

Statistical Mechanical Ensembles[†]

1. Microscopic Origin of Entropy

A common definition of entropy from a macroscopic thermodynamic viewpoint is $\Delta S = Q^{\text{rev}}/T$, where Q^{rev} is the amount of heat exchanged at temperature T between a system and its environment. But what is the origin of entropy in the microscopic, molecular world? Qualitatively, we expect that irreversible processes (which result in an increase in total entropy) also increase the degree of “disorder” in a system. It turns out that a quantitative measure of disorder, specifically the number of microscopic states Ω available to a system at a given internal energy U and for specified number of molecules (or moles) N and volume V , can be quantitatively linked to the entropy. A distinct microscopic state (*microstate*) is defined by all microscopic degrees of freedom – e.g. positions and velocities of molecules in a gas. A set of microstates with specified common properties (e.g. number of particles N , volume V and energy U) defines an *ensemble*. The expression of S in terms of microstates is provided by the famous 1872 entropy formula of Ludwig Boltzmann,

$$S = k_{\text{B}} \ln \Omega(N, V, U) \quad (1)$$

Boltzmann’s entropy formula is the foundation of statistical mechanics, connecting macroscopic and microscopic points of view. It allows calculations of macroscopic thermodynamic properties by determining properties of microscopic configurations. The constant k_{B} is called Boltzmann’s constant; it is the ratio of the ideal-gas constant to Avogadro’s number, or equivalently the gas constant on a per molecule basis: $k_{\text{B}} = R/\mathcal{N}_{\text{A}} = 1.38065 \cdot 10^{-23}$ J/K.

The concept of microstates arises naturally in quantum mechanics, but can be also introduced in classical systems, if positions and velocities can be grouped so that, when their values differ less than some selected (but arbitrary) small value, then they are considered to be equal. This quantization effect also arises naturally in computations, which are performed with a finite numerical accuracy, so that two quantities cannot differ by less than the machine precision.

[†] Draft material from “Statistical Thermodynamics” © 2014, A. Z. Panagiotopoulos

To illustrate the concept of counting microstates, we will use the simple system shown in Fig. 1. In the general case, it consists of N slots, each containing a ball that can be at energy levels 0, +1, +2, +3, ..., measured in units of $k_B T_0$, where T_0 is a reference temperature. The specific case of $N=10$ and 3 energy levels is shown in the figure. The concept of discrete energy levels arises very naturally in quantum mechanics. For this simple system, we can count states by using the combinatorial formula giving the number of ways we can pick M specific distinguishable objects out of N total objects:

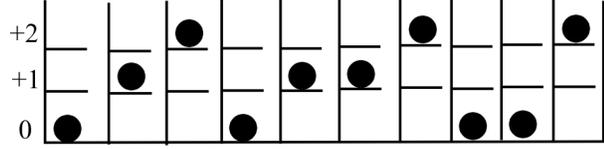


Figure 1 A system of 10 spheres with 3 energy levels.

$$\binom{N}{M} = \frac{N!}{M!(N-M)!} = \frac{N \cdot (N-1) \cdots (N-M+1)}{1 \cdot 2 \cdots M} \quad (2)$$

There is only 1 state with internal energy $U=0$. States with $U=1$ have one ball at level +1 and the others at level 0, so for $N=10$, there are 10 such states. States with energy $U=2$ may have 1 ball at +2 and the others at 0, or 2 balls at +1, so there are:

$$\binom{10}{2} + 10 = 55$$

such states. We can similarly obtain $\Omega(3)=220$; $\Omega(4)=715$, $\Omega(5)=2002$, and so on. Note that the number of microstates increases rapidly with the total energy of the system. This is generally the case for most systems.

Now we are in a position to show that S defined microscopically from Eq. 1 has the two key properties associated with the entropy of classical thermodynamics:

1. S is extensive: For two independent subsystems, A and B,

$$S_{A+B} = k_B \ln(\Omega_{A+B}) = k_B \ln(\Omega_A \cdot \Omega_B) = k_B \ln \Omega_A + k_B \ln \Omega_B$$

The reason is that each microstate of system A can be combined with a microstate of system B to give a microstate of the combined system. This is clearly true for the simple system illustrated on the previous page. However, when mixing two gases or liquids, we only get the above expression if we assume that the particles in the systems are indistinguishable. If particles are distinguishable, additional states are 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 entropy an extensive property.

2. 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 smaller than the number of microstates after the constraints are removed.

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

To demonstrate this second property, consider the box with particles of the example above, and think of any constraint to the system at a given total energy (say $U=+2$). 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.

One clear distinction between macroscopic and microscopic definitions of entropy is that the former is physically meaningful only as a difference of entropy between specified states, while the latter appears to provide a measure of *absolute* entropy. This apparent discrepancy results from the inability of classical physics to define uniquely when two nearby states (e.g. positions of a particle in free space differing by a fraction of a nm) are sufficiently different to justify distinguishing them from each other. Quantum mechanical methods, on the other hand, provide precise ways to count states.

At low temperatures, the number of microstates available to any physical system decreases rapidly. At the limit of absolute zero temperature, $T \rightarrow 0$, most systems adopt a unique "ground state" for which $\Omega = 1 \Rightarrow S = k_B \ln \Omega = 0$. This is the basis of the "Third Law of thermodynamics" postulated by Nerst in the early 1900's. The NIST Chemistry WebBook lists absolute entropies for pure components and chemical elements in the thermochemistry data section. However, using entropy values calculated with respect to an arbitrary reference state gives the same results as absolute entropies for heat and work amounts.

Example 1 – Entropy of a lattice chain

A common model for polymers is the Flory lattice model, which represents chains as "self-avoiding random walks" on a lattice (grid). Self-avoiding means that two beads cannot occupy the same position on the lattice. When two non-bonded beads occupy adjacent positions, they have energy of interaction equal to $-k_B T_0$, where T_0 is a reference temperature. Obtain the number of microstates Ω for a two-dimensional square-lattice chain of 5 beads, as a function of the energy U of the chain.

Fig. 2 shows the number of configurations for a square-lattice chain of 5 beads. Without loss of generality, we have fixed the configuration of the first two beads of

the chain to be in the horizontal direction, with the second bead to the right of the first. This reduces the total number of configurations by a factor of 4; such a multiplicative factor simply shifts the value of S obtained from Eq. 1 by a constant factor, akin to the reference state for the entropy. The last bond is shown in multiple configurations (arrows), along with their number and energy: $2 \times (-1)$ for the top left image means there are 2 configurations, each of energy -1 .

Overall, counting configurations of the same energy:

$$\Omega(U=0) = 3+2+2+3+2+2+3=17 \quad ; \quad \Omega(U=-1) = 2+1+1+1+1+2=8$$

The number of microscopic configurations and energy levels increases rapidly with chain length. Theoretical and Monte Carlo computer simulation techniques are used for determining the properties of models of this type for longer chains.

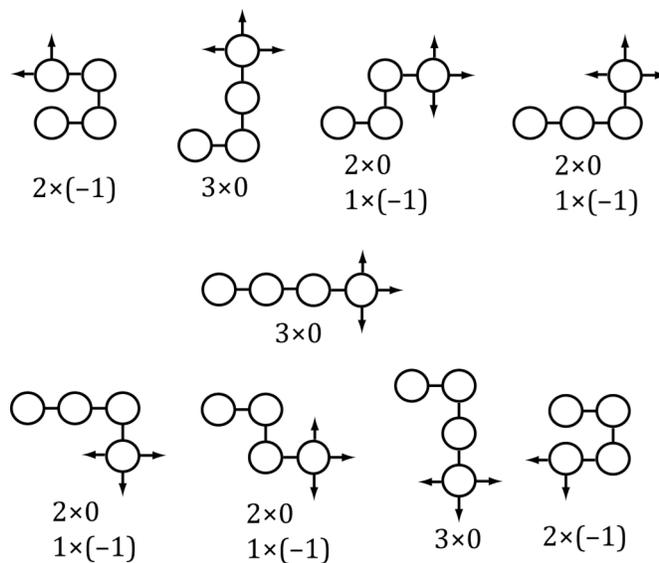


Figure 2 Configurations for a two-dimensional chain of 5 beads

A key issue in statistical mechanics is the frequency of occurrence of different microstates when the overall constraints of a thermodynamic system are specified. A basic postulate, comparable in importance to the postulate about the existence of equilibrium states in classical thermodynamics, is that *all* microstates of a system at a given U , N and V are equally probable.

The basic postulate of statistical mechanics implies that the probability of any microstate ν , \mathcal{P}_ν , is the same as that of any other microstate in the constant N, V, U ensemble:

$$\mathcal{P}_v = \frac{1}{\Omega} \text{ at constant } N, V \text{ and } U \quad (3)$$

From this postulate, we can now simply derive another famous expression, the *Gibbs entropy formula*, by substituting Eq. 3 into Eq. 1:

$$S = -k_B \sum_{\text{all micro-states } v} \mathcal{P}_v \ln \mathcal{P}_v \quad (4)$$

Gibbs
entropy
formula

The Gibbs entropy formula can be shown to be valid even for systems *not* at constant energy U , volume V , and number of particles N . This is in contrast to Eq. 1, which is only valid at constant for microstates at constant U , V and N . For example, in § 3 we prove Eq. 4 for systems at constant N , V , and T .

2. Phase Space and Statistical Mechanical Ensembles

Boltzmann's entropy formula links a macroscopic thermodynamic quantity, the entropy, to microscopic attributes of a system at equilibrium. In this section, we introduce many similar relationships for systems under constraints different than (N, V, U) , in a manner quite analogous to the introduction of fundamental equations in different variables developed in the previous chapter through the Legendre transform formalism.

The branch of physical science that aims to connect macroscopic properties and microscopic information about a system is called *Statistical Mechanics*. 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? Unlike classical thermodynamics, statistical mechanics does require input information on the microscopic constitution of matter of interest, as well as the interactions active among the microscopic building blocks. The advantage of the approach is that quantitative predictions of macroscopic properties can then be obtained, rather than simply relationships linking different properties to each other. Another important difference between statistical mechanics and macroscopic thermodynamics is that fluctuations, which are absent by definition in the thermodynamics, can be quantified and analyzed through the tools of statistical mechanics. Fluctuations are temporary deviations of quantities such as the pressure or energy of a system from their mean values, and are important in small systems – e.g. those studied by computer simulations, or present in modern nanoscale electronic devices and biological organelles.

Postulate I states that macroscopic systems at equilibrium can be fully characterized by $n+2$ independent thermodynamic variables. For a 1-component isolated system, these variables can always be selected to be the total mass N , total volume V and total energy U . However, at the microscopic level, molecules are in constant motion. Adopting temporarily a classical (rather than quantum mechanical) point of view, we can describe this motion through the instantaneous positions and velocities of the molecules. Examples of microscopic and macroscopic variables are given below for N molecules of a one-component monoatomic gas obeying classical mechanics.

Microscopic variables (in 3 dimensions)	Macroscopic variables
$3N$ position coordinates (x, y, z) $3N$ velocity components (u_x, u_y, u_z)	3 independent thermodynamic variables, e.g., $N, V,$ and U .

Given that N is of the order of Avogadro's number [$\mathcal{N}_A = 6.0221 \times 10^{23} \text{ mol}^{-1}$] for macroscopic samples, there is a huge reduction in the number of variables required for a full description of a system when moving from the microscopic to the macroscopic variables. Moreover, the microscopic variables are constantly changing with time, whereas for a system at equilibrium, all macroscopic thermodynamic quantities are constant. The multidimensional space defined by the microscopic variables of a system is called the *phase space*. This is somewhat confusing, given that the word "phase" has a different meaning in classical thermodynamics – the term was introduced by J. Willard Gibbs, no stranger to the concept of macroscopic phases.

In general, for a system with N molecules in 3 dimensions, phase space has $6N$ independent variables

$$(r^N, p^N) \equiv (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N, \mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_N) \quad (5)$$

where bold symbols indicate vectors, \mathbf{r}_i is the position and \mathbf{p}_i the momentum of the i -th molecule, ($\mathbf{p}_i = m_i \mathbf{u}_i$). The evolution of such a system in time is described by Newton's equations of motion:

$$\begin{aligned} \frac{d\mathbf{r}_i}{dt} &= \mathbf{u}_i \\ m_i \frac{d\mathbf{u}_i}{dt} &= - \frac{\partial \mathcal{U}(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)}{\partial \mathbf{r}_i} \end{aligned} \quad (6)$$

where \mathcal{U} is the total potential energy, which includes interactions among molecules and any external fields acting on the system.

For a much simpler system, a one-dimensional harmonic oscillator shown in Fig. 3, phase space is two-dimensional, with coordinates the position and the momentum variables. In the absence of friction, the oscillator undergoes harmonic motion, which can be represented as a circle in phase space if position and momentum are scaled appropriately. The diameter of the circle depends on the total energy (a constant of the motion), and the oscillator moves around the circle at a constant velocity.

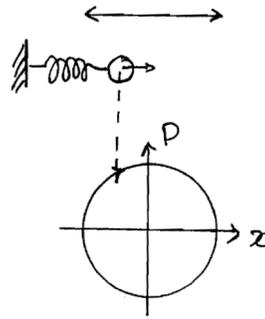


Figure 3 A one-dimensional harmonic oscillator (top), and the corresponding phase space (bottom).

A statistical mechanical *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 N molecules of gas in the container of volume V with a given total energy U is a statistical mechanical ensemble. For the frictionless one-dimensional harmonic oscillator, the ensemble of states of constant energy is the circular trajectory in position and momentum space shown in Fig. 3, bottom panel.

3. Molecular Chaos and Ergodic Hypothesis

What causes the huge reduction from $3N$ time-dependent coordinates needed to fully characterize a system at the molecular level to just a handful of time-independent thermodynamic variables at the macroscopic level? The answer turns out to be related to the chaotic behavior of systems with many coupled degrees of freedom.

To illustrate this concept, one can perform a thought experiment on the system shown in Fig. 4. In this system, a number of molecules of a gas are given identical velocities along the horizontal coordinate, and are placed in an insulated box with perfectly reflecting walls. Such a system would seem to violate the classical thermodynamics postulate of eventually approaching an equilibrium state. Since there is no initial momentum in the vertical direction, Newton's equations of motion would suggest that the molecules will never hit the top wall, which will thus experience zero pressure at all times, even though there is a finite density of gas in the system. In statistical mechanical terms, microstates with non-zero vertical momenta will never be populated in such a system. Of course, even tiny interac-

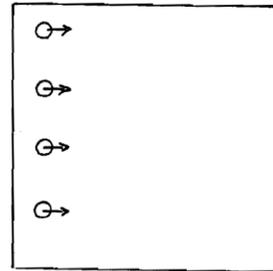


Figure 4 A conceptual experiment in an isolated system.

tions between the molecules, or minor imperfections of the walls, will eventually result in chaotic motion in all directions; thermodynamic equilibrium will be established over times significantly greater than the interval between molecular collisions.

Most molecular systems are able to “forget” their initial conditions and evolve towards equilibrium states, by sampling all available phase space. Only non-equilibrium (for example, glassy) systems violate this condition and have properties that depend on their history – such systems cannot be analyzed with the methods of equilibrium thermodynamics or statistical mechanics.

At a molecular level, the statement equivalent to Postulate I of classical thermodynamics is the *ergodic hypothesis*. The statement is as follows.

Ergodic hypothesis

For sufficiently long times, systems evolve through *all* microscopic states consistent with the external and internal constraints imposed on them.

Experimental measurements on any macroscopic system are performed by observing it for a finite period of time, during which the system samples a very large number of possible microstates. The ergodic hypothesis suggests that for “long enough” times, the entire ensemble of microstates consistent with the microscopic constraints on the system will be sampled. A schematic illustration of trajectories of ergodic and non-ergodic systems in phase space is shown in Fig. 5 – a two-dimensional representation of phase space is given, whereas we know that for realistic systems phase space has a very large number of dimensions. The interior of the shaded region is the phase space of the corresponding system. For the system



Figure 5 Schematic trajectories of ergodic (left) and non-ergodic (right) systems in phase space.

on the left, there are two connected regions of phase space, so the trajectory eventually passes from one to the other and samples the whole of phase space. By contrast, for the system on the right, the two regions of phase space are disconnected, so that the system cannot pass from one to the other. The system is non-ergodic.

There is a strong link between time scales and constraints; a system observed over a short time can appear to be isolated, while over longer times it may exchange energy and mass with its surroundings. Also, the “possible microstates” depend on the time scales of interest, and on internal constraints – for example, chemical reactions open up additional microstates, but can only occur over long time scales, or in the presence of a catalyst.

For systems satisfying the ergodic hypothesis, experimental measurements (performed by time averages) and statistical mechanical ensemble averages are equivalent. Of course, we have not specified anything up to this point about the relative probabilities of specific states; the ergodic hypothesis just states that *all* states will eventually be observed. A general property F can be formally written as:

$$F_{\text{observed}} = \sum \mathcal{P}_v \times F_v = \langle F \rangle \quad (7)$$

$\begin{array}{ccccccc} \text{time} & & & & & & \\ \text{average} & & \text{probability of finding} & & \text{value of property } F & & \text{ensemble} \\ & & \text{the system in microstate } v & & \text{in microstate } v & & \text{average} \end{array}$

The objective of the next few sections will be to determine the relative probabilities, \mathcal{P}_v , of finding systems in given microstates of ensembles under varying constraints. This will allow the prediction of properties by performing *ensemble averages*, denoted by the angle brackets of the rightmost side of Eq. 7.

4. Microcanonical Ensemble: Constant U , V , and N

The simplest set of macroscopic constraints that can be imposed on a system are those corresponding to isolation in a rigid, insulated container of fixed volume V . No energy can be exchanged through the boundaries of such a system, so Newton’s equations of motion ensure that the total energy U of the system is constant. For historical reasons, conditions of constant energy, volume, and number of particles (U, V, N) are defined as the *microcanonical* ensemble. In the present section (and the one that follows), we will treat one-component systems for simplicity; generalization of the relationships to multicomponent systems is straightforward.

How can we obtain the probabilities of microstates in a system under constant U, V , and N ? Consider for a moment two microstates with the same total energy, depicted schematically in Fig. 6. One may be tempted to say that the microstate on the left, with all molecules having the same velocity and being at the same horizontal position is a lot less “random” than the microstate on the right, and thus less likely to occur. This is, however, a misconception akin to saying that the number “111111” is less likely to occur than the number “845192” in a random sequence of 6-digit numbers. However, in a random (uniformly distributed) sample, all numbers are equally likely to occur, by definition. For molecular systems, it is not hard to argue that any specific set of positions and velocities of N particles in a

volume V that has a given total energy, should be equally probable as any other specific set. There are, of course, a lot more states that “look like” the right-hand side of Fig. 6 relative to the left-hand side, the same way that there are a lot more 6-digit numbers

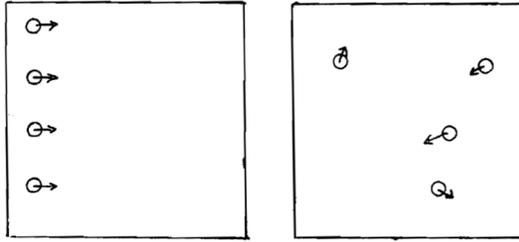


Figure 6 Two microstates in an isolated system.

with non-identical digits than there are with all digits equal. The statement that all microstates of a given energy are equally probable cannot be proved for the general case of interacting systems. Thus, we adopt it as the basic Postulate of statistical mechanics:

Basic Postulate
of Statistical
Mechanics

For an isolated system at constant U , V , and N , all microscopic states of a system are equally likely at thermodynamic equilibrium.

Just as was the case for the postulates of classical thermodynamics, the justification for this statement is that predictions using the tools of statistical mechanics that rely on this postulate are in excellent agreement with experimental observations for many diverse systems, provided that the systems are ergodic.

As already suggested, given that $\Omega(U, V, N)$ is the number of microstates with energy U , the probability of microstate v in the microcanonical ensemble, according to the postulate above, is:

$$\mathcal{P}_v = \frac{1}{\Omega(U, V, N)} \quad (8)$$

The function $\Omega(U, V, N)$ is called the *density of states*, and is directly linked to the entropy, via Boltzmann’s entropy formula, $S = -k_B \ln \Omega$. It is also related to the fundamental equation in the entropy representation, with natural variables (U, V, N) . Writing the differential form of the fundamental equation for a one-component system in terms of Ω , we obtain:

$$\frac{dS}{k_B} = d \ln \Omega = \beta dU + \beta P dV - \beta \mu dN \quad (9)$$

In Eq. 9, we have introduced for the first time the shorthand notation $\beta \equiv 1/(k_B T)$. This combination appears frequently in statistical mechanics, and is usually called the “inverse temperature,” even though strictly speaking it has units of inverse energy [J^{-1}]. Differentiation of Eq. 9 provides expressions for the inverse

temperature, pressure, and chemical potential, in terms of derivatives of the logarithm of the number of microstates with respect to appropriate variables:

$$\left(\frac{\partial \ln \Omega}{\partial U} \right)_{V,N} = \beta \quad (9)$$

$$\left(\frac{\partial \ln \Omega}{\partial V} \right)_{U,N} = \beta P \quad (10)$$

$$\left(\frac{\partial \ln \Omega}{\partial N} \right)_{U,V} = -\beta \mu \quad (11)$$

Example 2 – A system with two states and negative temperatures

Consider a system of N distinguishable particles at fixed positions, each of which can exist either in a ground state of energy 0, or in an excited state of energy ϵ . The system is similar to that depicted in Fig. 1, except it has only 2 (rather than 3) energy levels. Assuming that there are no interactions between particles, derive expressions for the density of states and the temperature as a function of the energy, at the thermodynamic limit, $N \rightarrow \infty$.

For a given total energy $U = M\epsilon$, the number of possible states is given by the ways one can pick M objects out of N total particles:

$$\Omega(U) = \binom{N}{M} = \frac{N!}{M!(N-M)!}$$

At the limit of large N , we can use Stirling's approximation, $\ln(N!) \approx N \ln N - N$:

$$\ln[\Omega(U)] \underset{N \rightarrow \infty}{\approx} N \ln N - N - M \ln M + M - (N-M) \ln(N-M) + (N-M) - (N-M) \Rightarrow$$

$$\boxed{\ln[\Omega(U)] = N \ln N - M \ln M - (N-M) \ln(N-M)}$$

The temperature as a function of U is obtained from Eq. 9, taking into account that the volume V is not a relevant variable for this system:

$$\beta = \left(\frac{\partial \ln \Omega}{\partial U} \right)_N = \left(\frac{\partial \ln \Omega}{\partial (M\epsilon)} \right)_N \Rightarrow$$

$$\beta = \frac{1}{\epsilon} \left(\frac{\partial}{\partial M} (N \ln N - M \ln M - (N-M) \ln(N-M)) \right)_N = \frac{1}{\epsilon} (-\ln M - 1 + \ln(N-M) + 1) \Rightarrow$$

$$\beta\varepsilon = \ln \frac{N-M}{M} \Rightarrow \beta\varepsilon = \ln \left(\frac{N}{M} - 1 \right) \Rightarrow \boxed{\frac{k_B T}{\varepsilon} = -\frac{1}{\ln \left(\frac{N}{M} - 1 \right)}}$$

The possible values of the normalized energy ratio, $U/U_{\max} = M/N$, range from 0 (every particle is in the ground state) to 1 (every particle is in the excited state). The relationship between M/N and the temperature is shown in Fig. 7. Low values of M/N correspond to low temperatures. Remarkably, the temperature approaches $+\infty$ as $M/N \rightarrow 1/2$ from below, and then returns from negative infinity to just below zero as $M/N \rightarrow 1$.

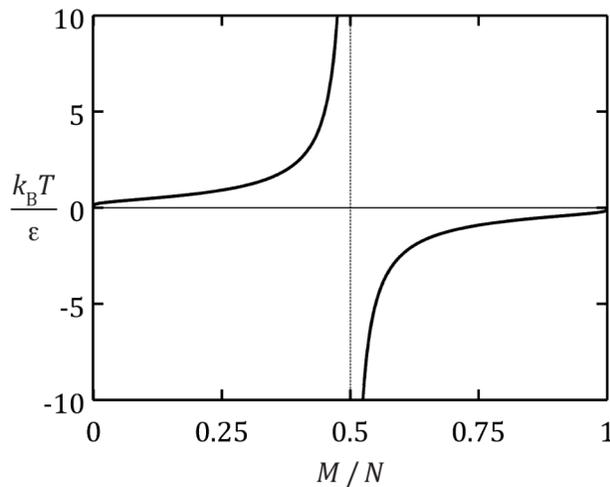


Figure 7 Temperature versus normalized energy.

Do these results make sense? Can negative temperatures exist in nature? It turns out that the existence of negative temperatures is entirely consistent with thermodynamics. The system of this example has a density of states that is a *decreasing* function of the total energy U when more than half the particles are in the excited state. Most physical systems (e.g. molecules free to translate in space) have a monotonic increase in the number of states at higher energies, and thus cannot exist at negative temperatures; however, some spin systems[‡] can closely approximate the system of this example. One needs to realize that negative temperatures are effectively *higher* than all positive temperatures, as energy will flow in the negative \rightarrow positive direction on contact between two systems of opposite sides of the dashed line $M/N = 1/2$, corresponding to $\beta = 0$, or $T = \pm\infty$.

[‡] For a recent example of such a system, see Braun, S. B., *et al.*, "Negative Absolute Temperature for Motional Degrees of Freedom," *Science*, **339**:52-55 (2013).

5. Canonical Ensemble: Constant N , V , and T

Just as in macroscopic thermodynamics, in statistical mechanics we are interested in developing relationships for systems subject to constraints other than constant U , V , and N . In classical thermodynamics, a change of variables is performed through the Legendre transform formalism, which will turn out to be highly relevant here as well. The first example to consider will be a system of fixed volume V and number of particles N , in thermal contact with a much larger reservoir, as shown in Figure 8. Because energy can be transferred between the small system and the reservoir without a significant change in the reservoir's properties, the small system is effectively at constant temperature, that of the reservoir. The set of microstates compatible with constant- NVT conditions is called the *canonical ensemble*. Going from the microcanonical (UVN) to the canonical (NVT) ensemble is akin to taking the first Legendre transformation of the fundamental equation in the entropy representation. Note that the order of variables ($UVN \rightarrow TVN$) is important when performing Legendre transformations; however, convention dictates that the canonical ensemble is referred to as the “ NVT ” ensemble – the ordering of variables is unimportant once a given transformation has been performed.

How do we derive the relative probabilities of microstates for the constant-temperature small system? The total system (small system + reservoir) is under constant- UVN conditions. In the previous sections of this chapter, we suggested that all microstates Ω of the total system, which is at constant energy, volume and number of particles, are equally probable. However, a given microstate ν of the small system with energy U_ν is consistent with many possible microstates of the reservoir – the only constraint is that the energy of the reservoir is $U_R = U - U_\nu$. The number of such microstates for the reservoir is

$$\Omega_R(U_R) = \Omega_R(U - U_\nu)$$

The probability of finding the small system in state ν is proportional to the number of such microstates,

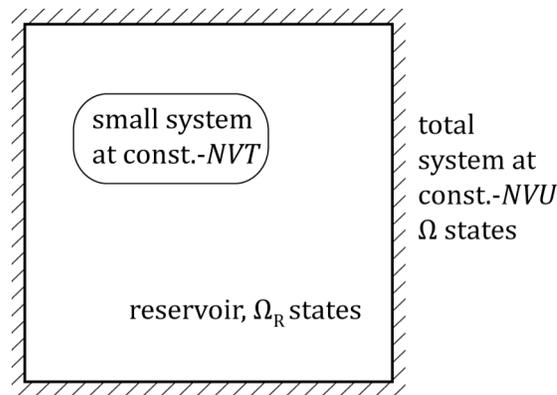


Figure 8 A small system in contact with a large reservoir.

$$\mathcal{P}_v \propto \Omega_R(U - U_v) = \exp\left(\ln\left[\Omega_R(U - U_v)\right]\right) \quad (12)$$

We can Taylor-expand $\ln\Omega_R$ around $\Omega_R(U)$ given that U_v is much smaller than U :

$$\ln\left[\Omega_R(U - U_v)\right] = \ln\left[\Omega_R(U)\right] - U_v \frac{\partial \ln\Omega_R}{\partial U} + \dots \quad (13)$$

Substituting 13 back in Eq. 12 and using Eq. 9, we can incorporate the term involving Ω_{Res} (that does not depend on the microstate v) into the constant of proportionality for \mathcal{P}_v :

$$\mathcal{P}_v \propto \exp(-\beta U_v) \quad \text{at constant } NVT \quad (14)$$

This is a very important result. The probability of each microstate in the canonical ensemble (constant NVT) decreases exponentially for higher energies. The probability distribution of Eq. 14 is known as the *Boltzmann distribution*.

In order to find the absolute probability of each microstate, we need to normalize the probabilities so that their sum is 1. The normalization constant is called the *canonical partition function*, Q and is obtained from a summation over all microstates,

canonical
partition
function

$$Q(N, V, T) = \sum_{\text{all microstates } v} \exp(-\beta U_v) \quad (15)$$

The probability of each microstate can now be written explicitly as an equality:

$$\mathcal{P}_v = \frac{\exp(-\beta U_v)}{Q} \quad \text{at constant } NVT \quad (16)$$

The probability of *all* microstates with a given energy U is the a sum of $\Omega(U)$ equal terms, each at the volume V and number of molecules N of the system:

$$\mathcal{P}(U) = \frac{\Omega(U) \exp(-\beta U)}{Q} \quad \text{at constant } NVT \quad (17)$$

An important implication of these derivations is that the energy of a system at constant temperature is strictly fixed. Instead it fluctuates as the system samples different microstates. This is in direct contrast with the postulate of classical thermodynamics that three independent variables (N , V , and T in this case) fully characterize the state of a system, including its energy. As will be analyzed in detail in the following chapter, fluctuations of quantities such as the energy are present in all finite systems, but their relative magnitude decreases with increasing system size. For macroscopic systems fluctuations are negligible for all practical purposes, except near critical points. In any statistical mechanical ensemble, however, we

need to make a clear distinction between quantities that are strictly constant (constraints, or independent variables in the language of Legendre transformations), and those that fluctuate (derivatives).

Any thermodynamic property \mathcal{B} can be obtained from a summation over microstates of the value of the property at a given microstate times the probability of observing the microstate:

$$\langle \mathcal{B} \rangle = \sum_{\text{all microstates } v} \mathcal{B}_v \cdot \mathcal{P}_v \quad \text{for any ensemble} \quad (18)$$

For example, the ensemble average energy $\langle U \rangle$ in the canonical ensemble is given by:

$$\langle U \rangle = \frac{1}{Q} \sum_v U_v \exp(-\beta U_v) \quad \text{at constant } NVT \quad (19)$$

Let us calculate the temperature derivative of the canonical partition function Q . We have:

$$\frac{\partial \ln Q}{\partial \beta} = \frac{1}{Q} \sum_v \frac{\partial \exp(-\beta U_v)}{\partial \beta} = - \sum_v \frac{U_v \exp(-\beta U_v)}{Q} = -\langle U \rangle \quad (20)$$

The fundamental equation for S/k_B , Eq., has U , V and N as its variables and β , βP , and $-\beta\mu$ as its derivatives. Its first Legendre transform with respect to U is:

$$\frac{S}{k_B} - \beta U = \frac{S}{k_B} - \frac{U}{k_B T} = -\frac{U - TS}{k_B T} = -\beta A \quad (21)$$

This is a function of β , V , and N , with:

$$\left(\frac{\partial(-\beta A)}{\partial \beta} \right)_{V,N} = -U \quad (22)$$

Comparing Eqs. 20 and 22, we see that the former (obtained from statistical mechanics) gives the ensemble average energy, recognizing that the energy fluctuates under constant- NVT conditions. The latter expression, obtained from thermodynamics using Legendre transformations, does not involve averages of fluctuating quantities. At the thermodynamic limit, $N \rightarrow \infty$, we can set the two expressions equal, and obtain a direct connection between the canonical partition function and the first Legendre transform of S/k_B ,

$$-\beta A = \ln Q \quad (23)$$

Eq. 23 relates a thermodynamic quantity, the Helmholtz energy A , to a microscopic one, the partition function Q . This also allows us to confirm the Gibbs en-

tropy formula for the case of a system at constant- NVT , in the thermodynamic limit $N \rightarrow \infty$:

$$\begin{aligned} -\sum_v \mathcal{P}_v \ln \mathcal{P}_v &\stackrel{\text{Eq. 5.14}}{=} -\sum_v \mathcal{P}_v (-\ln \mathcal{Q} - \beta U_v) = \ln \mathcal{Q} \sum_v \mathcal{P}_v + \beta \sum_v \mathcal{P}_v U_v = \\ &= \ln \mathcal{Q} + \beta U = \frac{-A + U}{k_B T} = \frac{S}{k_B} \end{aligned} \quad (24)$$

Given that $\ln \mathcal{Q}$ is the first Legendre transformation of $\ln \Omega$, we can now express all the first derivatives of the canonical partition function \mathcal{Q} , analogous to Eqs. 9–11 for the derivatives of the microcanonical partition function Ω :

$$\left(\frac{\partial \ln \mathcal{Q}}{\partial \beta} \right)_{V,N} = -U \quad (25)$$

$$\left(\frac{\partial \ln \mathcal{Q}}{\partial V} \right)_{T,N} = \beta P \quad (26)$$

$$\left(\frac{\partial \ln \mathcal{Q}}{\partial N} \right)_{T,V} = -\beta \mu \quad (27)$$

These expressions are strictly true only in the thermodynamic limit $N \rightarrow \infty$; for finite systems, for which we need to preserve the distinction between fluctuating and strictly constant quantities, the proper expressions involve ensemble averages; for example, the correct version of Eq. 25 is:

$$\left(\frac{\partial \ln \mathcal{Q}}{\partial \beta} \right)_{V,N} = -\langle U \rangle \quad (28)$$

Example 3 – A system with two states in the NVT ensemble

Consider the system of N distinguishable particles at fixed positions, each of which can exist either in a ground state of energy 0, or in an excited state of energy ε , introduced in Example 2. Determine the mean energy $\langle U \rangle$ as a function of temperature in the canonical ensemble, and compare the result to the microcanonical ensemble calculation of Example 2.

We denote the state of each particle $i = 1, 2, \dots, N$ by a variable l_i which can take the values 0 or 1, denoting the ground or excited state. The total energy is

$$U = \varepsilon \sum_{i=1}^N l_i$$

The partition function in the canonical ensemble is:

$$\ln Q = \ln \sum_{\nu} e^{-\beta U_{\nu}} \Rightarrow Q = \sum_{l_1, l_2, \dots, l_N=0,1} \exp\left(-\beta \sum_{i=1}^N \epsilon l_i\right) = \sum_{l_1, l_2, \dots, l_N=0,1} \prod_{i=1}^N \exp(-\beta \epsilon l_i)$$

Now we can use a mathematical identity that will turn out to be useful in all cases in which a partition function has contributions from many independent particles. The sum contains 2^N terms, each consisting of a product of N exponentials. We can regroup the terms in a different way, in effect switching the order of the summation and multiplication:

$$\sum_{l_1, l_2, \dots, l_N=0,1} \prod_{i=1}^N \exp(-\beta \epsilon l_i) = \prod_{i=1}^N \sum_{l_i=0,1} e^{-\beta \epsilon l_i} = (1 + e^{-\beta \epsilon})^N$$

You can easily confirm that the “switched” product contains the same 2^N terms as before. The final result for the partition function is:

$$\ln Q = N \ln(1 + e^{-\beta \epsilon})$$

The ensemble average energy $\langle U \rangle$ is

$$\langle U \rangle = \left(\frac{\partial \ln Q}{\partial (-\beta)} \right)_{N,V} = N \frac{\epsilon e^{-\beta \epsilon}}{1 + e^{-\beta \epsilon}} = \frac{N \epsilon}{1 + e^{\beta \epsilon}}$$

The microcanonical ensemble result can be written as:

$$\ln \left(\frac{N}{M} - 1 \right) = \frac{\epsilon}{k_B T} = \beta \epsilon \Rightarrow \frac{N}{M} - 1 = e^{\beta \epsilon} \Rightarrow M = \frac{N}{1 + e^{\beta \epsilon}} \Rightarrow M \epsilon = U = \frac{N \epsilon}{1 + e^{\beta \epsilon}}$$

The only difference is that in the canonical ensemble the energy is a fluctuating (rather than a fixed) quantity. Note also that the canonical ensemble derivation does not entail any approximations; the same result is valid for any N . By contrast, the microcanonical energy was obtained through the use of Stirling’s approximation, valid as $N \rightarrow \infty$. Small differences between the two ensembles are present for finite N .

6. Generalized Ensembles and Legendre Transforms

The previous two sections illustrated how, starting from the fundamental equation in the entropy representation and Boltzmann’s entropy formula, one can obtain relationships between macroscopic and microscopic quantities, first in the UVN (microcanonical) ensemble, with key quantity the density of states, Ω . A first Legendre transformation of U to T resulted in the canonical ensemble, with partition function Q . This process can be readily generalized to obtain relationships between microscopic and macroscopic quantities and partition functions for a system under arbitrary constraints.

In our derivation, we start from the multicomponent version of the fundamental equation in the entropy representation,

$$\frac{dS}{k_B} = d \ln \Omega = \beta dU + \beta P dV - \sum_{i=1}^n \beta \mu_i dN_i \quad (29)$$

We have already seen the first Legendre transformation:

$$\mathbf{y}^{(1)} = \frac{S}{k_B} - \beta U = \beta(TS - U) = -\beta A \quad (30)$$

The relationships between the original function and the first transform are depicted in the table below.

$\mathbf{y}^{(0)} = S/k_B = \ln \Omega$		$\mathbf{y}^{(1)} = -A/k_B T = \ln \mathcal{Q}$	
Variable	Derivative	Variable	Derivative
U	$1/(k_B T) = \beta$	$1/(k_B T) = \beta$	$-U$
V	$P/(k_B T) = \beta P$	V	$P/(k_B T) = \beta P$
N_i	$-\mu_i/(k_B T) = -\beta \mu_i$	N_i	$-\mu_i/(k_B T) = -\beta \mu_i$

These relationships link microscopic to macroscopic properties and are strictly valid at the thermodynamic limit, $N \rightarrow \infty$. One should keep in mind that in each ensemble, the variables of the corresponding transform are strictly held constant, defining the external constraints on a system, while the derivatives fluctuate – they take different values, in principle, for each microstate of the ensemble.

Probabilities of microstates in two statistical ensembles have already been derived. Microstate probabilities are all equal in the microcanonical ensemble, from the basic postulate of statistical mechanics; they are equal to the Boltzmann factor, $\exp(-\beta U)$, normalized by the partition function for the canonical ensemble (Eq. 16). Note that the factor $-\beta U$ that appears in the exponential for microstate probabilities in the canonical ensemble is exactly equal to the difference between the basis function and its first Legendre transform, $-\mathcal{X}_1 \xi_1$.

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, pressure, chemical potential, etc. In general, the k^{th} transform of the basis function is:

$$\mathbf{y}^{(0)} = \frac{S}{k_B} \quad ; \quad \mathbf{y}^{(k)} = \frac{S}{k_B} - \xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k \quad (31)$$

where ξ_i is the derivative of $\mathbf{y}^{(0)}$ with respect to variable x_i . The probability of a microstate in the ensemble corresponding to the k^{th} transform is given by

$$\mathcal{P}_v \propto \exp(-\xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k) \quad (32)$$

where the variables x_i and derivatives ξ_i refer to the original function $\mathbf{y}^{(0)}$. The normalization factor (partition function) of the ensemble corresponding to the transformed function, $\mathbf{y}^{(k)}$ is:

$$\Xi = \sum_{\text{all microstates } v} \exp(-\xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k) \quad (33)$$

Using the partition function Ξ , the probability \mathcal{P}_v can be written as an equality:

$$\mathcal{P}_v = \frac{\exp(-\xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k)}{\Xi} \quad (34)$$

As was the case for the canonical ensemble, the partition function Ξ is simply related to the transformed function, $\mathbf{y}^{(k)}$:

$$\ln \Xi = \mathbf{y}^{(k)} \quad (35)$$

Example 4 – Gibbs Entropy Formula

The Gibbs entropy formula,

$$\underline{S} = -k_B \sum_v \mathcal{P}_v \ln \mathcal{P}_v$$

was derived earlier for the microcanonical ensemble. Show that this relationship is valid for all statistical ensembles.

We use the expression for the probability of microstates, \mathcal{P}_v , in a generalized ensemble, Eq. 34:

$$\sum_v \mathcal{P}_v \ln \mathcal{P}_v = \sum_v \mathcal{P}_v \ln \frac{\exp(-\xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k)}{\Xi} = \sum_v \mathcal{P}_v (-\xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k - \ln \Xi)$$

Now recall that the variables for the k -th transform are $\xi_1, \xi_2, \dots, \xi_k, x_{k+1}, \dots, x_{n+2}$, which are strictly constant in the corresponding ensemble, while the derivatives, $x_1, x_2, \dots, x_k, \xi_{k+1}, \dots, \xi_{n+2}$ fluctuate. We can rewrite the equation above taking this into account:

$$\begin{aligned}
\sum_{\nu} \mathcal{P}_{\nu} (-\xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k - \ln \Xi) &= \\
= -\xi_1 \sum_{\nu} \mathcal{P}_{\nu} x_1 - \xi_2 \sum_{\nu} \mathcal{P}_{\nu} x_2 - \dots - \xi_k \sum_{\nu} \mathcal{P}_{\nu} x_k - \ln \Xi &= \\
= -\xi_1 \langle x_1 \rangle - \xi_2 \langle x_2 \rangle - \dots - \xi_k \langle x_k \rangle - \ln \Xi &
\end{aligned}$$

From Eqs. 31 and 35,

$$\ln \Xi = \mathcal{Y}^{(k)} = \frac{S}{k_B} - \xi_1 x_1 - \xi_2 x_2 - \dots - \xi_k x_k$$

At the thermodynamic limit, $N \rightarrow \infty$, so there is no distinction between ensemble averages and thermodynamic properties, $\langle x_i \rangle \equiv x_i$. Replacing $\ln \Xi$ and simplifying,

$$\sum_{\nu} \mathcal{P}_{\nu} \ln \mathcal{P}_{\nu} = -\frac{S}{k_B}, \text{ QED}$$

Example 5 – Grand Canonical (μVT) Ensemble

The grand canonical (constant- μVT) is frequently used in computer simulations. Derive the partition function, probability of microstates, and derivative relationships in this ensemble for a 1-component system.

Starting from the fundamental equation in the entropy representation with ordering of variables $\mathcal{Y}^{(0)} = S(U, N, V) / k_B = \ln \Omega$, the grand canonical ensemble partition function corresponds to:

$$\ln \Xi = \mathcal{Y}^{(2)} = \frac{S}{k_B} - \beta U + \beta \mu N = \frac{TS - U + \mu N}{k_B T}$$

The microstates possible in this ensemble include all possible particle numbers from 0 to ∞ , and all possible energy levels.

The Euler-integrated form of the fundamental equation is $U = TS - PV + \mu N$, so that the partition function of the grand canonical ensemble can be linked to the following thermodynamic property combination:

$$\ln \Xi = \frac{PV}{k_B T} = \beta PV$$

$\mathcal{Y}^{(0)} = S/k_B = \ln\Omega$		$\mathcal{Y}^{(2)} = \beta PV = \ln\Xi$	
Variable	Derivative	Variable	Derivative
U	β	β	$-U$
N	$-\beta\mu$	$-\beta\mu$	$-N$
V	βP	V	βP

For example, the average number of molecules in the system is given by:

$$\left(\frac{\partial \ln \Xi}{\partial (-\beta\mu)} \right)_{\beta, V} = -\langle N \rangle \Rightarrow k_B T \left(\frac{\partial \ln \Xi}{\partial \mu} \right)_{\beta, V} = \langle N \rangle$$

The probability of microstates in this ensemble is:

$$\mathcal{P}_v = \frac{\exp(-\beta U_v + \beta\mu N_v)}{\Xi},$$

$$\text{where } \Xi = \sum_v \exp(-\beta U_v + \beta\mu N_v) = \sum_{N=1}^{\infty} Q(N, V, T) \exp(\beta\mu N)$$

Example 6 – Constant-pressure (NPT) Ensemble

The constant-pressure (NPT) ensemble is also frequently used in computer simulations. Derive the partition functions, probability of microstates, and derivative relationships in this ensembles for a 1-component system.

We start from the fundamental equation in the entropy representation with ordering of variables $\mathcal{Y}^{(0)} = S(U, V, N)/k_B$, and obtain the second transform:

$$\ln \Xi = \mathcal{Y}^{(2)} = \frac{S}{k_B} - \beta U - \beta PV = \frac{TS - U - PV}{k_B T} = -\beta\mu N = -\beta G$$

The microstates possible in this ensemble include all possible volumes from 0 to ∞ , and all possible energy levels.

The derivative table is shown on the next page.

For example, the average volume is given by:

$$\left(\frac{\partial \ln \Xi}{\partial (\beta P)} \right)_{\beta, N} = -\langle V \rangle \Rightarrow k_B T \left(\frac{\partial \ln \Xi}{\partial P} \right)_{\beta, N} = \langle V \rangle$$

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

The probability of microstates in this ensemble is:

$$\mathcal{P}_v = \frac{\exp(-\beta U_v - \beta P V_v)}{\Xi}, \text{ where } \Xi = \sum_v \exp(-\beta U_v - \beta P V_v)$$