3.3: Canonical Ensemble

Equilibrium thermodynamics describes systems that are in thermal equilibrium. In an ensemble picture, this can be considered by assuming that the system is in contact with a very large— for mathematical purposes infinitely large— heat bath. Because of this, the individual systems in the ensemble can differ in energy. However, the probability density distribution in phase space or state space must be consistent with constant temperature T, which is the temperature of the heat bath. In experiments, it is the temperature of the environment.

> **♻ Concept 3.3.1: Canonical Ensemble**
>
> An ensemble with a constant number N of particles in a constant volume V and at thermal equilibrium with a heat bath at constant temperature T can be considered as an ensemble of microcanonical subensembles with different energies ϵ_i. The energy dependence of probability density conforms to the Boltzmann distribution. Such an ensemble is called a *canonical ensemble*.

> **📌 Note**
>
> Because each system can exchange heat with the bath and thus change its energy, systems will transfer between subensembles during evolution. This does not invalidate the idea of microcanonical subensembles with constant particle numbers N_i. For a sufficiently large ensemble at thermal equilibrium the N_i are constants of motion.

There are different ways of deriving the Boltzmann distribution. Most of them are rather abstract and rely on a large mathematical apparatus. The derivation gets lengthy if one wants to create the illusion that we know why the constant β introduced below *always* equals $1/k_\mathrm{B}T$, where $k_\mathrm{B} = R/N_\mathrm{Av}$ is the Boltzmann constant, which in turn is the ratio of the universal gas constant R and the Avogadro constant N_Av. Here we follow a derivation that is physically transparent and relies on a minimum of mathematical apparatus that we have already introduced.

Boltzmann Distribution

Here we digress from the ensemble picture and use a system of N particles that may exist in r different states with energies ϵ_i with $i = 0 \ldots r-1$. The number of particles with energy ϵ_i is N_i. The particles do not interact, they are completely independent from each other. We could therefore associate theses particles with microcanonical subensembles of a canonical ensemble, but the situation is easier to picture with particles. The probability $P_i = N_i/N$ to find a particle with energy ϵ_i can be associated with the probability density for the microcanonical subensemble at energy ϵ_i. The difference between this simple derivation and the more elaborate derivation for a canonical ensemble is thus essentially the difference between discrete and continuous probability theory. We further assume that the particles are classical particles and thus distinguishable.

To compute the probability distribution $P_i = N_i/N$, we note that

$$\sum_{0}^{r-1} N_i = N \tag{3.3.1}$$

and

$$\sum_{0}^{r-1} N_i \epsilon_i = E , \tag{3.3.2}$$

where E is a constant total energy of the system. We need to be careful in interpreting the latter equation in the ensemble picture. The quantity E corresponds to the energy of the whole canonical ensemble, which is indeed a constant of motion, if we consider a sufficiently large number of systems in contact with a thermal bath. We can thus use our simple model of N particles for guessing the probability density distribution in the canonical ensemble.

What we are looking for is the most likely distribution of the N particles on the r energy levels. This is equivalent to putting N distinguishable balls into r boxes. We did already solve the problem of distributing N objects to 2 states when considering the binomial distribution in Section [binomial_distribution]. The statistical weight of a configuration with n objects in the first state and $N-n$ objects in the second state was $\binom{N}{n}$. With this information we would already be able to solve the problem of a canonical

ensemble of N spins $S = 1/2$ in thermal contact with the environment, disregarding for the moment differences between classical and quantum statistics (see Section [section:quantum_statistics]).

Coming back to N particles and r energy levels, we still have $N!$ permutations. If we assign the first N_0 particles to the state with energy ϵ_0, the next N_1 particles to ϵ_1 and so on, we need to divide each time by the number of permutations $N_i!$ in the same energy state, because the sequence of particles with the same energy does not matter. We call the vector of the occupation numbers N_i a **configuration**. The configuration specifies one particular macrostate of the system and the relative probability of the macrostates for distinguishable particles and non-degenerate states is given by their statistical weights,

$$\Omega = \frac{N!}{N_0! N_1! \ldots N_{r-1}!} \ . \tag{3.3.3}$$

The case with degenerate energy levels is treated in Section [sec:Maxwell-Boltzmann].

The most probable macrostate is the one with maximum statistical weight Ω. Because of the peaking of probability distributions for large N, we need to compute only this most probable macrostate; it is representative for the whole ensemble. Instead of maximizing Ω we can as well maximize $\ln\Omega$, as the natural logarithm is a strictly monotonous function. This allows us to apply Stirling's formula,

$$\ln\Omega = \ln N! - \sum_{i=0}^{r-1} \ln N_i! \tag{3.3.4}$$

$$\approx N \ln N - N + 1 - \sum_{i=0}^{r-1} N_i \ln N_i + \sum_{0}^{r-1} N_i - r \ . \tag{3.3.5}$$

By inserting Equation 3.3.1 we find

$$\ln\Omega \approx N \ln N - \sum_{i=0}^{r-1} N_i \ln N_i + 1 - r \ . \tag{3.3.6}$$

Note that the second term on the right-hand side of Equation 3.3.6 has some similarity to the entropy of mixing, which suggests that $\ln\Omega$ is related to entropy.

At the maximum of $\ln\Omega$ the derivative of $\ln\Omega$ with respect to the N_i must vanish,

$$0 = \delta \sum_i N_i \ln N_i = \sum_i \left(N_i \delta \ln N_i + \delta N_i \ln N_i \right) = \sum_i \delta N_i + \sum_i \ln N_i \delta N_i \ . \tag{3.3.7}$$

In addition, we need to consider the boundary conditions of constant particle number, Equation 3.3.1,

$$\delta N = \sum_i \delta N_i = 0 \tag{3.3.8}$$

and constant total energy, Equation 3.3.2,

$$\delta E = \sum_i \epsilon_i \delta N_i = 0 \ . \tag{3.3.9}$$

It might appear that Equation 3.3.8 could be used to cancel a term in Equation 3.3.7, but this would be wrong as Equation 3.3.8 is a constraint that must be fulfilled separately. For the constrained maximization we can use the method of Lagrange multipliers.

The maximum or minimum of a function $f(x_1 \ldots, x_n)$ of n variables is a stationary point that is attained at

$$\delta f = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} \right)_{x_k \neq x_i} \partial x_i = 0 \ . \tag{3.3.10}$$

We now consider the case where the possible sets of the n variables are constrained by c additional equations

$$g_j(x_1, x_2, \ldots, x_n) = 0 \ , \tag{3.3.11}$$

where index j runs over the c constraints ($j = 1 \ldots c$). Each constraint introduces another equation of the same form as the one of Equation 3.3.10,

$$\delta g_j = \sum_{i=1}^{n} \left(\frac{\partial g_j}{\partial x_i} \right)_{x_k \neq x_i} \partial x_i = 0 \ . \tag{3.3.12}$$

The constraints can be introduced by multiplying each of the c equations by a multiplier λ_j and subtracting it from the equation for the stationary point without the constraints,

$$\delta \mathcal{L} = \sum_{i=1}^{n} \left[\left(\frac{\partial f}{\partial x_i} \right)_{x_k \neq x_i} - \sum_{j=1}^{c} \lambda_j \left(\frac{\partial g_j}{\partial x_i} \right)_{x_k \neq x_i} \right] \partial x_i \ . \tag{3.3.13}$$

If a set of variables $\{x_{0,1} \ldots, x_{0,n}\}$ solves the constrained problem then there exists a set $\{\lambda_{0,1} \ldots \lambda_{0,r}\}$ for which $\{x_{0,1}, x_{0,2}, \ldots, x_{0,n}\}$ also corresponds to a stationary point of the Lagrangian function $\mathcal{L}(x_1, \ldots, x_n, \lambda_1, \ldots \lambda_r)$. Note that not all stationary points of the Lagrangian function are necessarily solutions of the constrained problem. This needs to be checked separately. [concept:Lagrangian_multipliers]

With this method, we can write

$$0 = \sum_i \delta N_i + \sum_i \ln N_i \delta N_i + \alpha \sum_i \delta N_i + \beta \sum_i \epsilon_i \delta N_i \tag{3.3.14}$$

$$= \sum_i \delta N_i \left(1 + \ln N_i + \alpha + \beta \epsilon_i \right) \ . \tag{3.3.15}$$

The two boundary conditions fix only two of the population numbers N_i. We can choose the multipliers α and β in a way that $(1 + \ln N_i + \alpha + \beta \epsilon_i) = 0$ for these two N_i, which ensures that the partial derivatives of $\ln \Omega$ with respect to these two N_i vanishes. The other $r - 2$ population numbers can, in principle, be chosen freely, but again we must have

$$1 + \ln N_i + \alpha + \beta \epsilon_i = 0 \tag{3.3.16}$$

for all i to make sure that we find a maximum with respect to variation of any of the r population numbers. This gives

$$N_i = \gamma e^{-\beta \epsilon_i} \tag{3.3.17}$$

with $\gamma = e^{-(1+\alpha)}$. We can eliminate γ by using Equation 3.3.1,

$$\sum_i N_i = \gamma \sum_i e^{-\beta \epsilon_i} = N \ , \tag{3.3.18}$$

giving

$$\gamma = \frac{N}{\sum_i e^{-\beta \epsilon_i}} \ , \tag{3.3.19}$$

and finally leading to

$$P_i = \frac{N_i}{N} = \frac{e^{-\beta \epsilon_i}}{\sum_i e^{-\beta \epsilon_i}} \ . \tag{3.3.20}$$

For many problems in statistical thermodynamics, the Lagrange multiplier α is related to the chemical potential by $\alpha = \mu/(k_B T)$. The Lagrange multiplier β must have the reciprocal dimension of an energy, as the exponent must be dimensionless. As indicated above, we cannot at this stage prove that β is the same energy for all problems of the type that we have posed here, let alone for all of the analogous problems of canonical ensembles. The whole formalism can be connected to phenomenological thermodynamics *via* Maxwell's kinetic gas theory (see also Section [subsection:equipartition]). For this problem one finds

$$\beta = \frac{1}{k_B T} \ . \tag{3.3.21}$$

> ### ⚇ Concept 3.3.2: Boltzmann Distribution
>
> For a classical canonical ensemble with energy levels ϵ_i the probability distribution for the level populations is given by the *Boltzmann distribution*

$$P_i = \frac{N_i}{N} = \frac{e^{-\epsilon_i/k_B T}}{\sum_i e^{-\epsilon_i/k_B T}} \ . \tag{3.3.22}$$

The sum over states

$$Z(N, V, T) = \sum_i e^{-\epsilon_i/k_B T} \tag{3.3.23}$$

required for normalization is called *canonical partition function*.[10] The partition function is a thermodynamical state function.

For the partition function, we use the symbol Z relating to the German term *Zustandssumme*("sum over states"), which is a more lucid description of this quantity.

Equipartition Theorem

Comparison of Maxwell's kinetic theory of gases with the state equation of the ideal gas from phenomenological thermodynamics provides a mean kinetic energy of a point particle of $\langle \epsilon_{kin} \rangle = 3k_B T/2$. This energy corresponds to

$$\epsilon_{trans} = \frac{1}{2} m v^2 = \frac{1}{2m} p^2 \ , \tag{3.3.24}$$

i.e., it is quadratic in the velocity coordinates of dynamic space or the momentum coordinates of phase space. Translational energy is distributed via three degrees of freedom, as the velocities or momenta have components along three pairwise orthogonal directions in space. Each quadratic degree of freedom thus contributes a mean energy of $k_B T/2$.

If we accept that the Lagrange multiplier β assumes a value $1/k_B T$, we find a mean energy $k_B T$ of an harmonic oscillator in the high-temperature limit . Such an oscillator has two degrees of freedom that contribute quadratically in the degrees of freedom to energy,

$$\epsilon_{vib} = \frac{1}{2} \mu v^2 + \frac{1}{2} f x^2 \ , \tag{3.3.25}$$

where μ is the reduced mass and f the force constant. The first term contributes to kinetic energy, the second to potential energy. In the time average, each term contributes the same energy and assuming ergodicity this means that each of the two degrees of freedom contributes with $k_B T/2$ to the average energy of a system at thermal equilibrium.

The same exercise can be performed for rotational degrees of freedom with energy

$$\epsilon_{rot} = \frac{1}{2} I \omega^2 \ , \tag{3.3.26}$$

where I is angular momentum and ω angular frequency. Each rotational degree of freedom, being quadratic in ω again contributes a mean energy of $k_B T/2$.

Based on Equation 3.3.20 it can be shown that for an energy

$$\epsilon_i = \eta_0 + \eta_1 + \eta_2 + \ldots = \sum_{k=1}^{f} \eta_k \ , \tag{3.3.27}$$

where index k runs over the individual degrees of freedom, the number of molecules that contribute energy η_k does not depend on the terms η_j with $j \neq k$. It can be further shown that

$$\langle \eta_k \rangle = \frac{1}{2\beta} \tag{3.3.28}$$

for all terms that contribute quadratically to energy.[11]

This result has two consequences. First, we can generalize $\beta = 1/k_B T$, which we strictly knew only for translational degrees of freedom, to any canonical ensemble for which all individual energy contributions are quadratic along one dimension in phase space. Second, we can formulate the

Each degree of freedom, whose energy scales quadratically with one of the coordinates of state space, contributes a mean energy of $k_B T/2$.

The equipartition theorem applies to all degrees of freedom that are *activated*. Translational degrees of freedom are always activated and rotational degrees of freedom are activated at ambient temperature, which corresponds to the high-temperature limit of rotational dynamics. To vibrational degrees of freedom the equipartition theorem applies only in the high-temperature limit. In general, the equipartition theorem fails for quantized degrees of freedom if the quantum energy spacing is comparable to $k_\mathrm{B}T/2$ or exceeds this value. We shall come back to this point when discussing the vibrational partition function.

Internal Energy and Heat Capacity of the Canonical Ensemble

The internal energy u of a system consisting of N particles that are distributed to r energy levels can be identified as the total energy E of the system considered in Section ([subsection:Boltzmann]). Using Eqs. 3.3.2 and 3.3.22 we find

$$u = N \frac{\sum_i \epsilon_i e^{-\epsilon_i/k_\mathrm{B}T}}{\sum_i e^{-\epsilon_i/k_\mathrm{B}T}} = N \frac{\sum_i \epsilon_i e^{-\epsilon_i/k_\mathrm{B}T}}{Z} \ . \tag{3.3.29}$$

The sum in the numerator can be expressed by the partition function, since

$$\frac{\mathrm{d}Z}{\mathrm{d}T} = \frac{1}{k_\mathrm{B}T^2} \sum_i \epsilon_i e^{-\epsilon_i/k_\mathrm{B}T} \ . \tag{3.3.30}$$

Thus we obtain

$$u = Nk_\mathrm{B}T^2 \cdot \frac{1}{Z} \cdot \frac{\mathrm{d}Z}{\mathrm{d}T} = Nk_\mathrm{B}T^2 \frac{\mathrm{d}\ln Z}{\mathrm{d}T} \ . \tag{3.3.31}$$

Again the analogy of our simple system to the canonical ensemble holds. At this point we have computed one of the state functions of phenomenological thermodynamics from the set of energy levels. The derivation of the Boltzmann distribution has also indicated that $\ln \Omega$, and thus the partition function Z are probably related to entropy. We shall see in Section [section:state_fct_partition_fct] that this is indeed the case and that we can compute *all* thermodynamic state functions from Z.

Here we can still derive the heat capacity c_V at constant volume, which is the partial derivative of internal energy with respect to temperature. To that end we note that the partition function for the canonical ensemble relates to constant volume and constant number of particles.

$$c_V = \left(\frac{\partial u}{\partial T}\right)_V = N \frac{\partial}{\partial T}\left(k_\mathrm{B}T^2 \frac{\partial \ln Z}{\partial T}\right)_V = -N\frac{\partial}{\partial T}\left(k_\mathrm{B} \frac{\partial \ln Z}{\partial 1/T}\right)_V \tag{3.3.32}$$

$$= -Nk_\mathrm{B}\left(\frac{\partial\left[\partial \ln Z/\partial 1/T\right]}{\partial T}\right)_V = \frac{Nk_\mathrm{B}}{T^2}\left(\frac{\partial\left[\partial \ln Z/\partial 1/T\right]}{\partial 1/T}\right)_V \tag{3.3.33}$$

$$= \frac{k_\mathrm{B}}{T^2}\left(\frac{\partial^2 \ln z}{\partial(1/T)^2}\right)_V \ . \tag{3.3.34}$$

In the last line of Equation 3.3.34 we have substituted the molecular partition function Z by the partition function for the whole system, $\ln z = N \ln Z$. Note that this implies a generalization. Before, we were considering a system of N identical particles. Now we implicitly assume that Equation 3.3.34, as well as $u = k_\mathrm{B}T^2\frac{\mathrm{d}\ln z}{\mathrm{d}T}$ will hold for any system, as long as we correctly derive the system partition function z.

We note here that the canonical ensemble describes a closed system that can exchange heat with its environment, but by definition it cannot exchange work, because its volume V is constant. This does not present a problem, since the state functions can be computed at different V. In particular, pressure p can be computed from the partition function as well (see Section [section:state_fct_partition_fct]). However, because the canonical ensemble is closed, it cannot easily be applied to all problems that involve chemical reactions. For this we need to remove the restriction of a constant number of particles in the systems that make up the ensemble.