# STATISTISK MEKANIK 

FYSA21
Hösten 2009


Lars Gislén

Department of Theoretical Physics
University of Lund

# STATISTICAL MECHANICS 

Lars Gislén, Theoretical physics, University of Lund

## 1. Introduction

Classical thermodynamics is based on a few principles derived from: The first and second law. The first law is essentially a law stating the conservation of energy: The change of the internal energy of a body is the result of mechanical work and / or a change of its "heat". The second law can, in one of its formulations, be expressed as that "heat" spontaneously flows from warm to cold. In classical thermodynamics you also introduce the concept of entropy in connection with studying heat engines and the so-called Carnot process. This way of studying thermodynamics is essentially historically-technically and was the result of the need to understand and improve the heat engines that was used in the early mining

 technique, like Newcomen's steam engine.

In classical thermodynamics we use macroscopic quantities like temperature, mass, pressure, volume, density.


The method we will use was developed by the Austrian physicist Ludwig Boltzmann at the end of the 19th century. It derives the thermodynamic laws and macroscopic properties of a system starting from a microscopic description. The advantage of using this approach is that we can, starting from a few, very simple axioms, derive the entire classical thermodynamics. Another advantage is that the concept of entropy that is rather abstract and hard to understand in classical thermodynamics gets a very simple interpretation using statistical mechanics.

We will avoid using the word "heat" in our text as this word in everyday language is connected with several different concepts in thermodynamics like temperature, internal energy, transfer of energy using a temperature difference and also with entropy.

### 1.1 Fundament definitions



If we wait long enough the system will reach thermodynamical equilibrium (TE). Macroscopic parameters then have well defined and constant values. The system has no "memory" of past states.
-Example:


Diathermic means that energy (heat) can pass the wall. After having waited long enough we have TE.

The 0-th law of thermodynamics:
If system $A$ is in thermodynamic equilibrium with $B$ and $C$ is in thermodynamic equilibrium with B this implies that C is in thermodynamic equilibrium with A.


- Example: This can be use to (preliminarily) define temperature by having B above be a "thermometer". If the temperature of A is the same as the temperature of $B$ and the temperature of $C$ is the same as the temperature of $B$ then A and C have the same temperature. We can now determine if two bodies have the same temperature.

What we need is a temperature scale. Imagine a number of reference systems $B_{i} ; i=1,2, \ldots$ We can check which of them is in TE with A. The number $i$ then is a measure of the temperature of $A$.

- Example: In a standard thermometer, $i$ is the length of the liquid capillary of the thermometer.

We can calibrate a thermometer by using special media: ice, boiling water. We can assign the length of the thermometer capillary by 100 for boiling water and 0 for ice and then we divide the interval between them in 100 parts. The gives us the so-called Celcius scale but historically we also have other calibrations, the Fahrenheit scale, the Réaumur scale and the Rankin scale.

Calibrated thermometers of course agree at the calibration points but can deviate in other points. We would like to define temperature in a more
consistent way. One possibility is, as is done in classical thermodynamics, to define temperature as proportional to the pressure of an "ideal" gas wit fixed volume. We will return to this later. We will show that in statistical mechanics, we can define temperature theoretically and that this definition agrees with the ideal temperature definition.

### 1.2 Different kinds of equilibrium

We illustrate this using three pictures.


Thermal equilibrium


Mechanical equlilibrium


Chemical equilibrium

Thus we can have thermal equilibrium, mechanical equilibrium and chemical equilibrium. There are, in fact, several other possibilities; imagine that we have electrically charged particles and / or particles with magnetic moments and external electric and magnetic fields. We must then wait for electric and magnetic equilibrium and so on. We will for the moment ignore such effects on our system. To have thermodynamic equilibrium, we demand that we have (if possible) all these kind of equilibrium at the same time.

### 1.3 Functions of state

At thermodynamic equilibrium it turns out that the properties of the system are simple; they only depend on a few macroscopic parameters; we have what is called a well-defined macroscopic state. We then have certain relations between some of the macroscopic parameters. Let $T$ be temperature, $p$ pressure and $V$ volume.

- Example: $T=T(p, V) \quad p=p(T, V) \quad V=V(p, T)$


## - Example:

For an ideal gas at constant temperature we have experimentally $p V=$ const, Boyle's law. It the temperature is not constant we have (for an ideal gas) $p V=n R T$, the general law of state. Observe that this implies that in classical thermodynamics we can introduce the ideal gas scale by

$$
T=\frac{p V}{n R}
$$

We can measure temperature by measuring the pressure of an ideal gas in a container with constant volume: the gas thermometer.

- Example. We can model a non-ideal gas using the equation of van der Waal:

$$
\left(p+\frac{a}{V^{2}}\right)(V-b)=n R T
$$

where $a$ and $b$ are suitable constants.
Consider an arbitrary (nice) function

$$
G=G(x, y, \ldots)
$$

where $G$ is such that its value is uniquely determined by the arguments $x, y$, We make a small change of $G$ by changing the arguments. We can compute the change in $G$ using a Taylor expansion:

$$
d G=\frac{\partial G}{\partial x} d x+\frac{\partial G}{\partial y} d y=A(x, y) d x+B(x, y) d y
$$

Notice that if $G$ is "nice" enough, in practice always, we have

$$
\frac{\partial^{2} G}{\partial y \partial x}=\frac{\partial^{2} G}{\partial x \partial y}=\frac{\partial A(x, y)}{\partial y}=\frac{\partial B(x, y)}{\partial x}
$$

that is a very stringent condition on the functions $A$ and $B$.

### 1.4 Internal energy $E$

Change the temperature of a system from $T_{1}$ to $T_{2}$. To do this we have to perform work, either using thermal work (cooling or heating the system) or by using mechanical work (for instance by friction) or by adding chemical, electric, magnetic... energy, we neglect the last ones. The sum of all these works change the energy of the system and this change only depends on the start and end states of the system. We can retain the principle of energy conservation if we assume that the added energy is stored as internal energy, $E$, in the system:

$$
\Delta E=Q+W=\text { thermal }+\quad \text { mechanical wort }
$$

This is the first law of thermodynamics.
The internal energy is a function of state of the system. This means that it is, as a function, uniquely determined by a handful of macroscopic parameters.

Notice that what often is called "heat" that is transfer of energy using a temperature difference here is given a special name, thermal work.

- Example: For a gas we have $E=E(p, V, T, N)$. (Besides for an ideal gas it turns out that the internal energy only depends on the temperature.)

BUT: $W$ and $Q$ that is mechanical and thermal work are not functions of state! Explain why?

We rewrite the first law in terms of infinitesimal changes and also put in the specific expression for the mechanical work in case we have a gas:

$$
d E=t Q+t W=t Q-p \cdot d V
$$

The lines over the d:s mark that they are not "proper" differentials, that is do not correspond to changes in functions of state. Can you explain the minus sign in the formula?

The first law is often written:

$$
t Q=d E+p \cdot d V
$$

### 1.5 Heat capacities

We define the heat capacity $C$ by $C=\frac{d Q}{d T}$ that is the ratio of thermal work and temperature change. In most cases the heat capacity will be different if we measure it at constant pressure or at constant volume. We denote the heat capacity at constant volume or pressure by an index $V$ and $p$ respectively on $C$. It turns out that a physically more interesting quantity is the molar heat capacity, the heat capacity per mol. We denote molar heat capacities by a lower case $c$.

Study the following table that shows molar heat capacities for a number of gases:

| Gas | $c_{p}$ | $c_{V}$ | $c_{p}-c_{V}$ | $\gamma=c_{p} / c_{V}$ |
| :---: | :---: | :---: | :---: | :---: |
| He | 20.9 | 12.6 | 8.3 | 1.66 |
| Ar | 20.9 | 12.5 | 8.4 | 1.67 |
| Hg | 20.9 | 12.5 | 8.4 | 1.67 |
| $\mathrm{O}_{2}$ | 29.3 | 20.9 | 8.4 | 1.40 |
| CO | 29.3 | 21.0 | 8.3 | 1.40 |
| $\mathrm{Cl}_{2}$ | 34.1 | 25.1 | 9.0 | 1.36 |
| $\mathrm{SO}_{2}$ | 40.6 | 31.4 | 9.2 | 1.29 |
| $\mathrm{C}_{2} \mathrm{H}_{6}$ | 51.9 | 43.1 | 8.8 | 1.20 |

You can note several interesting facts in the table. The difference between heat capacities at constant pressure and volume is more or less constant. Further, the ration between these heat capacities in the last column is very close to rational numbers $5 / 3 \approx 1.67,7 / 5=1.4,9 / 7 \approx 1.29$.

We will understand and explain these observations later on.
Now study a gas at constant volume.

$$
d Q=d E+p \cdot d V=d E \quad(d V=0)
$$

This implies

$$
C_{V}=\left(\frac{d Q}{d T}\right)_{V}=\left(\frac{d E}{d T}\right)_{V} \Rightarrow c_{V}=\frac{1}{n}\left(\frac{d E}{d T}\right)_{V}
$$

At constant pressure we have

$$
\begin{aligned}
& d Q=d E+p \cdot d V \\
& C_{p}=\left(\frac{d Q}{d T}\right)_{p}=\left(\frac{d E}{d T}\right)_{p}+p \cdot \frac{d V}{d T}
\end{aligned}
$$

For an ideal gas we have $p V=n R T$. If the pressure $p$ is constant this implies

$$
p \frac{d V}{d T}=n R
$$

Thus

$$
C_{p}=\left(\frac{d E}{d T}\right)_{p}+n R \Rightarrow c_{p}=\frac{1}{n}\left(\frac{d E}{d T}\right)_{p}+R
$$

For an ideal gas the internal energy only depends on the temperature which means

$$
\left(\frac{d E}{d T}\right)_{p}=\left(\frac{d E}{d T}\right)_{V}
$$

or

$$
c_{p}=c_{V}+R
$$

where $R$ is the gas constant with value $8.3143 \mathrm{~J} /(\mathrm{mol} \cdot \mathrm{K})$. This agrees very well with the third column in the table above.

Exercise: Explain why $c_{p}$ is larger than $c_{v}$ ?

### 1.6 Adiabatic process, ideal gas

An adiabatic process is defined as a process in which the thermal work is zero (no "heat" is added or subtracted from the system) and if we consider 1 mol of gas we have

$$
d E=-p d V=c_{V} d T
$$

For 1 mol of an ideal gas we also have

$$
p V=R T
$$

If we differentiate this equation we get

$$
d p V+p d V=R d T=-\frac{R}{c_{V}} p d V
$$

or

$$
0=d p V+p d V\left(1+\frac{R}{c_{V}}\right)=d p V+p d V \frac{c_{V}+R}{c_{V}}=d p V+p d V \gamma
$$

or

$$
0=\frac{d p}{p}+\gamma \frac{d V}{V}
$$

We integrate

$$
\text { const }=\ln p+\gamma \ln V=\ln \left(p V^{\gamma}\right)
$$

or

$$
p V^{\gamma}=\text { const }
$$

### 1.7 Some terminology

Certain macroscopic parameters like volume, mass, internal energy have the property that if we make a new system by uniting two systems, these parameters will simply be the sum of the parameters of the original systems. Such parameters are additive and are called extensive. Other parameters like pressure, temperature and density behave differently. Such parameters are called intensive.

## Exercise problems. Chapter 1

1. The gas law can be written in several different ways:

$$
p V=N k_{B} T \quad p V=n R T \quad \rho=\frac{M p}{R T} \quad n=\frac{N}{N_{A}}
$$

eher $N$ is the number of particles, $n$ number of mols, $M$ mol mass, $\rho$ density, and $R$ the gas constant. $N_{\wedge}$ is Avogadro's constant. Show that these formulations are equivalent and express gas constant in more fundamental constants of nature.
2. You compress air in a bicycle pump rapidly to 1 / 10 of the original volyme. What kind of process is this? What is the final temperature and pressure of the air? $\gamma=1.4$. If you instead compress the air very slowly, what is the final pressure? Explain! Practical use?
3. Newcomen's steam engine worked like this:
a) Steam of atmospheric pressure was let into the cylinder of the engine.
b) A small amount of cold water was injected in the cylinder causing the steam to condensate.
c) Steam takes up a volume that is about 1700 times the volume of liquid water. This means that a vacuum was essentially created in the cylinder. The piston was pressed down by the atmospheric pressure. This was the work phase of the engine.
d) The cycle was repeated from a).

Problem: Compute the theoretical efficiency for this process. Hint: How much energy was needed


Fig. 602. Newcomens ángmaskin. to transform water of $100^{\circ} \mathrm{C}$ to steam of $100 \mathrm{C}^{\circ}$ ?
How large is the work done by the atmospheric pressure?

## 2. About probabilities

### 2.1 Introduction

Before we enter statistical mechanics, we will repeat some concepts from probability science.

For instance throwing a die is called an experiment. The result is called an event.

We can enumerate the events with an index $i$. To each event $i$ we connect a (real) number $P_{i} \in[0,1]$ that we call the probability of the event. We can plot the possible events as points in an abstract space.

- Example: Throwing a die

- Example: Throwing two dice


We can also study a complex event where we choose a group of points in the diagram above and say that the complex event occurs if any event occurs that belongs to the group of events. An example would be in the case of throwing two dice that the sum of the two dice is 5 .

We chose the probabilities such that $\sum_{i} P_{i}=1$, normalization.

### 2.2 Classical probability

Choose $P_{i}$ such that $P_{i}=\frac{1}{\Omega}$ where $\Omega$ is the total number of possible events.

- Example: Heads or tails. $\Omega=2, P_{\text {krona }}=P_{\text {klave }}=\frac{1}{2}$
- Example: Throwing one die. $\Omega=6, P_{i}=\frac{1}{6}, \quad i=1 . .6$
- Example: Throwing two dice. $\Omega=36, P_{i}=\frac{1}{36}, \quad i=1.36$

In the real world it is not always true that the different probabilities are equal, a dice can be prepared. But if we don't have much information of a system the assumption of equal probability is reasonable if we want to explore the system.

### 2.3 Statistical probability

Make $N$ experiments. If an event $i$ occurs $n_{i}$ times, we define the statistical probability as the limit

$$
P_{i}=\lim _{N \rightarrow \infty} \frac{n_{i}}{N}
$$

In practice we can of course not make an infinite number of experiments but have to be satisfied by "many". Another problem is that we cannot be sure that the limit exists.

- Example: The value of a share on the stock market.
2.4 Probability postulates

1) $P_{i} \in[0,1], \sum_{i} P_{i}=1$
2) $P(i \vee j \vee k \ldots)=P_{i}+P_{j}+P_{k}+\ldots$ if $i, j, k \ldots$ are mutually exclusive events.
3) $P(i \wedge j \wedge k \ldots)=P_{i} \cdot P_{j} \cdot P_{k} \cdot \ldots$ if $i, j, k \ldots$ are independent events.

Mutually exclusive means that if one of the events occurs none of the other can occur. If you throw a die you can only get one of the events $1,2,3,4,5$ or 6. Independent events means that they cannon influence each other. If you throw two dice, the result of one die does not influence the result of the other die.

- Example: What is the probability of getting at least one six in three throws with a die?
The throws are independent events. The probability of not getting any six is $\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6}$ (Postulate 3). The probability of not getting this result that is getting at least one six is then $1-\frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \approx 0.42$ (Postulate 1)


### 2.5 Permutations

A permutation of $N$ different objects $=$ the number of ways that you can order $N$ different objects in a row. Some reflection tells us the number is

$$
N \cdot(N-1) \cdot(N-2) \cdot(N-3) \cdot \ldots \cdot 1=N!
$$

If $n$ objects are identical we get $\frac{N!}{n!}$ permutations. We have to compensate for the $n!$ permutations that are alike.

If we have several kinds of identical objects with $n_{1}$ of one kind, $n_{2}$ of another kind and so on, the number of permutations is

$$
\frac{N!}{n_{1}!\cdot n_{2}!\cdot \cdots}=\frac{N!}{\prod_{i} n_{i}!} \text { where } N=\sum_{i} n_{i}
$$

Note that $0!=1$ by definition.

### 2.6 Distributions

Example:


Geiger counter

Count the number of pulses in the Geiger counter during a certain time, say 10 seconds. Repeat many times. Denote the number of measured pulses in experiment $i$ by $x_{i}$. Plot the result in a diagram that may look like this:

$n_{i}$ is the frequency that is the number of times we measured $x_{i}$ pulses.
We now define the average (mean) number of pulses or the expectation value of the number of pulses as:

$$
\langle x\rangle=\frac{\sum_{i} n_{i} x_{i}}{N}=\sum_{i} \frac{n_{i}}{N} x_{i}=\sum_{i} P_{i} x_{i}
$$

where the last equality follows of the number of measurements, $N$, is large.
Note that the expectation value in general is NOT the same as the most probable value, the number of pulses in the maximum bar in the diagram.

We will often have a continuous distribution where $\rho(x) d x$ is the probability of finding $x$ in the interval $[x, x+d x]$. For this case it is natural to define

$$
\langle x\rangle=\int x \rho(x) d x
$$

$\rho(x)$ is called distribution in probability density. Evidently we have a normalization condition $\int \rho(x) d x=1$ corresponding to $\sum_{i} P_{i}=1$ in the discrete case.

Example. Quantum mechanics where $\rho(x)=|\Psi(x)|^{2}$ with $\Psi(x)=$ the wave function.

An important statistical quantity is the standard deviation or scattering, $\sigma$, defined by the variance

$$
\sigma^{2}=\sum_{i} P_{i}(x-\langle x\rangle)^{2}
$$

Note! $x$ can represent any physically interesting variable like position, speed, energy ...

## Exercise problems. Chapter 2

1. In how many ways can you permute 4 girls and 5 boys?
2. What is the probability of getting either 7 or 6 by throwing two dice?
3. Show that we can write $\sigma^{2}(x)=\left\langle x^{2}\right\rangle-\langle x\rangle^{2}$. This is very useful as it simplifies the number of steps in the computation of the standard deviation.
4. A neutron that moves in a piece of uranium- 235 can hit an uranium nucleus and start a chain reaction. Assume that the probability for a neutron to hit a nucleus when it moves a distance $d x$ is $p \cdot d x$.
a) What is the probability the neutron does not hit a nucleus when it moves a distance $d x$ ?
b) What is the probability that the neutron moves $N$ steps $d x$ without hitting a nucleus and then hits a nucleus in the next step?
c) Assume that $N$ steps correspond to a total distance $x$. Use the relation $\lim _{N \rightarrow \infty}\left(1+\frac{z}{N}\right)^{N}=e^{z}$ to rewrite the expression you got in b ) in a simpler way.
d) Compute the average distance a neutron travels before it hits a nucleus.

## 3. Statistical mechanics

We will now construct the statistical mechanics, which derives the classical thermodynamics from a few simple postulates.

We assume that we study isolated systems with a large number, $N$, of identical, weakly interacting particles in a volume $V$. The particles have a total energy $E$. The assumption of weak interactions means that the total energy is the sum of one-particle energies.

### 3.1 Macrostates and microstates



A given macrostate can be realised by an enormous number of microstates, even worse, if we consider one mol of gas it changes microstate $10^{2 \pi}$ times each second! We can look at the air in this lecture hall that looks the same in spite of an enormous number of collisions each moment between the molecules that then change their velocities and by that the microstate. In principle we could describe the microstates if we knew the three-dimensional positions and velocities of every molecule. This is of course impossible in practice and we will see that we can manage quite well by using statistical methods to describe the microstates.

We denote the number of accessible microstates by $\Omega$. This number is of course very large.

How can we motivate this postulate? Simply by saying that it is the simplest assumption. We don't know anything about these probabilities thus we assume they are the same. If we would assume they were different we would immediately have to face a more complicated problem: what then are they? Besides, it turns out that the thermodynamical laws that we get from this postulate agree very well with experiment.

Our statistical methods have the following steps:

1) Solve the one-particle problem. This means solving a quantum mechanical problem that gives us the energy levels $\varepsilon_{i}, i=1,2,3 \ldots$ and states of the particle. We assume that we have solved this problem.
2) What distributions $\left\{n_{i}\right\}$ of particles can we have in the energy levels given the constraints $N=\sum_{i} n_{i}$ and $E=\sum_{i} n_{i} \varepsilon_{i}$ ?
3) How many microstates $t\left\{n_{i}\right\}$ are there in each distribution?
4) Determine the average distribution.

### 3.2 Entropy, S

We define the entropy of a system by

$$
S=k_{B} \ln \Omega
$$

where $k_{B}=1.38 \cdot 10^{23} \mathrm{~J} / \mathrm{K}$, Boltzmann's constant, determines the scale and dimension of the entropy. As you can see, the entropy is simply a measure of the number of accessible microstates for the system. We have once above used the word accessible. All microstates are not in general accessible to the system. For instance we demand that a gas should be confined in a certain volume which means that microstates where a molecule is outside the volume are forbidden. Further we want that the system shall have a certain internal energy that puts a constraint on the possible energies of the particles; the sum of their energies must have a fixed value.

Why not having the entropy be just $\Omega$ ? There are several evident advantages by using a logarithm.

1) $\Omega$ is an ENORMOUS number! By using the logarithm of large numbers we get numbers that are more manageable.
2) Consider two systems with $\Omega_{1}$ and $\Omega_{2}$ microstates respectively. The combined system has evidently $\Omega=\Omega_{1} \cdot \Omega_{2}$ microstates. But with our definition the entropy of the combined system is

$$
S=k_{B} \ln \Omega=k_{B} \ln \left(\Omega_{1} \cdot \Omega_{2}\right)=k_{B} \ln \Omega_{1}+k_{B} \ln \Omega_{2}=S_{1}+S_{2}
$$

This is a very nice property; entropy is an additive or extensive quantity. As we will soon see the entropy has several other nice and useful properties.

- Example: Place $N$ particles in a volume $V$. Divide $V$ in a number of small "compartments" each with the fixed volume $\Delta V$. Place the particles randomly in the compartments. There are $\frac{V}{\Delta V}$ compartments, thus the number of possible ways of placing the particles $=$ the number of accessible microstates is

$$
\Omega=\left(\frac{V}{\Delta V}\right)^{V}
$$

The entropy is

$$
S=k_{B} \ln \Omega=k_{B} \ln \left(\frac{V}{\Delta V}\right)^{N}=N k_{B} \ln \frac{V}{\Delta V}=N k_{B}(\ln V-\ln \Delta V)
$$

It looks like we have a problem here. The value of the entropy depends on the volume of the compartments. Just now we can avoid this problem by saying that we normally are only interested in changes of the entropy and in such cases the offending term disappears. Later on we will see that using quantum mechanics we actually get a definite size of the compartments. But this was a problem for Boltzmann who lived at a time when quantum mechanics hadn't been invented

- Example: Detailed computation for a "Mickey Mouse system" with 4 particles. We use our statistical method outlined above.

Suppose that we have equidistant energy levels (again quantum mechanics!) $0, \varepsilon, 2 \varepsilon, 3 \varepsilon \ldots$ Suppose that the total energy of the system is $4 \varepsilon$. We also suppose that the particles are distinguishable, that is can be thought of as having labels $A, B, C .$. such that we can tell them apart. We can then place the particles in the levels in five different ways (distributions) that all give the same given total energy:


We now count the number of permutations for each distribution. Particles in one level are not permuted. We get the following result:

$$
\frac{4!}{3!1!}=4 \quad \frac{4!}{1!1!2!}=12 \quad \frac{4!}{1!2!1!}=12 \quad \frac{4!}{2!2!}=6 \quad \frac{4!}{4!}=1
$$

(In general you have $\frac{N!}{\prod_{i} n_{i}!}$ permutations.)
In total we have $\Omega=4+12+6+12+1=35$ microstates.
In this case the entropy is $S=k_{B} \ln 35 \approx 3.56 \cdot k_{B}$
The probabilities of the respective distributions are $4 / 35,12 / 35,6 / 35,12 / 35$, $1 / 35$. Observe that certain, rather few distributions dominate the scene. The average number of particles in level 0 is

$$
\left\langle n_{0}\right\rangle=3 \cdot \frac{4}{35}+2 \cdot \frac{12}{35}+1 \cdot \frac{6}{35}+2 \cdot \frac{12}{35}+0 \cdot \frac{1}{35} \approx 1.71
$$

For the other levels we easily calculate the corresponding averages

$$
\left\langle n_{1}\right\rangle \approx 1.14\left\langle n_{2}\right\rangle \approx 0.69 \quad\left\langle n_{3}\right\rangle \approx 0.34\left\langle n_{4}\right\rangle \approx 0.11
$$

It is interesting to plot the result


The curve has some similarity with an exponentially decreasing function. We will return to this fact and see that our supposition is true.

We divide the system in two subsystems $A$ and $B$ each with 2 particles and energies $E_{A}=4 \varepsilon$ and $E_{B}=0$. You easily show that $\Omega_{A}=5$ and $\Omega_{s}=1$, that is the total number of microstates is $\Omega=\Omega_{A} \cdot \Omega_{8}=5$. If we bring the systems together and allow them to reach thermodynamical equilibrium we will have the situation we studied before: as we approach thermodynamical equilibrium the entropy increases.

### 3.3 The second law. Use of the entropy

Postulate: (Second law) At thermodynamical equilibrium the entropy of an isolated system takes its maximum value (given the constraints on the system like internal energy, volume, number of particles and so on).

The postulate means that at thermodynamical equilibrium (TE) the system exploits all accessible microstates with the same probability.

This implies that at TE, $S$ has a fixed value determined by the parameters $E, V$, $N$. This implies that $S$ is a function of state.
Finally this implies that for small changed in the parameters we have

$$
d S=\frac{\partial S}{\partial E} d E+\frac{\partial S}{\partial V} d V+\frac{\partial S}{\partial N} d N
$$

### 3.3.1 Definition of temperature

Consider a system
diathermal wall


Only the internal energy $E_{1}$ and $E_{2}$ and of course the entropy can vary, all other parameters are fixed.
At TE the entropy has a maximum which means that $d S=0$ when we vary the internal energy:

$$
0=d S=\frac{\partial S_{1}}{\partial E_{1}} d E_{1}+\frac{\partial S_{2}}{\partial E_{2}} d E_{2}
$$

We use that the energy is conserved $d E_{2}=-d E_{1}$ and get

$$
0=\left(\frac{\partial S_{1}}{\partial E_{1}}-\frac{\partial S_{2}}{\partial E_{2}}\right) d E_{1} \Rightarrow \frac{\partial S_{1}}{\partial E_{1}}=\frac{\partial S_{2}}{\partial E_{2}}
$$

In this case we evidently have thermal equilibrium and the temperature must be the same in the two subsystems. The partial derivative has dimension inverse temperature. This leads us to define temperature by

$$
\frac{\partial S}{\partial E}=\frac{1}{T}
$$

We have earlier seen that we can (at least in principle, just count the number of accessible microstates) compute the entropy of a system. Given the entropy $S$ we can then compute the temperature $T$. It turns out that the temperature that we get in this way is identical with the one in classical thermodynamics that you get from the ideal gas thermometer.

Now assume that the to subsystems are NOT in thermodynamical equilibrium but that $T_{1}>T_{2}$. When we let the two systems exchange energy, the entropy will increase towards a maximum, with other words $d S>0$. Then we have

$$
d S=\frac{\partial S_{1}}{\partial E_{1}} d E_{1}+\frac{\partial S_{2}}{\partial E_{2}} d E_{2}=\left(\frac{\partial S_{1}}{\partial E_{1}}-\frac{\partial S_{2}}{\partial E_{2}}\right) d E_{1}=\left(\frac{1}{T_{1}}-\frac{1}{T_{2}}\right) d E_{1}>0
$$

The factor in front of $d E_{1}$ is less than zero, thus $d E_{1}<0$ which we interpret that system 1 is losing energy and system 2 is gaining energy, energy flows
spontaneously from a warmer system to a colder one. This is one of the alternative formulations of the second law and agrees with physical common sense.

### 3.3.2. Definition of pressure

We start again but now with a movable, diathermal wall between the subsystems.

| $\boldsymbol{E}_{1}$ | $\boldsymbol{E}_{2}$ |
| :---: | :---: |
| $V_{1}$ | $V_{2}$ |
| $\mathrm{~N}_{1}$ | $\mathrm{~N}_{2}$ |
| $\mathrm{~S}_{1}$ | $\mathrm{~s}_{2}$ |

Here $E_{1}, E_{2}, V_{1}$, and $V_{2}$ can vary.
At TE the entropy has a maximum:
$0=d S=\frac{\partial S_{1}}{\partial V_{1}} d V_{1}+\frac{\partial S_{2}}{\partial V_{2}} d V_{2}+\frac{\partial S_{1}}{\partial E_{1}} d E_{1}+\frac{\partial S_{2}}{\partial E_{2}} d E_{2}$
We know that for thermal equilibrium the sum of the first two terms is zero, thus the sum of the last two terms must be zero. The total volume is constant or $d V_{2}=-d V_{1}$ which implies

$$
\left(\frac{\partial S_{1}}{\partial V_{1}}-\frac{\partial S_{2}}{\partial V_{2}}\right) d V_{1}=0 \Rightarrow \frac{\partial S_{1}}{\partial V_{1}}=\frac{\partial S_{2}}{\partial V_{2}}
$$

Now we have also mechanical equilibrium and again by dimensional reasoning it seem to be a good idea to define pressure, $p$, by

$$
\frac{\partial S}{\partial V}=\frac{p}{T}
$$

because if we use that we have thermal equilibrium and then the same temperature in the two subsystems we get

$$
p_{1}=p_{2}
$$

This is intuitively correct, in this kind of equilibrium both temperature and pressure are equal in the subsystems.

In the same way as before we can now show that if we have thermal equilibrium but not mechanical equilibrium the subsystem with the higher pressure will expand at the expense of the volume of the other subsystem.

## *3.3.3. Chemical potential $\mu$

Finally we study a permeable wall that allows particles to pass, is diathermal and movable.


As a concrete example you can think of having a gas in 1 and a liquid in 2 and that the wall is the interface between liquid and gas.

At TE the entropy is maximal:
$0=d S=\frac{\partial S_{1}}{\partial N_{1}} d N_{1}+\frac{\partial S_{2}}{\partial N_{2}} d N_{2}+\frac{\partial S_{1}}{\partial V_{1}} d V_{1}+\frac{\partial S_{2}}{\partial V_{2}} d V_{2}+\frac{\partial S_{1}}{\partial E_{1}} d E_{1}+\frac{\partial S_{2}}{\partial E_{2}} d E_{2}$
As we have thermal and mechanical equilibrium the sum of the first four terms is zero. In the same way as before, using that the total number of particles is conserved, we have

$$
\frac{\partial S_{1}}{\partial N_{1}}=\frac{\partial S_{2}}{\partial N_{2}}
$$

We then have chemical equilibrium and define the chemical potential, $\mu$, by

$$
\frac{\partial S}{\partial N}=-\frac{\mu}{T}
$$

The sign is chosen such that we later get consistent results. The chemical potential has dimension energy. Our procedures above are actually very general, the partial derivative of the entropy with respect to an extensive variable gives us an intensive parameter divided by temperature.

In summary we give an alternative formulation of the second law: In an isolated system the change of the entropy is always larger than or equal to zero. (Either the system is at TE and the entropy has attained its maximum value or it is on its way to equilibrium and the entropy is increasing.)

### 3.3.4 Rewards

Earlier we had

$$
d S=\frac{\partial S}{\partial E} d E+\frac{\partial S}{\partial V} d V+\frac{\partial S}{\partial N} d N \equiv \frac{1}{T} d E+\frac{p}{T} d V-\frac{\mu}{T} d N
$$

We rewrite

$$
d E=T d S-p d V+\mu d N
$$

This looks familiar! We rediscover the first law; actually not very exciting as we have used that the internal energy is conserved. The interesting thing is that the first term looks different. The first term evidently corresponds to thermal work, the second is mechanical work, the thirst "chemical" work, if we add particles to the system the carry some kind of chemical energy $\mu$ into the system. (Just now we are not interested in such processes but this term will be important when we study chemical processes or equilibrium problems for a liquid-gas interface.)

Identifying the first term with thermal energy we have $d Q=T d S$ or $d S=\frac{d Q}{T}$. This is the original, historical definition of entropy. If we integrate between two (macro)states $A$ and $B$ we have

$$
\Delta S_{A B}=\int_{A}^{P} \frac{d Q}{T}
$$

Further $d Q=C \cdot m \cdot d T \Rightarrow \Delta S_{A B}=\int_{T_{1}}^{T_{1}} \frac{C \cdot m \cdot d T}{T}$
We have found a simple macroscopic way of computing entropy changes when we heat a body!
$\bullet$ Example: What is $\Delta S$ when 1 kg water melts?
The temperature is constant $=273 \mathrm{~K}$. Thus $\Delta S=C_{m} / 273=334 / 273 \mathrm{~kJ} / \mathrm{K}$ This shows how to compute the entropy change of a phase change. Note that the entropy increases, the molecules in the water can access many more microstates than they have in ice.

A further reward!
We return to our volume that we divided in small compartments with fixed volume

$$
S=N k_{B}(\ln V-\ln \Delta V)
$$

We can now compute the pressure

$$
\frac{p}{T} \equiv \frac{\partial S}{\partial V}=N k_{B} \frac{1}{V} \quad \Rightarrow \quad p V=N k_{B} T
$$

We have in a very simple way derived the gas law for an ideal! This also shows that our temperature definition and the ideal gas thermometer temperature are equivalent. Note that we only have used our two simple postulates and out definitions of pressure and temperature. Statistical mechanics is an extremely powerful tool!

### 3.4. We dig deeper and find the Boltzmann factor

We will now make a more advanced calculation on a system that is more realistic than our earlier Mickey Mouse system. We assume that we have $N$ particles where $N$ is really large, of order $10^{32}$. We assume that we have a number of energy levels $\varepsilon_{i} ; i=1,2, \ldots$ not necessarily equidistant. Assume that we in one of the possible distributions have $n_{i} ; i=1,2, \ldots$ particles in the respective levels. Here we exploit an important fact. At TE the entropy is maximal. It then turns out (se below) that if the number of particles is large, only ONE distribution will dominate in probability over all the others. We saw this tendency already in the Mickey Mouse system. The number of microstates in this distribution will, if the number of particles is large, be almost the same as the total number of microstates. This means that we only have to study one distribution and in this distribution arrange the particles such that the entropy is maximised given the constraints that the total energy and number of particles is constant. The number of microstates in this distribution is

$$
\Omega^{*}=\frac{N!}{\prod_{i} n_{i}!}
$$

and the entropy

$$
S=k_{B} \ln \frac{N!}{\prod_{i} n_{i}!}
$$

When $n$ is large we can use Stirling's approximation: $\ln n!\approx n \ln n-n$. This approximation is very good even for rather reasonable $n$.

Example: Study a large number $(N)$ of tosses of a penny. For each toss we can have either of two events thus we have $2^{*}$ "microstates" in total, each with the same probability. In a distribution with $n$ heads and $m=N-n$ tails we have $t=\frac{N!}{n!(N-n)!}$ microstates. This expression has evidently a maximum when $n=N / 2$ or

$$
t_{\max }=\frac{N!}{\left(\frac{N}{2}\right)!\left(\frac{N}{2}\right)!}
$$

| $N$ | $t_{\max }$ | $\ln 2^{N}$ | Relative <br> error <br> $(\%)$ |
| ---: | ---: | ---: | ---: |
| 2 | 0.69 | 1.39 | 50.00 |
| 4 | 1.79 | 2.77 | 35.38 |
| 6 | 3.00 | 4.16 | 27.97 |
| 8 | 4.25 | 5.55 | 23.38 |
| 10 | 5.53 | 6.93 | 20.23 |
| 20 | 12.13 | 13.86 | 12.52 |
| 30 | 18.86 | 20.79 | 9.30 |
| 40 | 25.65 | 27.73 | 7.49 |
| 50 | 32.47 | 34.66 | 6.31 |
| 100 | 66.78 | 69.31 | 3.65 |
| 200 | 135.75 | 138.63 | 2.07 |
| 500 | 343.24 | 346.57 | 0.96 |
| 1000 | 689.47 | 693.15 | 0.53 |
| 10000 | 6926.64 | 6931.47 | 0.07 |

If we use Sterling's approximation we have

$$
\begin{gathered}
\ln t_{\max }=N \ln N-N-2\left(\frac{N}{2} \ln \frac{N}{2}-\frac{N}{2}\right)=N \ln N-N \ln \frac{N}{2}= \\
N \ln N-N \ln N+N \ln 2=N \ln 2=\ln 2^{N}
\end{gathered}
$$

When $N$ is large, the number of microstates in the most probable distribution approaches the total number of microstates.

We return to our main problem. For the entropy we then have

$$
\begin{aligned}
& S=k_{B}\left(\ln N!-\ln \prod_{i} n_{i}!\right)=k_{B}\left(\ln N!-\sum_{i} \ln n_{i}!\right)= \\
& =k_{B}\left(N \ln N-N-\sum_{i}\left(n_{i} \ln n_{i}-n_{i}\right)\right)
\end{aligned}
$$

We now want to maximise $S$ by varying the different $n_{i}: s$. This is a bit problematic as the $n_{i}:$ s are not independent. We have two constraints on $n_{i}$ :

$$
N=\sum_{i} n_{i} \text { and } E=\sum_{i} n_{i} \varepsilon_{i},
$$

the number of particles and the internal energy are given and constant Mathematically we can handle such a situation by inserting the constraints via Lagrange multiplicators and instead maximise the function

$$
f=\frac{S}{k_{B}}+\alpha\left(N-\sum_{i} n_{i}\right)+\beta\left(E-\sum_{i} n_{i} \varepsilon_{i}\right)
$$

Through this trick we can treat the $n_{i}:$ s as if they were independent and get

$$
0=\frac{\partial f}{\partial n_{i}}=-\ln n_{i}-1+1-\alpha-\beta \varepsilon_{i}
$$

or

$$
n_{i}=e^{-\alpha} e^{-\beta \varepsilon_{i}}
$$

We determine the first factor that contains $\alpha$ by the condition

$$
\sum_{i} n_{i}=\sum_{i} e^{-\alpha} e^{-\beta \varepsilon_{i}}=e^{-\alpha} \sum_{i} e^{-\beta \varepsilon_{i}}=N \quad \Rightarrow e^{-\alpha}=\frac{N}{\sum e^{-\beta \varepsilon_{i}}} \equiv \frac{N}{Z}
$$

where we have defined $Z=\sum_{i} e^{-\beta \varepsilon_{i}}$, the partition function that will soon prove very useful.

The occupation number of each level is then given by $n_{i}=\frac{N}{Z} e^{-\beta \varepsilon_{i}}$
$e^{-\beta \varepsilon_{i}}$ is the so-called Boltzmann factor. We now see what we guessed in the case of the Mickey Mouse system, that the number of particles in the levels decrease exponentially as the energy increases.

### 3.5 What is $\beta$ ?

We plug in our result in the entropy

$$
\begin{aligned}
& S=k_{B}\left(N \ln N-N-\sum_{i}\left(n_{i} \ln n_{i}-n_{i}\right)\right)= \\
& k_{B} N \ln N-k_{B} \sum_{i}\left(n_{i} \ln n_{i}\right)=k_{B} N \ln N-k_{B} \frac{N}{Z} \sum_{i}\left(e^{-\beta \varepsilon_{i}} \ln \left(\frac{N}{Z} e^{-\beta \varepsilon_{i}}\right)\right)= \\
& k_{B} N \ln N-k_{B} \frac{N}{Z} \sum_{i}\left(e^{-\beta \varepsilon_{i}}\left(\ln N-\ln Z-\beta \varepsilon_{i}\right)\right)= \\
& k_{B} N \ln N-k_{B} \frac{N}{Z}\left(Z \ln N-Z \ln Z-\beta \sum_{i} \varepsilon_{i} e^{-\beta \varepsilon_{i}}\right)= \\
& k_{B} N \ln Z+k_{B} \beta \sum_{i} \varepsilon_{i} \frac{N}{Z} e^{-\beta \varepsilon_{i}}=k_{B} N \ln Z+k_{B} \beta \sum_{i} \varepsilon_{i} n_{i}=k_{B} N \ln Z+k_{B} \beta E
\end{aligned}
$$

To sum up we have: $S=k_{B} N \ln Z+k_{B} \beta E$
Finally use the temperature definition $\frac{\partial S}{\partial E}=\frac{1}{T}=k_{B} \beta$ or $\beta=\frac{1}{k_{B} T}$.
We collect our results expressed in more familiar quantities, the positions of the energy levels, the temperature and the internal energy, all of them measurable or computable quantities:
$n_{i}=\frac{N}{Z} e^{-\beta \varepsilon_{i}}$ where $Z=\sum_{i} e^{-\beta \varepsilon_{i}}$ and $\beta=\frac{1}{k_{B} T}$
$S=k_{B} N \ln Z+\frac{E}{T}$

Remember for the future that the probability to find a particle (or a system) in level $i$ is $\frac{n_{i}}{N}=\frac{1}{\mathrm{Z}} e^{-\beta \varepsilon_{i}} \propto e^{-\beta \varepsilon_{i}}$, the Boltzmann probability.

### 3.6 Energy reservoir and subsystem

Consider a subsystem in contact and in thermal equlibrium with an energy reservoir with temperature $T$. We assume that the subsystem is small and that the energy reservoir is large. The energy reservoir and the subsystem are isolated ffromthe environment and have the total (and constant) energy $E$. We want the probability $p_{i}$ that the subsystem delsystemet is in the state with $E_{i}$. The energy reservoir the has the energy $E-E_{i}$. The entropy of the energy reservoir is a function of its energy and is $S_{i}\left(E-E_{i}\right)$. The number of microstates of the energy reservoir then is $e^{s_{i}\left(E-E_{i}\right) / k_{B}}$. For the subsystem and the energy reservoir we then have in total $1 \cdot e^{s_{i}\left(E-E_{i}\right) / k_{B}}$ microstates. The probability that the subsystem is in a state with energy $E_{i}$ is proportional to the number of mirostates in the combines system, thus we have

$$
p_{i}=A e^{s_{i}\left(E-E_{i}\right) / k_{B}}
$$

As the energy reservoir is large we have $E_{i} \ll E$ and can Taylor expand:

$$
S_{i}\left(E-E_{i}\right)=S_{i}(E)-E_{i} \frac{\partial S_{i}}{\partial E}=S_{i}(E)-\frac{E_{i}}{T}
$$

that implies

$$
p_{i}=A e^{\left(s_{i}(E)-\frac{E_{i}}{T}\right) / k_{B}}=C e^{-E_{i} / k_{B} T}
$$

As the sum of the probabilities of the subsystem has to be we have

$$
1=\sum_{k} p_{k}=C \sum_{k} e^{-E_{k} / k_{B} T}=C Z
$$

giving us $p_{i}=\frac{e^{-E_{i} \beta}}{Z}$,
a result that should look familiar.

### 3.7 The useful $Z$, the partition function

We have

$$
E=\sum_{i} \varepsilon_{i} n_{i}=\frac{N}{Z} \sum_{i} \varepsilon_{i} e^{-\beta \varepsilon_{i}}=-N \frac{1}{Z} \frac{\partial}{\partial \beta} \sum_{i} e^{-\beta \varepsilon_{i}}=-N \frac{1}{Z} \frac{\partial Z}{\partial \beta}=-N \frac{\partial \ln Z}{\partial \beta}
$$

This is important. Once we know the energy levels of a system, a problem we solve in quantum mechanics, we know the partition function. We then can compute the internal energy in a simple way without having to use the maybe unfamiliar entropy. (If we want the entropy it can simply be computed from $S=k_{B} N \ln Z+\frac{E}{T}$.) Once we know the internal energy we can compute other measurable quantities like heat capacities.

In many systems the energy levels are degenerate that is there are several different states having the same energy. We must then in the partition function count the term that corresponds to a degenerate level as many times as the degeneracy (multiplicity). If we assume that the energy levels $\varepsilon_{,} i=1,2$, 3... have multiplicities $g$ we get a partition sum

$$
Z=\sum_{i} g_{i} e^{-\beta \varepsilon_{i}}
$$

More about this later.

### 3.8 Energy fluctuations

We now study a large number ( $N$ ) of identical systems in contact with a heat bath with temperature $T$. As an example you can think of the atoms or molecules in a gas. Such a set of systems is called the canonical ensemble. Each system can be in one of its energy levels $E_{\text {.. }}$ The probability that this happens is according to what we have seen above

$$
p_{r}=\frac{e^{-E_{t} \beta}}{\sum_{i} e^{-E_{i} \beta}}=\frac{e^{-E_{r} \beta}}{Z}, \quad Z=\sum_{i} e^{-E_{i} \beta}
$$

The average energy of the systems then is

$$
\langle E\rangle=\sum_{i} p_{i} E_{i}=\frac{1}{Z} \sum_{i} E_{i} e^{-E_{i} \beta} .
$$

The average of the square of the energy is

$$
\left\langle E^{2}\right\rangle=\sum_{i} p_{i} E_{i}^{2}=\frac{1}{Z} \sum_{i} E_{i}^{2} e^{-E_{i} \beta}
$$

We are interested in the fluctuation in energy or the variance of the energy that is

$$
\Delta E^{2}=\left\langle E^{2}\right\rangle-\langle E\rangle^{2}=\frac{1}{Z} \sum_{i} E_{i}^{2} e^{-E_{i} \beta}-\frac{1}{Z^{2}}\left(\sum_{i} E_{i} e^{-E_{i} \beta}\right)^{2} \text { (See the end of chapter }
$$

2) 

We have

$$
\begin{aligned}
& \Delta E^{2}=\frac{1}{Z} \sum_{i} E_{i}^{2} e^{-E_{i} \beta}-\frac{1}{Z^{2}}\left(\sum_{i} E_{i} e^{-E_{i} \beta}\right)^{2}= \\
& \frac{1}{Z} \frac{\partial^{2} Z}{\partial \beta^{2}}-\frac{1}{Z^{2}}\left(\frac{\partial Z}{\partial \beta}\right)^{2}=\frac{\partial}{\partial \beta}\left(\frac{1}{Z} \frac{\partial Z}{\partial \beta}\right)= \\
& -\frac{\partial\langle E\rangle}{\partial \beta}=-\frac{\partial\langle E\rangle}{\partial T} \frac{d T}{\partial \beta}
\end{aligned}
$$

Now $\beta=\frac{1}{k_{B} T} \Rightarrow T=\frac{1}{k_{B} \beta} \Rightarrow \frac{d T}{d \beta}=-\frac{1}{k_{B} \beta^{2}}=-k_{B} T^{2}$
This gives the total variance

$$
\Delta E_{\text {total }}^{2}=N \Delta E^{2}=\frac{\partial N\langle E\rangle}{\partial T} k_{B} T^{2}=\frac{\partial E_{\text {total }}}{\partial T} k_{B} T^{2}
$$

Further $\frac{\partial E_{\text {total }}}{\partial T}=n_{\text {mol }} c_{V}=\frac{N}{N_{A}} c_{V}$, implies $\Delta E_{\text {total }}^{2}=\frac{N k_{B} T^{2}}{N_{A}} \Rightarrow \Delta E_{\text {total }} \propto \sqrt{N}$ The relative fluctuation $\frac{\Delta E_{\text {total }}}{E_{\text {total }}} \propto \frac{1}{\sqrt{N}}$ can obviously be neglected if $N$ is of order $10^{33}$.

### 3.9 What is entropy intuitively?

I think that you sometimes have heard entropy be described as a measure of disorder. This is only half of the truth though. If we add that entropy also is freedom we get a rather good intuitive description of the entropy concept. When the entropy is maximised it means that the system tries to gain access to all accessible microstates. This is the freedom. Each microstate is then occupied with the same probability. This is the disorder. Also be careful with the condition in the second law: The entropy increases (or is maximum) in a closed (isolated) system but it can decrease locally. Living creatures is an example of regions with very low local entropy and when we arrange bricks in very ordered patterns to build a house we create a very low entropy locally. This is possible because we have an easily accessible source of low entropy nearby, the Sun. If we include the Sun in our system we will have a good approximation of a closed system and the total entropy in this, larger system is increasing. Living creatures with low entropy do not violate the laws of physics or need some supernatural interaction! Finally, in our theory here we have focused on systems in thermodynamical equilibrium. Living creatures are very far from being in thermodynamical equilibrium, which is precisely one of the properties that make them living. In modern advanced thermodynamics you study systems that are not in thermodynamical equilibrium.

### 3.10.1 The Boltzmann factor, Mount Everest, and the use of fridges

Consider a flat earth with an atmosphere above.


Assume that the atmosphere is isothermal that is the temperature is the same everywhere. The probability of finding an air molecule at height $x$ then is

$$
P(x) \propto e^{-\beta \varepsilon(x)}
$$

where $\varepsilon(x)$ is the energy of the molecule. This energy is the sum of the kinetic and potential energy. The kinetic energy is on average the same everywhere as the temperature is the same, independent of the height. The potential energy is $M g x$ where $M$ is the mass of the molecule. This gives

$$
P(x) \propto e^{-\beta \varepsilon_{k}} e^{-\beta \varepsilon_{p}(x)} \propto e^{-\beta M g x}=e^{-\frac{M g x}{k_{B} T}}
$$

Now, the density of the air is evidently proportional to the probability of finding a molecule, and the pressure in turn is proportional to the density. This implies

$$
p(x)=p(0) e^{-\frac{M g x}{k_{B} T}}
$$

where we have normalised the pressure with $p(0)$, the ground pressure. This is the well-known barometric formula used by for instance aviators. Putting in numerical values we get

$$
p(x)=p(0) e^{\frac{x}{8000[m]}}
$$

where 8000 m is the so-called scale height.
Another application of the Boltzmann factor is connected with why we have fridges and freezers. Assume that we have some kind of foodstuff that has a probability $P_{\text {rum }}$ to become stale during let us say a day in room temperature. To get stale means some kind of chemical change, in most cases caused by bacteria. Chemical changes typically deal with energy changes of order $E \approx 1$ eV . The probability of a change at room temperature then is proportional to the Boltzmann probability with this energy in the exponent:

$$
P_{r u m}=C \cdot e^{-\frac{E}{k_{B} T_{n u m}}} \text { where } C \text { is some constant. }
$$

The probability in a fridge with temperature $T_{k y l}$ then is

$$
P_{k y l}=C \cdot e^{-\frac{E}{k_{B} T_{k y l}}}
$$

and in a freezer

$$
P_{f r y s}=C \cdot e^{-\frac{E}{k_{B} T_{f y s}}} .
$$

We then have

$$
P_{k y l} / P_{r u m}=e^{-\frac{E}{k_{B}}\left(\frac{1}{T_{k l}} \frac{1}{T_{n u m}}\right)}
$$

If we use numerical values, say $T_{\text {rum }}=295 \mathrm{~K}$ and $T_{\text {kyl }}=280 \mathrm{~K}$ we get $P_{k y l} / P_{\text {rum }}=0.12$ which means that the foodstuff will remain fresh about 10 times longer than in room temperature. If we use $T_{\text {frys }}=255$, freezer temperature, we instead get $P_{\text {fyys }} / P_{\text {rum }}=0.0014$, which means that the foodstuff will remain fresh about 700 times longer than at room temperature, i.e. for months! The very rapid change is the result of the Boltzmann factor being exponential.

## *3.10.2. The equipartition theorem. Derivation

Suppose we have a system described by its (generalised) coordinates $q_{i} ; i=1,2, \ldots N$ and (generalised) momenta $p_{i} ; i=1,2, \ldots N$. The energy of the system is a function of these variables.

$$
E=E\left(q_{i}, p_{i}\right)
$$

The probability that the coordinates of the system are in the interval [ $\left.q_{i}, q_{i}+d q_{i}\right]$ and the momenta in $\left[p_{i}, p_{i}+d p_{i}\right]$ then is
$C \cdot d q_{1} d q_{2} \ldots d q_{N} \cdot d p_{1} d p_{2} \ldots d p_{N} e^{-\beta E\left(q_{i}, p_{i}\right)}$
We have the normalisation condition

$$
1=\iiint \ldots \int C \cdot d q_{1} d q_{2} \ldots d q_{N} \cdot d p_{1} d p_{2} \ldots d p_{N} e^{-\beta E\left(q_{i}, p_{i}\right)}
$$

that determines the constant

$$
C=\frac{1}{\iiint \ldots \int d q_{1} d q_{2} \ldots d q_{N} \cdot d p_{1} d p_{2} \ldots d p_{N} e^{-\beta E\left(q_{i}, p_{i}\right)}}
$$

The average value (the expectation value) of the energy then is

$$
\langle E\rangle=\frac{\iiint \ldots \int d q_{1} d q_{2} \ldots d q_{N} \cdot d p_{1} d p_{2} \ldots d p_{N} E\left(q_{i}, p_{i}\right) e^{-\beta E\left(q_{i}, p_{i}\right)}}{\iiint \ldots \int d q_{1} d q_{2} \ldots d q_{N} \cdot d p_{1} d p_{2} \ldots d p_{N} e^{-\beta E\left(q_{i}, p_{i}\right)}}
$$

This looks quite nasty but we will use the same trick as we used for the internal energy and the partition function. We can rewrite the monster integral to something more palatable

$$
\langle E\rangle=-\frac{\partial}{\partial \beta} \ln \iiint \ldots \int d q_{1} d q_{2} \ldots d q_{N} \cdot d p_{1} d p_{2} \ldots d p_{N} e^{-\beta E\left(q_{i}, p_{i}\right)}
$$

We now assume that the energy is a quadratic function of the coordinates and the momenta. This is very often true:

$$
E\left(q_{i}, p_{i}\right)=a_{1} q_{1}^{2}+a_{2} q_{2}^{2}+\ldots b_{1} p_{1}^{2}+b_{2} p_{2}^{2}+\ldots
$$

which means

$$
e^{-\beta E}=e^{-\beta a_{1} q_{1}^{2}} \cdot e^{-\beta a_{2} q_{2}^{2}} \cdot \ldots \cdot e^{-\beta b_{1} p_{1}^{2}} \cdot e^{-\beta b_{2} p_{2}^{2}} \cdot \ldots
$$

We can now rewrite the integral more simply as

$$
\langle E\rangle=-\frac{\partial}{\partial \beta} \ln \left(\prod_{i} \int d q_{i} e^{-\beta a_{i} q_{i}^{2}} \cdot \prod_{i} \int d q_{i} e^{-\beta b_{i} p_{i}^{2}}\right)
$$

or

$$
\langle E\rangle=-\frac{\partial}{\partial \beta}\left(\sum_{i}\left(\ln \int d q_{i} e^{-\beta q_{i} q_{i}^{2}}+\ln \int d q_{i} e^{-\beta b_{i} p_{i}^{2}}\right)\right)
$$

All the terms in the sums have the same structure and we only consider one of them, say the first one

$$
\int d q_{i} e^{-\beta_{i} q_{i}^{2}} \underset{t=q_{1} \sqrt{\beta a_{1}}}{=}\left(\beta a_{1}\right)^{-1 / 2} \int d t e^{-t^{2}}=\left(\beta a_{1}\right)^{-1 / 2} D_{1}
$$

The integration is over all allowable values of this coordinate and the remaining integral is just some number that we call $D_{1}$. All the other terms in the sum give similar contributions. Thus we have

$$
\begin{aligned}
& \langle E\rangle=-\frac{\partial}{\partial \beta}\left(\ln \left(D_{1} a_{1}^{-1 / 2} \beta^{-1 / 2}\right)+\ln \left(D_{1} a_{1}^{-1 / 2} \beta^{-1 / 2}\right)+\ldots\right)= \\
& -\frac{\partial}{\partial \beta}\left(\ln D_{1}-\frac{1}{2} \ln a_{1}-\frac{1}{2} \ln \beta+\ldots\right)=\frac{1}{2 \beta}+\frac{1}{2 \beta}+\frac{1}{2 \beta}+\ldots= \\
& \frac{1}{2} k_{B} T+\frac{1}{2} k_{B} T+\frac{1}{2} k_{B} T+\ldots
\end{aligned}
$$

We get the extremely simple result:

Each quadratic term in the expression of the total energy of a particle system implies a contribution to the internal energy by $\frac{1}{2} k_{B} T$.

This is the equipartition theorem.

### 3.10.3. The equipartition theorem. Applications.

Consider an ideal, monoatomic gas of $N$ particles. Monoatomic means that the gas particles only can have translational movement; ideal means that there is no interaction between the particles. This is a fairly good model of a noble gas at normal pressure and temperature. The total energy of a particle then is

$$
\frac{1}{2} M v_{x}^{2}+\frac{1}{2} M v_{y}^{2}+\frac{1}{2} M v_{z}^{2}
$$

We have 3 quadratic terms for each atom. Thus the internal energy is

$$
E=N \cdot 3 \cdot \frac{1}{2} k_{B} T=\frac{3}{2} N k_{B} T=\frac{3}{2} n R T
$$

The molar heat capacity at constant volume then is

$$
c_{V}=\frac{1}{n} \frac{d E}{d T}=\frac{3}{2} R \approx 12.5 \mathrm{~J} /(\mathrm{mol} \cdot \mathrm{~K})
$$

We have explained the heat capacities of the noble gases that we studied in the table on page 4 !

Now consider a solid. We can model its atoms as a system where the atoms are connected by springs with spring constants $k$ in a three dimensional grid:


The energy of an atom is $\frac{1}{2} M v_{x}^{2}+\frac{1}{2} M v_{y}^{2}+\frac{1}{2} M v_{z}^{2}+\frac{1}{2} k x^{2}+\frac{1}{2} k y^{2}+\frac{1}{2} k z^{2}$
We have 6 quadratic terms per atom and the internal energy is

$$
E=N \cdot 6 \cdot \frac{1}{2} k_{B} T=3 N k_{B} T=3 n R T
$$

And the molar heat capacity

$$
c_{V}=\frac{1}{n} \frac{d E}{d T}=3 R \approx 25
$$

This is called Dulong-Petit's law and it works very well for most solids, see the diagram below. There are two evident exceptions, graphite that has a heat capacity that is precisely $1 / 3$ of what it should have and diamond that also have a value that is too low. We will later be able to explain this exception further on.


Encouraged by these results we try to apply our theory on diatomic gases like oxygen, nitrogen and hydrogen. A diatomic molecule has more possible ways of moving than just a translation in three dimensions. It can vibrate along the connection line between the atoms and it can rotate around two axes perpendicular to this line. We can easily write down the total energy for such a molecule

$$
\frac{1}{2} M v_{x}^{2}+\frac{1}{2} M v_{y}^{2}+\frac{1}{2} M v_{z}^{2}+\frac{1}{2} \mu v^{2}+\frac{1}{2} k x^{2}+\frac{1}{2} I_{1} \omega_{1}^{2}+\frac{1}{2} I_{2} \omega_{2}^{2}
$$

The first three terms correspond to translational energy, the two following correspond to the vibration energy and the two last ones correspond to the rotational energies. In total we have 7 quadratic terms that gives an internal energy

$$
E=N \cdot 7 \cdot \frac{1}{2} k_{B} T=\frac{7}{2} N k_{B} T=\frac{7}{2} n R T
$$

that implies a molar heat capacity at constant volume

$$
c_{V}=\frac{1}{n} \frac{d E}{d T}=\frac{7}{2} R \approx 29
$$

Unfortunately this does not agree at all with values that you get from experiment. These give a value close to $\frac{5}{2} R \approx 21$. The physics of the $19^{\text {tit }}$ century could not explain this evident failure of classical thermodynamics. As we will se we will need quantum mechanics to solve the problem.

Also later, when the electron was discovered, there were problems. A very good model of a metal is that you have a gas of free electrons that can move in the lattice of the solid. For the solid itself we have as before

$$
c_{V, \text { lattice }}=3 R
$$

Foe the electron gas we expect the result from a monoatomic gas

$$
c_{V, e l}=\frac{3}{2} R
$$

The total heat capacity is $c_{V, e l}=\frac{9}{2} R$, a metal should have a heat capacity that is $50 \%$ larger than for a non-metal. But experimentally you find that the molar heat capacities for metals and non-metals are essentially the same. Why?

Finally we will point on a problem that also is connected with heat capacities and entropy. Earlier we saw that we had

$$
\Delta S_{A B}=\int_{A}^{T} \frac{d Q}{T}=\int_{T_{1}}^{T} \frac{C \cdot m \cdot d T}{T}=C \cdot m \cdot \ln \frac{T_{2}}{T_{1}}
$$

We can se that we have a problem when $T_{1}=0$, the entropy change gets singular! One way of solving this would be if the heat capacity $C$ goes to zero suitably fast when the temperature goes to zero. Why would this happen? It turns out that this also can be explained by quantum mechanics.

We want to make this very clear: Simple an uncontroversial experimental measurements of heat capacities show that the classical (non-quantum mechanical) thermodynamics is WRONG! This was a great problem at he beginning of the $20^{\text {m }}$ century. We will see that we need to use quantum mechanics to get results that agree with experiment. Besides quantum mechanics will turn out to describe the microcosmos in a new and exciting way.

## Exercise problems. Chapter 3

1. We return to the Mickey Mouse system. Compute $\Omega$ for total internal energies $E=0, \varepsilon, 2 \varepsilon, 3 \varepsilon, 5 \varepsilon, 6 \varepsilon$ for the 4 particles. The compute $s=S / k_{B}$ (renormalised entropy) and plot $s(E)$. Sketch a curve through the points and estimate relatively the temperature for different internal energies.
Qualitatively sketch the relation $E(T)$. Hint: There are respectively 1, 1, 2, 3, 5, 6 , and 9 different distributions.
2. Compute $\Delta S$ when 1 kg water of $100^{\circ} \mathrm{C}$ is transformed from liquid to steam. Then compute the change in entropy when you heat 1 kg water from 0 ${ }^{\circ} \mathrm{C}$ to $100^{\circ} \mathrm{C}$. Comments?
3.1 kg vatten of $0{ }^{\circ} \mathrm{C}$ is put in contact with a large heat source that can be assumed to have a constant temperature of $100^{\circ} \mathrm{C}$. What is the entropy change in the system water + heat source when the water has reached its final temperature?
3. We want to determine the extremum of the function $f(x, y)=x^{2}+y^{2}$ given the constraint $x+y=1$. Do this in two ways,
a) By eliminating for instance $y$ from the first function using the constraint. The result is a function of only one variable that can easily be handled.
b) By adding the constraint using a Lagragian multiplier and then put derivatives to zero.
Show that both methods give the same result.
