and has trace one. Calculate the entropy as a function of x and R, and find the values of x and R that make the entropy maximal. Note that you still need to check the condition that the matrix is positive! Also, show that it is a maximum!

The entropy is given by:

$$S = -k_B Tr\rho\log(\rho)$$

Use

$$\frac{\partial}{\partial X}Tr\rho\log(\rho) = Tr\left(\frac{\partial\rho}{\partial X}\right)\log(\rho) + Tr\left(\frac{\partial\rho}{\partial X}\right)$$

for any variable X. Using R, R^*, and x for X and setting the result to zero gives:

$$0 = Tr \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \log(\rho) + Tr \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

$$0 = Tr \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \log(\rho) + Tr \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

$$0 = Tr \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \log(\rho) + Tr \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

or

$$(\log(\rho))_{21} = 0$$

$$(\log(\rho))_{12} = 0$$

$$(\log(\rho))_{11} - (\log(\rho))_{22} = 0$$

Therefore

$$\log(\rho) = \begin{pmatrix} y & 0 \\ 0 & y \end{pmatrix}$$

and hence

$$\rho = \begin{pmatrix} z & 0 \\ 0 & z \end{pmatrix}$$

Using the fact that the trace is one we finally get:

$$\rho = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$$

or $R = 0$, $x = \frac{1}{2}$.
The same can be obtained by finding the eigenvalues of ρ.
The second order derivatives are all negative or zero. Use result (9,33) for any of the elements and change to coordinates R, R^*, and x.

PROBLEM 6.

A quantum mechanical system is described by a simple Hamiltonian H, which obeys $H^2 = 1$. Evaluate the partition function for this system. Calculate the internal energy for $T \to 0$ and $T \to \infty$.

$$\mathcal{Z} = Tr \sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n \beta^n H^n$$

Split in odd/even terms:

$$\mathcal{Z} = Tr(\sum_{k=0}^{\infty} \frac{1}{(2k)!} \beta^{2k} - \sum_{k=0}^{\infty} \frac{1}{(2k+1)!} \beta^{2k+1} H)$$

$$\mathcal{Z} = Tr(\cosh(\beta) - H \sinh(\beta))$$

The internal energy is $U = -\frac{\partial}{\partial \beta} \log(\mathcal{Z})$*, which gives:*

$$U = -\frac{Tr(\sinh(\beta) - H \cosh(\beta))}{Tr(\cosh(\beta) - H \sinh(\beta))}$$

In the limit $T \to \infty$ *we find*

$$U = \frac{TrH}{Tr1}$$

which makes sense, it is the average of all energies.
In the other limit $T \to 0$ *we use*

$$U = -\frac{Tr(e^{\beta}(1 - H) + e^{-\beta}(-1 - H))}{Tr(e^{\beta}(1 - H) + e^{-\beta}(1 + H))}$$

$$U = -\frac{e^{\beta}Tr(1 - H) - e^{-\beta}Tr(1 + H)}{e^{\beta}Tr(1 - H) + e^{-\beta}Tr(1 + H)}$$

which gives $U = -1$ unless $TrH = Tr1$, in which case all energy values are $+1$ and only the second terms survive to give $U = 1$.

SOLUTION TO PROBLEM 7.

As stated the problem is easy. Since the sum of the eigenvalues is 1 there is exactly one eigenvalue equal to 1, the rest is zero. The system is in a specific state and the entropy is zero.
The question should have been that the eigenvalues are either 0 or $\frac{1}{N}$ in which case we have exactly N eigenvalues equal to $\frac{1}{N}$ and the entropy is $k_B \log(N)$.
Part C is simple: $\rho = \rho_1 \rho_2 \cdots$ where ρ_1 only acts on the part of the wave function pertaining to particle 1, and gives 1 for everything else.