

STATISTICAL MECHANICAL ENSEMBLES¹

MICROSCOPIC AND MACROSCOPIC VARIABLES

The central question in Statistical Mechanics can be phrased as follows: If particles (atoms, molecules, electrons, nuclei, or even living cells) obey certain *microscopic* laws with specified interparticle interactions, what are the observable properties of a *macroscopic* system containing a large number of such particles?

Examples of microscopic and macroscopic variables are given below for 1/6th mole of an n -component monoatomic gas ($\approx 10^{23}$ molecules) obeying classical mechanics.

Microscopic variables	Macroscopic variables
3×10^{23} positions (x,y,z) 3×10^{23} velocities (x,y,z)	$n+2$ independent thermodynamic variables - e.g. for $n=1$, N, \underline{V}, T

PHASE SPACE

The multidimensional space defined by the microscopic variables of a system. In the example above, it would be a 6×10^{23} -dimensional space, with independent variables

$$(\mathbf{r}^N, \mathbf{p}^N) = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N; \mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3, \dots, \mathbf{p}_N)$$

where bold symbols indicate vectors, \mathbf{r} is the position and \mathbf{p} the momentum ($\mathbf{p} = m\mathbf{u}$). Momenta are used rather than velocities, because the classical and, especially, the quantum equations of change are more elegantly written in terms of momenta. The evolution of such a system in time is described by Newton's laws:

$$\frac{d\mathbf{r}_i}{dt} = \mathbf{u}_i ; \quad m_i \frac{d\mathbf{u}_i}{dt} = - \frac{\partial U(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \dots, \mathbf{r}_N)}{\partial \mathbf{r}_i}$$

where U is the total potential energy of the system, the gradient of which is minus the force.

For a much simpler system, a one-dimensional harmonic oscillator, phase space is two dimensional, with coordinates the position and the momentum variables.

ENSEMBLES

An ensemble is a collection of all microstates of a system, consistent with the constraints with which we characterize a system macroscopically. For example, a collection of all possible states of

¹ Some material in this section is derived from Chap. 3, D. Chandler, "Introduction to modern statistical mechanics", Oxford University Press, 1987.

the 10^{23} molecules of gas in the container of volume \underline{V} with a given total energy \underline{U} is a statistical mechanical ensemble.

For the one-dimensional harmonic oscillator with a given energy, the phase space is a circular trajectory in (properly normalized) position and momentum space.

ERGODIC HYPOTHESIS

Experimental measurements on any macroscopic system are performed by observing a system for a finite period of time, during which the system samples a very large number of possible microstates. In order to connect the measured properties with the properties calculated from statistical mechanics, we have to assume that:

For sufficiently long times, a macroscopic system will evolve through (or will come arbitrarily close to) *all* microscopic states consistent with the macroscopic constraints we impose in order to control the system. In other words, experimental measurements (performed by time averages) and ensemble averages are equivalent.

The "ergodic hypothesis" is more than just a hypothesis. It is a general property of almost all real systems composed of a large number of particles. The ergodic hypothesis is equivalent to the postulate of classical thermodynamics that there exist stable equilibrium states fully characterized by $n+2$ independent macroscopic variables.

Note that the ergodic hypothesis does not state anything about the relative probabilities of observing given states - it just states that *all* states will eventually be observed. A general property F will thus obey the following rule:

$$F_{\text{observed}} = \sum \mathcal{P}_v \times \underline{F}_v = \langle F \rangle \quad [1]$$

-----	-----	-----
probability of finding the system in microstate v	value of property F in microstate v	ensemble average

MICROCANONICAL ENSEMBLE: CONSTANT $\underline{U}, \underline{V}, N$

Basic Postulate of Statistical Mechanics: For an isolated system at constant \underline{U} , \underline{V} and N , all microscopic states of a system are equally likely at thermodynamic equilibrium.

If $\Omega(N, \underline{V}, \underline{U})$ is number of microstates with energy \underline{U} , then the probability of microstate v , according to the postulate above, is

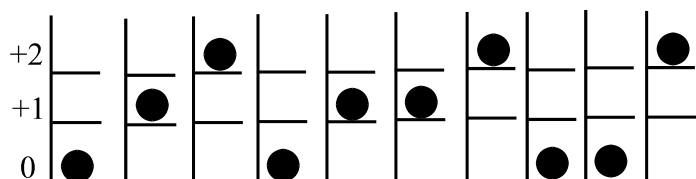
$$\mathcal{P}_v = 1/\Omega(N, \underline{V}, \underline{U}) \quad [2]$$

The existence of a finite number of microstates for a given energy is a concept fully justified in quantum mechanics. In classical statistical mechanics, we can introduce a "density of states" by

assuming that, say, two possible positions must be apart by more than a small distance in order to be counted as “different.” The (arbitrary) scales chosen to discretize positions and velocities determine how many states one counts in equation [2], within a *multiplicative* constant. In other words, if one doubles the volume of a container, there are twice as many states for molecular positions.

Example: Balls in a box

Let us consider a very simple example, namely a box with 10 slots, each containing a ball that can be at the bottom of the slot (with energy 0), or at one of two higher levels, with energies +1 and +2, respectively.



The number of microstates for this system depends on the overall energy of the box. There are:

1 state with energy 0; 10 states with energy +1; $\binom{10}{2} + 10 = 55$ states with energy +2 and so on.

As you can see, the number of microstates for even a simple system increases rapidly with the energy of the system.

Definition of Entropy

Let us define a quantity \underline{S} , such that:

$$\underline{S} = k_B \ln \Omega(N, \underline{V}, \underline{U}) \quad [3]$$

where k_B is a constant (to be later identified as Boltzmann's constant, $k_B = R/N_A = 1.380 \times 10^{-23} \text{ J/K}$, where N_A is Avogadro's Number, $N_A = 6.023 \times 10^{23} \text{ mol}^{-1}$). \underline{S} has the following properties:

1. \underline{S} is extensive: If we have two *independent* subsystems, A and B, then

$$\underline{S}_{A+B} = k_B \ln(\Omega_{A+B}) = k_B \ln(\Omega_A \cdot \Omega_B) = k_B \ln \Omega_A + k_B \ln \Omega_B = \underline{S}_A + \underline{S}_B \quad [4]$$

The reason for this is that, for *independent* subsystems, each microstate of system A can be combined with a microstate of system B to give a microstate of the combined system. For mixing two fluids, we only get the above expression if we assume that the particles in the systems are indistinguishable. If particles were distinguishable, additional "states" would be available to the combined system resulting from the possibility of exchanging the “labels” of particles. Although the indistinguishability of particles is really of quantum mechanical origin, it was introduced *ad hoc* by Gibbs before the development of quantum mechanics, in order to make statistical-mechanical entropy an extensive property.

2. \underline{S} is maximized at equilibrium: For a system with internal constraints (e.g. internal rigid walls or barriers to energy transfer), the number of possible microstates is always less than the number of microstates after the constraints are removed.

$$\underline{S}(N, \underline{V}, \underline{U}) > \underline{S}(N, \underline{V}, \underline{U}; \text{internal constraints})$$

To see this second property, consider the box with particles of the example above, and think of any constraint to the system at a given energy (say $\underline{U}=+2$). An example of a "constraint" would be to have that the first five slots have exactly 1 unit of energy. The number of microstates in this case is ($5 \times 5 = 25$), less than the 55 states available to the unconstrained system.

In conclusion, \underline{S} has the same properties as the entropy. Statement (2) above is the microscopic statement of the Second Law of thermodynamics. From the Fundamental Equation of thermodynamics,

$$d\underline{U} = Td\underline{S} - Pd\underline{V} + \sum_i \mu_i dN_i \Rightarrow d\underline{S} = \frac{1}{T} d\underline{U} + \frac{P}{T} d\underline{V} - \sum_i \frac{\mu_i}{T} dN_i \quad [5]$$

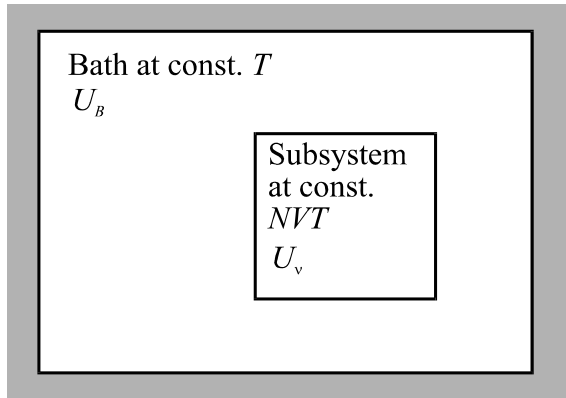
The second form of the fundamental equation (the "Entropy Representation") is the most useful for Statistical Mechanics.

$$\text{Since (from equation 5)} \quad \left(\frac{\partial \underline{S}}{\partial \underline{U}} \right)_{N, \underline{V}} = \frac{1}{T} \Rightarrow \left(\frac{\partial \ln \Omega}{\partial \underline{U}} \right)_{N, \underline{V}} = \frac{1}{k_B T} = \beta \quad [6]$$

The symbol β is commonly used in statistical mechanics to denote the inverse temperature.

Example (cont.): Balls in a box

In the example above, we can now define the temperature of the box from equation 1. Initially, as the energy increases, the temperature increases, as expected. However, something strange happens to the system at energies greater than +10 - can you guess what that is? What is the sign of the temperature for energies $> +10$? Is this physically reasonable?

CANONICAL ENSEMBLE: CONSTANT N, \underline{V}, T 

Total System: isolated,
fixed $N, V, U=U_B+U_v$

Let us first obtain the equilibrium condition for two systems, I and II, that are placed in thermal contact, so that they can exchange energy. The number of microstates available to the combined system must be a maximum at equilibrium, since the combined system is under conditions of constant N, \underline{V} and \underline{U} discussed previously. Mathematically, the condition for equilibrium is that

$$\underline{S}^{\text{total}} \text{ is maximum} \Rightarrow \delta \underline{S}^{\text{total}} = 0 \Rightarrow \delta \underline{S}^{\text{I}} + \delta \underline{S}^{\text{II}} = 0 \Rightarrow \frac{\partial \underline{S}^{\text{I}}}{\partial \underline{U}^{\text{I}}} \delta \underline{U}^{\text{I}} + \frac{\partial \underline{S}^{\text{II}}}{\partial \underline{U}^{\text{II}}} \delta \underline{U}^{\text{II}} = 0$$

Since the total system is isolated, $\delta \underline{U}^{\text{I}} + \delta \underline{U}^{\text{II}} = 0$. Combining with the above condition, we obtain

$$\frac{\partial \underline{S}^{\text{I}}}{\partial \underline{U}^{\text{I}}} = \frac{\partial \underline{S}^{\text{II}}}{\partial \underline{U}^{\text{II}}} \Rightarrow T^{\text{I}} = T^{\text{II}} \quad [7]$$

In other words, systems that can exchange energy much have the same temperature at equilibrium.

Now, let us assume that one of the two systems is much larger than the other, so that it effectively acts as a "constant-temperature bath." The total system (small system + bath) is again considered under $N\underline{V}\underline{U}$ conditions.

Now consider the small system at a *given* microstate v with energy \underline{U}_v . The energy of the bath is $\underline{U}_B = \underline{U} - \underline{U}_v$. The bath can be in any of $\Omega(\underline{U}_B) = \Omega(\underline{U} - \underline{U}_v)$ microstates. Since the probability of the total system being in any particular combination of microstates with a given total energy is the same, the probability of finding the small system in state v is

$$\mathcal{P}_v \propto \Omega(\underline{U} - \underline{U}_v) = \exp(\ln(\Omega(\underline{U} - \underline{U}_v))) \quad [8]$$

We can expand $\ln \Omega$ around $\Omega(\underline{U})$ given that \underline{U}_v is much smaller than \underline{U} :

$$\ln(\Omega(\underline{U} - \underline{U}_v)) = \ln(\Omega(\underline{U})) - \underline{U}_v \frac{\partial \ln \Omega}{\partial \underline{U}} + \text{higher order terms}$$

and substituting back in the expression for P_v using $\partial \ln \Omega / \partial U = \beta = 1/k_B T$,

$$\mathcal{P}_v \propto \exp(-\beta U_v) \quad [9]$$

This is a very important result. In words, we are finding that the probability of each microstate in the canonical ensemble (constant N, V, T) is proportional to the exponential of the energy divided by the temperature. In order to find the *absolute* probability of each microstate, we need to make sure that the sum of all the probabilities is one. The normalization constant for this is called the "canonical partition function," Q .

$$Q(N, V, T) = \sum_{\text{all microstates } i} \exp(-\beta U_i) \quad [10]$$

The summation over microstates is performed over *all* energies and particle positions. For the simple example of the box with particles we have been following, the partition function at a given inverse temperature β would be:

$$Q(\beta) = 1 + 10\exp(-\beta) + 55\exp(-2\beta) + \dots$$

Where the first term in the summation comes from the single energy state with $U = 0$, the second from the 10 states with $U = +1$ and so on. For this simple system, both volume and number of particles are fixed, so they do not appear in the summation - however, in the general case the partition function would be a function of both N and V .

Once the partition function is defined, the probability of each microstate can now be written explicitly:

$$\mathcal{P}_v = \frac{\exp(-\beta U_v)}{Q} \quad [11]$$

Therefore, in the canonical ensemble, a general property F is given by

$$\langle F \rangle = \frac{\sum_{\text{all microstates } i} F_i \exp(-\beta U_i)}{\sum_{\text{all microstates } i} \exp(-\beta U_i)} \quad [12]$$

It is possible to relate derivatives of $\ln Q$ to thermodynamic properties. For example,

$$\left(\frac{\partial \ln Q}{\partial \beta} \right)_{N, V} = \frac{(\partial Q / \partial \beta)_{N, V}}{Q} = - \frac{\sum U_v \exp(-\beta U_v)}{\sum \exp(-\beta U_v)} = - \langle U \rangle \quad [13]$$

One can also calculate the averaged squared fluctuation of energy in the canonical ensemble:

$$\begin{aligned} \langle (\delta \underline{U})^2 \rangle &= \langle (\underline{U} - \langle \underline{U} \rangle)^2 \rangle = \langle \underline{U}^2 \rangle - \langle \underline{U} \rangle^2 = \sum P_v \underline{U}_v^2 - \left(\sum P_v \underline{U}_v \right)^2 = \\ &= \frac{(\partial^2 Q / \partial \beta^2)_{N, \underline{V}}}{Q} - \left(\frac{(\partial Q / \partial \beta)_{N, \underline{V}}}{Q} \right)^2 = \left(\frac{\partial^2 \ln Q}{\partial \beta^2} \right)_{N, \underline{V}} = - \left(\frac{\partial \langle \underline{U} \rangle}{\partial \beta} \right)_{N, \underline{V}} \end{aligned}$$

using the definition of the heat capacity, $\underline{C}_V = \left(\frac{\partial \langle \underline{U} \rangle}{\partial T} \right)_{N, \underline{V}}$, we get

$$\langle (\delta \underline{U})^2 \rangle = k_B T^2 \underline{C}_V \quad [14]$$

This is a remarkable result! We have obtained a relationship between the thermodynamic parameters of a system and the size of the spontaneous fluctuations. It is interesting to note that the heat capacity of the system, \underline{C}_V , is an extensive variable (grows linearly with the size of the system). This implies that the *relative* magnitude of the spontaneous fluctuations grows as:

$$\frac{\langle (\delta \underline{U})^2 \rangle^{1/2}}{\langle \underline{U} \rangle} = \frac{(k_B T^2 \underline{C}_V)^{1/2}}{\langle \underline{U} \rangle} \propto \frac{N^{1/2}}{N} = \frac{1}{\sqrt{N}} \quad [15]$$

For a macroscopic system, $N = O(10^{23})$, this is a very small number: the energy of an ideal gas system at equilibrium with a room-temperature bath is constant to roughly 1 part in 10^{11} . However, for small systems typically used in simulations, $N \approx 100-1,000$, so that typical fluctuations are 10%-3%. Equation 15 is not valid near critical points, at which the heat capacity (and fluctuations) diverge.

Q can be identified with a familiar thermodynamic function. To do this, let us write

$$k_B \ln Q = k_B \ln \left(\sum_v \exp(-\beta \underline{U}_v) \right) \quad [16]$$

As just shown, the relative fluctuations in energy for a macroscopic system are very small. We can approximate the sum in equation 16 by summing just the dominant terms. All of these will have $\underline{U}_v = \langle \underline{U} \rangle$. There are $\Omega(\langle \underline{U} \rangle)$ such terms (since this is the number of microstates at that \underline{U} , and thus

$$k_B \ln Q \approx k_B \ln(\Omega(\langle \underline{U} \rangle) \exp(-\beta \langle \underline{U} \rangle)) = k_B \ln \Omega(\langle \underline{U} \rangle) - \langle \underline{U} \rangle / T = \underline{S} - \underline{U} / T = -\underline{A} / T \quad [17]$$

This should have been expected. In the microcanonical (const. $N\underline{V}\underline{U}$) ensemble the important function Ω is such that $k_B \ln \Omega = \underline{S}$. Recall that \underline{S} is the function maximized at const. $N\underline{V}\underline{U}$. In the canonical ($N\underline{V}T$) ensemble, the thermodynamic function being minimized is \underline{A} , or equivalently $-\underline{A} / T$ is maximized. We see a similar relationship of $-\underline{A} / T$ with $k_B \ln Q$.

The approximation is exact at the "thermodynamic limit" (for an infinite system). The table below summarizes the connections between "classical" and "statistical" thermodynamic properties discussed thus far.

Const. N , \underline{V} and \underline{U}		Const. N , \underline{V} , and T	
Classical	Statistical	Classical	Statistical
Entropy, \underline{S} is maximized at equil.	$k_B \ln \Omega = \underline{S}$	Helmholtz Energy, \underline{A} is minimized at equil.	$k_B \ln Q = -\underline{A} / T$

GENERALIZED ENSEMBLES AND LEGENDRE TRANSFORMS

Starting from the fundamental equation in terms of \underline{S}/k_B

$$\frac{1}{k_B} d\underline{S} = \frac{1}{k_B T} d\underline{U} + \frac{P}{k_B T} d\underline{V} - \sum \frac{\mu_i}{k_B T} dN_i \quad [18]$$

First Legendre Transform: $y^{(1)} = \underline{S}/k_B - \underline{U}/k_B T = -\underline{A}/k_B T$ (\underline{A} is the Helmholtz free energy)

$y^{(0)} = \underline{S}/k_B = \ln \Omega$		$y^{(1)} = -\underline{A}/k_B T = \ln Q$	
Variable	Derivative	Variable	Derivative
\underline{U}	$1/(k_B T) = \beta$	\underline{U}	$-1/(k_B T) = -\beta$
\underline{V}	$P/(k_B T) = \beta P$	\underline{V}	$P/(k_B T) = \beta P$
N	$-\mu/(k_B T) = -\beta \mu$	N	$-\mu/(k_B T) = -\beta \mu$

We then have all the relationships we need for the thermodynamic functions:

$$\frac{1}{k_B T} = \beta = \left(\frac{\partial \ln \Omega}{\partial \underline{U}} \right)_{\underline{V}, N} \quad [19]$$

$$\underline{U} = \left(\frac{\partial \ln Q}{\partial \beta} \right)_{\underline{V}, N} \quad [20]$$

$$\frac{P}{k_B T} = \beta P = \left(\frac{\partial \ln Q}{\partial \underline{V}} \right)_{T, N} = \left(\frac{\partial \ln \Omega}{\partial \underline{V}} \right)_{\underline{U}, N} \quad [21]$$

$$-\frac{\mu}{k_B T} = -\beta\mu = \left(\frac{\partial \ln Q}{\partial N} \right)_{T, V} = \left(\frac{\partial \ln \Omega}{\partial N} \right)_{\underline{U}, \underline{V}} \quad [22]$$

One can continue this with Legendre transforms of higher order. The probabilities of microstates in the corresponding ensembles can be derived in a way completely analogous to the derivation for the canonical ensemble, involving a subsystem and bath of constant temperature. In general, starting from

$$y^{(0)} = \underline{S}/k_B ; \quad y^{(n)} = y^{(0)} - \xi_1 X_1 - \xi_2 X_2 - \dots - \xi_n X_n \quad [23]$$

where ξ_i is the i -th derivative of $y^{(0)}$ with respect to variable X_i . The probability of a microstate v in any given ensemble is proportional to

$$\mathcal{P}_v \propto \exp(-\xi_1 X_1 - \xi_2 X_2 - \dots - \xi_n X_n) \quad \{\text{where the } \xi_i\text{'s and } X_i\text{'s refer to } y^{(0)}\} \quad [24]$$

with normalization factor

$$\Xi = \sum_v \exp(-\xi_1 X_1 - \xi_2 X_2 - \dots - \xi_n X_n) \quad [25]$$

Ξ is the partition function for the corresponding statistical ensemble. Using Ξ , the probability \mathcal{P}_v is now equal to:

$$\mathcal{P}_v = \exp(-\xi_1 X_1 - \xi_2 X_2 - \dots - \xi_n X_n) / \Xi \quad [26]$$

As for the canonical ensemble, Ξ has the property that :

$$\ln \Xi = y^{(n)} = y^{(0)} - \xi_1 X_1 - \xi_2 X_2 - \dots - \xi_n X_n \quad [27]$$

Using the expression above, it easy to prove the famous "Gibbs entropy formula":

$$\underline{S} = -k_B \sum_v P_v \ln P_v \quad [28]$$

Proof:

$$\begin{aligned} \sum_v P_v \ln P_v &= \sum_v P_v (-\ln \Xi - \xi_1 X_1 - \xi_2 X_2 - \dots - \xi_n X_n) \quad \{\text{from equation 26}\} \\ &= -\ln \Xi - \xi_1 \langle X_1 \rangle - \xi_2 \langle X_2 \rangle - \dots - \xi_n \langle X_n \rangle = -y^{(0)} \\ &\quad \{\text{since all } \xi_i\text{'s are constant in this ensemble}\} \\ &= -\underline{S} / k_B \end{aligned}$$

APPLICATION: GRAND CANONICAL ENSEMBLE (CONST. μ , \underline{V} , T)

The probability of microstates in this ensemble is

$$\mathcal{P}_v = \exp(-\beta \underline{U}_v + \beta \mu N_v) / \Xi \quad [29]$$

where

$$\ln \Xi = \underline{S}/k_B - \underline{U}/k_B T + \mu N/k_B T = P\underline{V}/k_B T$$

The derivative relationships for the Grand Canonical partition function Ξ can be obtained from a Legendre transform table:

$y^{(0)} = \underline{S}/k_B = \ln \Omega$		$y^{(2)} = P\underline{V}/k_B T = \ln \Xi$	
Variable	Derivative	Variable	Derivative
\underline{U}	$1/(k_B T) = \beta$	$1/(k_B T) = \beta$	$-\underline{U}$
N	$-\mu/(k_B T) = -\beta \mu$	$-\beta \mu$	$-N$
\underline{V}	$P/(k_B T) = \beta P$	\underline{V}	$P/(k_B T) = \beta P$

e.g.,

$$\frac{\partial \ln \Xi}{\partial(-\beta \mu)} = -\langle N \rangle \quad [30]$$

Fluctuation formulae in the grand canonical ensemble are analyzed in the same fashion as in the canonical ensemble. For example,

$$\langle (\delta N)^2 \rangle = \langle (N - \langle N \rangle)^2 \rangle = \langle N^2 \rangle - \langle N \rangle^2 = \sum_v \mathcal{P}_v N_v^2 - \sum_v \sum_{v'} \mathcal{P}_v N_v \mathcal{P}_{v'} N_{v'} =$$

$$= \left(\frac{\partial^2 \ln \Xi}{\partial(\beta \mu)^2} \right)_{\beta, \underline{V}} = \left(\frac{\partial \langle N \rangle}{\partial(\beta \mu)} \right)_{\beta, \underline{V}} \quad [31]$$