

where $s_0(T_0, p_0)$ is an arbitrary dimensionless function of the reference state (T_0, p_0) . In the statistical treatment we will directly calculate values for $s_0(T_0, p_0)$.

Up to now we have considered the entropy from a purely thermodynamic point of view and have assumed some facts known from experience without substantiating them. The meaning of entropy, however, becomes immediately obvious, if we consider its microscopic interpretation.

Insertion: Microscopic interpretation of entropy and of the second law

Statistical considerations are of enormous importance for gaining deeper insight into the purely phenomenological relationships of thermodynamics. As we will see, it is just the second law which establishes a very close connection between the statistical and the phenomenological point of view. The second law formulates the well-known fact that all isolated physical systems converge toward an equilibrium state, where the state quantities do not change anymore after a certain relaxation time. Furthermore, it asserts that this process never reverses itself. A nice example for this fact is a gas which one spontaneously offers a larger volume and which will homogeneously fill this volume after a certain time period. A spontaneous gathering of the gas in one corner of the volume, however, has never been observed, although this would not contradict the law of energy conservation.

The entropy is the state quantity that uniquely characterizes this tendency. Processes which happen spontaneously and lead to equilibrium are connected with an increase in entropy. In equilibrium the entropy thus assumes its maximum value and does not change anymore. Ludwig Boltzmann was the first to show, in his famous *H-theorem* (pronounced Eta-Theorem, H =greek *Eta*) in 1872, that this tendency can also be founded on a statistical description and classical mechanics. In mathematical statistics one can uniquely assign to an arbitrary random event a measure for the predictability of this event. This function is commonly denoted by H and is called uncertainty function (see Chapter 6). Boltzmann was able to show that the uncertainty associated with an arbitrary nonequilibrium velocity distribution can only grow or at least remain equal as a function of time. The Maxwell-Boltzmann velocity distribution (equilibrium distribution) is characterized by a maximum of the uncertainty function. This means that for the Maxwell-Boltzmann distribution the prediction of the momentum of a particle in a gas at given temperature is associated with the largest uncertainty. Analogously, a homogeneous distribution of the particles in coordinate space is associated with the largest uncertainty with respect to the prediction of the coordinates. Other distribution functions (e.g., all particles gathered in a corner of the box) allow for a better prediction. In Chapter 6 the close relationship between the uncertainty function and the entropy is more extensively discussed (both are simply proportional to each other).

An important consequence of the H -theorem is that an arbitrary (nonequilibrium) distribution of particles changes, after a sufficiently long period of time, into the Maxwell-Boltzmann velocity distribution presented in Example 1.2 and that the latter is the only possible equilibrium distribution (cf. also *Quantum Theory: Special Problems*, Exercise 6.5). Many renowned physicists of Boltzmann's day were of the opinion that this seemed to contradict invariance with respect to time reversal, which is a known principle of classical mechanics. This principle guarantees that if one reverses the momenta of all particles in the final state, the process should as well be reversed. In other words, if the diffusion of a gas into a larger volume can be understood in purely mechanical terms, the reverse process, namely the gathering of the gas in a small part of the container, should also be possible, if all momenta of the final state are reversed. This seeming contradiction is resolved by statistical investigation of the problem.

In classical mechanics a state of motion of N particles is uniquely determined by the $3N$ coordinates and $3N$ momenta (q_v, p_v) of the N particles at a certain time. The set (q_v, p_v) is also called the *microstate* of the system, which of course varies with time. Each microstate (q_v, p_v) corresponds to *one* point in a $6N$ -dimensional space, the *phase space*. The set (q_v, p_v) i.e., the microstate, can therefore be identified with a point in phase space. Let us now consider the diffusion of a gas in an initial state $(q_v(t_0), p_v(t_0))$ from a smaller into a larger volume. If one is really able to reverse all momenta in the final state $(q_v(t_f), p_v(t_f))$ and to prepare a state $(q_v(t_f), -p_v(t_f))$, the process would in fact be reversed. From a statistical point of view, however, this is an event with an incredibly small probability. For there is only one point (microstate) in phase space which leads to an exact reversal of the process, namely $(q_v(t_f), -p_v(t_f))$. The great majority of microstates belonging to a certain macrostate, however, lead under time reversal to states which cannot be distinguished macroscopically from the final state (i.e., the equilibrium or Maxwell-Boltzmann distribution). The fundamental assumption of statistical mechanics now is that *all* microstates which have the same total energy can be found with equal probability. This, however, means that the microstate $(q_v, -p_v)$ is only one among very many other microstates which all appear with the same probability.

Let us illustrate this. Although the microstate of the system drastically changes with time also, in macroscopic equilibrium we nevertheless always observe the same macroscopic state quantities such as pressure, temperature, volume, etc. This means that many microstates lead to the same macroscopic state quantities. We will denote by Ω the number of these microstates "compatible" with a given macrostate. However, we can readily convince ourselves that the number of available microstates for a gas which homogeneously occupies a volume V is overwhelmingly larger than the number of available states compatible with a smaller volume. If we characterize the macrostate by the volume V available to our N particles, the number of microstates $\Omega(V)$ available to *one* particle is proportional to V . In fact, a single particle may assume all possible momenta $-\infty < p_v < \infty$ and all coordinates q_v in the volume V . However, the number of coordinate cells is just proportional to the volume V . For N particles which are independent from each other, the numbers of microstates available to each particle have to be multiplied, so that $\Omega(V) \propto V^N$. If we now compare $\Omega(V)$ with the number of microstates in a volume half the size, then $\Omega(V/2)$ is proportional to $(V/2)^N$. Thus, this case is less probable by a factor $(1/2)^N$ than

the case where the gas would homogeneously occupy the whole volume. However, for a macroscopic number of particles, e.g., of the order of Avogadro's number $N_A \approx 10^{23}$, this is an incredibly small number: $(1/2)^{10^{23}}$.

Thus, we can interpret the statement derived from the second law—that the gathering of gas particles in one half of the container is impossible—in statistical terms and assert that it may be possible, but unimaginably improbable. In particular, we thus realize that thermodynamics is a special case of statistical mechanics for very many particles ($N \rightarrow \infty$), since then we have $(1/2)^N \rightarrow 0$, and the gathering of all particles in one half of the container has zero probability. Note that this holds also for any other arbitrary partial volume of the container, since $q^N \rightarrow 0$ for $q < 1$ and $N \rightarrow \infty$.

Despite this simple explanation of the seeming contradiction between classical mechanics and the H -theorem, Boltzmann had to fight against sharp criticism of his ideas during his lifetime. Besides the criticism concerning the time reversal invariance of classical mechanics, presented mainly by Loschmidt, there were objections by Zermelo, who emphasized the quasiperiodical behavior of isolated systems (Poincaré cycles).

Objections were raised by Ostwald and Mach, who were pure “thermodynamicists,” even against the foundation of Boltzmann's theory, the molecular or atomistic kinetic theory of gases. This last criticism was evidently defeated by Einstein's and Smoluchowski's investigations of brownian motion in 1905.

As we have already realized with the help of a simple consideration, the number of microstates Ω compatible with a given macrostate is a quantity very similar to the entropy of this macrostate. The larger Ω , the more probable is the corresponding macrostate, and *the macrostate with the largest number Ω_{\max} of possible microscopic realizations corresponds to thermodynamic equilibrium*. Hence, if the particles of a system have arbitrary given coordinates and momenta in the beginning (e.g., a gas which gathers in one corner of a box), (q_v, p_v) assume some other values in the course of time, i.e., another possible microstate. All of the many possible microstates correspond to only one macrostate, namely the equilibrium state: thus, after a long period of time, one observes only the equilibrium state, with a probability that is very near to unity. In particular, in this statistical point of view deviations from the equilibrium state are, for a finite number of particles, not impossible (as thermodynamics claims), but only extremely improbable. Especially for very small systems or systems in extreme situations (e.g., gases at the critical point) one may indeed observe deviations from the equilibrium state in the form of *fluctuations* which have, of course, appreciable size only under certain conditions.

We are now able to go one step further and try to establish a connection between the entropy and the number of microstates compatible with a macrostate. For two statistically independent systems the total number of compatible microstates Ω_{tot} is obviously the product of the numbers for the individual systems, namely $\Omega_{\text{tot}} = \Omega_1 \Omega_2$. We have seen that entropy is an extensive state quantity which is simply added for both partial systems: $S_{\text{tot}} = S_1 + S_2$. Furthermore, according to our conclusions after Equation (2.37), the entropy S has to assume a maximum for the equilibrium state, just as the number of compatible microstates Ω assumes a maximum in equilibrium. If we now assume that there is a one-to-one correspondence between entropy and Ω , for instance $S = f(\Omega)$, there is only one mathematical function which simultaneously fulfills $S_{\text{tot}} = S_1 + S_2$ and $\Omega_{\text{tot}} = \Omega_1 \Omega_2$:

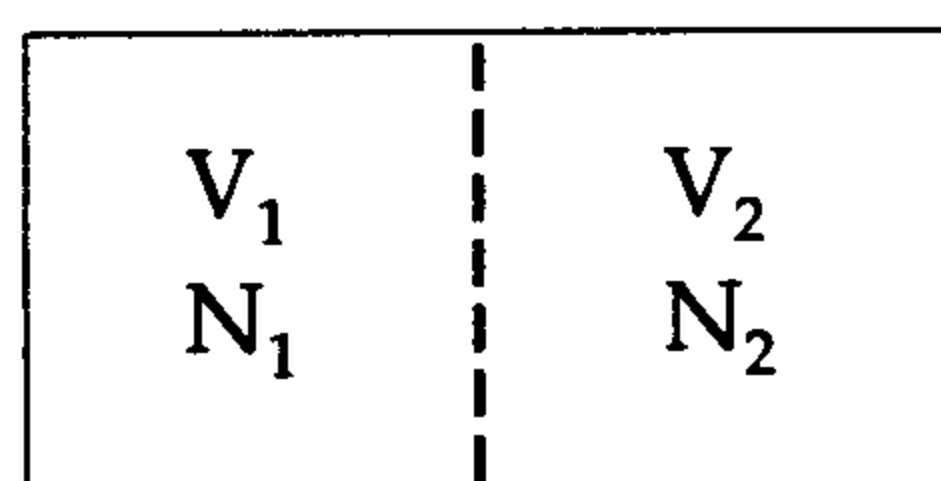
the logarithm. Therefore it must hold that $S \propto \ln \Omega$, and this is in fact the fundamental relationship between thermodynamics and statistical mechanics, which we will extensively discuss in the course of this book. In particular, we recognize that the equilibrium state of thermodynamics with a maximum in entropy is the state with the largest number of microscopic possibilities of realizing this state.

With the help of a simple example we can show that even small deviations from the equilibrium state reduce the number Ω of compatible microstates so drastically that they may be observed in nature only as small fluctuations around the equilibrium state.

Example 2.4: Microstates in a simple system

Again we consider a container with volume V , which is homogeneously filled with N particles of a gas in equilibrium. We now imagine the container to be divided into two compartments with volumes V_1 and V_2 , where $V_1 + V_2 = V$, and with N_1 and N_2 particles ($N_1 + N_2 = N$), respectively. We set $V_1 = pV$ and $V_2 = qV$. For the fractions p and q of the total volume we of course have $p + q = 1$.

The total number of microstates compatible with N particles and the total volume V has to be proportional to V^N , as we have already seen. According to the binomial theorem we have



$$\Omega_{\text{tot}}(N, V) \propto V^N = (V_1 + V_2)^N = \sum_{N_1=0}^N \binom{N}{N_1} V_1^{N_1} V_2^{N-N_1} \quad (2.41)$$

Figure 2.4. Model system.

where $N - N_1 = N_2$. On the other hand,

$$\Omega_1(N_1, V_1) \propto V_1^{N_1}$$

and

$$\Omega_2(N_2, V_2) \propto V_2^{N_2}$$

are just the numbers of microstates compatible with the partial volumes V_1 and V_2 and the particle numbers N_1 and N_2 , respectively. Equation (2.41) therefore attains a very illustrative meaning. Let us consider a macrostate with N_1 particles in the partial volume V_1 . Obviously there are exactly $\binom{N}{N_1}$ possibilities of selecting these particles from a total of N *enumerated* particles. The number of corresponding microstates is proportional to $V_1^{N_1}$, while for the remaining $N_2 = N - N_1$ particles in the volume V_2 there are still $V_2^{N_2}$ microstates which can be assumed. Each term in the sum is therefore the total number of microstates, if there are N_1 particles in the volume V_1 . One now obtains the total number of *all* possible microstates by summing over all possible particle numbers in V_1 . Therefore, we may interpret the expression

$$\Omega(V_1, V_2, K, N) = \binom{N}{K} V_1^K V_2^{N-K}$$

as the number of microstates of the situation where K particles are in V_1 and $N - K$ particles in V_2 . Then we can furthermore interpret

$$p_K = \frac{\Omega(V_1, V_2, N, K)}{\Omega_{\text{tot}}(V, N)} = \frac{1}{V^N} \binom{N}{K} (pV)^K (qV)^{N-K} = \binom{N}{K} p^K q^{N-K} \quad (2.42)$$

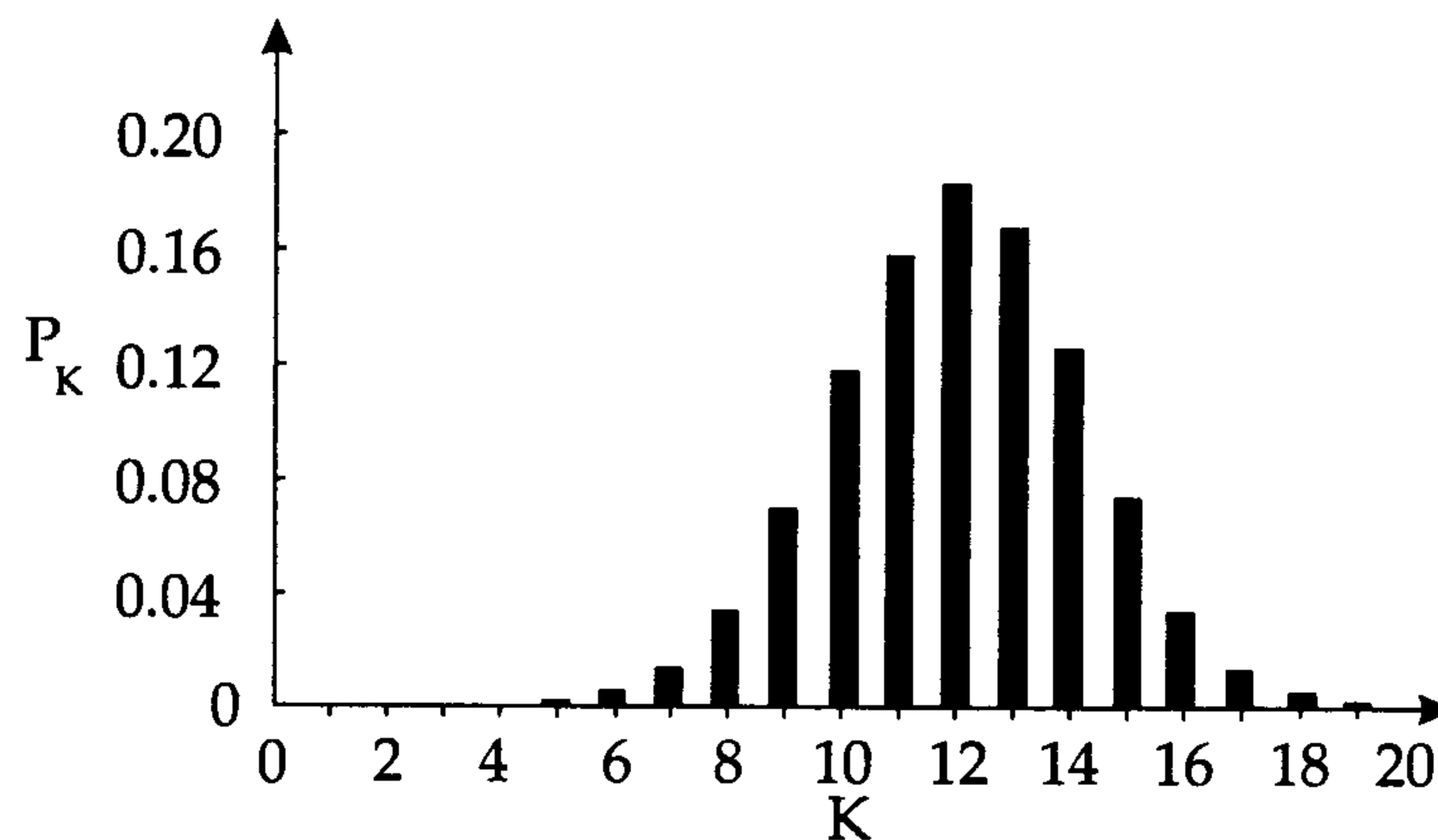


Figure 2.5. Probability of finding K particles in the volume $V_1 = 0.6V$. The total number of particles is $N = 20$.

as the probability that there are just K particles in the fictitious volume V_1 and $N - K$ particles in V_2 . With Equation (2.42) we can now immediately find the average particle number in volume V_1 . This number is by definition

$$\bar{K} \equiv \sum_{K=0}^N p_K K = \sum_{K=0}^N \binom{N}{K} K p^K q^{N-K} \quad (2.43)$$

We now apply a trick often used in the calculation of such average values and write $K p^K$ as $p \frac{\partial}{\partial p} (p^K)$, so that

$$\bar{K} = p \frac{\partial}{\partial p} \sum_{K=0}^N \binom{N}{K} p^K q^{N-K} = p \frac{\partial}{\partial p} (p + q)^N = N p (p + q)^{N-1}$$

Since $p + q = 1$ we have $\bar{K} = Np$, or $\bar{K}/N = V_1/V$. On the average, there will be just as many particles in the fictitious volume V_1 as corresponds to the fraction V_1/V . This is of course obvious, since equilibrium corresponds to a homogeneous distribution of particles. Furthermore, we can even calculate the *fluctuations around this value* and ask for the probability of finding a value that deviates from $\bar{K} = Np$. To this end we simply plot the probabilities p_K as a function of K (Figure 2.5). The maximum in the figure just corresponds to the average ($K_{\max} = \bar{K}$).

For our example, with a total of $N = 20$ particles and a volume $V_1 = 0.6V$ it is therefore still quite probable to find 11 or 13 particles in V_1 instead of $\bar{K} = 12$ particles. A definite measure for the size of these deviations from the average value is the mean square deviation. This is defined as

$$\overline{(\Delta K)^2} \equiv \overline{(K - \bar{K})^2} = \sum_{K=0}^N p_K (K - \bar{K})^2 = \sum_{K=0}^N p_K K^2 - \bar{K}^2 \quad (2.44)$$

since

$$\sum_{K=0}^N 2p_K K \bar{K} = 2\bar{K}^2$$

and

$$\sum_{K=0}^N p_K = 1$$

The calculation of Equation (2.44) proceeds as in Equation (2.43), but now we have to put

$$K^2 p^K = \left(p \frac{\partial}{\partial p}\right)^2 p^K$$

and we have

$$\begin{aligned} \overline{(\Delta K)^2} &= \left(p \frac{\partial}{\partial p}\right)^2 (p+q)^N - (pN)^2 \\ &= p \frac{\partial}{\partial p} (pN(p+q)^{N-1}) - (pN)^2 \\ &= p \left[N(p+q)^{N-1} \right. \\ &\quad \left. + pN(N-1)(p+q)^{N-2} \right] - (pN)^2 \end{aligned}$$

If we again insert $p+q=1$, it follows that

$$\overline{(\Delta K)^2} = pN + p^2 N(N-1) - (pN)^2 = pN - p^2 N = Np(1-p) = Npq$$

This means that the width of the distribution, measured via

$$\Delta^* K = \sqrt{\overline{(\Delta K)^2}}$$

increases with \sqrt{N} . The *relative width*, i.e., the width $\Delta^* K$ with respect to the average particle number in the volume V_1 , is then

$$\frac{\Delta^* K}{K} = \frac{\sqrt{Npq}}{Np} = \sqrt{\frac{q}{p}} \frac{1}{\sqrt{N}}$$

This represents the fluctuation relative to the (mean) particle number. Here we clearly see that the relative deviation from the equilibrium distribution decreases as $N^{-1/2}$ and is very small for macroscopic particle numbers $N_A \approx 10^{24} \Rightarrow N_A^{-1/2} \approx 10^{-12}$. Hence macroscopic fluctuations (e.g., all particles suddenly in V_1) are extremely improbable. However, small deviations in small spatial regions are quite natural. If we choose a very small partial volume ($p \rightarrow 0$), fluctuations can nevertheless be quite large. This corresponds, for example, to local density fluctuations in gases, which are, however, of appreciable size only on a microscopic scale (10^{-8} – 10^{-10} m). As already mentioned, these fluctuations can become observable in special situations. For instance, the density fluctuations of a gas in the vicinity of the critical point are especially large. They can be observed by a change in the scattering of light in the gas (critical opalescence). It is very difficult to understand this phenomenon without the help of the statistical point of view.

Now we want to critically review the arguments leading to the important formula $S \propto \ln \Omega$. We will see that two fundamental problems arise, which up to now we implicitly have taken for granted. We started from the example of a gas gathered in a part of the volume of a container, with values $(q_v(t_0), p_v(t_0))$ for the coordinates and momenta. We then assumed that in the course of time the coordinates and momenta statistically attain some

values and that we therefore find, by counting *all* possible values (Ω), the macrostate with Ω_{\max} as the equilibrium state. Exactly at this point, however, we assumed that all microstates (q_v, p_v) compatible with our macrostate have “equal rights,” i.e., equal probability. In an isolated system the macrostate is characterized by the total energy, the particle number and the volume of the system. We determine Ω by counting all microstates compatible with these values of E , N and V . It may of course be possible that these microstates do not have equal probabilities. Let us again consider an example:

Example 2.5: Concerning the equal probability of all microstates

We imagine a box-shaped container and initialize all gas particles at time t_0 in the vicinity of the point P with *parallel* momenta (Figure 2.6). Obviously, this is a microstate which is far from equilibrium. However, if our system behaves ideally, it will *never* reach equilibrium! For if each particle is reflected at the walls in an ideal way, the trajectories will forever stay parallel, and certain regions of coordinate space will never be reached.

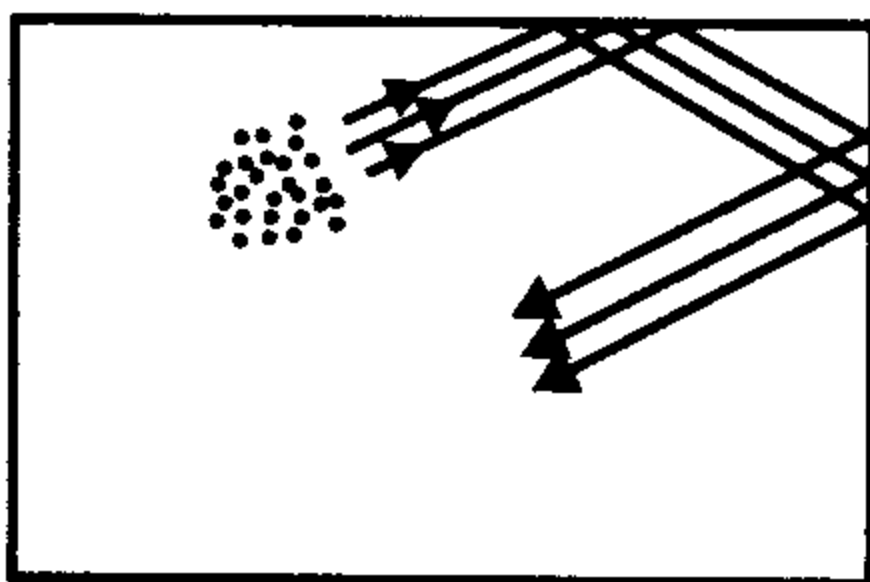


Figure 2.6. Model system.

In addition, the velocities of the particles will change (for pointlike particles without mutual influence) only with respect to the direction, but not with respect to their absolute values. Thus, the Maxwell-Boltzmann distribution will never be reached. This, however, means that we may not assume that all microstates for this system have equal probability. A more detailed consideration shows us that this system is much too idealized. The fundamental objection is that, because of the quantum mechanical uncertainty relation $\Delta x \Delta p > \hbar$, one cannot *in principle* construct such an initial state, so that the particles move randomly in all directions even after a short period of time.

But even in the framework of classical mechanics, where such an initial state is possible, there are actually no ideally reflecting walls and pointparticles without mutual interactions. Even the microscopic structure of a real wall is sufficient to distribute the particles, after a short time, homogeneously over the whole container (not to mention the thermal motion of the particles of the wall and the mutual collisions of the gas particles). Moreover, even the initial conditions $(q_v(t_0), p_v(t_0))$ can only be determined up to a certain measurement uncertainty, which also leads to diffraction of the beam.

Anyhow, we see that the equal probability of all compatible microstates is an assumption which can only be justified by an experimental examination of the consequences. *Up to now, no one has invented an experiment which disproves this assumption.* We are now likewise able to illustrate another fundamental problem of statistical mechanics with the help of our example. We have stated that a system which is initialized at time t_0 in a nonequilibrium state assumes, with high probability after a certain relaxation time, only microstates $(q_v(t), p_v(t))$ which belong to the macroscopic equilibrium state, since their number is by far larger than the number of all other microstates. However, here we presuppose that in the course of time the system comes at least very close to each microstate (q_v, p_v) *which is compatible with the equilibrium state*. For if certain microstates are never reached (not even approximately),

we actually must not count them in the calculation of Ω ! We return to this problem in Part II of this book.

In summary we may therefore say: for every macroscopic thermodynamic state there exists a large number of microscopic realization possibilities. The fact that an isolated system assumes an equilibrium state after a sufficiently long period of time can be traced back to the fact that more microstates belong to the equilibrium state than to a nonequilibrium state. Therefore we may understand the effort of a system to reach a maximum in entropy also as a transition into the most probable state, i.e., into the state with the largest number of microscopic realization possibilities. In particular, the entropy is thus a measure for the number of possible microstates of a system when a macrostate is given.

In a manner similar to the first law we can now formulate also the second law of thermodynamics in different ways.

With respect to the energy law we may for example say:

U_1	U_2
S_1, T_1	S_2, T_2
V_1, p_1	V_2, p_2
N_1, μ_1	N_2, μ_2
...	...

Figure 2.7. The considered system.

a) There is no perpetuum mobile of the second kind. A perpetuum mobile of the second kind is an engine which does nothing but perform work while cooling a heat bath. Thus it is an engine which transforms heat into work with 100% efficiency.

From the microscopic point of view the following formulation is especially instructive:

b) Each isolated macroscopic system wants to assume the most probable state, i.e., the state which is characterized by the largest number of microscopic realization possibilities.

Now we want to use both laws to derive some consequences for the state variables T, p, μ, ϕ, \dots in an isolated system in equilibrium.

To this end we imagine the completely isolated system to be divided into two parts. Let the total system be characterized by the state variables S, V , and N , where the intrinsic energy U is a function of these variables. *Since the total system is isolated, all these state variables are constant.* There is no exchange of work or heat with the surroundings. Both partial systems, however, can mutually exchange all forms of work or heat.

Therefore the state variables U_i, S_i, V_i and N_i with $i = 1, 2$ do not have constant values; however, it must be the case that

$$\begin{aligned} U_1 + U_2 &= U = \text{const.} & S_1 + S_2 &= S = \text{const.} \\ V_1 + V_2 &= V = \text{const.} & N_1 + N_2 &= N = \text{const.} \\ && \dots & \end{aligned} \tag{2.45}$$

Let us now remember the first law for a reversible change of state for both partial systems:

$$\begin{aligned} dU_1 &= T_1 dS_1 - p_1 dV_1 + \mu_1 dN_1 + \dots \\ dU_2 &= T_2 dS_2 - p_2 dV_2 + \mu_2 dN_2 + \dots \end{aligned} \tag{2.46}$$

Here T_i, p_i , and μ_i are the temperatures, pressures and chemical potentials in the two partial systems. Because of Equation (2.45) we now have $dU_1 + dU_2 = 0$. If we therefore add both Equations (2.46), it follows, with $dS_1 = -dS_2, dV_1 = -dV_2, \dots$, that

$$0 = (T_1 - T_2) dS_1 - (p_1 - p_2) dV_1 + (\mu_1 - \mu_2) dN_1 + \dots \tag{2.47}$$

Since the change of the variables S_1, V_1, N_1, \dots in system 1 underlies no restrictions, Equation (2.47) is true only if it separately holds that

$$T_1 = T_2, \quad p_1 = p_2, \quad \mu_1 = \mu_2, \quad \dots \quad (2.48)$$

These are the necessary conditions for thermodynamic equilibrium. Since the imagined partition of the isolated system was arbitrarily chosen, we can conclude that *if an isolated system is in equilibrium, it has everywhere the same constant temperature, the same pressure and the same chemical potential, etc.* If there is, however, a real wall instead of the fictitious wall separating the partial systems which, for instance, prohibits a change of volume or of the particle number — $dN_1 = 0, dV_1 = 0$ — then only the condition

$$T_1 = T_2 \quad (2.49)$$

remains. Correspondingly, the conditions (2.48) hold also separately or in combination, if the partition is permeable only for certain changes of the state variables. One speaks of thermal, mechanical, chemical, etc., equilibria. If the isolated total system is not in equilibrium, e.g., $T_1 \neq T_2$ and $p_1 \neq p_2$, the relaxation times for reaching equilibrium are in general different for the different variables T, p, μ, \dots ! It may therefore happen that a system is already to good approximation in thermal equilibrium ($T_1 = T_2$), but not in chemical equilibrium. In most cases, pressure balance occurs most quickly, i.e., mechanical equilibrium is followed by thermal equilibrium. Establishing chemical equilibrium may take very long in some cases. This depends on the speed of the chemical reactions which cause a change dN_i in the particle number.

Global and local equilibrium

If a system is in thermodynamic equilibrium, i.e., if it has everywhere the same temperature, the same pressure and the same chemical potential, one speaks of global equilibrium. However, thermodynamic notions (intensive state quantities) are not restricted to such systems in their applicability. If one can divide the whole system into small partial systems which still contain very many particles and which are individually approximately in thermodynamic equilibrium, these partial systems also can be described by thermodynamic state quantities. However, these quantities will vary from partial system to partial system. The differences in temperature, pressure, and chemical potential affect heat flow (from hotter to cooler parts), volume changes (regions of high pressure expand at the expense of regions with low pressure), and particle fluxes. These fluxes are driven by the corresponding potential differences and cause a compensation of these potentials which leads, for an isolated system, to global equilibrium during the course of time. If the total system can be divided into such parts one speaks of local equilibrium. However, it is of crucial importance that the partial systems can be chosen large enough (having sufficiently many particles) for a statistical description to be reasonable, but on the other hand that in each partial system the (intensive) thermodynamic state quantities assume definite constant values and do not vary too strongly from one partial system to another (small gradients). Then the relationships derived in this book for total systems can be often applied locally in a system which is