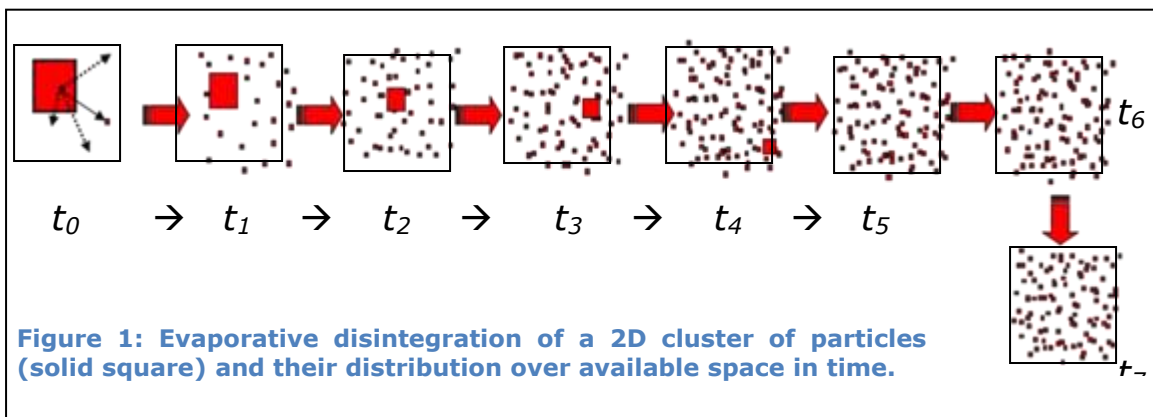


Statistical Ensembles and Partition Functions

In a previous section, for the behavior of either the Boltzmann “H-function” or the state variable “entropy,” a criterion was developed for estimating the extent of a process in its approach to asymptotic equilibrium. The entropy S turned out to be simply a measure of the spread of an ensemble of systems (e.g., 10^6 N -particle assemblies over a given spatial volume V , and with a fixed total energy) over all Ω system states available at the given fixed energy E . The effect is simply due to the number of equivalent pathways: As exemplified in Fig. 2, there are multitudes more pathways further away from the initial configuration compared to just a few trajectories back to the origin. And therefore, while there is a non-zero probability for a complex system to return to its starting point in time, such a reversal will essentially never be observed. This is true in particular, when the initial configuration of the complex is highly organized, i.e., is very different from a random association, and a larger number of more complex states are accessible to the system.



As an illustration, Fig. 1 depicts schematically the time dependent disintegration of a cluster of two-dimensional particles populating the states available in a planar box of a fixed volume (area). Initially, at $t=t_0$ the N particles are spatially highly correlated and form a compact cluster. The cluster is assumed to be unstable and can emit ("evaporate") successively its constituent particles, which populate the previously empty locations (single-particle *s.p.* states) in the planar box. All empty *s.p.* states are accessible to all particles. The progress of the disintegration is followed by the snapshots of the N -particle configuration taken at successive time intervals t_n . In the sequence shown in Fig. 1, a cluster remnant is still discernible as a small square up to $t = t_4$. For $t > t_4$, it is no longer in existence. The configurations for t_5, t_6, t_7 are very similar in that they correspond to slightly different distributions of the N particles over the box of volume V . For these 3 configurations, a time evolution is no longer obvious. The configurations all look random and are essentially equivalent realizations of the final asymptotic equilibrium distribution of the cluster particles. **The configurations for t_5, t_6, t_7 are different microstates of the N -particle system in equilibrium. They and all equivalent microstates give rise to the same equilibrium macrostate of the system.** Since the particles can move, the system will change its appearance from one of the Ω_{NVE} microstates to another. No microstate is excluded, since they are all equivalent, i.e., all have the same probability

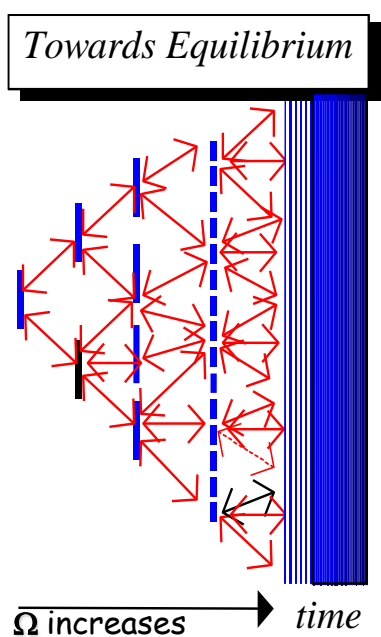
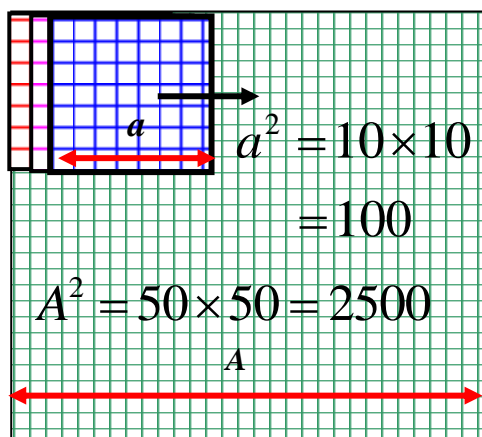


Figure 2: Illustration of the time evolution of an unbiased transport process towards regions of higher state densities (Ω).

since they are all equivalent, i.e., all have the same probability

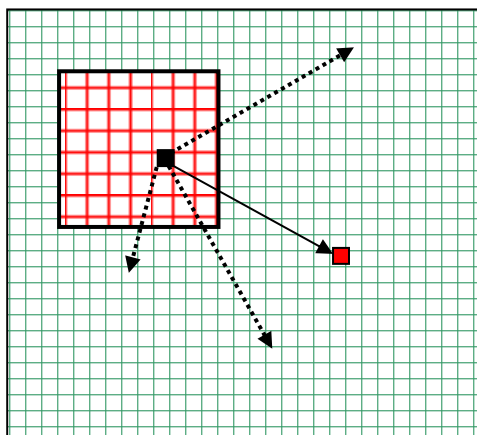
$$P(\{x_i, y_i | i=1, \dots, N\}) \propto \frac{1}{\Omega_{NVE}} \quad (1)$$

In order to appreciate the very different magnitudes of the various types of states of interest it is useful to examine a numerical example for the above 2D cluster decay.



Assume a 2D box of $50 \times 50 = 2500$ x-y bins and $A=100$ particles on this grid. As shown in Fig. 3 (top), in the initial configuration, the particles form a compact 2D cluster of square area of

$$a^2 = 10 \times 10 \quad (2)$$



The number of states $\Omega_{cluster}$ of the cluster in the box is given by the ways the cluster square can be fit into the box by sliding it into the different possible positions,

$$\Omega_{cluster} = (N - a)^2 = 40^2 = 1.6 \cdot 10^3 \quad (3)$$

Figure 3: 100-particle cluster configurations (100 intrinsic states): square of side length $a = 10$ bin widths in a volume of 2500 s.p. states.

From this, one can now easily calculate the number of states of the configurations where the cluster has evaporated one of its 100 particles, as illustrated at the bottom panel of Fig.3. The evaporated particle has access to $(2500 - 100) = 2400$ open s.p. states (bins), since (approximately) only 100 are blocked by the cluster. Therefore, the number of states for such a configuration is,

$$\Omega_{cluster-1p,1p} = \left[100 \cdot \Omega_{cluster-1p} \right] \times \Omega_{1p} = 1.6 \cdot 10^5 \times 2.4 \cdot 10^3 = 3.84 \cdot 10^8 \quad (4)$$

where the factor of 100 represents the number of choices for the evaporated particle. Therefore, there are already 2.4×10^5 as many states (configurations) possible with one evaporated particle than are accessible to the original cluster. For a completely dissociated cluster, where all particles are randomly distributed over the 2500 bins, this number increases to the enormous

$$\Omega_{NVE} = \Omega_{chaotic} = \binom{N^2}{N} = \frac{N^2!}{N!(N^2 - N)!} = \frac{2500!}{100!2400!} \approx 10^{182} \quad (5)$$

of states (arrangements) of the 100 -particle system in the two dimensional box. The probability for any of these equivalent microstates to be occupied in equilibrium m is then,

$$f_i = \frac{1}{\Omega_{NVE}} \quad \{i = 1, 2, \dots, \Omega_{NVE}\} \quad (6)$$

In a **series of repeated measurements of A similarly prepared systems**, these quantities would experimentally be obtained from the number (A_i) of ensembles ending up in microstate state i , via

$$f_i = A_i / A \quad (7)$$

Conversely, the equilibrium macrostate defined as a superposition of all Ω_{NVE} microstates occurs with a yet unnormalized probability of

$$P \approx \Omega_{NVE} = e^{S(\Omega_{NVE})/k_B} \quad (8)$$

which is a very large number compared to $P \approx \Omega_{cluster}$ given by Equ. 4. The total probability defined in Equ. 8 is also called **microcanonical partition function**. It is generating function for the probability distribution of any associated system observables. The variable S is the entropy belonging to the equilibrium state density,

$$S = k_B \cdot \ln \Omega_{NVE} \quad (9)$$

It is clear that once it is in the chaotic equilibrium state, the system will essentially never find back to its completely assembled cluster state. The time that it can be estimated to take, the so-called **Poincaré recurrence time** is proportional to the number (density) of states in equilibrium.

According to an earlier section, one can follow the time evolution of the spontaneous cluster disintegration process through the H (eta)-function or, equivalently, the entropy of the N -particle system. Along the t -dependent path of system evolution, one observes an increase in S from an initially few units (of k_B) to several hundred units. This increase is significant because of the exponential dependence of the probability on entropy (Equ. 9). Generally, it is clear that for a transition $i \rightarrow f$ between two macrostates of a "microcanonical" system (one with fixed numbers of particles N , fixed volume V and fixed energy E) to occur **spontaneously, the entropy has to increase**

$$\Delta S = S_f - S_i = k_B \cdot \ln \frac{\Omega_f}{\Omega_i} > 0 \quad (10)$$

The type of **microcanonical systems for fixed particle number, volume and energy discussed above are isolated systems** such as individual atoms which have no contact with their surroundings and exchange neither energy nor particles with any other system. However, many systems encountered in practice allow at least some exchange of energy with an outside world. Of particular interest are relatively small systems that can easily be manipulated in a laboratory environment but are capable of heat energy exchange with a much larger environment. The latter type of very large or energetic system is sometimes called "heat bath," because it forces the smaller system of interest to acquire a macrostate consistent with the larger environment. For example, the environment can maintain a small system at a fixed average particle energy $\langle \varepsilon \rangle$, earlier identified with the temperature T . The small **systems with fixed number of particle (N), fixed volume (V) and a given temperature (T) are called "canonical."** The corresponding ensembles of many similarly prepared systems are called **canonical** ensembles. The small system taken together with its entire environment produces of course again a microcanonical system, because the combination of system and environment (= universe) is isolated by definition. The combined system must conserve energy (**1st Law of Thermodynamics**), while the small canonical system preserves this only on average (average energy \rightarrow temperature) but is **capable of thermal fluctuations about that average.**

As is true for a microcanonical stochastic system, a canonical system is not precisely predictable (not deterministic). Statements have validity for such stochastic system only when averages are performed over ensembles of many instances (representations), where "many" means $A \gg \Omega$. In practice, because the typically immense number Ω of available states, this task is literally

impossible to fulfill. In principle, such averages (or means) can experimentally be produced by performing a time average of a single measurement of the system, but over long enough times, such that a large number of microstates of the system are sampled in succession. The alternative is to repeat the same kind of measurement many times over, but with different systems that have been equally prepared. For example, one can make a pressure measurement on one mole of a gas over times long compared to thermal relaxation times or determine the pressure for a large number of independent samples, each of one mole of the gas. Since the same microstates are sampled in each of the two measurements, the two experiments give the same result (***Ergodic Hypothesis***). Clearly, it is more practical to perform a time averaged measurement which comes close to the task at hand only very approximately.

Similar to the probabilities f_i of microstates of microcanonical systems, those for canonical systems have to add up to unity:

$$\sum_i f_i = 1 \quad (11)$$

where the sum runs over all Ω microstates. However, while for a microcanonical system, $E = \sum_i \varepsilon_i = \text{const.}$, the corresponding ***energy conservation law applies only on average***,

$$\langle E \rangle = \sum_i \varepsilon_i f_i = \text{const.} \quad (12)$$

The most likely canonical equilibrium states are those with the highest entropies S . This variable is defined like in Equ. IV.126, i.e.,

$$S = -k_B \sum_{i=1}^{\Omega} f_i \cdot \ln f_i \quad (13)$$

where Ω is the number of microstates compatible with the macrostate considered. There are now more states i fulfilling the weaker (mean) energy conservation stated in Equ. 12 than permitted by a rigorous energy conservation law. Equations 11 and 12 represent boundary conditions, which have to be fulfilled for all macrostates, including those defining the asymptotic equilibrium. Mathematically, the equilibrium state distribution is found again by searching for a combination of f_i giving the maximum entropy ($\partial S / \partial f_i = 0$) but under the conditions defined in Eqs. 11 and 12. Since both sums are constant, the respective derivatives with respect to f_i are zero,

$$\frac{\partial}{\partial f_n} \sum_i f_i = 0 \quad \text{and} \quad \frac{\partial}{\partial f_n} \sum_i \varepsilon_i f_i = 0 \quad \{n, i = 1, 2, \dots, \Omega\} \quad (14)$$

Therefore, under these and only under these conditions, one can add any linear combination of the two equations in (13) to the condition of maximum entropy,

$$\begin{aligned} 0 &= \frac{\partial}{\partial f_n} \left\{ S + \lambda_1 \sum_i f_i + \lambda_2 \sum_i \varepsilon_i f_i \right\} = 0 \quad \{n = 1, 2, \dots, \Omega\} \quad (15) \\ &= -k_B [\ln f_n + 1] + \lambda_1 + \lambda_2 \varepsilon_n \end{aligned}$$

This maximization method is called the **Method of Lagrange multipliers**, which are the two constants λ_1 and λ_2 still to be determined. From Equ. 15 follows,

$$\ln f_n = \frac{1}{k_B} [\lambda_1 + \lambda_2 \varepsilon_n] - 1 \quad (16)$$

or

$$f_n = e^{\frac{1}{k_B} [\lambda_1 + \lambda_2 \varepsilon_n] - 1} \quad (17)$$

Summing over all microstates,

$$\sum_{n=1}^{\Omega} f_n = 1 = e^{\lambda_1/k_B - 1} \sum_{n=1}^{\Omega} e^{\lambda_2 \varepsilon_n / k_B} \quad (18)$$

one obtains the **partition function**

$$Z = e^{-\lambda_1/k_B + 1} = \sum_{n=1}^{\Omega} e^{\lambda_2 \varepsilon_n / k_B} \quad (19)$$

which is also an unnormalized probability and a generating function. Equation 18 can be written as

$$\sum_{n=1}^{\Omega} f_n = \frac{1}{Z} \sum_{n=1}^{\Omega} e^{\lambda_2 \varepsilon_n / k_B} \quad (20)$$

Obviously, a normalized probability

$$f_n = \frac{1}{Z} e^{\lambda_2 \varepsilon_n / k_B} \quad (21)$$

for each microstate with energy ε_n would explain the above relations, e.g., Equ. 20. This result applies to different types of ensembles. For example, it applies to a **canonical** ensemble of non-interacting particles in **molecular chaos**, characterized by a **constant temperature T** , which is also associated with the mean

particle energy. The probability f of finding a particle with energy ε in such a chaotic ensemble is governed by the Boltzmann factor $\exp\{-\varepsilon/k_B T\}$. Therefore, it is plausible to postulate

$$f_n = \frac{1}{Z} e^{-\varepsilon_n/k_B T} \quad \text{with} \quad Z = \sum_n e^{-\beta \cdot \varepsilon_n} = \sum_n e^{-\varepsilon_n/k_B T} \quad (22)$$

and to identify the Lagrange multiplier with $\lambda_2 = -\beta \cdot k_B = -1/T$. For other ensembles, the parameter β is connected to other statistical or to dynamical properties of the ensemble of particles.

At this point, and assuming a canonical ensemble, it is clear that the function Z is definitively an explicit function of the parameter $\beta = 1/k_B T$, i.e., $Z = Z(\beta)$. The dependence on other system parameter is less obvious, since the partition sum runs over the entire microstate energy spectrum ε_n . Nevertheless, the functional dependence of Z on the energy spectrum is specific for each system and may even reveal the characteristics of the system. For example, the energy terms of a harmonic oscillator with fundamental frequency could be $\varepsilon_n = (n + 1/2)\hbar\omega$, while that for a free particle n with mass m and velocity u_n could be $\varepsilon_n = mu_n^2/2$. Therefore, the actual specific physical information about a system lies in the structure of this function, e.g., $Z(\beta, \omega, \dots)$ or $Z(\beta, m, \dots)$, as given by the underlying structure of its energy spectrum ε_n .

In the expression (Equ. 22) for the canonical partition function Z , each of the quantities ε_n represents the energy of one microstate consistent with the macroscopic state of the system. Many microstates may have the same energy, and hence there can be many equal terms in the above sum representation of Z . If $\Omega(\varepsilon)$ is

number of microstates with energy ε , also called energy “level density,” then the sum over index n in Equ. 22 can be replaced by a sum over levels at energy ε ,

$$Z(\beta) = \sum_n e^{-\beta \cdot \varepsilon_n} = \sum_{\varepsilon} \Omega(\varepsilon) \cdot e^{-\beta \cdot \varepsilon} = \sum_{\varepsilon} \Omega(\varepsilon) \cdot e^{-\varepsilon/k_B T} \quad (23)$$

Therefore, the canonical partition function $Z(\beta)$, which is valid for a fixed temperature, can be written as a weighted sum over microcanonical partition functions $\Omega(\varepsilon)$, each of which is valid for a fixed energy. Of course, particle number and volume are also constant. Note that, except for the macroscopic temperature, which is the coupling parameter to the surroundings of the system, the partition function Z is given entirely in terms of microscopic system properties (Ω and ε).

The practical value of the canonical partition function $Z(\beta)$ is derived from the fact that it can be regarded as a generating function for all macroscopic variables X . These variables are expressed as expectation values or statistical means $\langle X \rangle$ with respect to the distribution f_n . For example, one obtains the mean total energy from

$$-\frac{\partial}{\partial \beta} \ln Z = \frac{-1}{Z} \frac{\partial Z}{\partial \beta} = \sum_n \varepsilon_n \frac{e^{-\beta \cdot \varepsilon_n}}{Z} = \sum_n \varepsilon_n f_n = E \quad (24)$$

Here, the quantity $E := \langle E \rangle$ **is the total internal energy** of the system residing in various microcanonical energy states, as given

by the weighted average, calculated with the T dependent Boltzmann weighting factors $f_n = e^{-\beta \cdot \varepsilon_n} / Z$.

Using the second of Eqs. 14, multiplying by f_n and summing over all states, one obtains another equation determining the Lagrange multipliers for the equilibrium probabilities:

$$\begin{aligned} 0 &= -k_B \sum_n f_n \ln f_n - k_B \sum_n f_n + \lambda_1 \sum_n f_n + \lambda_2 \sum_n f_n \varepsilon_n \\ &= S - k_B + \lambda_1 + \lambda_2 E = S - k_B \ln Z + \lambda_2 E \end{aligned} \quad (25)$$

Rearranging the terms leads to

$$-(k_B T) \cdot \ln Z = E - T \cdot S =: A \quad (26)$$

and another, closed-form expression for the canonical partition function Z ,

$$Z = \exp \left\{ - \frac{[E - T \cdot S]}{k_B T} \right\} = e^{-\frac{A}{k_B T}} \quad (27)$$

The function

$$A = E - T \cdot S \quad (28)$$

introduced above is known as the **Helmholtz free energy**. According to Equ. 27, A is equivalent to Z and contains the entire physical information about the system. Both functions are entirely given in terms of macroscopic state functions of temperature and entropy. The Helmholtz free energy function A is an important relation used extensively in the formalism of macroscopic equilibrium thermodynamics for **canonical systems of constant particle**

number N and constant volume V . These systems are confined but not thermally isolated, allowing therefore energy exchange with their surroundings.

The partition function, in particular in its macroscopic representation (cf. Equ. 27), has an innocuous appearance that hides a potentially rather complicated implicit dependence on the microscopic coordinates $\{x_1, x_2, x_3, \dots, x_M\}$ of the system it represents. For example, an N -particle system has $M=6N$ continuous position and momentum (velocity) degrees of freedom. For continuous degrees of freedom, where the energy is a function of the set of coordinates, $\varepsilon(x_1, x_2, x_3, \dots, x_M)$, the partition sum of Equ. 23 is transformed into a partition integral,

$$Z = \sum_n e^{-\beta \cdot \varepsilon_n} \rightarrow Z = \int dx_1 \cdots dx_M e^{-\beta \cdot \varepsilon(x_1, \dots, x_M)} \quad (29)$$

The corresponding expression for the mean energy is now

$$E = -\frac{\partial}{\partial \beta} \ln Z = \frac{-1}{Z} \frac{\partial Z}{\partial \beta} = \frac{1}{Z} \int dx_1 \cdots dx_M \varepsilon(x_1, \dots, x_M) \cdot e^{-\beta \cdot \varepsilon(x_1, \dots, x_M)} \quad (30)$$

which is equivalent to Equ. 24. If the individual degrees of freedom are decoupled, i.e., independent of each other, the energy is a sum over the individual energies,

$$\varepsilon(x_1, \dots, x_M) = \sum_{i=1}^M \varepsilon(x_i) \quad (31)$$

For free particles, only the momentum (velocity) coordinates contribute non-zero energy terms to the sum. If they are confined in a position dependent potential, then the energies depend also on

spatial coordinates. Consequently, the partition function is a product of partition functions for individual degrees of freedom:

$$Z = \int dx_1 \cdots dx_M e^{-\beta \cdot \sum_{i=1}^M \varepsilon(x_i)} = \prod_{i=1}^M \int dx_i e^{-\beta \cdot \varepsilon(x_i)} = \prod_{i=1}^M Z_i \quad (32)$$

The average energy for any particle degree of freedom i is then calculated as

$$E_i = -\frac{\partial}{\partial \beta} \ln Z_i = -\frac{\partial}{\partial \beta} \ln \left[\int dx_i e^{-\beta \cdot \varepsilon(x_i)} \right] \quad (33)$$

Assuming particle i is a free particle, its energy function is given by $\varepsilon(u) = (m/2)u^2$. Here, the velocity u represents one of the above coordinates x_i . For simplicity, set $u=x$. One then calculates from Equ. 33,

$$E = -\frac{\partial}{\partial \beta} \ln Z = -\frac{\partial}{\partial \beta} \ln \left[\int dx e^{-\beta \cdot \varepsilon(x)} \right] = -\frac{\partial}{\partial \beta} \ln \left[\int du e^{-\beta \cdot \frac{m}{2} u^2} \right] \quad (34)$$

Using $v := \sqrt{\beta} u$, the integral on the r.h.s. can be transformed into

$$Z = \left[\int du e^{-\beta \cdot \frac{m}{2} u^2} \right] = \left[\frac{1}{\sqrt{\beta}} \int dv e^{-\frac{m}{2} v^2} \right] \quad (35)$$

Noting that the integral does not explicitly depend on β , one finds

$$\begin{aligned}
 E = \langle E \rangle &= -\frac{\partial}{\partial \beta} \ln Z = -\frac{\partial}{\partial \beta} \left[\ln \frac{1}{\sqrt{\beta}} + \ln \int d\mathbf{v} e^{-\frac{m}{2}v^2} \right] \quad (36) \\
 &= -\frac{\partial}{\partial \beta} \ln \frac{1}{\sqrt{\beta}} = \frac{1}{2\beta} = \frac{1}{2} k_B T
 \end{aligned}$$

This result shows that the average energy of free particles in equilibrium is equal to $k_B T/2$. The same conclusion was reached in the specific discussion of ideal gas kinetics. The present derivation is more general and proves the

Equipartition Theorem:

Any degree of freedom i with a quadratic energy dependence carries in thermal equilibrium a mean energy of $\langle E_i \rangle = (1/2)k_B T$.

A posteriori, this result justifies the choice made earlier for the Lagrange multiplier λ_2 .