5.3 The aim of statistical mechanics

The aim of statistical mechanics is the evaluation of the laws of classical thermodynamics for macroscopic systems using the properties of its atomic particles. This we already did in the last sections but we need three different approaches to solve for the three fundamental distribution functions. In what follows we will introduce an even more general concept which allows to solve for all distribution functions in one approach. The introduction of an information function will as well give a deeper understanding of entropy and in addition to the classical TD the statistical approach provides information on the nature of statistical errors and variations of thermodynamic parameters.

macro state	micro state
e.g.- is characterized by $T, p, V, N, ...$	• For each particle at least 6 variables (3 for position, 3 for momentum) • e.g. the spin, rotational momentum, magnetic moments are additional parameters of each atom (molecule) • all properties which characterize a system • many micro states belong to one macro state • statistical ensemble • averages describe state of macroscopic system

Question: Which **weight** has a micro state in a macro state?

Most general principle: Maximize "Degree of Uncertainty"

within the restrictions of the macro state

"Degree of uncertainty" = "Thermodynamic entropy"

Event i with probability p_i (Note: in contrast to the last sections i indicates micro states, not energy levels)

$$1 \geq p_i \geq 0 \quad \text{and} \quad \sum_i p_i = 1 \qquad . \tag{5.32}$$

The degree of uncertainty is defined by the contents of information of a statement.
For a function I to be an information $I(p_i)$ of a statement with a probability p_i we need several properties:

1. $I(1) = 0$ (From a certain statement to be fulfilled we get no new information)
2. $I(p_i)$ monotonously increasing with $1/p_i$
3. $I(p_i p_j) = I(p_i) + I(p_j)$ for two independent events the information just adds up

These three properties are fulfilled using the function

$$I(p_i) = -k \ln(p_i) \qquad . \tag{5.33}$$

with k: (arbitrary) Measurement unit for information.
For our information function we find:

- $I(0) = \infty$ (This is at least plausible)

To calculate the **average information** we must multiply the information of a statement with its weight (probability) of occurrence.
Thus we get

$$S' = \sum_i p_i I(p_i) = -k \sum_i p_i \ln(p_i) \qquad . \tag{5.34}$$

For the equilibrium of a physical system the degree of uncertainty S' must be maximized:
The mathematical effort is to find

$$S = \max S' \tag{5.35}$$

within the restrictions of the macro state.
The entropy S is therefore just the maximum average information.

Justification of this principle: The results describe all experiments: Macro states are dominated by micro states with large probabilities.

As we will learn in the next sections for classical particles in an isolated system the maximum of S' (cf. Eq. (5.34)) is found if all states i are occupied with the same probability $p_i = p$, i.e.

$$1 = \sum_{i=1}^{W} p_i = Wp \qquad \text{resp.} \qquad p = \frac{1}{W} \quad . \tag{5.36}$$

Inserting this result into Eq. (5.34) we find the famous equation

$$S = \max S' = -k \sum_{i=1}^{W} p \ln(p) = -k \ln\left(\frac{1}{W}\right) = k \ln(W) \tag{5.37}$$

which clarifies the relation to the first definition of entropy in Eq. (5.11).
The relation between "average information" and "degree of uncertainty" may be somewhat counter intuitive; so we will discuss it for an example:
Let us assume a set of classical particles. All particles shall occupy state 1, i.e. $p_1 = 1$ and $p_i = 0$ for $i > 1$.
Since

$$\lim_{x \to 0} [x \ln(x)] = 0 \qquad \text{and} \qquad [1 \ln(1)] = 0 \tag{5.38}$$

we find for the average information $S' = 0$. We know "everything" about the occupation of the states i; therefore the degree of uncertainty is 0. Maximizing the information for **one** state therefore minimizes the average information for the ensemble. Since p_i and $I(p_i)$ are positive numbers in fact $S' = 0$ is the global minimum of S'.
A second possible misunderstanding concerns the relation between W and the particle number N. W is incomparably larger than N. Let us discuss this for a (most) simple example with only two different possible states, e.g. left half, right half of a box. Each particle therefore has two possibilities for occupying states leading to $W = 2^N$ possible arrangements, i.e. micro states. Including this into Eq. (5.37) we get

$$S = k \ln\left(2^N\right) = kN \ln(2) \quad . \tag{5.39}$$

Eq. (5.39) demonstrates that of course the entropy is an extensive parameter; it scales with the size of the system. For a thermodynamic system N is already a large number, but W is much larger.
For typical thermodynamic systems each particle can occupy many different states, so the factor of 2 in our above example must be replaced typically by numbers in the order of 10^{20} and thus $W \approx 10^{(20\,N)}$ which is indeed a huge number.