

Computational Molecular Physics – Lecture Notes

Petra Imhof

19th May 2014

Contents

1	Introduction	2
1.1	Examples	2
1.1.1	Random Walker	2
1.1.2	1-dimensional harmonic oscillator	2
1.1.3	Lennard-Jones fluid	2
1.1.4	Classical water	2
1.1.5	Ising model	3
1.1.6	Diffusion	3
2	Statistical Mechanics/Statistical Thermodynamics	5
2.1	Microstates and Macrostates	5
2.2	Ensembles	7
2.2.1	The Microcanonical Ensemble (N,V,E)	7
2.2.2	The most probable distribution in the microcanonical ensemble	9
2.2.3	Equilibrium	11
2.2.4	Entropy and Temperature	11
2.2.5	The Canonical Ensemble (N,V,T)	13
2.2.6	The most probable distribution in the canonical ensemble	15
2.2.7	Isobaric-Isothermal Ensemble (N,p,T)	18
2.3	Thermodynamic functions derived from the partition function	19
2.3.1	Internal energy, pressure and heat capacity	19
2.3.2	Entropy	21
2.3.3	Helmholtz Free energy	21
2.3.4	Chemical equilibrium	22
2.4	Ensemble Averages	22

1 Introduction

Computational Molecular Physics or Molecular Simulations are “somewhere” between theory and experiment. Simulations are often understood as the “third pillar” of science, between theory and experiment.

Simulations are used to obtain numerical solutions to theories that are analytically not solvable or too tedious to be solved “manually”. Such problems are e.g. many-body systems with many interacting particles. Thereby theories can be tested.

Moreover simulations allow to mimick experiments, Such computer experiments are useful to explain real experiments because of the additional or complementary information contained in the simulation, or to predict/replace experiments that are difficult or even impossible to perform in real. Think of experiments which have to be conducted a te.g. very high pressure or temperature or other extreme or unrealistic conditions. One can even simulate the behaviour of yet unmade molecules.

Just like real experiments, computer experiments, simulations, can help to test and to develop theories.

This course deals with the questions how to perform simulation by introducing some of the numerical methods and how to reate the simulation to theory and/or experiment.

1.1 Examples

There are a number of classical example systems that will reoccur in this course to show various different concepts.

1.1.1 Random Walker

(to be completed...)

1.1.2 1-dimensional harmonic oscillator

(to be completed...)

1.1.3 Lennard-Jones fluid

A Lennard-Jones potential is a means to introduce interaction between particles. It is a good model for weakly interacting particles such as noble gases (e.g. Argon).

The ineration is described by a pair potential as a function of the interatomic distances r

$$U^{LJ}(r) = 4\varepsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (1.1.1)$$

σ is the optimal (lowest energy) distance and ε is the well depth at the optimal distance.

1.1.4 Classical water

Water can be modelled classically by a Lennard-Jones potential plus electrostatic interactions

$$U = U^{LJ} + U^{coulomb}$$

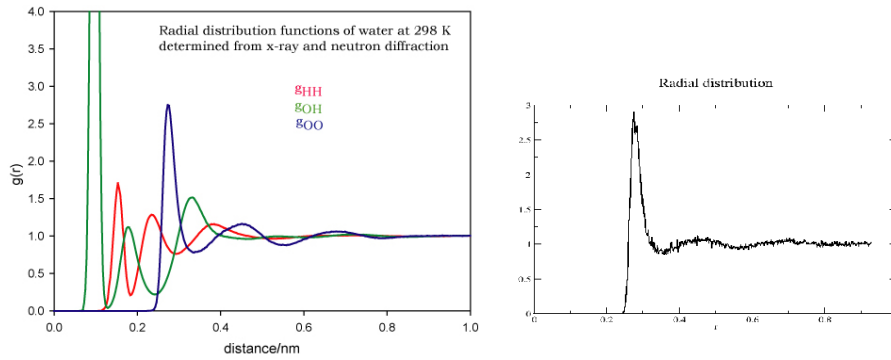


Figure 1: Radial distribution function of oxygen atoms in water. a) experiment (blue) b) simulation

One way to relate the molecular dynamics or Monte Carlo simulation of water is via the radial distribution function $g(r)$. This function describes the distribution of e.g. oxygen-oxygen distances in bulk water, cf. 1. From this distribution the water density can be obtained.

Integrating over the first solvation shell (up to a distance of 3.5\AA) gives the average number of water molecules (strictly speaking oxygen atoms) within that distance.

$$n = 4\pi\rho \int_0^{3.5} g_{OO}(r) r^2 dr = 4$$

1.1.5 Ising model

The Ising model describes the interaction of N spins on a lattice, such as



The energy is given by

$$E_r = - \sum_{i=1}^N H\mu_{s_i} - J \sum_{i,j}^{n.n.} s_i s_j$$

with $s_{i,j} = \pm 1$.

1.1.6 Diffusion

Diffusion can be described by Fick's law

$$j = -D\nabla c$$

where the flux j is proportional to the gradient of concentration ∇c and the diffusion coefficient D . To conserve the total amount of (labelled) material we introduce

$$\frac{\partial c(r,t)}{\partial t} + \nabla j(r,t) = 0$$

combining the two

$$\frac{\partial c(r,t)}{\partial t} - D\nabla^2 c(r,t) = 0 \quad (1.1.2)$$

By introducing the boundary condition $c(r, 0) = \delta(r)$ we get

$$c(r, t) = \frac{1}{(4\pi Dt)^{d/2}} \exp\left(-\frac{r^2}{4Dt}\right)$$

We need only the second moment $\langle r^2(t) \rangle = \int d\mathbf{r} c(r, t) r^2$ using $\int d\mathbf{r} r c(r, t) = 1$
 Multiplying eq 1.1.3 by r^2 and integrate

$$\frac{\partial}{\partial t} \int d\mathbf{r} \cdot r^2 c(r, t) = D \int d\mathbf{r} \cdot r^2 \nabla^2 c(r, t)$$

solve by integration by parts

$$\begin{aligned} \frac{\partial}{\partial t} \int d\mathbf{r} \cdot r^2 c(r, t) &= D \int d\mathbf{r} \cdot r^2 \nabla^2 c(r, t) \\ \frac{\partial}{\partial t} \langle r^2(t) \rangle &= D \int d\mathbf{r} \nabla \cdot (r^2 \nabla c(r, t)) - D \int d\mathbf{r} \nabla r^2 \cdot \nabla c(r, t) \\ &= D \int d\mathbf{S} (r^2 \nabla c(r, t)) - 2D \int d\mathbf{r} \mathbf{r} \cdot \nabla c(r, t) \\ &= 0 - 2D \int d\mathbf{r} (\nabla \cdot \mathbf{r} c(r, t)) + 2D \int d\mathbf{r} (\nabla \cdot \mathbf{r}) c(r, t) \\ &= 0 + 2dD \int d\mathbf{r} c(r, t) \\ &= 2dD \end{aligned} \tag{1.1.3}$$

Equation 1.1.3 relates the macroscopic diffusion coefficient D to the mean-squared distance over which particles have moved in time. This distance can be measured in computer simulations where we can monitor for every particle the position and hence the distance travelled in time t

$$\langle \Delta r(t)^2 \rangle = \frac{1}{N} \sum_{i=1}^N \Delta \mathbf{r}_i(t)^2 \tag{1.1.4}$$

2 Statistical Mechanics/Statistical Thermodynamics

The relation between the very detailed information we can obtain in simulations and the “real world” which allows for a comparison with experiment is given by statistical mechanics. Computing the individual properties we are now interested in the properties of systems.

Table 1: Molecular properties and properties of systems. Note that this is not a one-to-one relation

Mechanical properties of molecules n with atoms i	Thermodynamic properties of systems
positions / coordinates x_i, y_i, z_i	Temperature T
velocities / momenta $p_{x_i}, p_{y_i}, p_{z_i}$	pressure p
individual masses m_i	total mass m
potential energy U_n	Internal Energy U
kinetic energy E_{kin_n}	Entropy S
	Enthalpy H
	Free Energy A, G

For a comprehensive presentation of statistical mechanics the reader is referred to the appropriate text books [?, ?]. The following brief introduction is only meant to outline the terms, ideas and concepts necessary for the following chapters.

2.1 Microstates and Macrostates

Let us start the definition of a **microstate** by an illustration. If you toss a coin it will show head or tail (and in very rare cases it will stand on the edge but those cases will be neglected here). Then head or tail are two possible states for the coin. These are microstates. Another example is a particle with spin, say an electron. That particle can have two spin states, spin up or spin down. Without an external field these two states are both equally probable. A third example is a die that can show numbers 1 . . . 6. For a fair die all numbers are equally probable. All of those examples are about one particle and we further note that in all those examples the possible outcomes, the **microstate have equal a priori probability**. Let us now consider systems that contain many particles.

Two coins: The possible outcomes are head or tail for each of the two coins leading to following combinations.

Coin 1	Coin 2	“Value”
H	H	2
H	T	1
T	H	1
T	T	0

If we assign a value to each of the two cases, say 1 for head and 0 for tail, the result can be assigned a total value.

Four electrons with spin: The spin of an individual electron is either $s = +\frac{1}{2}$ or $s_z = -\frac{1}{2}$. In a magnetic field B we can assign an energy value for each possible realisation r by $E_r = 2\mu_B B \sum_{\nu} s_{z,\nu}$ (μ_B : Bohr's magneton). Setting $\uparrow = +\frac{1}{2}$ and $\downarrow = -\frac{1}{2}$ we have the following microstates

Realisations						Value
(↑↑↑↑)						$E_r = 4\mu_B B$
(↑↑↑↓)	(↑↑↓↑)	(↑↓↑↑)	(↓↑↑↑)			$E_r = 2\mu_B B$
(↑↑↓↓)	(↑↓↓↑)	(↓↑↑↓)	(↓↓↑↑)	(↓↑↓↑)	(↑↓↓↑)	$E_r = 0$
(↓↓↓↑)	(↓↓↑↓)	(↓↑↓↓)	(↑↓↓↓)			$E_r = -2\mu_B B$
(↓↓↓↓)						$E_r = -4\mu_B B$

Already with these simple examples we notice a couple of things:

1. The total value depends on the values of the individual outcomes.
2. Different combinations of individual values can lead to the same total value.
3. The numbers of individual combinations realising a total value differ.

The total value of the system here defines its macrostate. More general a **macrostate** of a system is defined by its (macromolecular) parameters, e.g. an energy value, a conformational fold (α -helix or β -sheet) of a protein, being solid or fluid, bound or unbound, before or after a reaction etc.

As we can see from above, different macrostates can be realised by different numbers of microstates which determines the probability of the macrostates.

Two dice: For one die the probability for each individual number 1,2...6 is $\frac{1}{6}$. All numbers have equal *a priori* probability and all possible counts 1–6 have equal probability. Throwing now the two dice and adding the numbers to give a count, the individual numbers per die still have the same probability. However, the probabilities of the counts 2, 3, 4,...12 show a different distribution:

	Realisations																										
	1+1	2+1	3+1	4+1	5+1	6+1	5+2	6+2	4+3	5+3	6+3	3+4	4+4	5+4	6+4	2+5	3+5	4+5	5+5	6+5	1+6	2+6	3+6	4+6	5+6	6+6	
Number of Realisations	1	2	3	4	5	6	5	4	3	2	1	6	5	4	3	2	1	2	3	4	5	6	5	4	3	2	1
Probability	$\frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$
Count	2	3	4	5	6	7	8	9	10	11	12	7	8	9	10	11	12	8	9	10	11	12	11	10	9	8	7

The most probable outcome is "7" which has six realisation possibilities and hence a probability of $\frac{6}{36}$. This is the macrostate with the highest number of microstates.

Let us summarise:

The probability of a microstate r is given by

$$P_r = \frac{1}{\sum_r}$$

and the probability of a macrostate (with value E) depends on the number of realisations representing it

$$\{P_r \text{ with } E_r=E\} = (P_1, P_2 \dots)$$

The microstate, i.e. the states (position, velocity, spin, ...) of each particle in a system is known in a simulation, hence the macrostate of the total system determined by those microstates is also known.

We remember a definition of probability as the relative frequency of an event i according to

$$P_i = \lim_{N \rightarrow \infty} \frac{N_i}{N}$$

This shows that we need many experiments or many equivalent systems to obtain reasonable probabilities. Consider M many equivalent systems, then the probability to be in microstate r is given by M_r , the number of systems in state r divided by the total number of systems.

$$P_r = \lim_{M \rightarrow \infty} \frac{M_r}{M} \approx \frac{M_r}{M}$$

for M large. Those M systems are called a **statistical ensemble**.

2.2 Ensembles

In this section we will look at different ensembles, distributions therein and arrive at relations to thermodynamics properties.

2.2.1 The Microcanonical Ensemble (N,V,E)

Imagine an isolated system with a fixed number of particles N , fixed volume V and fixed total energy E . The Hamilton function $\hat{H}(q, p, x)$ with $x = N, V$ and q :positions; p :momenta, defines the energy of microstates $E_r = E_r(x)$ either direct (classical) or as eigenvalues of the Hamiltonian (quantum mechanical)

$$\begin{aligned} \mathcal{H}(q, p, x) &= E_r(x) \\ \hat{H}(x) |r\rangle &= E(x) |r\rangle \end{aligned} \tag{2.2.1}$$

Example: 1dim Harmonic Oscillator For the one-dimensional harmonic oscillator the classical Hamilton function is

$$\mathcal{H}(q, p) = \underbrace{\frac{p^2}{2m}}_{\text{kinetic}} + \underbrace{\frac{1}{2}kq^2}_{\text{potential}}$$

with m : mass; k : spring constant.

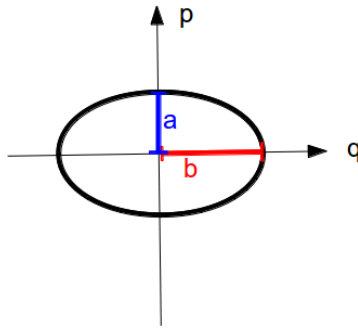


Figure 2: Elliptic orbit of an one-dimensional harmonic oscillator.

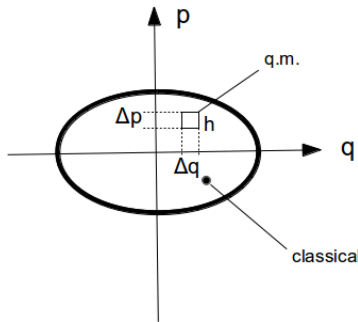


Figure 3: Space space spanned by positions q and momenta p . Classically points in phase space represent microstates, uqnatum mechanically these are small cells of size h^ν .

The total energy

$$E_{total} = E_{kin} + E_{pot}$$

does not depend on time. Therefore the total energy $E_{total} = \mathcal{H}(q, p)$ is a **constant of motion**.

Figure 2 shows all combinations of position and momenta the harmonic oscillator can reach lie on an ellipse. The positions q and momenta p span the **phase space**.

The ellipse's principal axes a, b (see Fig. 2) are given by

a: zero potential energy $E = \frac{p^2}{2m}$ hence $p = \sqrt{2mE}$ and

b: zero kinetic energy $E = \frac{1}{2}kq^2$ hence $q = \sqrt{\frac{2E}{k}}$.

That is, for a given system (given mass and spring constant) all points on the ellipse have the same total energy. Each point in the ellipse represents one state of motion with total energy E . Each point is a microstate.

Note that according to the uncertainty principle $\Delta p \Delta q \geq \frac{\hbar}{2}$ position and momenta cannot be exactly defined at the same time. To account for that, we imagine phase space to be divided into small cells of size h^ν where ν is the dimension. Each cell represents a microstate. In the classical picture each point in phase space is a microstate (see Fig.3).

We can define an energy hypersurface with area $\sigma(E) = \int_{E=\mathcal{H}(q,p)} dqdp$. For the one-dimensiona harmonic oscillator this area is $\sigma = \pi ab$. The size of the hypersurface can be used as a measure for the number of microstates.

Let us now be a bit more tolerant and allow states with energy $E + \delta E$. Then all allowed states are on an energy shell, cf. Fig. 4.

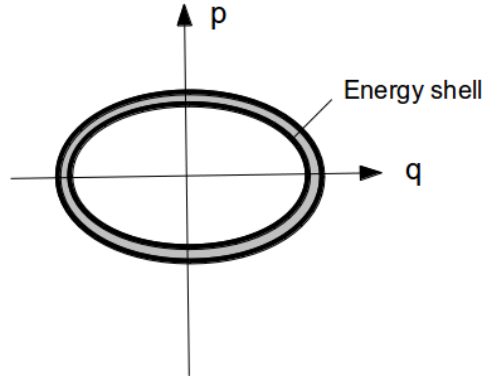


Figure 4: Energy shell with $E + \delta E$

If we sum over all points that are compliant with the condition $E_r = E + \delta E$ we obtain the **microcanonical partition function**

$$\sum_{r: E \leq E_r(N, V) \leq E + \delta E} 1 = \Omega(N, V, E) \quad (2.2.2)$$

The probabilities for the microstates are

$$P_r(N, V, E) = \begin{cases} \frac{1}{\Omega(N, V, E)} & \text{for } E \leq E_r(N, V) \leq E + \delta E = \text{const.} \\ 0 & \text{otherwise} \end{cases} \quad (2.2.3)$$

In a continuous way one can likewise define the probability to find a system in a phase space element

$$\rho(N, V, E) = \frac{1}{\Omega(N, V, E)} \delta[E - \mathcal{H}(q, p)] \quad (2.2.4)$$

where the δ -function ensures that only points on the energy surface contribute.

The statistical ensemble of systems with fixed $N, V, E = \text{const.}$ is called a **microcanonical ensemble**.

2.2.2 The most probable distribution in the microcanonical ensemble

Consider an ensemble of M systems, all with the same with $N, V, E = \text{const.}$ Each of the M systems is at time t in a microstate (q_ν, p_ν) . Microstates can be different, but they are all on the same energy surface. Now partition the surface in elements such that element i has area $\Delta\sigma_i$ and contains m_i systems (cf, Fig. 5). All systems are distributed over the energy surface such that $M = \sum_i m_i$. Then m_i represents the statistical weight of the system in the ensemble and $\frac{m_i}{M}$ is the probability of system to be in σ_i .

There are many possible distributions of the M systems over the surface elements. For example 5 systems can be distributed such that $m_1 = 2; m_2 = 2; m_3 = 1; m_4 = 0$ or $m_1 = 1; m_2 = 1; m_3 = 1; m_4 = 2$, What is now the most probable distribution?

The probability for a certain distribution (set of $\{m_i\}$) is given by

$$W \{m_i\} = \frac{M!}{\prod_i m_i!} \quad (2.2.5)$$

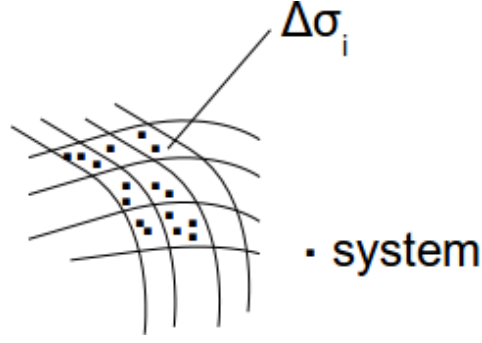


Figure 5: Energy surface partitioned into elements of area $\Delta\sigma_i$ with n_i systems (dots)

There are $M!$ possibilities to relabel the M systems, however those $m_i!$ relabelled within the same cell do not contribute to the total number of possibilities. Hence the total number of relabelling possibilities is reduced (divided) by all the combinations (the product) of all relabellings within the cells.

In order to find the most probable distribution $\{m_i\}^*$ we have to determine the maximum $\max\{W_{tot}\{m_i\}\}$. Since we are dealing with a smooth function, we will instead search for the maximum of $\ln(W_{tot}\{m_i\})$ which will have the same $\{m_i\}$. The logarithm enables us to make use of Strling's formula such that we don't have to deal with the factorials

$$\begin{aligned}\ln W_{tot} &= \ln M! - \sum_i \ln m_i! & (2.2.6) \\ &= M \ln M - M - \sum_i (m_i \ln m_i - m_i)\end{aligned}$$

To consider the boundary condition $M = \sum_i m_i = const$ we use a Lagrangian multiplier α and maximise the Lagrangian \mathcal{L} , i.e. we take the derivative with respect to m_i :

$$\begin{aligned}\mathcal{L} &= -\ln W_{tot} + \alpha \left(\sum_i m_i - M \right) \\ \frac{\partial \mathcal{L}}{\partial m_i} &= -\frac{\partial}{\partial m_i} \left[\ln W_{tot} + \alpha \left(\sum_i m_i - M \right) \right] \\ 0 &= -\frac{\partial}{\partial m_i} \ln W_{tot} + \alpha \\ 0 &= -\frac{\partial}{\partial m_i} \left[M \ln M - M - \sum_i (m_i \ln m_i - m_i) \right] + \alpha \\ 0 &= \frac{\partial}{\partial m_i} \left[\sum_i (m_i \ln m_i - m_i) \right] + \alpha \\ 0 &= \ln m_i + \frac{m_i}{m_i} - 1 + \alpha \\ 0 &= \ln m_i + \alpha\end{aligned}$$

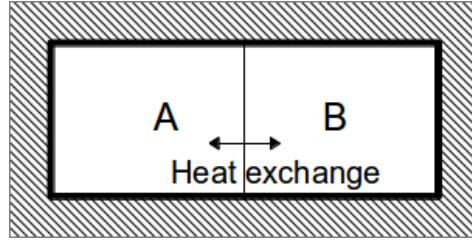


Figure 6: An isolated system with constant N, V, E , subdivided into two subsystems A and B.

Arriving at

$$m_i = \exp(-\alpha) \quad (2.2.7)$$

Equation 2.2.7 must hold for all m_i . The most probable distribution is a constant distribution with all cells having the same number of systems. For the microcanonical ensemble, we obtain the same probability we have already seen

$$P_i = \frac{m_i}{M_i} = \begin{cases} \text{const} & \text{for } \mathcal{H} = E \\ 0 & \text{otherwise} \end{cases} \quad (2.2.8)$$

or for a small energy shell in continuous form

$$\rho_{N,V,E} = \begin{cases} \text{const} & \text{for } E \leq \mathcal{H}(q_\nu, p_\nu) \leq E + \delta E \\ 0 & \text{otherwise} \end{cases} \quad (2.2.9)$$

2.2.3 Equilibrium

From experience we know that any systems, left alone, evolves to a stable state whose macroscopic properties do not change any more. We call this macrostate **equilibrium**. We can furthermore define the equilibrium as the most probable macrostate.

2.2.4 Entropy and Temperature

Consider an isolated system that is divided into two sub-systems. The total system has $N, V, E = \text{const}$ but between the two subsystems we allow heat exchange.

The total energy, the sum of the two subsystem's energies is constant $E_{total} = E_A + E_B$. We know that in the total system all microstates are equally probable with probability $P_r = \frac{1}{\Omega_{total}(N, V, E)}$. What is now the probability $W(E_1)$ for subsystem A to have energy E_1 (and hence for subsystem B to have energy $E_2 = E_{tot} - E_1$)? We sum over all microstates that are compliant with the partition $E_{total} = E_1 + E_2$:

$$\begin{aligned} W(E_1) &= \sum_{r: E_1, E_2} P_r \\ &= \sum_{r: E_A = E_1} \sum_{r: E_B = E_2} \frac{1}{\Omega_{total}(N, V, E)} \\ &= \frac{\Omega_A(E_1) \cdot \Omega_B(E_2)}{\Omega_{total}(E)} \end{aligned} \quad (2.2.10)$$

The partition functions, i.e., statistical weights of the two subsystems multiply $\Omega = \Omega_A \cdot \Omega_B$. In order to get the number of realisations of the total system (statistical weight), each realisation of subsystem A has to be combined with every realisation of subsystem B.

To find the most probable energy E_1 , i.e. the energy for which $W(E_1)$ is maximal we take the derivative w.r.t. E_1 and set to zero

$$\frac{d}{dE_1} W(E_1) = 0$$

Since the partition function of the total system $\Omega_{total}(E)$ is constant we leave this here out of our calculation. We are thus left with

$$\begin{aligned} \frac{d\Omega_A(E_1) \cdot \Omega_B(E_2)}{dE_1} &= \frac{d\Omega_A(E_1)}{dE_1} \cdot \Omega_B(E_2) + \frac{d\Omega_B(E_2)}{dE_1} \cdot \Omega_A(E_1) \\ \frac{d\Omega_A(E_1)}{dE_1} \cdot \Omega_B(E_2) &= -\frac{d\Omega_B(E_2)}{dE_1} \cdot \Omega_A(E_1) \\ \frac{\frac{d\Omega_A(E_1)}{\Omega_A(E_1) dE_1}}{\frac{d\ln \Omega_A(E_1)}{dE_1}} &= \frac{\frac{d\Omega_B(E_2)}{d\Omega_B(E_2) dE_1}}{\frac{d\ln \Omega_B(E_2)}{dE_1}} \\ \frac{d\ln \Omega_A(E_1)}{dE_1} &= -\frac{d\ln \Omega_B(E - E_1)}{d(E - E_1)} \\ \frac{d\ln \Omega_A(E_1)}{dE_1} &= \frac{d\ln \Omega_B(E_2)}{dE_2} \end{aligned} \tag{2.2.11}$$

At this point it is useful to define a quantity that is related to $\ln \Omega$. We define the **entropy** S

$$S = k_B \ln \Omega \tag{2.2.12}$$

The constant k_B is called the Boltzmann constant. We see that while the partition functions of the two subsystems multiply to give the total partition function, the entropy is additive

$$\begin{aligned} \Omega_{total} &= \Omega_A \cdot \Omega_B \\ \ln \Omega_{total} &= \ln \Omega_A + \ln \Omega_B \\ S_{total} &= S_A + S_B \end{aligned} \tag{2.2.13}$$

Entropy is a fundamental quantity in statistical mechanics. Its partial derivative with respect to the variables determining the microcanonical ensemble define important thermodynamic quantities. These are **temperature** T

$$\left(\frac{\partial S}{\partial \bar{E}} \right)_{N,V} = \frac{1}{T} \tag{2.2.14}$$

pressure p

$$\left(\frac{\partial S}{\partial V} \right)_{N,E} = \frac{p}{T} \tag{2.2.15}$$

and the **chemical potential**

$$\left(\frac{\partial S}{\partial N} \right)_{V,E} = -\frac{\mu}{T} \tag{2.2.16}$$

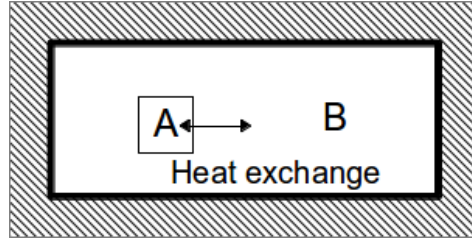


Figure 7: A small subsystem A embedded in a larger subsystem B that can function as a heat bath and hence allow heat exchange. The total system is isolated with $N, V, E = \text{const}$.

With those definitions we can rewrite the condition for the most probable state as

$$\begin{aligned}
 \frac{d \ln \Omega_A(E_1)}{dE_1} &= \frac{d \ln \Omega_B(E_2)}{dE_2} \\
 \frac{dk_B S_A}{dE_1} &= \frac{dk_B S_B}{dE_2} \\
 \frac{k_B}{T_A} &= \frac{k_B}{T_B} \\
 T_B &= T_A
 \end{aligned} \tag{2.2.17}$$

and see that the most probable state represents a **thermal equilibrium** where the temperature is the same in both subsystems. This also means that the change in entropy is the same in both subsystems. At equilibrium $\ln \Omega$ is maximal

$$d \ln \Omega = d \ln \Omega_A + d \ln \Omega_B = 0$$

and therefore also the entropy is maximal

$$dS = dS_A + dS_B = 0 . \tag{2.2.18}$$

2.2.5 The Canonical Ensemble (N, V, T)

Let us consider again an isolated system that is divided into two subsystems A and B between which we allow heat exchange. Let A be much smaller than B, B can be regarded as a heat bath for A, resulting in constant temperature $T = T_A = T_B$. Furthermore the total energy, which is constant is composed of the energies of the two subsystems $E_{total} = E_A + E_B$ but we have $E_A \ll E_B$ such that $E_{total} \approx E_B$. Again all $\Omega(E)$ microstates of the total system are equally probable.

Let now subsystem A have energy $E_A = E_r$, corresponding to $\Omega_A(E_r)$ states. Then only $\Omega_B(E - E_r)$ states, that are equally probable, are left for the total system. The probability for that scenario is

$$P_r = \frac{\Omega_B(E - E_r)}{\Omega(E)} \tag{2.2.19}$$

We can already guess that for larger E_r this ratio, and hence the probability becomes smaller. In order to calculate the probability we make use of the fact that $E_r \ll E$ and expand $\Omega_B(E - E_r)$ in

powers of E_r around E . Actually, we expand $\ln \Omega_B(E - E_r)$:

$$\begin{aligned} \ln \Omega_B(E - E_r) &\approx \ln \Omega_B(E) - \frac{\partial}{\partial E} \ln \Omega_B(E) \cdot E_r + \dots \\ &= \ln \Omega_B(E) - \frac{\partial}{\partial E} \frac{S_B}{k_B} \cdot E_r + \dots \\ &= \ln \Omega_B(E) - \frac{E_r}{k_B T} + \dots \end{aligned} \quad (2.2.20)$$

Neglecting all higher order terms we get

$$\Omega_B(E - E_r) = \Omega_B(E) \cdot \exp\left(-\frac{E_r}{k_B T}\right) \quad (2.2.21)$$

and hence for the probability

$$P_r = \frac{\Omega_B(E)}{\Omega(E)} \cdot \exp\left(-\frac{E_r}{k_B T}\right) \quad (2.2.22)$$

$$= \frac{1}{Z} \cdot \exp\left(-\frac{E_r}{k_B T}\right) \quad (2.2.23)$$

The exponential term is called the **Boltzmann factor** and defines the relative probability of a state with energy E_r at a given temperature. The pre-exponential factor $\frac{1}{Z}$ can be determined from the normalisation, i.e. the sum of all probabilities must be equal to one $\sum P_r = 1$, as

$$Z = \sum_r \exp\left(-\frac{E_r}{k_B T}\right) \quad (2.2.24)$$

$Z(N, V, T)$ is called the **canonical partition function**.

In order to arrive at a general expression (which will turn out to look like eq. 2.2.22) for distributing many systems over different energy levels, we will first look at a simple, explicit example.

Example three systems Let us consider three molecules A, B, and C. These molecules can be at energy levels $\varepsilon_0, \varepsilon_1, \varepsilon_2, \varepsilon_3$ where $\varepsilon_0 = 0$, $\varepsilon_2 = 2 \cdot \varepsilon_1$, and $\varepsilon_3 = 2 \cdot \varepsilon_2$.

The total energy $\sum_i N_i \cdot \varepsilon_i = E = 3 \cdot \varepsilon_1$ and the total number of particles (molecules) is $\sum_i N_i = N = 3$ are fixed where N_i is the number of molecules at energy level ε_i . Distributing the molecules over the available energy levels we obtain three different macrostates I, II and III with the following occupation numbers :

- I: $N_0 = 2$, $N_1 = 0$, $N_2 = 0$, and $N_3 = 1$
- II: $N_0 = 1$, $N_1 = 1$, $N_2 = 1$, and $N_3 = 0$
- III: $N_0 = 0$, $N_1 = 3$, $N_2 = 0$, and $N_3 = 0$

The realisations are as follows

Macrostate	I			II					III
ε_3	A	B	C						
ε_2				C	B	C	A	B	A
ε_1				B	C	A	C	A	B
ε_0	BC	AC	AB	A	A	B	B	C	C
Number of microstates	$3 = \frac{3!}{2! \cdot 0! \cdot 0! \cdot 1!}$			$6 = \frac{3!}{0! \cdot 1! \cdot 1! \cdot 1!}$					$1 = \frac{3!}{2! \cdot 0! \cdot 0! \cdot 0!}$

The probability, i.e. the statistical weight W_n is the largest for state II, because it has most microstates.

When counting the realisations, the microstates, the following rules apply:

- exchanging two particles within the same energy level, does not generate a new microstate, e.g. exchanging the order of writing from AB to BA at energy level ε_0 of state I
- exchanging two particles between two different energy levels generate a new microstate, e.g. swapping A and B in the first realisation (microstate) of state I generates the second microstate
- putting one particle at a different energy level generates a new macrostate, e.g. putting A of the first microstate of state I from ε_3 to ε_2 would generate a macrostate that is neither I, nor II nor III.

The statistical weight W_n can be calculated by

$$W_n = \frac{N!}{N_0!N_1!\dots N_m!} \quad (2.2.25)$$

with $N!$ is the number of possibilities (permutations) to distribute N particles over the energy levels, divided by the product of all $N_i!$ to account for the permutation of particles within the same energy level (N_i =number of particles at energy level i). This is equivalent to how we calculated the statistical weight in Sec. 2.2.2 for a microcanonical ensemble.

2.2.6 The most probable distribution in the canonical ensemble

Consider now a system with N particles. We are again looking for the most probable distribution, e.g. the set of N_i which maximises

$$W_n = \frac{N!}{N_0!N_1!\dots N_m!} \quad (2.2.26)$$

under the constraint that the total number of particles (molecules)

$$\sum_i N_i = N \quad (2.2.27)$$

and the total energy

$$\sum_i N_i \cdot \varepsilon_i = E \quad (2.2.28)$$

are fixed.

Instead of searching for the maximum of W_n we will again search the maximum of $\ln W$. Rewrite $\ln W = \ln N! - \sum \ln_i N_i!$. Using Stirling's equation for large N this can be approximated to

$$\ln W = N \cdot \ln N - N - \sum_i N_i \ln N_i + \sum_i N_i$$

Taking the derivative w.r.t. N_j the term $N \cdot \ln N - N$ vanishes. We have

$$\begin{aligned}
\frac{\partial \ln W}{\partial N_j} &= \frac{\partial}{\partial N_j} \left(N \cdot \ln N - N - \sum_i N_i \ln N_i + \sum_i N_i \right) \\
\frac{\partial \ln W}{\partial N_j} &= -\frac{\partial}{\partial N_j} \sum_i N_i \ln N_i + \frac{\partial}{\partial N_j} \sum_i N_i \\
\frac{\partial \ln W}{\partial N_j} &= -\sum_i \ln N_i \frac{\partial N_i}{\partial N_j} - \sum_i N_i \frac{\partial \ln N_i}{\partial N_j} + \frac{\partial}{\partial N_j} \sum_i N_i \\
\frac{\partial \ln W}{\partial N_j} &= -\sum_i \ln N_i \frac{\partial N_i}{\partial N_j} - \sum_i N_i \frac{1}{N_i} \frac{\partial N_i}{\partial N_j} + \frac{\partial}{\partial N_j} \sum_i N_i
\end{aligned} \tag{2.2.29}$$

of the sum only element $i = j$ remains, hence

$$\frac{\partial \ln W}{\partial N_j} = -\ln N_j \tag{2.2.30}$$

We will account for the two boundary conditions eq.2.2.27 and eq.2.2.28 by introducing the lagrangian multipliers α and β

and minimise the Langrangian \mathcal{L} , i.e. take the derivative w.r.t. N_j and set to zero:

$$\begin{aligned}
\mathcal{L} &= -\ln W + \alpha \left(\sum_i N_i - N \right) + \beta \left(\sum_i N_i \varepsilon_i - E \right) \\
\frac{\partial \mathcal{L}}{\partial N_j} &= \ln N_j + \alpha + \beta \varepsilon_j = 0 \\
\ln N_j &= (-\alpha - \beta \varepsilon_j) \\
N_j &= \exp(-\alpha) \exp(-\beta \varepsilon_j)
\end{aligned} \tag{2.2.31}$$

We know need to know the values of the Langrange multipliers. The one for α can be determined from the boundary condition $\sum_i N_i = N$. Using eq 2.2.31 (and changing the index to i) yields

$$\begin{aligned}
\sum_i \exp(-\alpha - \beta \varepsilon_i) &= N \\
\exp(-\alpha) \sum_i \exp(-\beta \varepsilon_i) &= N \\
\exp(-\alpha) &= \frac{N}{\sum_i \exp(-\beta \varepsilon_i)}
\end{aligned} \tag{2.2.32}$$

plugging this expression 2.2.32 for α now into eq 2.2.31 yields (again using index i):

$$\begin{aligned}
N_i &= \frac{N}{\sum_i \exp(-\beta \varepsilon_i)} \exp(-\beta \varepsilon_i) \\
\frac{N_i}{N} &= \frac{\exp(-\beta \varepsilon_i)}{\sum_i \exp(-\beta \varepsilon_i)}
\end{aligned} \tag{2.2.33}$$

We recognise this as the same relation for the probability in eq. 2.2.22 and identify $\beta = \frac{1}{k_B T}$.

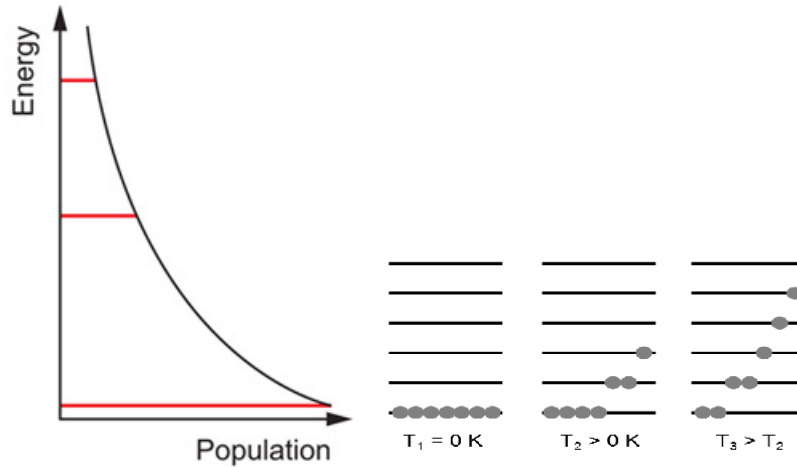


Figure 8: a) According to Boltzmann's law, states with high energy have are significantly less populated (are less likely) than states with low energy. b) At higher temperature, more and more higher energy states are populated.

This relation is called **Boltzmann's law**

$$\frac{N_i}{N} = \frac{\exp(-\varepsilon_i/k_B T)}{\sum_i \exp(-\varepsilon_i/k_B T)} \quad (2.2.34)$$

and applies for all N_i . Hence the most probable distribution is the **Boltzmann distribution** where the relative occupation N_i/N depends exponentially on the energy ε_i of level i and the inverse temperature. This means that the occupation number of a high energy level will be (very) low unless the temperature is high (see Fig. 8).

For degenerate energy levels ε_i with a degeneracy of g_i eq. 2.2.34 changes to

$$\frac{N_i}{N} = \frac{g_i \exp(-\varepsilon_i/k_B T)}{\sum_i g_i \exp(-\varepsilon_i/k_B T)} \quad (2.2.35)$$

Let us look at the temperature dependence of the partition function

$$Z = \sum_i \exp\left(-\frac{1}{k_B T} \varepsilon_i\right) \quad (2.2.36)$$

For a temperature close to zero $\frac{1}{k_B T}$ will be close to infinity and all terms except that for $\varepsilon_0 = 0$ vanish ($e^{-x} \rightarrow 0$ for $x \rightarrow \infty$) because $\varepsilon_0/k_B T = 0$ at all temperatures. On the other hand, if the temperature is very high ($T \rightarrow \infty$) then $\varepsilon_i/k_B T \rightarrow 0$ and $e^{-x} \rightarrow 1$. Each term in the sum contributes by 1, the sum becomes equal to the number of states (usually also close to ∞). The partition function essentially tells which states are accessible at a given temperature. At $T = 0$ only the lowest energy state (ground state) is populated, at higher temperatures more and more higher energy states are populated, and at very high temperatures almost all states are accessible and Z will become large.

In continuous form the partition function reads

$$Z = \frac{1}{h^\nu} \int_{\text{phase space}} \exp[\mathcal{H}(p, q)/k_B T] \quad (2.2.37)$$

and the phase space density

$$\rho_{N,V,T}(q, p) = \frac{1}{h^\nu} \frac{\exp[\mathcal{H}(p, q)/k_B T]}{\int dpdq \exp[\mathcal{H}(p, q)/k_B T]}$$

2.2.7 Isobaric-Isothermal Ensemble (N,p,T)

We can derive many other ensembles in the same manner we obtained the canonical ensemble. Let us do this briefly on more time.

We consider a system, divided in two subsystems and allow energy and volume exchange. One subsystem is small, the other functions as a heat and a “volume bath”. The probability that the small subsystem has energy E_r and volume V_r is

$$P_r = \frac{\Omega_B(E - E_r, V - V_r)}{\Omega(E, V)} \quad (2.2.38)$$

We again expand $\ln \Omega_B$ in a power series

$$\begin{aligned} \ln \Omega_B(E - E_r, V - V_r) &= \ln \Omega_B(E, V) - \frac{\partial}{\partial E} \ln \Omega_B(E, V) \cdot E_r - \frac{\partial}{\partial V} \ln \Omega_B(E, V) \cdot V_r + \dots \\ &= \ln \Omega_B(E, V) - \frac{\partial}{\partial E} S_B \cdot \frac{E_r}{k_B} - \frac{\partial}{\partial V} S_B \cdot \frac{V_r}{k_B} + \dots \\ &= \ln \Omega_B(E, V) - \frac{E_r}{k_B T} - \frac{pV_r}{k_B T} + \dots \end{aligned} \quad (2.2.39)$$

and neglecting all higher order term we get

$$\Omega_B(E - E_r, V - V_r) = \Omega_B(E, V) \cdot \exp\left(-\frac{E_r}{k_B T}\right) \cdot \exp\left(-\frac{pV_r}{k_B T}\right) \quad (2.2.40)$$

and for the probability

$$P_r = \frac{\Omega_B(E, V)}{\Omega(E, V)} \cdot \exp\left(-\frac{E_r}{k_B T}\right) \cdot \exp\left(-\frac{pV_r}{k_B T}\right) \quad (2.2.41)$$

The pre-exponential factor can again be determined from the normalisation, and this is the isothermal-isobaric partition function

$$\Delta(N, p, T) = \sum_V \sum_r \exp(-\beta E - \beta pV) \quad (2.2.42)$$

such that

$$P_r = \frac{\exp\left(-\frac{E_r}{k_B T}\right) \cdot \exp\left(-\frac{pV_r}{k_B T}\right)}{\Delta(N, p, T)} \quad (2.2.43)$$

We can also rewrite the new partition function as

$$\begin{aligned}\Delta(N, p, T) &= \sum_V Z(N, V, T) \cdot \exp\left(-\frac{pV_r}{k_B T}\right) \\ &= \sum_V \sum_E \Omega(N, V, E) \cdot \exp\left(-\frac{E_r}{k_B T}\right) \cdot \exp\left(-\frac{pV_r}{k_B T}\right)\end{aligned}\quad (2.2.44)$$

The corresponding thermodynamic potential is the Gibbs' free energy

$$G = -k_B T \ln \Delta(N, p, T) \quad (2.2.45)$$

which is also known from thermodynamics as

$$G = H - TS \quad (2.2.46)$$

Generalisation We notice a scheme: The partition function of the microcanonical ensemble can be transformed to other partition functions by summing over the macroscopic variables X that are allowed to fluctuate and weight by a Boltzmann factor $\exp\left(-\frac{X}{k_B} \cdot \frac{\partial S}{\partial X}\right)$. With the same recipe we can obtain the probability of a state r for a given value of the macroscopic properties that are allowed to fluctuate by multiplying the Boltzmann factors $\exp\left(-\frac{X}{k_B} \cdot \frac{\partial S}{\partial X}\right)$ of all fluctuating variable and dividing by the partition function of that ensemble.

Grand-Canonical Ensemble We use the Grand-Canonical ensemble as an example. We can again consider a system, divided into a small subsystem and a larger system that now functions as a heat bath and a particle reservoir. That means, we allow the energy and the number of particles to fluctuate. According to recipe the grand-canonical partition function is

$$\begin{aligned}\Xi &= \sum_N \sum_E \Omega(N, V, E) \cdot \exp\left(-\frac{E_r}{k_B} \cdot \frac{\partial S}{\partial E}\right) \cdot \exp\left(-\frac{N_r}{k_B} \cdot \frac{\partial S}{\partial N}\right) \\ \Xi &= \sum_N \sum_E \Omega(N, V, E) \cdot \exp\left(-\frac{E_r}{k_B T}\right) \cdot \exp\left(-\frac{\mu N_r}{k_B T}\right)\end{aligned}\quad (2.2.47)$$

$$\Xi = \sum_N Z(N, V, T) \cdot \exp\left(-\frac{\mu N_r}{k_B T}\right) \quad (2.2.48)$$

The probability is then

$$P_r(E_r, N_r) = \frac{\exp\left(-\frac{E_r}{k_B T}\right) \cdot \exp\left(-\frac{\mu N_r}{k_B T}\right)}{\Xi} \quad (2.2.49)$$

2.3 Thermodynamic functions derived from the partition function

2.3.1 Internal energy, pressure and heat capacity

Knowing the partition function, we can calculate several thermodynamic properties.

For example the average energy as the arithmetic mean of energies.

$$\langle E_i \rangle = \sum_i P_i \cdot \varepsilon_i = \frac{\sum_i N_i \varepsilon_i}{\sum_i N_i} \quad (2.3.1)$$

This is the thermodynamic **internal energy** $U(N, V, T) = \langle E \rangle$.

From the internal energy we can furthermore calculate pressure

$$p = \left(\frac{\partial U}{\partial V} \right)_N \quad (2.3.2)$$

enthalpy

$$H = U + pV \quad (2.3.3)$$

heat capacity at constant volume C_v and constant pressure C_p

$$C_v = \left(\frac{\partial U}{\partial T} \right)_{V,N}$$

$$C_p = \left(\frac{\partial H}{\partial T} \right)_{p,N}$$

and the **Helmholtz Free energy**

$$A = U + TS \quad (2.3.4)$$

We can compute $\langle E \rangle$ and hence the internal energy from the partition function by taking the derivative w.r.t. temperature T

$$\frac{\partial Z}{\partial T} = \frac{\partial \sum_i \exp(\varepsilon_i/k_B T)}{\partial T} \quad (2.3.5)$$

$$\frac{\partial Z}{\partial T} = \frac{1}{k_B T^2} \sum_i \varepsilon_i \exp(-\varepsilon_i/k_B T)$$

rearrange

$$k_B T^2 \frac{\partial z}{\partial T} = \sum_i \varepsilon_i \exp(-\varepsilon_i/k_B T) \quad (2.3.6)$$

and divide by Z which is the same as dividing by $\sum_i \exp(-\varepsilon_i/k_B T)$:

$$k_B T^2 \frac{1}{z} \frac{\partial Z}{\partial T} = \frac{\sum_i \varepsilon_i \exp(-\varepsilon_i/k_B T)}{\sum_i \exp(-\varepsilon_i/k_B T)}$$

$$k_B T^2 \frac{\partial \ln Z}{\partial T} = \langle E \rangle \quad (2.3.7)$$

and this yields

$$U = k_B T^2 \left(\frac{\partial \ln Z}{\partial T} \right)_{V,N} \quad (2.3.8)$$

Likewise the heat capacity can be calculated taking the derivative w.r.t. temperature

$$C_v = \frac{d}{dT} \left[k_B T^2 \left(\frac{\partial \ln Z}{\partial T} \right)_{V,N} \right]$$

$$= 2k_B T^2 + \left(\frac{\partial^2 \ln Z}{\partial T^2} \right)_{V,N}$$

2.3.2 Entropy

The canonical partition function $Z(N, V, T) = \sum_r \exp\left(-\frac{E_r}{k_B T}\right)$ is related to the microcanonical partition function $\Omega(N, V, E) = \sum_r 1$ by the Boltzmann factor

$$Z = \Omega \cdot \exp\left(-\frac{E_r}{k_B T}\right) \quad (2.3.9)$$

While all microstates on one energy surface are equally probable, the different energy surfaces are Boltzmann-weighted.

We can also write

$$\begin{aligned} \ln Z &= \ln \left[\Omega \cdot \exp\left(-\frac{E_r}{k_B T}\right) \right] \\ \ln Z &= \ln \Omega - \frac{E_r}{k_B T} \\ k_B \ln Z &= k_B \ln \Omega - \frac{E_r}{T} \end{aligned} \quad (2.3.10)$$

We can now replace E_r by the most probable energy E^* or the average energy $\langle E \rangle = U$.

$$\begin{aligned} k_B \ln Z &= S - \frac{U}{T} \\ -k_B T \ln Z &= U - TS \end{aligned} \quad (2.3.11)$$

The thermodynamic entropy in the canonical ensemble is thus

$$S = \frac{U}{T} + k_B \ln Z \quad (2.3.12)$$

and using the above equation 2.3.8 for the internal energy

$$S = k_B T \left(\frac{\partial \ln z}{\partial T} \right)_{N, V} + k_B \ln z \quad (2.3.13)$$

We also note that the Helmholtz free energy

$$\begin{aligned} A &= U - TS \\ A &= -k_B T \ln Z \end{aligned} \quad (2.3.14)$$

can be directly obtained from the canonical partition function. The Helmholtz free energy is thus the macroscopic **thermodynamic potential** of the canonical ensemble.

2.3.3 Helmholtz Free energy

Let us now consider two states M and L with m and l microstates of energy ε_m and ε_l , respectively. Then the free energy of state M is

$$\begin{aligned} A_M &= -k_B T \ln \sum_i^m \exp(-\varepsilon_i/k_B T) \\ A_M &= -k_B T \ln [m \cdot \exp(-\varepsilon_m/k_B T)] \\ A_M &= -m \cdot k_B T - \ln [\exp(-\varepsilon_m/k_B T)] \cdot k_B T \\ A_M &= \varepsilon_m - k_B T \ln m \end{aligned} \quad (2.3.15)$$

and for state L accordingly

$$A_L = \varepsilon_l - k_B T \ln l \quad (2.3.16)$$

For the free energy difference we get

$$\begin{aligned} \Delta A &= A_M - A_L & (2.3.17) \\ \Delta A &= \varepsilon_m - k_B T \ln m - \varepsilon_l + k_B T \ln l \\ \Delta A &= \varepsilon_m - \varepsilon_l - k_B T \ln \frac{m}{l} \end{aligned}$$

At equilibrium the free energy difference is zero, i.e.

$$\begin{aligned} \varepsilon_m - \varepsilon_l &= k_B T \ln \frac{m}{l} \\ \exp\left(\frac{\varepsilon_m - \varepsilon_l}{k_B T}\right) &= \frac{m}{l} \end{aligned} \quad (2.3.18)$$

Thus, if the two energies ε_m and ε_l are equal, the number of microstates m and l are also equal. If one energy is larger than the other, e.g. $\varepsilon_M > \varepsilon_L$, then the number of microstates m must be much larger than l to be in equilibrium, i.e. to have equally stable macrostates M and L . However, with increasing temperature the exponential term on the left hand side decreases and the ratio of m/l needs to be smaller and smaller. For $T \rightarrow \infty$ the exponent tends to zero and macrostates with equal number of microstates are equally stable.

2.3.4 Chemical equilibrium

Consider a system in equilibrium that is divided in two subsystems. WE allow exchange of heat, volume and particles such that

$$\begin{aligned} E &= E_A + E_B = \text{const} \\ V &= V_A + V_B = \text{const} \\ N &= N_A + N_B = \text{const} \end{aligned} \quad (2.3.19)$$

at equilibrium the entropy s maximal

$$S(E_A, V_A, N_A) = S_A(E - E_A, V - V_A, N - N_A) + S_B(E - E_A, V - V_A, N - N_A)$$

And from the optimality condition $\frac{\partial S}{\partial E_A} = 0$ we get $T_A = T_B$; from $\frac{\partial S}{\partial V_A} = 0$ we get $\frac{p_A}{T_A} = \frac{p_B}{T_B}$ hence $p_A = p_B$ and finally from $\frac{\partial S}{\partial N_A} = 0$ we get $\frac{\mu_A}{T_A} = \frac{\mu_B}{T_B}$ that is

$$\mu_A = \mu_B$$

the two subsystems are also in **chemical equilibrium**.

2.4 Ensemble Averages

In an experiment often a series of measurements is carried out at certain time intervalls τ to probe an observable O with $O(q, p)$ where q are the positions and p are the momenta of the articles

(atoms) in the system. Let us assume we know how O evolves in time. Then we can calculate the time-average of the observable O of a system (in the N,V,E ensemble) as

$$\overline{O(q,p)} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t d\tau O(q,p) \quad (2.4.1)$$

The time average does not depend on the initial conditions for long times t , hence the time average does not change if we average over different initial conditions q_0, p_0

$$\overline{O(q,p)} = \frac{\sum_{\text{initial conditions}} \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t d\tau O(q,p,q_0,p_0,\tau)}{\text{Number of initial conditions}} \quad (2.4.2)$$

and we can also swap the summation with the integration

$$\overline{O(q,p)} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t d\tau \frac{\sum_{\text{initial conditions}} O(q,p,q_0,p_0,\tau)}{\text{Number of initial conditions}} \quad (2.4.3)$$

For the limit case that the sum runs over all initial conditions that are compatible with the N,V,E of the ensemble we can write

$$\frac{\int_E dqdp O(q,p)}{\int_E dqdp} = \langle O \rangle_{NVE} \quad (2.4.4)$$

where $\langle \dots \rangle$ denotes an ensemble average and the integration index E is used to indicate that only points in phase space with the correct energy are taken into account (integration over the correct “shell” of phase space). This can be alternatively written as

$$\frac{\int dqdp O(q,p) \delta(\mathcal{H}(q,p) - E)}{\int dqdp \delta(\mathcal{H}(q,p) - E)} = \langle O \rangle_{NVE} \quad (2.4.5)$$

where the density of states $\rho_{NVE} = \delta(\mathcal{H}(q,p) - E)$ is a delta function, that is non-vanishing only for points that have the correct energy $\mathcal{H}(q,p) = E$.

Using the notation for the ensemble average we can rewrite eq. 2.4.3 as

$$\overline{O(q,p)} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t d\tau \langle O(q,p, q_0, p_0, \tau) \rangle \quad (2.4.6)$$

Since the ensemble average does not depend on time τ the expression simplifies to

$$\overline{O(q,p)} = \langle O(q,p) \rangle_{NVE} \quad (2.4.7)$$

which means that averaging over all initial phase space coordinates is equivalent to averaging over time-evolved phase space coordinates (for infinitely long times). This is the **ergodic hypothesis**, which can also be understood as: if a time series visits all points in phase space, the average over the time series is equivalent to the ensemble average.

We can therefore generate one or the other average. Generating an ensemble would be “the Monte-Carlo way”, whereas generating a time series is “the Molecular-Dynamics way”.

Note: In the canonical ensemble with the density of states $\rho = \exp(-\beta\mathcal{H}(q,p))$ the thermal average is

$$\frac{\int dqdp O(q,p) \exp(-\beta\mathcal{H}(q,p))}{\int dqdp \exp(-\beta\mathcal{H}(q,p))} = \langle O \rangle_{NVT} \quad (2.4.8)$$