

Chapter 7

Equilibrium statistical physics

7.1 Introduction

The movement and the internal excitation states of atoms and molecules constitute the *microscopic* basis of thermodynamics. It is the task of *statistical physics* to connect the microscopic theory, viz the classical and quantum mechanical equations of motion, with the *macroscopic* laws of thermodynamics.

Thermodynamic limit The number of particles in a macroscopic system is of the order of the Avogadro constant $N_a \sim 6 \cdot 10^{23}$ and hence huge. An exact solution of 10^{23} coupled equations of motion will hence not be possible, neither analytically,^{*} nor numerically.[†] In statistical physics we will be dealing therefore with probability distribution functions describing averaged quantities and their fluctuations. Intensive quantities like the free energy per volume, F/V will then have a well defined *thermodynamic limit*

$$\lim_{N \rightarrow \infty} \frac{F}{V} . \quad (7.1)$$

The existence of a well defined thermodynamic limit (7.1) rests on the law of large numbers, which implies that fluctuations scale as $1/\sqrt{N}$. The statistics of intensive quantities reduce hence to an evaluation of their mean for $N \rightarrow \infty$.

Branches of statistical physics. We can distinguish the following *branches* of statistical physics.

- Classical statistical physics. The microscopic equations of motion of the particles are given by classical mechanics. Classical statistical physics is incomplete in the sense that additional assumptions are needed for a rigorous derivation of thermodynamics.
- Quantum statistical physics. The microscopic equations of quantum mechanics provide the complete and self-contained basis of statistical physics once the statistical ensemble is defined.

^{*} In classical mechanics one can solve only the two-body problem analytically, the Kepler problem, but not the problem of three or more interacting celestial bodies.

[†] A computer made of 10^{23} atoms will not be able to simulate 10^{23} differential equations.

- Statistical physics in equilibrium. The macroscopic variables are *time-independent* and the macroscopic world can be described in terms of microscopic average values, distribution functions or probabilities. The derived relations between the macroscopic quantities correspond then to the laws of thermodynamics.
- Statistical physics in non-equilibrium. The macroscopic variables are *time-dependent*. Advanced approximate methods are needed in this case for a microscopic treatment.

In this course, we will restrict ourselves to classical and quantum statistical physics in *equilibrium*.

7.2 Statistical mechanics fundamentals

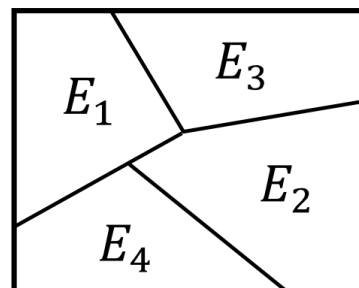
Our objective is to describe the macroscopic properties of $N \gg 1$ particles in thermodynamic equilibrium.

Conserved quantities. The properties of a system in *thermodynamic equilibrium* do not change with time. This implies, as discussed further in Sect. 7.2.1, that there exist conserved quantities like energy and momentum. Conservation of energy and particle number will be of particular importance for the definition of statistical ensembles.

Additivity. Macroscopic extensive variables must add when a thermodynamic system is partitioned into subsystems. The total energy E of a system subdivided into subsystems with energies E_i should hence obey

$$E = \sum_i E_i$$

with negligible boundary effects. Additivity is present when the correlations between the distinct subsystems vanish in the thermodynamic limit. In this course we will not discuss the derivation of additivity, which is based microscopically on finite *correlation lengths*.



Energy vs. internal energy. With energy being both conserved and additive we can identify it with the internal energy $U = U(S, V, N)$. The total differential (5.14) of the entropy $S(U, V, N) = S(E, V, N)$ then allows us to evaluate with

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E} \right)_{V,N}, \quad \frac{P}{T} = \left(\frac{\partial S}{\partial V} \right)_{E,N}, \quad -\frac{\mu}{T} = \left(\frac{\partial S}{\partial N} \right)_{E,V} \quad (7.2)$$

in particular the temperature T . Above relation can be hence viewed as one of the possible definition of the temperature.

Defining the entropy. The relations (7.2) can be evaluated once the entropy S and the energy E have been computed for a systems of N particles confined to a volume V . This is not a problem for the energy, for which well-established microscopic expressions are at hand both for classical and for quantum mechanics. In order to proceed with statistical mechanics a microscopic definition of the entropy is hence necessary.

7.2.1 Classical mechanics

The state of a particle at any instant of time is specified within classical mechanics by its position \vec{q} and momentum \vec{p} . A N -particle system has therefore $6N$ degrees of freedom specified by

$$(\vec{q}_i, \vec{p}_i), \quad i = 1, \dots, N. \quad (7.3)$$

Hamiltonian. The motion of the particles are governed by the canonical equations

$$\dot{\vec{p}}_i = -\frac{\partial H}{\partial \vec{q}_i}, \quad \dot{\vec{q}}_i = \frac{\partial H}{\partial \vec{p}_i}, \quad H = H(\vec{q}_1, \dots, \vec{q}_N, \vec{p}_1, \dots, \vec{p}_N), \quad (7.4)$$

where

$$H = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} + \frac{1}{2} \sum_{i \neq j} \Phi(\vec{r}_i - \vec{r}_j) \quad (7.5)$$

is the Hamiltonian of the system. In (7.5) we have used the real-space \vec{r}_i coordinates for the generalized coordinates \vec{q}_i . The interaction between the particles is given in (7.5) by the interaction potential $\Phi(\vec{r}_i - \vec{r}_j)$.

Noether's theorem. A corollary of Emmy Noether's theorem is that energy is conserved when the Hamiltonian is independent of time. For a thermodynamic setting, where particles are confined to a volume V , one needs to add a one-particle potential

$$\Phi_V(\vec{r}) = \begin{cases} 0 & (\vec{r} \in V) \\ \infty & (\vec{r} \notin V) \end{cases} \quad (7.6)$$

to the Hamiltonian (7.5). Energy is therefore not guaranteed to be conserved for a volume change, which corresponds to a time-dependent confining potential Φ_V and therefore also to a time-dependent Hamiltonian.

Ideal gas limit. The ideal gas limit corresponds to a negligible inter-particle interaction $\Phi(\Delta\vec{r}) \rightarrow 0$ in (7.5). For most real gases the density of molecules is small enough that this condition is approximately fulfilled, with the residual contribution of $\Phi(\Delta\vec{r})$ giving rise to corrections, as captured e.g. by the van der Waals theory discussed in Sect. 6.4.

7.2.2 Microstates

The complete description of an isolated system with N particles is given by the generalized coordinates

$$q = (q_1, \dots, q_{3N}), \quad p = (p_1, \dots, p_{3N}), \quad (7.7)$$

which span the *phase space*. The generalized coordinates (\vec{q}_i, \vec{p}_i) of the individual particles are lumped together as components in (7.7). Alternatively one may consider as in (7.3) directly the set of all individual (\vec{q}_i, \vec{p}_i) .[‡]

Microstates. In dynamical system theory one is interested in evaluating the *trajectories* in the phase space, viz the *time evolution* $(\vec{q}_j(t), \vec{p}_j(t))$ for $j = 1, \dots, N$. In contrast one is interested in statistical mechanics in the statistical properties of the system. One then defines a given state (7.7) in phase space as a *microstate*.

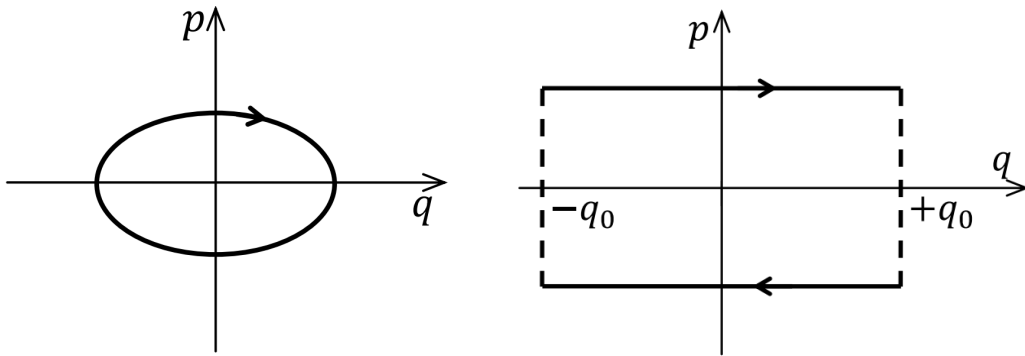
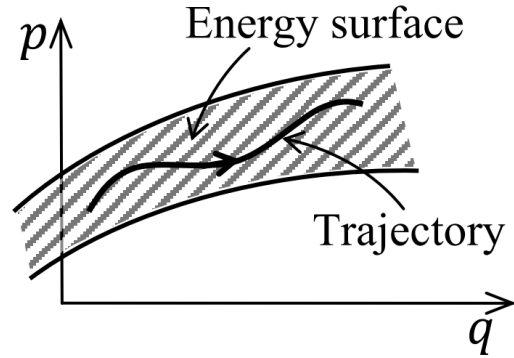
[‡] Eq. (7.7) is sometimes called the Γ -representation of phase space and Eq. (7.3) the μ -representation.

Energy surface. Energy is a conserved quantity if H does not depend explicitly on time. The trajectories are then confined to the *energy surface*.

As an example consider the one-dimensional harmonic oscillator ($N = 1$) with the Hamiltonian

$$H = \frac{p^2}{2m} + \frac{1}{2}m\omega^2 q^2 = E .$$

The trajectory of this system in the Γ -space is given by the ellipse $E = \text{const.}$



The energy surface is on the other side discontinuous for a particle in a potential box with ideal reflecting walls, as specified by $\vec{p} \leftrightarrow (-\vec{p})$ and (7.6).

7.2.3 Probability distribution functions

Integrating over the microstates (7.7) corresponds to the differential

$$d^{3N}q d^{3N}p = dq_1, \dots, dq_{3N}, dp_1, \dots, dp_{3N} .$$

Probability distribution function. We define the *distribution function* $\rho(q, p)$ (also called probability density), with

$$\int \rho(q, p) d^{3N}q d^{3N}p = 1, \quad \rho(q, p) \geq 0 , \quad (7.8)$$

which gives the probability to find the system in the state $(q, p) = (q_1, \dots, q_{3N}, p_1, \dots, p_{3N})$.

Ideal-gas limit. The constituent particles do not interact in the ideal-gas limit. Defining with

$$\int f(\vec{p}, \vec{q}) d\vec{q} d\vec{p} = 1, \quad f(\vec{q}, \vec{p}) \geq 0$$

the probability that a given particle has coordinates \vec{q} and momentum \vec{p} leads then to

$$\rho(q, p) = \prod_{i=1}^N f(\vec{p}_i, \vec{q}_i) , \quad (7.9)$$

where i runs now over all particles. This relation just states that the probability of a combination of independent events (like throwing a dice several times) is just the product of the individual probabilities.

Expectation values. We use in our notation E for the total energy of the system and

$$\frac{E}{N} = \int \frac{\vec{p}^2}{2m} f(\vec{q}, \vec{p}) d\vec{q} d\vec{p} = \int \frac{\vec{p}^2}{2m} f(\vec{p}) d\vec{p} \quad (7.10)$$

for the energy per particle, where we have used in the last step that $f(\vec{q}, \vec{p}) = f(\vec{p})$ when the system is uniform. In this case the properties do not depend on where they are measured. Note that only the kinetic energy did enter (7.10), as we are working in the ideal-gas limit.

7.3 Liouville theorem

Let us consider a system with the distribution function as defined by (7.8). We are now interested in the time dependence, viz in $\rho = \rho(q, p, t)$,

Phase space flow. The velocity \vec{v}_{ps} in phase space (PS),

$$\vec{v}_{ps} = \frac{d}{dt}(q, p) = (\dot{q}_1, \dot{q}_2, \dots, \dot{q}_{3N}, \dot{p}_1, \dots, \dot{p}_{3N}), \quad (7.11)$$

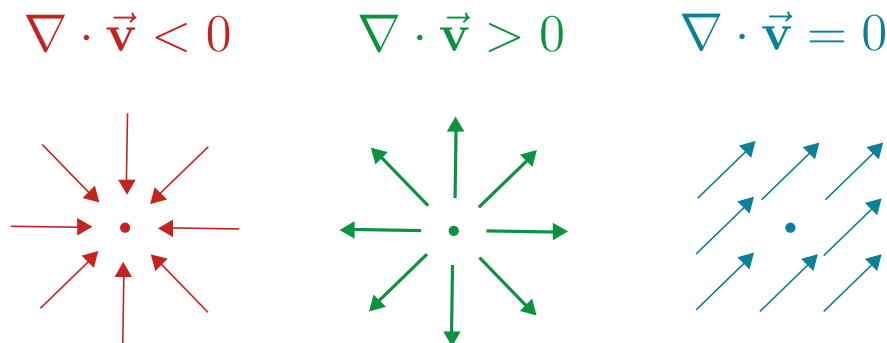
is called the *flow*. Its divergence

$$\nabla \cdot \vec{v}_{ps} = \sum_{j=1}^{3N} \left(\frac{\partial \dot{q}_j}{\partial q_j} + \frac{\partial \dot{p}_j}{\partial p_j} \right), \quad \nabla \cdot \vec{v}_{ps} \equiv 0, \quad (7.12)$$

vanishes generically for Hamiltonian systems governed by (7.4), since.

$$\frac{\partial \dot{q}_j}{\partial q_j} = \frac{\partial}{\partial q_j} \frac{\partial H}{\partial p_j}, \quad \frac{\partial \dot{p}_j}{\partial p_j} = -\frac{\partial}{\partial p_j} \frac{\partial H}{\partial q_j}.$$

Any vector field with a vanishing divergence, like the phase space velocity \vec{v}_{ps} , is volume-conserving. Repeating the arguments leads to *Liouville's theorem* (see below).



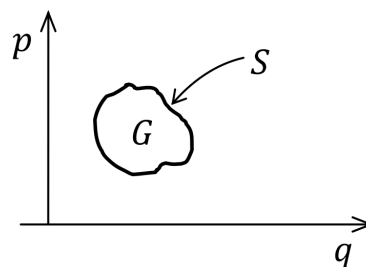
Phase space current. The flow (7.11) defines also the phase-space current density \vec{j}_{ps} , viz

$$\vec{j}_{ps} = \rho(q, p, t) \vec{v}_{ps} .$$

Remember that the electrical current is defined in analogy as the product of charge, particle density and velocity.

Continuity equation in integral form. Consider a region G in phase space with surface S . The number of trajectories that cross the surface S per unit time is given by

$$\boxed{\int_{S(G)} d\vec{S} \cdot \vec{j}_{ps}} , \quad d\vec{S} = dS \cdot \vec{n} ,$$



where \vec{n} is a vector orthonormal to the surface S . The flux of trajectories through the surface $S(G)$ obeys

$$\int_{S(G)} d\vec{S} \cdot \vec{j}_{ps} = -\frac{\partial}{\partial t} \int_G d^{3N}q d^{3N}p \rho(q, p, t) , \quad (7.13)$$

as there are no sources or sinks for the trajectories. Eq. (7.13) is just the continuity equation in integral form.

Gauss' law. By making use of the Gauss law we can replace the integral over the surface in (7.13) by a volume integral,

$$\boxed{\int_G d^{3N}q d^{3N}p \left[\frac{\partial}{\partial t} \rho + \vec{\nabla} \cdot \vec{j}_{ps} \right] = 0} , \quad (7.14)$$

with

$$\nabla \equiv \left(\frac{\partial}{\partial q_1}, \dots, \frac{\partial}{\partial q_{3N}}, \frac{\partial}{\partial p_1}, \dots, \frac{\partial}{\partial p_{3N}} \right)$$

and

$$\nabla \cdot \vec{j}_{ps} = \rho \nabla \cdot \vec{v}_{ps} + (\vec{v}_{ps} \cdot \nabla) \rho \equiv (\vec{v}_{ps} \cdot \nabla) \rho , \quad (7.15)$$

where we have used (7.12), namely that the divergence $\nabla \cdot \vec{v}_{ps}$ of the flow vanishes for Hamiltonian systems.

Liouville equation. With the continuity equation in integral form (7.14) being valid for all G we obtain

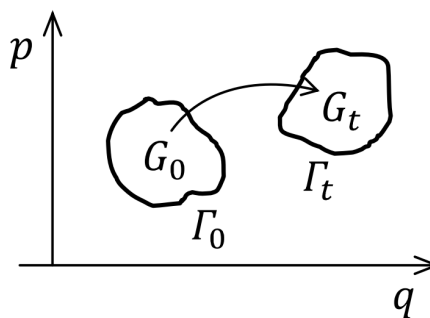
$$\frac{\partial}{\partial t} \rho(q, p, t) + (\vec{v}_{ps} \cdot \nabla) \rho(q, p, t) = 0, \quad \boxed{\frac{d}{dt} \rho(q, p, t) = 0} , \quad (7.16)$$

where we have used (7.15) for the divergence of \vec{j}_{ps} and that the left-hand side of the first term in (7.16) corresponds to the total differential of $\rho(q, p, t)$ with respect to the time. Eq. (7.16) is denoted *Liouville theorem*.

Conservation of phase space volume. The Liouville equation (7.16) tells us that the total time derivative $d\rho/dt$ of the density $\rho(q, p, t)$ of trajectories vanishes. It implies that

$$\rho(q(t), p(t), t) \equiv \rho(q(0), p(0), 0)$$

is a constant of motion. Phase space volume is hence conserved. This result holds whenever the divergence $\nabla \cdot \vec{v}_{ps}$ of the flow vanishes, see (7.12).



7.4 Relaxing towards thermal equilibrium

One mol of an ideal gas occupies a molar volume of

$$V_0 = 2.24 \cdot 10^4 \text{ cm}^3, \quad R = \frac{P_0 V_0}{T_0} = 8.3 \frac{\text{J}}{\text{mol K}},$$

where we have used $T_0 = 273.15 \text{ K}$ and $P_0 = 1 \text{ bar}$, corresponding to $10^5 \text{ Pa} = 10^5 \text{ N/m}^2$. The density of the gas is then

$$n = \frac{N_a}{V_0} = 2.7 \cdot 10^{19} \text{ atoms/cm}^3, \quad N_a = 6.022 \cdot 10^{23}.$$

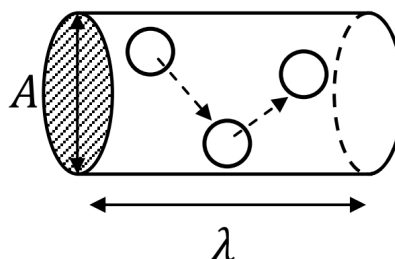
Scattering cross section. The thermal equilibrium is reached in a gas because of particle collisions, viz due to the inter-particle interaction $\Phi(\Delta\vec{r})$ in (7.5). The scattering cross section

$$\sigma = \pi r_0^2, \quad r_0 \approx 2 \cdot 10^{-8} \text{ cm}$$

is for short-ranged inter-particle interactions $\Phi(\Delta\vec{r})$ of the order of the area of the particles. Note that the Bohr radius is $5.29 \cdot 10^{-11} \text{ m} = 0.529 \text{ \AA}$.

Mean free path. We define the *mean free path* λ as the average distance traveled by a particle between two successive collisions. Consider a cylinder with a cross section A containing a gas of density n .

The length of the cylinder is of the mean-free path λ when A is approximately equal to the cumulative cross-sections of the contained atoms, viz when



$$A \approx \underbrace{A \lambda n}_N \sigma, \quad \lambda = \frac{1}{n\sigma} \sim 10^{-5} \text{ cm}.$$

The average distance between two collisions is hence much larger than the typical size of a molecule, $\lambda \gg r_0$, which implies that the molecules in real gases interact only weakly.

Average velocity. We did evaluate in Sect. 3.4.1 the internal energy per particle of an ideal gas as $3k_B T/2$. Then, we can obtain the average velocity v of the constituent particles as

$$\frac{1}{2}mv^2 = \frac{3}{2}k_B T, \quad \boxed{v = \sqrt{\frac{3k_B T}{m}}}$$

The mean velocity is at room temperature ($T = 300$ K) of the order of $v \approx 10^5$ cm/s, viz of the order of one km/s.

Relaxation time. The average time between successive collisions τ_c ,

$$\boxed{\tau_c = \frac{\lambda}{v}} \approx 10^{-10} \text{ s},$$

corresponds then to the time the gas needs to relax to thermal equilibrium. The *relaxation time* τ_c is essentially instantaneous with regard to the time scales macroscopic objects need to move.

7.4.1 Ergodicity timescales

A point

$$(q, p) = (q_1, \dots, q_{3N}, p_1, \dots, p_{3N})$$

in phase space corresponds to a state of the N -particle system. It follows as a function of time a trajectory on the energy surface. After a certain time, which should be long enough compared to the collision time $\tau_c \approx 10^{-10}$ s (for an ideal gas at a temperature of $T = 300$ K and a pressure of $P = 1$ bar), the trajectory has lost essentially all information about its initial state. Averaging further over time then corresponds to a sampling of the thermal state.

Ergodicity. A dynamical system is said to be *ergodic* when it eventually explores the entire available phase space.

ERGODICITY Given sufficient time, the representative point of an isolated system will come arbitrarily close to any given point of the energy surface.

A system needs to be ergodic in order to correspond to a uniquely defined thermal state. Its time-averaged properties would otherwise depend on the initial conditions.

Liouville theorem. Liouville theorem (7.16) states that phase space volume is invariant with respect to Hamiltonian flows like (7.4). With the energy surface having a finite volume this implies that the flow visits eventually every point in phase space in the sense that it will approach it arbitrarily close, time given. The involved recurrence times may however be exceedingly long.[§]

[§] A famous example is the teacup experiment. Try to stare at your teacup and to wait for a drop of tea to be pushed (by coordinated movements of the water molecules) out into the open air.

Chaotic states. The continuous scattering of the particle results in unpredictable and hence chaotic motions. The loss in predictability results from the fact that small variations in the distance of near miss collisions may result in large changes of the trajectories of the two colliding particles. Systems showing [deterministic chaos](#) sample phase space effectively.

7.4.2 Time- vs. statistical averaging

Classical experiment, measuring thermodynamic quantities like the temperature, typically need a certain time. They hence perform averages over *time intervals* that are *small* on the macroscopic scale but *large* compared to the collision time τ_c .

POSTULATE In statistical mechanics it is assumed that *time averages* can be replaced by averages over a suitably chosen collection of systems called *statistical ensemble*.

Observables. To see what this means let us consider an observable $O = O(q, p)$,

$$O(q, p) = O(q_1, \dots, q_{3N}, p_1, \dots, p_{3N}) ,$$

which is classical just a function on phase space. A well known observable is the kinetic energy

$$E_{kin} = \sum_{i=1}^N \frac{\vec{p}_i^2}{2m} .$$

Measurements. A *measuring process* provides via

$$\bar{O}(t_0) = \frac{1}{\tau} \int_0^\tau dt O(q(t_0 + t), p(t_0 + t)) \quad (7.17)$$

a time average \bar{O} of the observable. t_0 is here initial time and τ the duration of the experiment.

Ergodic hypothesis. It is impossible to calculate $\bar{O}(t_0)$ from the the solution of the Hamiltonian equations for $N = 10^{23}$ molecules. For an chaotic and hence ergodic system we can however assume that the system samples the chaotic attractor within the duration τ of the measurement. It is then possible to replace the time average (7.17) with

$$\langle O \rangle \equiv \int d^{3N}q \int d^{3N}p \rho(q, p) O(q, p) , \quad (7.18)$$

viz by the corresponding expectation value over the phase-space density $\rho(q, p)$.

STATISTICAL ENSEMBLE A *statistical ensemble* is defined as the infinite collection of identical copies of the system characterized by a distribution function $\rho(q, p, t)$.

Interpretation. The time average \bar{O} of a function $O = O(q, p)$ on phase space and the ensemble average $\langle O \rangle$, as defined respectively by (7.17) and (7.18), coincide for systems in thermal equilibrium.

7.5 Postulate of equal a priori probabilities

In order to connect the microscopic equations of motion to the laws of thermodynamics one needs to define in addition either the temperature T or the entropy S , as done in the following chapters. A key step in this direction is to understand the structure of $\rho(q, p)$.

Constant energy. The total energy E is defined in the ideal gas limit as

$$E = \frac{1}{2m} \sum_{j=1}^N (\vec{p}_j)^2 \quad (7.19)$$

This equation defines the surface of constant energy as a hyper-sphere in phase space. The question regards then the form the density of states $\rho(q, p)$ takes on this hypersphere.

Constant velocities. It could be, as a matter of principle, that

$$\rho(q, p) = \prod_{i=1}^{3N} \delta(p_i - \bar{p}), \quad \bar{p} = \sqrt{2mE/(3N)}. \quad (7.20)$$

In this case all particles would have identical and constant momenta and hence velocities. Eq. (7.20) is however not what is observed.

All allowed microstates are equally likely. The *postulate of equal a priori probabilities* states that all allowed microstates are equally likely. Eq. (7.20) is hence as wrong as it gets, with $\rho(q, p)$ being on the contrary constant on the hypersphere defined by (7.19).

The postulate of equal a priori probabilities can only be motivated but not proven. We will discuss it further in the next chapter. It has the stance of a *law of nature*.

All laws of nature are *postulates* in the sense that they need to be verified by experiments. They cannot be derived.