1 Foundations of statistical physics

1.1 Density operators

In quantum mechanics we assume that the state of a system is described by some vector $|\Psi\rangle$ belonging to a Hilbert space $\mathcal{H}$. If we know the initial state then we can use the Schrödinger equation to work out the state at a subsequent time (if we are working in the Schrödinger picture). We can then use that to calculate the probability that measuring an observable A gives a particular result.

In statistical physics, the individual quantum states of a system are called *microstates*. For a macroscopic system, consisting of $N \sim 10^{23}$ particles, there is no way we could possibly know which microstate the system is in. Instead we refer to the *macrostate* of a system, where we specify a few macroscopically observable quantities e.g. energy, volume, temperature etc. In general, there will be many different microstates corresponding to a particular macrostate e.g. switching the position of a couple of molecules does not change the macroscopic behaviour of a system. Let $|\Psi_i\rangle$ be the microstates corresponding to a particular macrostate. We assume these microstates to be normalized but not necessarily orthogonal. For a given macrostate, not all of these microstates need be equally likely, so let P_i be the probability that the state is $|\Psi_i\rangle$. Note that this probability has got nothing to do with quantum uncertainty, it simply reflects our ignorance of the system.

A standard way of thinking about this is using an *ensemble* of $W \gg 1$ (fictitious) copies of our system. Each copy is in some definite microstate that is consistent with the known macrostate. If n_i denotes the number of members of the ensemble that are in the state $|\Psi_i\rangle$ then we say that the probability that the actual system is in the state $|\Psi_i\rangle$ is $P_i = n_i/W$. So specifying the probability distribution P_i is equivalent to specifying an ensemble.

Mathematically, we describe a macrostate using a *density operator*

$$\rho = \sum_i P_i |\Psi_i\rangle\langle\Psi_i| \tag{1.1}$$

You met the concept of a density operator last term in *Principles of Quantum Mechanics*. The term "density operator" is a bit stupid (what is it the density of?) but it is standard so we are stuck with it. The components of a density operator w.r.t. a basis are called a *density matrix*.

From the density operator we can calculate the probability that an observable A takes a particular value. Let $|a\rangle$ denote the normalized eigenstate of A with eigenvalue

a. Then, using the rules of conditional probability

$$\begin{aligned} P(A=a) &= \sum_i P(A=a\big|\text{state is } |\Psi_i\rangle)P(\text{state is } |\Psi_i\rangle) \\ &= \sum_i |\langle a|\Psi_i\rangle|^2 P_i = \sum_i \langle a|\Psi_i\rangle\langle\Psi_i|a\rangle P_i = \langle a|\rho|a\rangle \end{aligned} \tag{1.2}$$

Some simple properties of ρ are obvious: ρ is hermitian and positive (semi)-definite. To see the latter note that, for any state $|\Psi\rangle$,

$$\langle\Psi|\rho|\Psi\rangle = \sum_i P_i|\langle\Psi|\Psi_i\rangle|^2 \geq 0 \tag{1.3}$$

If we let $\{|i\rangle\}$ be an orthonormal basis for $\mathcal{H}$ then we can calculate the trace of ρ over $\mathcal{H}$:

$$\begin{aligned} \mathrm{Tr}(\rho) &\equiv \sum_i \rho_{ii} = \sum_i \langle i|\rho|i\rangle = \sum_{i,j} P_j\langle i|\Psi_j\rangle\langle\Psi_j|i\rangle = \sum_{i,j} P_j\langle\Psi_j|i\rangle\langle i|\Psi_j\rangle \\ &= \sum_j P_j\langle\Psi_j|\Psi_j\rangle = \sum_j P_j = 1 \end{aligned} \tag{1.4}$$

Based on the above, we now make a definition:

Definition. A density operator is a Hermitian, positive semi-definite operator ρ such that $\mathrm{Tr}(\rho) = 1$.

Given a density operator ρ let $|i\rangle$ be the normalized eigenvectors of ρ, with eigenvalues p_i. Since ρ is Hermitian, we can assume that $\{|i\rangle\}$ form an orthonormal basis. Positive semi-definiteness implies $p_i \geq 0$ and $\mathrm{Tr}(\rho) = 1$ implies $\sum_i p_i = 1$. Furthermore the density operator takes then form

$$\rho = \sum_i p_i|i\rangle\langle i| \tag{1.5}$$

which is the same form as (1.1). Hence we can interpret p_i as the probability that, in a macrostate described by ρ, the microstate is $|i\rangle$. The eigenvalues p_i are not the same as the P_i unless the states $|\Psi_i\rangle$ are orthogonal.

Note that (using $p_i^2 \leq p_i$)

$$\mathrm{Tr}(\rho^2) = \sum_i p_i^2 \leq 1 \tag{1.6}$$

with equality if, and only if, $p_i = \delta_{ij}$ for some j, in which case we know that the microstate is definitely $|j\rangle$. So $\mathrm{Tr}(\rho^2) = 1$ if, and only if, we know the microstate of the system. Under such circumstances, we say that ρ describes a *pure state*. In the general case $\mathrm{Tr}(\rho^2) < 1$, the system could be in one of several different states; we then say that ρ describes a *mixed state*.

Let's now discuss the dynamics of mixed states. Consider a system which at time $t = 0$ is in a mixed state described by the density operator $\rho(0)$. Let $\{|i;0\rangle\}$ be the normalized eigenvectors of $\rho(0)$ so $\rho(0)$ has the form (1.5) with $|i\rangle$ replaced by $|i;0\rangle$. Now, at time t, the state $|i;0\rangle$ evolves to $e^{-iHt/\hbar}|i;0\rangle$ (where H is the Hamiltonian operator) and so the density operator at time t is

$$\rho(t) = \sum_i p_i e^{-iHt/\hbar}|i;0\rangle\langle i;0|e^{iHt/\hbar} \tag{1.7}$$

From this it follows that ρ satisfies the equation of motion (exercise)

$$i\hbar\frac{\partial\rho}{\partial t} = [H, \rho] \tag{1.8}$$

This is the Schrödinger equation for mixed states.

In statistical physics we will be interested in systems that have "settled down to equilibrium". By this, we mean that the probability of any observable taking a particular value is independent of time. Equation (1.2) implies that $\rho(t)$ must be independent of time so, from (1.8), ρ must commute with H. So a macrostate described by a density operator ρ is an equilibrium state iff ρ commutes with H. This is true iff we can simultaneously diagonalize ρ and H. Hence, in equilibrium, the eigenstates $|i\rangle$ of ρ can be assumed to be energy eigenstates, with energies E_i.

In equilibrium, ρ must commute with H. How restrictive is this? Obviously any function of H will commute with H. Furthermore, if the system has appropriate continuous symmetries then there exist other operators that commute with H. For example, if the system is invariant under translations then the momentum operator will commute with H. If it is invariant under rotations then the angular momentum operator will commute with H. So, in equilibrium, the density operator must be a function of these various operators. In practice, we are usually interested in systems for which boundary conditions break translational and rotational symmetry e.g. a gas in a cubic box. In such a situation, any operator which commutes with H must be a function of H. So, in equilibrium, the density operator must be a function of the Hamiltonian operator H.

1.2 The microcanonical ensemble

We've seen how a density operator describes a system about which we have imperfect knowledge. But *which* density operator should we use? In other words, what determines the probabilities p_i?

The fundamental assumption of statistical physics. For an isolated system, all states consistent with our partial knowledge of the system are equally probable.

In other words, in the absence of any other information, we simply assume that the non-vanishing p_i all take the the same value. Here "partial knowledge" means that we know something about the system but we don't know enough to determine its quantum state exactly. "Isolated" means that the system is not interacting with anything else. Think of a gas in a vacuum flask.

In this course we will be studying the properties of systems in equilibrium. By this we mean that the probabilities of measurements do not depend on time, i.e, the density matrix is time independent. Many systems in nature appear to "settle down" to equilibrium over time. To achieve this, some damping mechanism e.g. friction is often important (a harmonic oscillator without damping would never settle down to rest). However, once equilibrium has been achieved, a system will remain in equilibrium if it is completely isolated. It is the equilibrium of isolated systems that we will focus on.

From what we said above, the density matrix of an equilibrium state must commute with the Hamiltonian hence the eigenstates $|i\rangle$ of ρ must be energy eigenstates. Assume that we know that the energy of an isolated system lies in the range $(E, E+\delta E)$. Then, from our fundamental assumption we know that the density operator has the form (1.5) where $|i\rangle$ are energy eigenstates with energies E_i and

$$p_i = \begin{cases} c \text{ if } E_i \in (E, E+\delta E) \\ 0 \text{ else} \end{cases} \tag{1.9}$$

for some constant c. From the normalization condition $\sum_i p_i = 1$ we have

$$c = \frac{1}{\Omega} \tag{1.10}$$

where Ω is the number of energy eigenstates with $E_i \in (E, E+\delta E)$. Here we are assuming that the energy eigenstates are discrete so we can count them. In practice, we can arrange that this is true by putting our system in a box of fixed volume V: you are familiar e.g. from the infinite square well that this renders the energy levels discrete. If V is large compared to microscopic length scales then the spacing of the energy levels will be very small, so our experimental accuracy δE will be much larger than the spacing between energy levels.

The above probability distribution is called the *microcanonical distribution*. The corresponding ensemble is called the *microcanonical ensemble*. Note that the microcanonical ensemble is appropriate for an *isolated* system. We'll see below how other ensembles describe equilibrium states of certain kinds of non-isolated system.

Note that Ω depends on E and δE. It also depends on N, the number of particles, and V (the volume) because the E_i depend on V. So we write $\Omega(E, V, N; \delta E)$. If there are other experimentally variable macroscopic quantities present in the system (e.g. a magnetic field) then Ω will also depend on them. We'll refer to such quantities as "external parameters".

1.3 Boltzmann entropy

Consider an isolated system in equilibrium, described with the microcanonical ensemble. Define the *Boltzmann entropy*

$$S = k \log \Omega \tag{1.11}$$

where k is a positive constant which we'll keep arbitrary for now.

In principle S depends on E, V, N and δE. However, the dependence on δE is negligible in practice. To see this, let $G(E, V, N)$ denote the number of energy eigenstates with energy less than E. We then have, for $\delta E \ll E$, (suppressing the dependence on V, N)

$$\Omega(E; \delta E) = G(E + \delta E) - G(E) \approx \frac{\partial G}{\partial E}\delta E \tag{1.12}$$

and so

$$S \approx k \log\left(E\frac{\partial G}{\partial E}\right) + k \log\frac{\delta E}{E} \tag{1.13}$$

For a macroscopic system, the second term is absolutely tiny compared to the first term. For example, for a system of N particles we would typically expect, $G \sim cE^{\alpha N}$ for some constants c and α where α is dimensionless and of order 1 (examples sheet 1). Now $E\partial G/\partial E = \alpha N G$ and

$$S \approx k \log G + k \log(\alpha N) + k \log\frac{\delta E}{E} \tag{1.14}$$

Note that the first term above is proportional to N. For $N \sim 10^{23}$ this is much much larger than the second term, which scales as $\log N$. For the final term to be comparable to the first term we'd need $\delta E/E \sim e^{-10^{23}}$, and no practical measurement could possibly achieve such fantastic accuracy! So, for a macroscopic system, the second and third terms above are always negligible compared to the first term. We can therefore neglect the dependence of the entropy on δE and write simply $S(E, V, N)$. Of course if there are additional external parameters (e.g. magnetic field) then S will depend on them too.

A final comment: we've seen that $S \sim kN$, hence $\log \Omega \sim N$. For $N \sim 10^{23}$ this implies that Ω is a fantastically colossal, gigantic, enormous number!

1.4 Thermal equilibrium

Consider two macroscopic systems which are coupled together weakly so that they are free to exchange energy but the coupling does not significantly change the energy levels of each system. We assume that the volume and number of particles of each system is fixed and that the combined system is isolated, i.e., not interacting with anything else. An example of such a set-up would be an insulating flask divided into two regions by a very thin wall, with gas on each side of the wall. The two systems are the two separate regions of gas, and they can exchange energy by molecules of the first gas colliding with the wall molecules, which then collide with molecules of the second gas. However, the volume and number of particles of each gas is fixed. The *total* energy is also fixed.

Let V_1 and V_2 be the volumes of the two systems and N_1, N_2 be the numbers of particles in each system. These quantities cannot vary so we will often not write them in the equations below. Let E_1 and E_2 be the energies of the two systems.[1] Since the systems can exchange energy, E_1 and E_2 can vary so only the total energy $E = E_1 + E_2$ is fixed.

Let's assume that the combined system is in statistical equilibrum. By the fundamental assumption it is described by the microcanonical ensemble. We can now ask: what is the probability that E_1 takes a particular value? Let $\Omega_1(E_1)$ be the number of microstates of the first system with energy in $(E_1, E_1 + \delta E)$ and define $\Omega_2(E_2)$ similarly for the second system. The number of microstates for the combined systems such that the first system has energy in $(E_1, E_1 + \delta E)$ and the second has energy in $(E_2, E_2 + \delta E)$ is then $\Omega_1(E_1)\Omega_2(E_2) = \Omega_1(E_1)\Omega_2(E - E_1)$.

In the microcanonical ensemble any given microstate with energy in the range $(E, E + \delta E)$ has probability $1/\Omega(E)$ where here $\Omega(E)$ refers to the combined system. Hence the probability that the first system has energy in the range $(E_1, E_1 + \delta E)$, with the second in the range $(E - E_1, E - E_1 + \delta E)$ is

$$p(E_1) = \frac{\Omega_1(E_1)\Omega_2(E - E_1)}{\Omega(E)} \tag{1.15}$$

As we saw above, $\Omega_1(E_1)$ is an incredibly rapidly increasing function of E_1. Similarly $\Omega_2(E - E_1)$ is an incredibly rapidly decreasing function of E_1. This implies that their product will have a very sharply defined maximum at some value E_1^*, and decrease incredibly rapidly moving away from this maximum. Hence E_1^* is overwhelmingly the most probable value of E_1: the probability of observing a different value of E_1 is negligible. On examples sheet 1 it is shown that the probability is negligible outside a

---

[1] Hopefully you will not confuse E_1 and E_2 with the energy eigenvalues of a single system, which we previously denoted E_i.

region of size $E/\sqrt{N}$ around E_1^*, where $N_1, N_2 \sim N$, the total number of particles. For $N \sim 10^{23}$ we see that indeed the maximum is very sharp. For a macroscopic object, the width of this region is usually much less than δE, the accuracy with which we can measure E.

To determine the value of E_1^* we just demand that the derivative w.r.t. E_1 vanishes. It's easiest to do this by taking a log first, giving

$$0 = \left(\frac{\partial S_1}{\partial E_1}\right)_{E_1=E_1^*} - \left(\frac{\partial S_2}{\partial E_2}\right)_{E_2=E-E_1^*} \tag{1.16}$$

here we write a *partial* derivative w.r.t. E_1 because S_1 depends also on V_1 and N_1, which are held fixed when we take the derivative, and similarly for S_2. We see that, in equilibrium, it is overwhelmingly likely that the two systems will have equal values for $(\partial S/\partial E)_{V,N}$.

1.5 The second law

We can conclude more from the above discussion. Since $p(E_1)$ is negligible unless $E_1 = E_1^*$ (to accuracy δE), it follows from (1.15) and the fact that probabilities must sum to 1 that we have

$$\Omega(E) \approx \Omega_1(E_1^*)\Omega_2(E - E_1^*) \tag{1.17}$$

Hence, in equilibrium, the entropy of the combined system is

$$S(E) = S_1(E_1^*) + S_2(E - E_1^*) \tag{1.18}$$

This is simply the sum of the entropies of the individual systems: entropy is *additive*.

Notice that E_1^* was defined to maximize $\Omega_1(E_1)\Omega_2(E - E_1)$ and hence to maximize

$$S(E, E_1) \equiv k \log\left[\Omega_1(E_1)\Omega_2(E - E_1)\right] = S_1(E_1) + S_2(E - E_1) \tag{1.19}$$

Hence $S(E) = S(E, E_1^*)$, the maximum value of $S(E, E_1)$. Let's now discuss the interpretation of $S(E, E_1)$ for $E_1 \neq E_1^*$. Consider a situation in which our two systems are initially non-interacting, with each system separately in equilibrium with energies E_1 and $E - E_1$ respectively. So system 1 has entropy $S_1(E_1)$ and system 2 has entropy $S_2(E_2)$. Now starting at time $t = 0$ we let the two systems interact. (In our example of two gases in a box, the "non-interacting" set-up could correspond to an insulating wall as well as the thin wall; we then remove the insulating wall to let the gases interact.) If $E_1 \neq E_1^*$ then, from the above analysis, we know that the combined system is not in equilibrium at $t = 0$. Now so far, we have only defined entropy for systems in equilibrium. But now it is very natural to *define* the entropy of the non-equilibrium state at $t = 0$ to be simply $S(E, E_1)$ since this is the sum of the entropies of the two individual systems.

What happens next? From the above analysis, we know that for $E_1 \neq E_1^*$, the macrostate at $t = 0$ is a very improbable state. So it "seems very likely" that the combined system will evolve to more probable macrostates and eventually "settle down" to the overwhelmingly most probably macrostate, namely the equilibrium state with $E_1 = E_1^*$. Notice in this process, the entropy $S(E, E_1)$ *increases* and eventually achieves its maximum value $S(E)$ when equilibrium is reached. This is a special case of the *second law of thermodynamics*.

Second Law of Thermodynamics. The entropy of an isolated system increases (or remains the same) in any physical process. In equilibrium, the entropy attains its maximum value.

Note the important word "isolated": if the system is not isolated, i.e., if it interacts with an environment, then its entropy can decrease. However, the *total* entropy of the system and its environment can only increase.

The second law lets us divide physical processes into two classes. For an isolated system in equilibrium, a process is *reversible* if it does not change the entropy. It is *irreversible*, if the entropy increases. If the system is not isolated, the same definitions can be made using the total entropy of the system and its environment.

This statement of the second law is a bit vague because we have not defined the entropy of a non-equilibrium macrostate of a general system. In fact no precise definition exists! The best one can do is to generalize the situation just discussed (with 2 interacting systems) and assume that the system can be chopped up into subsystems which are each large compared to the scale of microscopic physics, and such that each part is approximately in equilibrium (this is called *local* equilibrium). Think for example of fluid mechanics: the fluid can be moving in a fairly complicated way and yet still described by just a few functions e.g. density, velocity etc. The idea is that in this case the chunks are small enough that these functions are approximately constant on each chunk. One then defines the total entropy as the sum of the entropies of each part. However it's not possible to do this in a completely unambiguous way: it depends on how you do the chopping up into chunks.

A more precise version of the second law refers only to equilibrium states, for which the definition of entropy is unambiguous:

Second Law of Thermodynamics v2. Consider an isolated system which is initially in equilibrium, then undergoes some change of its external parameters, and settles down to equilibrium again. Then the entropy of the final state will be greater than (or equal to) the entropy of the initial state.

An example of such a process is the *free expansion of a gas into a vacuum.* Consider a box separated into two regions by a rigid partition. On one side of the partition is a gas in equilibrium, on the other side there is vacuum. Assume there is a valve in the partition, with a timer set to open the valve at some prescribed time. This is an isolated system. When the valve opens, the gas expands into the vacuum and then settles down to equilibrium again, with gas on both sides of the partition. In this process, the volume of the gas (an external parameter) changes. The initial and final states are equilibrium states, and the final state will have greater entropy than the initial state, so the change is irreversible.

The second law is motivated by probabilistic arguments: low entropy states are "atypical" macrostates with very low probability so we expect such states to evolve to "typical" macrostates with high probability and high entropy. The probabilistic nature of this reasoning suggests that the second law may not be an absolute law, and that there should be some small probability that one could observe a violation of the second law. This is indeed true, but for a macroscopic system, the probability is so small that one would have to wait many times the age of the Universe in order to observe such an unlikely event.

1.6 *Time asymmetry and the second law*

Note that in our two system example above we did not *derive* the second law because we invoked the *assumption* that the system evolves to a final state that maximizes the probability (it "seems very likely that..."). The second law is not a consequence of something else, it has the status of an *axiom* of statistical mechanics.

There have been attempts to derive the second law from the microscopic laws of physics. However this does not work because most microscopic laws of physics are invariant under time-reversal: if one is shown a film of particles moving according to microscopic laws then the same film shown backwards would also be consistent with those laws.[2] Therefore it is not possible to deduce a time-asymmetric statement such as the second law purely from microscopic laws.

To see the problem, consider applying time-reversal to the probabilistic motivation for the second law. Given a low entropy macrostate we can ask: where did it come from? If we evolve *backwards* in time, then what is the most probable state of the system at some earlier time? Well, "more probable", i.e., "more typical" states should correspond to macrostates of higher entropy. So at earlier time, the macrostate should have had higher entropy. We've deduced that the low entropy state formed from an

---

[2] Actually there are subatomic particles called kaons which do not possess this time-reversal symmetry. But these exotic particles are not present in most systems of interest in statistical physics.

initially higher entropy state, in violation of the second law! Note that this argument is possible because of the time-reversal symmetry of the microscopic laws, which means there is nothing in these laws which selects a preferred direction of time. Hence if the probabilistic argument is valid in one direction of time then it is equally valid in the opposite direction.

The resolution of this apparent paradox is that it neglects initial conditions: in the real-world, low entropy states occur because someone has prepared them in a lab! So we have to consider the entropy of the combined system plus its environment, which includes the experimentalist. This experimentalist is an example of a low-entropy phenomenon, namely life. Now one can ask what makes low entropy life possible. The answer is that the Sun supplies us with energy in a low entropy form, namely high energy photons, with the energy re-emitted by the Earth in a higher entropy form, namely a greater number of low-energy photons. This makes the total entropy increase, even accounting for life on Earth. Now one can ask where did the Sun come from etc. Tracing things back to earlier and earlier times, the ultimate reason that the second law is true appears to originate in cosmology: the initial condition of the Universe at the Big Bang was very special, i.e., very low entropy. The entropy of the Universe has been increasing ever since.

1.7 Temperature

Let's return to the discussion of two systems. We showed that, in equilbrium, the energy of system 1 is overwhelmingly likely to be $E_1 = E_1^*$ where E_1^* is determined by the condition (1.16). Hence equilibrium corresponds to equal values of $\partial S/\partial E$ for the two systems. We therefore give this quantity a name: (inverse) *temperature*. In general, the temperature of a system is defined by

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{V,N} \tag{1.20}$$

where the subscript denotes that V, N are held fixed when we take the partial derivative.

We have shown that a necessary condition for equilibrium of a system composed of two macroscopic bodies is that they have equal temperature:

$$T_1 = T_2 \tag{1.21}$$

We'll see later that this definition of T agrees with other standard definitions of temperature (e.g. using properties of a dilute gas) provided one chooses the constant k appropriately. The obvious choice would be $k = 1$, which would correspond to measuring temperature in the same units used to measure energy (i.e. Joules). Unfortunately, for historical reasons a different unit was chosen for temperature, namely the *Kelvin*. Because of this, in order for our definition of T to agree with earlier definitions we must take

$$k = 1.38 \times 10^{-23} \mathrm{JK}^{-1} \tag{1.22}$$

This constant is called *Boltzmann's constant*.

Assume that we prepare an initial state such that each system is individually in equilibrium, but they are not in equilibrium with each other, so we have $T_1 \neq T_2$. A short time later, the systems will have exchanged energy and the entropy has changed by

$$\delta S = \delta S_1 + \delta S_2 = \frac{\delta E_1}{T_1} + \frac{\delta E_2}{T_2} = \left(\frac{1}{T_1} - \frac{1}{T_2} \right) \delta E_1 \tag{1.23}$$

where the first equality follows from our definition of entropy in the non-equilibrium situation, the second equality is the definition of temperature, and the final equality uses the fact that the total energy is fixed so $\delta E_2 = -\delta E_1$. The second law of thermodynamics says that $\delta S \geq 0$. Hence if $T_1 > T_2$ then $\delta E_1 < 0$: energy flows from the system with the higher temperature to the system with the lower temperature, i.e., from the "hotter" system to the "colder" system.

For a system with $G \approx cE^{\alpha N}$ we have from (1.13)

$$S \approx k \log G = k\alpha N \log E + \text{const.} \tag{1.24}$$

and hence

$$kT = \frac{E}{\alpha N} \tag{1.25}$$

so kT is a measure of the energy per particle. We'll sharpen this statement later.

1.8 Heat capacity

Our definition of temperature gives $T = T(E, V, N)$. Inverting this relation gives $E = E(T, V, N)$, and we define the *heat capacity at constant volume*

$$C_V = \left(\frac{\partial E}{\partial T} \right)_{V.N} \tag{1.26}$$

This quantity is important because it can be measured experimentally by seeing how much energy is required to achieve a certain change in temperature. For a system with $G \approx cE^{\alpha N}$ we have $C_V \approx \alpha N$, i.e., constant heat capacity proportional to the number of particles. (Often one works with the *specific heat capacity*, which is the heat capacity of an object divided by its mass, and hence independent of the number of particles.)

Note that if we write $S = S(T, V, N)$ then we have

$$\left(\frac{\partial S}{\partial T}\right)_{V,N} = \left(\frac{\partial S}{\partial E}\right)_{V,N} \left(\frac{\partial E}{\partial T}\right)_{V,N} = \frac{C_V}{T} \tag{1.27}$$

and hence

$$S(T_2, V, N) - S(T_1, V, N) = \int_{T_1}^{T_2} \frac{C_V(T, V, N)}{T} dT \tag{1.28}$$

so by measuring the heat capacity as a function of temperature we can determine entropy *differences* experimentally.

Finally note that

$$\left(\frac{\partial^2 S}{\partial E^2}\right)_{V,N} = \left(\frac{\partial (T^{-1})}{\partial E}\right)_{V,N} = -\frac{1}{T^2 C_V} \tag{1.29}$$

Now apply this to our two systems in equilibrium. Recall that the second law requires that the non-equilbrium entropy $S(E, E_1) = S_1(E_1) + S_2(E - E_1)$ is maximized, which led us to the condition $E = E_1^*$ and hence $T_1 = T_2$. But $E_1 = E_1^*$ is only sufficient for the entropy to be extremized, we did not check that the extremum was a maximum! To do this we need to look at the second derivative:

$$\frac{\partial^2 S}{\partial E_1^2} = \left(\frac{\partial^2 S_1}{\partial E_1^2}\right)_{E_1 = E_1^*} + \left(\frac{\partial^2 S_2}{\partial E_2^2}\right)_{E_2 = E - E_1^*} = -\frac{1}{T_1^2 C_{V1}} - \frac{1}{T_2^2 C_{V2}} = -\frac{1}{T_1^2}\left(\frac{1}{C_{V1}} + \frac{1}{C_{V2}}\right) \tag{1.30}$$

For the extremum to be a maximum we need the RHS to be negative. This is indeed the case when both systems have positive heat capacity.

Negative heat capacity would indicate instability in a homogeneous system. This is because one could simply divide the system into two subsystems, each of which would have negative heat capacity (by homogeneity). Then the above argument would imply that the entropy is *minimized* when $T_1 = T_2$ so the second law predicts that fluctuations would make the temperatures unequal so the system would become inhomogeneous. However, this argument does not exclude stable *in*homogeneous systems with negative heat capacity. Black holes (and some stars) are examples of such systems!

1.9 Pressure and the first law

Consider again our two weakly coupled macroscopic systems, with each individual system in equilibrium but the two systems not necessarily in equilibrium with each other. Previously we assumed that they were able to exchange energy but the volume of each was fixed. But now let's assume that they can also exchange volume, with the total volume $V = V_1 + V_2$ fixed. For example, we could have an insulating flask divided into two regions by a very thin *moveable* wall, with gas on each side of the wall. They can exchange energy as before, but now they can also exchange volume. However, the total number of particles N_1, N_2 of each system is fixed so we will not write N_1, N_2 below.

The non-equilibrium entropy of the combined system is

$$S(E, E_1, V, V_1) = S_1(E_1, V_1) + S_2(E - E_1, V - V_1) \tag{1.31}$$

The second law of thermodynamics says that, in equilibrium, this must be maximized. E and V are fixed so the equilibrium values (E_1^*, V_1^*) of E_1 and V_1 are determined by solving $\partial S/\partial E_1 = 0$ and $\partial S/\partial V_1 = 0$. We've already shown that the first equation gives $T_1 = T_2$. The second equation gives

$$\left(\frac{\partial S_1}{\partial V_1}\right)_{E_1=E_1^*, V_1=V_1^*} - \left(\frac{\partial S_2}{\partial V_2}\right)_{E_2=E-E_1^*, V_2=V-V_1^*} = 0 \tag{1.32}$$

We now define the *pressure* of a system as

$$p = T\left(\frac{\partial S}{\partial V}\right)_{E,N} \tag{1.33}$$

Since $T_1 = T_2$, (1.32) implies that

$$p_1 = p_2 \tag{1.34}$$

So, in equilibrium, two systems which can exchange energy and volume must have equal temperatures and equal pressures. If $T_1 = T_2$ but $p_1 > p_2$ then an argument analogous to that of equation (1.23) shows that the second law predicts that V_1 will increase (so V_2 will decrease), in agreement with our intuitive idea of pressure.

For a single system with a fixed number N of particles, our definitions of temperature and pressure can be rewritten as

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV = \frac{1}{T}dE + \frac{p}{T}dV \tag{1.35}$$

so rearranging gives

$$dE = TdS - pdV \tag{1.36}$$

This is called the *fundamental thermodynamic relation*. It relates two infinitesimally nearby equilibrium states of a single system.

Let's see whether our definition of p agrees with the usual definition as "force per unit area". Consider a system (say a gas) in an insulating box, one wall of which is a moveable piston of area A. If we move the piston outwards through distance dx then the increase in volume of the system is $dV = Adx$. Assume that this can be done reversibly, so that S remains constant. Then (1.36) gives $dE = -pAdx$. Let F be the force holding the piston in place. Then the work done by the force is $-Fdx$ (positive if we compress the system), so the change in energy of the system in $dE = -Fdx$. Equating these expressions gives $F = pA$, so indeed our definition of pressure reproduces the more familiar definition as force per unit area. Furthermore, we see that the term $-pdV$ in the fundamental thermodynamic relation is the work done on the system.

Some words of caution: we made an assumption here, namely that the change was reversible. Only for a reversible change can we identify $-pdV$ with the work done on the system. For such changes, we identify TdS as the "heat supplied" to the system, which is zero for a reversible change of an isolated system, but can be non-zero for a reversible change of a non-isolated system. As an example of a irreversible change, consider a piston for which there is some friction. Then the work done to compress the system will be greater than $-pdV$. Another example is the free expansion of a gas into a vacuum, which we discussed above. Imagine the "vacuum" side of the container to be very small compared to the side which originally contains the gas, so that we can think of the final state as being infinitesimally close to the initial state. In this case, we have $dE = 0$ and $dV > 0$ but *no* work is done on (or by) the gas. In this case the fundamental thermodynamic relation implies $TdS = pdV > 0$ so the entropy increases, confirming that this is an irreversible change.

For a general change between two equilibrium states of a not-necessarily-isolated system we write

$$\Delta E = Q + W \tag{1.37}$$

where W is the work done on the system and this equation defines Q, the *heat supplied* to the system. This equation is the *first law of thermodynamics*: it simply expresses conservation of energy. As it stands, this is simply a definition of Q but it has content when combined with an additional statement: for any change of an *isolated* system no heat is supplied, i.e., $Q = 0$. This is really just another definition, in this case defining what we mean by an isolated system. In practice, a good approximation to an isolated system is a system surrounded by insulating walls e.g. a system in a vacuum flask. For an isolated system, the first law gives $\Delta E = W$, i.e., the work done to transform one equilibrium state to another is equal to the difference in energy between those states. We'll return to this later when we discuss classical thermodynamics.

For an infinitesimal change we write

$$dE = đQ + đW \tag{1.38}$$

The reason for writing $đ$ is the following. On the LHS, E is a *function of state*, i.e., it is uniquely defined by the macrostate of the system. However, there is no function of state corresponding to "heat" or "work": one cannot divide the total energy E of a macrostate into a certain amount of "heat energy" and a certain amount of "work energy". In a *change* between two macrostates, the heat supplied and the work done are well-defined quantities. But these quantities depend on *how* the change is achieved. Different changes resulting in the same final state can involve different amounts of heat supplied and work done. However, the change in E is independent of how the change is achieved: it depends just on the initial and final states. Since the heat supplied $\text{đ}Q$ is not the differential of any quantity we write đ instead of d, and similarly for $\text{đ}W$. From the discussion above we have

$$\textit{for a reversible change}: \qquad \text{đ}Q = TdS \qquad \text{đ}W = -pdV \tag{1.39}$$

Finally, let's return to heat capacity. Using the fundamental thermodynamic relation we can rewrite our definition of the heat capacity at constant volume as

$$C_V = \left(\frac{\partial E}{\partial T}\right)_{V,N} = T\left(\frac{\partial S}{\partial T}\right)_{V,N} \tag{1.40}$$

on the RHS we see the "heat supplied" TdS (at least for reversible changes), explaining the name "heat capacity". Often it is more convenient to perform experiments at constant pressure (e.g. atmospheric pressure) than at constant volume. This motivates defining the *heat capacity at constant pressure* as

$$C_p = T\left(\frac{\partial S}{\partial T}\right)_{p,N} \tag{1.41}$$

Note that this is *not* equal to $(\partial E/\partial T)_{p,N}$. Usually $C_p > C_V$. This is because at fixed volume, all of the heat supplied to a system goes into raising its temperature. However, at fixed pressure, most systems will expand when heat is supplied, so some of the energy supplied goes into doing work during this expansion. Therefore more heat has to be supplied to achieve a certain change in temperature.

1.10 Extensive and intensive variables

Consider a homogeneous system in equilibrium (e.g. a gas in a box). Let x be some thermodynamic parameter of the system (e.g. energy, pressure etc.). Now imagine dividing the system into two parts with a partition. Let x_1 and x_2 denote the values of our thermodynamic parameter for the subsystems on each side of the partition. Thermodynamic parameters can be classified according to how x is related to x_1 and x_2. We say that x is *extensive* if $x = x_1 + x_2$ and that x is *intensive* if $x = x_1 = x_2$.

Extensive variables scale with the size of the system: if the system is increased in size by a factor λ (e.g. by putting λ identical copies of the system next to each other) than an extensive variable increases by the same factor λ. However an intensive variable is unchanged by such a scaling.

Examples of extensive variables are the volume V, energy E, and number of particles N of a system. Entropy S is also extensive: for example when we put two systems together we found that the entropy of the combined system is $S_1 + S_2$. On the other hand, the definition of temperature involves a derivative of one extensive quantity w.r.t. another extensive quantity, which does not change if we scale the size of the system. Hence temperature is an intensive variable. Similarly, pressure is intensive.

Heat capacities are extensive since they involve derivatives of an extensive quantity w.r.t. an intensive one. However, *specific* heat capacities (heat capacity per unit mass) are intensive.

1.11 The canonical ensemble

In the microcanonical ensemble we consider an isolated system. However most systems are not isolated, they interact with their environment. We can deal with this by viewing the environment as a second system. We now have the situation of two interacting systems which we discussed previously, viewing the combination of the two systems as a single system which we could study using the microcanonical ensemble. But we now have an additional property, namely that the second system (the environment) is much larger than the original system. This implies that a small fluctuation in its energy has negligible effect on its temperature.

Let $\mathcal{S}$ be the system of interest. We refer to the second system as a *heat reservoir* $\mathcal{R}$. This is defined as a system so large compared to $\mathcal{S}$ that the energy of $\mathcal{S}$ is negligible compared to the energy of $\mathcal{R}$ so exchanges of energy between $\mathcal{S}$ and $\mathcal{R}$ do not change the temperature of $\mathcal{R}$. We assume that the volumes of $\mathcal{S}$ and $\mathcal{R}$, and the numbers of particles in $\mathcal{S}$ and $\mathcal{R}$, are all fixed.

We treat the combined system $\mathcal{S} + \mathcal{R}$ using the microcanonical ensemble with energy in the range $(E_{\text{total}}, E_{\text{total}} + \delta E)$. Let $|n\rangle$ be the energy eigenstates of $\mathcal{S}$, with energy eigenvalues E_n. Assume that $\mathcal{S}$ is in the state $|i\rangle$. Then the energy of $\mathcal{R}$ is in the range $(E_{\text{total}} - E_i, E_{\text{total}} - E_i + \delta E)$. The number of microstates of $\mathcal{R}$ with this property is

$$\begin{aligned} \Omega_R(E_{\text{total}} - E_i) &= \exp\left[k^{-1} S_R(E_{\text{total}} - E_i)\right] \\ &\approx \exp\left[k^{-1} S_R(E_{\text{total}}) - k^{-1} \left(\frac{\partial S_R}{\partial E}\right)_{V,N} E_i\right] \\ &= e^{k^{-1} S_R(E_{\text{total}})} e^{-\beta E_i} \end{aligned} \tag{1.42}$$

in the second line we used $E_i \ll E_{\text{total}}$ and the assumption that the volume of $\mathcal{R}$ is fixed. In the third line we used the definition of temperature and defined

$$\beta \equiv \frac{1}{kT} \tag{1.43}$$

where T is the temperature of the heat reservoir $\mathcal{R}$. Now, the total number of microstates of the combined system with energy in the range $(E_{\text{total}}, E_{\text{total}} + \delta E)$ is $\Omega(E_{\text{total}})$. In the microcanonical ensemble, all states with energy in this range have equal probability $1/\Omega(E_{\text{total}})$. It follows that the probability that $\mathcal{S}$ is in the state $|i\rangle$ is

$$p_i = \frac{\Omega_R(E_{\text{total}} - E_i)}{\Omega(E_{\text{total}})} \propto e^{-\beta E_i} \tag{1.44}$$

From $\sum_i p_i = 1$ we can fix the constant of proportionality on the RHS, giving

$$p_i = \frac{e^{-\beta E_i}}{Z} \tag{1.45}$$

where

$$Z = \sum_i e^{-\beta E_i} \tag{1.46}$$

Equation (1.45) is a probability distribution on the microstates of $\mathcal{S}$, called the *Boltzmann distribution*. All details of the reservoir have dropped out except for its temperature T. Of course, in equilibrium, T is also the temperature of $\mathcal{S}$. So we can regard (1.45) as defining a new kind of ensemble for the system $\mathcal{S}$ in which its temperature, rather than its energy, is held fixed. This is called the *canonical ensemble*.

Note that the probability decays exponentially for $E_i \gg kT = 1/\beta$ so only states with $E_i \lesssim kT$ have a reasonable probability. As $T \to 0$, only the ground state is occupied.

The corresponding density operator is

$$\rho = \frac{1}{Z} \sum_i e^{-\beta E_i} |i\rangle\langle i| = \frac{1}{Z} e^{-\beta H} \tag{1.47}$$

where H is the Hamiltonian of $\mathcal{S}$. The final equality can be checked by calculating matrix elements in the basis $\{|n\rangle\}$. Since ρ is a function of H, this is an equilibrium (i.e. time-independent) mixed state.

The quantity Z defined by (1.46) is called the *partition function*. It has the following important property. Consider two *non-interacting* systems, with energy levels $E_n^{(1)}$ and $E_n^{(2)}$ respectively. Then the partition function for the combined system factorizes:

$$Z(\beta) = \sum_{m,n} e^{-\beta(E_m^{(1)} + E_n^{(2)})} = \sum_m e^{-\beta E_m^{(1)}} \sum_n e^{-\beta E_n^{(2)}} = Z_1(\beta) Z_2(\beta) \tag{1.48}$$

In the canonical ensemble, the temperature, rather than the energy, is held fixed. We can calculate the mean value of the energy using the partition function

$$\langle E\rangle = \sum_i p_i E_i = \frac{1}{Z}\sum_i E_i e^{-\beta E_i} = -\left(\frac{\partial}{\partial\beta}\log Z\right)_{E_i} \tag{1.49}$$

where the derivative w.r.t. β on the RHS is taken with the energy levels E_i held fixed. These energy levels will depend on external parameters such as the volume V, magnetic field B etc, as well as the number of particles N. If V, N are the only external parameters then we can write

$$Z = Z(T,V,N) \qquad \langle E\rangle = -\left(\frac{\partial}{\partial\beta}\log Z\right)_{V,N} \tag{1.50}$$

Note that $\partial/\partial\beta = -kT^2\partial/\partial T$.

We can also calculate the variance of E, which measures how sharply peaked it is around the mean, i.e., the typical size of fluctuations. The variance is

$$\Delta E^2 \equiv \langle (E-\langle E\rangle)^2\rangle = \langle E^2\rangle - \langle E\rangle^2 \tag{1.51}$$

A simple calculation gives (exercise)

$$\Delta E^2 = \left(\frac{\partial^2}{\partial\beta^2}\log Z\right)_{V,N} = -\left(\frac{\partial\langle E\rangle}{\partial\beta}\right)_{V,N} = kT^2\left(\frac{\partial\langle E\rangle}{\partial T}\right)_{V,N} = kT^2 C_V \tag{1.52}$$

For a macroscopic system we expect $\langle E\rangle \propto N$ and $C_V \propto N$ and hence

$$\frac{\Delta E}{\langle E\rangle} \propto \frac{1}{\sqrt{N}} \tag{1.53}$$

Hence fluctuations in E are negligible in the *thermodynamic limit* $N \to \infty$. For a macroscopic system N is enormous so we are always in the themodynamic limit and fluctuations are negligible except in very special circumstances.[3] This means that the energy of our system is essentially fixed and so the canonical and microcanonical ensembles should agree. Therefore we will often write E instead of $\langle E\rangle$ when discussing the canonical ensemble.

1.12 Gibbs entropy

The Boltzmann definition of entropy applies to the microcanonical ensemble. How do wo define entropy for a different ensemble such as the canonical ensemble?

---

[3] We will return to this when we discuss critical phenomena.

Consider a system described by a density operator ρ with associated probabilities p_i. In the corresponding ensemble, we have $W \gg 1$ copies of the system and $p_i W$ of them are in the state $|i\rangle$. Now let's apply the Boltzmann definition of entropy to the entire ensemble. Define Ω to be the number of ways of choosing which copies of the system are in which states. A simple calculation gives

$$\Omega = \frac{W!}{\Pi_i (p_i W)!} \tag{1.54}$$

We now define the Boltzmann entropy of the entire ensemble as

$$S_{\text{ensemble}} = k \log \Omega \approx -kW \sum_i p_i \log p_i \tag{1.55}$$

where the second equality follows from Stirling's approximation ($n! \approx \sqrt{2\pi} n^{n+1/2} e^{-n}$ for $n \gg 1$) assuming W and $W p_i$ are large (exercise). This suggests that we define the entropy of a single copy of our system, in the mixed state described by ρ, as

$$S = -k \sum_i p_i \log p_i \tag{1.56}$$

This is the *Gibbs entropy* of the mixed state. In terms of the density operator it can be written[4]

$$S = -k \text{Tr}(\rho \log \rho) \tag{1.57}$$

To see this, evaluate the trace in the basis of eigenstates of ρ.

Let's calculate the Gibbs entropy of the microcanonical ensemble. This gives

$$S_{\text{micro}}(E, V, N) = -k \sum_{i: E \le E_i \le E + \delta E} \frac{1}{\Omega} \log \frac{1}{\Omega} = k \log \Omega = S_{\text{Boltzmann}} \tag{1.58}$$

So the Gibbs entropy of the microcanonical ensemble is the Boltzmann entropy.

Now consider the canonical ensemble. The Gibbs entropy is

$$\begin{aligned} S_{\text{canonical}}(T, V, N) &= -k \sum_i p_i \left(-\beta E_i - \log Z\right) = k\beta \langle E \rangle + k \log Z \\ &= k \left[\frac{\partial}{\partial T} (T \log Z) \right]_{V,N} \end{aligned} \tag{1.59}$$

---

[4] For general ρ this is also called the *von Neumann entropy*. Note that the equation of motion (1.8) implies that this quantity is constant under time evolution so it does not obey the second law. Therefore it does not solve the problem of defining thermodynamic entropy for non-equilibrium states. The term "Gibbs entropy" usually refers only to equilibrium states, i.e. ones for which ρ commutes with H.

We've seen that, in the canonical ensemble, fluctuations in E are negligible for a macroscopic body and so we would expect $S_{\text{canonical}}(T,V,N)$ to agree with $S_{\text{micro}}(\langle E\rangle,V,N)$. This is indeed the case. To prove this, note that the contribution to the partition function from states with energies in $(E,E+\delta E)$ is $\Omega(E)e^{-\beta E}$, assuming that $\delta E \ll E$ so that $e^{-\beta E}$ can be approximated as constant in this range. Summing over all such energy ranges gives

$$Z \approx \sum_E \Omega(E)e^{-\beta E} \tag{1.60}$$

Inside the sum, we have a very rapidly increasing function $\Omega(E)$ multipled by a very rapidly decreasing function $e^{-\beta E}$ and so the product will be negligible except close to a very sharply defined maximum at some value $E=E_*$. Hence we can write

$$Z \approx \Omega(E_*)e^{-\beta E_*} \tag{1.61}$$

The value of E_* is determined by

$$0 = \left[\frac{\partial}{\partial E}\left(\Omega(E)e^{-\beta E}\right)\right]_{E=E_*} \qquad \Rightarrow \qquad \frac{1}{\Omega(E_*)}\left(\frac{\partial \Omega}{\partial E}\right)_{E=E_*} = \beta \tag{1.62}$$

Using (1.61) in (1.49) gives

$$\langle E\rangle = -\frac{\partial}{\partial \beta}\left(\log\Omega(E_*) - \beta E_*\right) = -\frac{1}{\Omega(E_*)}\left(\frac{\partial \Omega}{\partial E}\right)_{E=E_*}\frac{\partial E_*}{\partial \beta} + \beta\frac{\partial E_*}{\partial \beta} + E_* \tag{1.63}$$

From (1.62) the first and second terms cancel giving

$$\langle E\rangle = E_* \tag{1.64}$$

and hence from (1.61)

$$Z \approx \Omega(\langle E\rangle)e^{-\beta\langle E\rangle} \tag{1.65}$$

Finally, plugging this into the RHS of the first line of (1.59) gives (reinstating V,N)

$$S_{\text{canonical}}(T,V,N) \approx k\log\Omega(\langle E\rangle,V,N) = S_{\text{Boltzmann}}(\langle E\rangle,V,N) \tag{1.66}$$

So, for a macroscopic system, the entropy is the same in the canonical and microcanonical ensembles.

The Gibbs entropy gives a different way of deriving the microcanonical and canonical ensembles. The idea is to maximize the Gibbs entropy subject to certain constraints. In other words, we want to maximize S_{Gibbs} as a function of the probabilities p_i. Obviously we have to do this subject to the constraint

$$\sum_i p_i = 1 \tag{1.67}$$

To obtain the microcanonical ensemble we assume that all systems in the ensemble have energy in the range $(E, E+\delta E)$, so we extremize S_{Gibbs} over all probability distributions such that $p_i = 0$ unless $E_i \in (E, E+\delta E)$. To impose the constraint (1.67) we introduce a Lagrange multiplier α and extremize $k^{-1}S_{\text{Gibbs}} + \alpha(\sum_i p_i - 1)$ over the non-zero p_i:

$$\frac{\partial}{\partial p_i}\left(-\sum_j p_j \log p_j + \alpha \sum_j p_j - \alpha\right) = 0 \qquad \Rightarrow \qquad p_i = e^{\alpha - 1} \tag{1.68}$$

Hence all states with $E_i \in (E, E+\delta E)$ are equally likely, so we've recovered is the microcanonical ensemble.

To obtain the canonical ensemble we follow the same procedure but now, instead of assuming that all systems in the ensemble have energy in the range $(E, E+\delta E)$, we assume that the *average* energy $\langle E \rangle$ takes some fixed value. This constraint introduces a second Lagrange multiplier β, which turns out to be the same quantity β (i.e. inverse temperature) that is used to define the canonical ensemble. (See examples sheet 1.)

1.13 Free energy

Let's investigate the consequences of the second law of thermodynamics for a system in contact with a heat reservoir. We'll want to consider non-equilibrium situations so we will denote the temperature of the reservoir as T_0 to distinguish it from the temperature T of the system. Let S be the entropy of the system and S_0 the entropy of the heat reservoir. The second law applied to the combined system then states that in any process, the entropy must increase:

$$\Delta(S + S_0) \geq 0 \tag{1.69}$$

We are assuming that the reservoir has fixed volume and so the fundamental thermodynamic relation applied to the reservoir gives $dE_0 = T_0 dS_0$ where E_0 is the energy of the reservoir. By definition, T_0 cannot change so integrating this relation gives $\Delta E_0 = T_0 \Delta S_0$ hence

$$\Delta S_0 = \frac{\Delta E_0}{T_0} = -\frac{\Delta E}{T_0} \tag{1.70}$$

where E is the energy of the system. In the final equality we've used the fact that the total energy $E + E_0$ is conserved. Substituting into (1.69) and rearranging gives

$$\Delta A \leq 0 \qquad \text{where} \qquad A = E - T_0 S \tag{1.71}$$

The quantity A is called the *availability*. We have shown that the availability must decrease in any physical process of a system in contact with a heat reservoir. In equilibrium the availability will be minimized.

Consider a physical process in which the initial and final states of the system both have temperature $T = T_0$, but the temperature of the system may change during the process. We then have $\Delta A = \Delta F$ where F is the *Helmholtz free energy*:

$$F = E - TS \tag{1.72}$$

So, for such a process, the second law implies $\Delta F \leq 0$ with F minimized in equilibrium. The advantange of using F rather than A is that the definition of F refers only to the system whereas A refers to both the system and the heat reservoir.

As an example of such a process, consider two substances which can undergo a chemical reaction. Assume that a container of fixed volume contains the two substances separated by a partition, with the container immersed in a heat reservoir. The partition is removed, allowing the substances to react. During the reaction the temperature may change but the system will eventually return to equilibrium with the heat reservoir. In this process the free energy of the system will decrease.

In the definition of F, since we are working in the canonical ensemble we should really write $\langle E \rangle$ but, as discussed above, fluctuations in E are expected to be negligible so we simply write E. The entropy S in this definition is the Gibbs entropy of the canonical ensemble, which, in equilibrium, we expect to be in very good agreement with the Boltzmann entropy. Note that F is an extensive variable.

If we consider two infinitesimally nearby equilibrium states then the fundamental thermodynamic relation $dE = TdS - pdV$ implies

$$dF = dE - TdS - SdT = -SdT - pdV \tag{1.73}$$

Note that the form of the RHS suggests that it is natural to regard F as a function of T and V (and N which is not varied here), which is indeed what we do in the canonical ensemble.

From the above relation we deduce

$$S = -\left(\frac{\partial F}{\partial T}\right)_{V,N} \qquad p = -\left(\frac{\partial F}{\partial V}\right)_{T,N} \tag{1.74}$$

Another useful relation relates F to the partition function:

$$F = -kT \log Z \qquad \text{i.e.} \qquad Z = e^{-\beta F} \tag{1.75}$$

To prove this, use $\partial/\partial\beta = kT^2 \partial/\partial T$ to write

$$\begin{aligned} F = E - TS = -\left(\frac{\partial \log Z}{\partial \beta}\right)_{V,N} - TS &= kT^2 \left(\frac{\partial \log Z}{\partial T}\right)_{V,N} - kT\left(\frac{\partial (T \log Z)}{\partial T}\right)_{V,N} \\ &= -kT \log Z \end{aligned} \tag{1.76}$$

1.14 The chemical potential

So far we have kept the total number of particles N fixed. But now let's consider what happens when N can vary. In analogy with how we defined temperature and pressure, we can consider two weakly interacting macroscopic systems, with each individual system in equilibrium, but with the two systems not necessarily in equilibrium with each other. Previously we've considered the situation in which the two systems can exchange energy and volume. Let's now assume that they can also exchange particles. For example, we could have an insulated flask divided into two regions by a thin moveable and permeable wall so that particles can cross through the wall from one side to the other. Now we have $E = E_1 + E_2$, $V = V_1 + V_2$ and $N = N_1 + N_2$ where E, V, N are fixed but E_i, V_i and N_i can vary. What are the conditions for equilibrium?

The non-equilibrium entropy of the combined system is

$$S(E, E_1, V, V_1, N, N_1) = S_1(E_1, V_1, N_1) + S_2(E - E_1, V - V_1, N - N_1) \tag{1.77}$$

The second law of thermodynamics requires that this is maximized in equilibrium. Hence the equilibrum values (E_1^*, V_1^*, N_1^*) of (E_1, V_1, N_1) are determined by solving $\partial S/\partial E_1 = 0$, $\partial S/\partial V_1 = 0$ and $\partial S/\partial N_1 = 0$ with fixed (E, V, N). The first two equations reproduce the conditions $T_1 = T_2$ and $p_1 = p_2$ we've discussed previously but now we have an additional condition for equilibrium:

$$\left(\frac{\partial S_1}{\partial N_1}\right)_{E_1=E_1^*,V_1=V_1^*,N_1=N_1^*} - \left(\frac{\partial S_2}{\partial N_2}\right)_{E_2=E-E_1^*,V_2=V-V_1^*,N_2=N-N_1^*} = 0 \tag{1.78}$$

We define the *chemical potential* of a system as

$$\mu = -T\left(\frac{\partial S}{\partial N}\right)_{E,V} \tag{1.79}$$

so we see that equilibrium of our two systems requires

$$\mu_1 = \mu_2 \tag{1.80}$$

so in equilibrium, the two systems must have equal temperatures, pressures and chemical potentials. Note that chemical potential is an intensive quantity.

From the definitions of T, p and μ we have

$$dS = \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{E,N} dV + \left(\frac{\partial S}{\partial N}\right)_{E,V} dN = \frac{1}{T}dE + \frac{p}{T}dV - \frac{\mu}{T}dN \tag{1.81}$$

and hence, when particle number can vary, the fundamental thermodynamic relation is

$$dE = TdS - pdV + \mu dN \tag{1.82}$$

As before, this equation relates two infinitesimally nearby equilibrium states of a single system. From this equation we have an alternative formula for μ

$$\mu = \left(\frac{\partial E}{\partial N}\right)_{S,V} \tag{1.83}$$

Equation (1.82) also reveals the physical interpretation of μ: it is the increase in energy if we add one particle to the system, at constant entropy and volume. In fact μ often turns out to be negative. This is because adding a particle at fixed E tends to increase S (there are more ways to share the energy around) and so to keep S fixed, E must decrease.

So far we've used the microcanonical ensemble, in which (E, V, N) are the independent variables. We can easily switch to the canonical ensemble, where (T, V, N) are the independent variables. To do this, we use the Helmholtz free energy $F = E - TS$. From (1.82) we have

$$dF = -SdT - pdV + \mu dN \tag{1.84}$$

and hence

$$\mu = \left(\frac{\partial F}{\partial N}\right)_{T,V} \tag{1.85}$$

If a system contains several different types of particle then its macrostate can be specified by $(E, V, N_1, N_2, \ldots)$ where N_i is the number of particles of type i. We can then define a chemical potential for each species of particle as

$$\mu_i = -T\left(\frac{\partial S}{\partial N_i}\right)_{E,V,N_{j\neq i}} \tag{1.86}$$

It then follows as above that

$$dE = TdS - pdV + \sum_i \mu_i dN_i \tag{1.87}$$

Equations (1.83) and (1.85) generalize in the obvious way.

This can be applied to situations in which the number of particles of each type is not conserved e.g. when reactions between different particles are possible. For example consider a system containing molecules of types A, B, C which can undergo the chemical reaction $A+B \leftrightarrow C$. In this case, generic values for N_i will not correspond to equilibrium, so when we write S as a function of (E, V, N_i) we are considering a non-equilibrium entropy. To determine the values of N_i in equilibrium we need to maximize this entropy over the allowed variations of the N_i. Not all variations are allowed: the equation of the chemical reaction implies that $dN_A = dN_B = -dN_C$. If we assume the system has constant E and V then maximizing S gives $dS = dE = dV = 0$ hence from (1.87) we have $\sum_i \mu_i dN_i = 0$, i.e., $(\mu_A + \mu_B - \mu_C)dN_A = 0$ for arbitrary dN_A. Hence the condition for equilibrium is $\mu_A + \mu_B - \mu_C = 0$. If the system has constant T and V then the same condition arises from minimizing F.

1.15 The grand canonical ensemble

Consider a system $\mathcal{S}$ in equilibrium with a *heat and particle reservoir* $\mathcal{R}$: a much larger system (e.g. "the environment" of a non-isolated system) with which it can exchange energy and particles. Since $\mathcal{R}$ is very large, we can neglect fluctuations in the temperature and chemical potential of $\mathcal{R}$, and regard these quantities as fixed. We can generalize the argument of section (1.11).

Treat the combined system $\mathcal{S} + \mathcal{R}$ using the microcanonical ensemble with energy in the range $(E_{\rm total}, E_{\rm total} + \delta E)$ and total number of particles $N_{\rm total}$. Label the energy eigenstates of $\mathcal{S}$ as $|n\rangle$ and let the state $|n\rangle$ have energy E_n and contain N_n particles. If $\mathcal{S}$ is in the state $|i\rangle$ then $\mathcal{R}$ contains $N_{\rm total} - N_i$ particles and has energy in the range $(E_{\rm total} - E_i, E_{\rm total} - E_i + \delta E)$. The number of microstates of $\mathcal{R}$ with these properties is

$$\begin{aligned}\Omega_R(E_{\rm total} - E_i, N_{\rm total} - N_i) &= \exp\left[k^{-1} S_R(E_{\rm total} - E_i, N_{\rm total} - N_i)\right] \\ &\approx \exp\left[k^{-1} S_R(E_{\rm total}, N_{\rm total}) - k^{-1}\left(\frac{\partial S_R}{\partial E}\right)_{V,N} E_i - k^{-1}\left(\frac{\partial S_R}{\partial N}\right)_{E,V} N_i\right] \\ &= e^{k^{-1} S_R(E_{\rm total}, N_{\rm total})} e^{-\beta(E_i - \mu N_i)}\end{aligned} \tag{1.88}$$

where in the second line we used $E_i \ll E_{\rm total}$ and $N_i \ll N_{\rm total}$ and the assumption that the volume of $\mathcal{R}$ is fixed. In the third line we used the definitions of temperature and chemical potential. Let $\Omega(E_{\rm total}, N_{\rm total})$ be the total number of microstates of $\mathcal{S} + \mathcal{R}$ with $N_{\rm total}$ particles, and energy in the range $(E_{\rm total}, E_{\rm total} + \delta E)$. In the microcanonical ensemble these are all equally likely so it follows that the probability that $\mathcal{S}$ is in the state $|i\rangle$ is

$$p_i = \frac{\Omega_R(E_{\rm total} - E_i, N_{\rm total} - N_i)}{\Omega(E_{\rm total}, N_{\rm total})} \propto e^{-\beta(E_i - \mu N_i)} \tag{1.89}$$

Using $\sum_i p_i = 1$ we obtain

$$p_i = \frac{e^{-\beta(E_i - \mu N_i)}}{\mathcal{Z}} \tag{1.90}$$

where

$$\mathcal{Z}(T, V, \mu) = \sum_i e^{-\beta(E_i - \mu N_i)} \tag{1.91}$$

The probability distribution (1.90) defines the *grand canonical ensemble* and is called the *Gibbs distribution*, or *grand canonical distribution*. It describes a system in equilibrium with a heat and particle reservoir with temperature T and chemical potential μ. Equation (1.91) defines the grand canonical partition function.

The density operator of the grand canonical ensemble is

$$\rho = \frac{1}{\mathcal{Z}} \sum_i e^{-\beta(E_i - \mu N_i)} |i\rangle\langle i| = \frac{1}{\mathcal{Z}} e^{-\beta(H - \mu\mathcal{N})} \tag{1.92}$$

where H is the Hamiltonian of $\mathcal{S}$ and $\mathcal{N}$ is the *number operator* of $\mathcal{S}$, whose eigenvalue gives the number of particles is a given state of $\mathcal{S}$, i.e., $\mathcal{N}|n\rangle = N_n|n\rangle$. $\mathcal{N}$ and H commute so ρ commutes with H, as expected in equilibrium.

We can derive formulae for the average energy, and average particle number, just as in the canonical ensemble. Note that

$$\langle E\rangle - \mu\langle N\rangle = \sum_i (E_i - \mu N_i) p_i = -\left(\frac{\partial \log Z}{\partial \beta}\right)_{V,\mu} \tag{1.93}$$

and

$$\langle N\rangle = \sum_i N_i p_i = \frac{1}{\beta}\left(\frac{\partial \log \mathcal{Z}}{\partial \mu}\right)_{T,V} \tag{1.94}$$

Plugging into the previous formula now determine $\langle E\rangle$. We can also calculate variances e.g. a simple calculation gives

$$\Delta N^2 \equiv \langle N^2\rangle - \langle N\rangle^2 = \frac{1}{\beta^2}\left(\frac{\partial^2 \log \mathcal{Z}}{\partial \mu^2}\right)_{T,V} = \frac{1}{\beta}\left(\frac{\partial \langle N\rangle}{\partial \mu}\right)_{T,V} \tag{1.95}$$

The RHS is extensive and usually proportional to $\langle N\rangle$ hence we have $\Delta N/\langle N\rangle \propto 1/\sqrt{\langle N\rangle}$ so we see that fluctuations in N are negligible in the thermodynamic limit $\langle N\rangle \to \infty$. The same is true for fluctuations in E. In the thermodynamic limit, the microcanonical, canonical and grand canonical ensembles all agree because fluctuations are negligible in this limit.

The (Gibbs) entropy of the grand canonical ensemble is

$$\begin{aligned} S_{\text{grand}} &= -k\sum_i p_i \log p_i = -k\sum_i p_i \left[-\beta(E_i - \mu N_i) - \log \mathcal{Z}\right] \\ &= k\beta\left(\langle E\rangle - \mu\langle N\rangle\right) + k\log \mathcal{Z} = -k\beta\left(\frac{\partial \log \mathcal{Z}}{\partial \beta}\right)_{V,\mu} + k\log \mathcal{Z} \\ &= k\left[\frac{\partial}{\partial T}\left(T\log \mathcal{Z}\right)\right]_{V,\mu} \end{aligned} \tag{1.96}$$

which has the same form as the expression for the entropy of the canonical ensemble. In the thermodynamic limit, this will agree with the entropy of the microcanonical and canonical ensembles.

1.16 Grand canonical potential

In the canonical ensemble we defined the Helmholtz free energy F. In the grand canonical ensemble we define the *grand canonical potential*

$$\Phi = F - \mu N = E - TS - \mu N \tag{1.97}$$

Strictly speaking we should write $\langle E\rangle$ and $\langle N\rangle$ in this expression and S refers to the Gibbs entropy of the grand canonical ensemble. But, as emphasized above, the different ensembles agree in the thermodynamic limit so we don't usually bother to emphasize these details.

The argument of section 1.13 can be generalized straightforwardly to show that, in physical changes of a system in contact with a heat and particle reservoir, the quantity $E - T_0 S - \mu_0 N$ must decrease, attaining a minimum in equilibrium. Here T_0 and μ_0 are the (constant) temperature and chemical potential of the reservoir. Hence for a physical process for which the initial and final temperatures of the system are the same, and the initial and final chemical potentials of the system are the same (although the temperature and chemical potential may vary during the process) we have $\Delta\Phi \leq 0$, with Φ minimized in equilbrium.

From the fundamental thermodynamic relation we obtain

$$d\Phi = -SdT - pdV - Nd\mu \tag{1.98}$$

which indicates that it is most natural to regard Φ as a function of (T, V, μ), the independent variables used to define the grand canonical ensemble. Hence Φ is a natural object to work with when we use this ensemble.

From (1.93) and (1.96) we obtain

$$\begin{aligned}\Phi &= -\left(\frac{\partial \log \mathcal{Z}}{\partial \beta}\right)_{V,\mu} - kT\left[\frac{\partial}{\partial T}(T\log\mathcal{Z})\right]_{V,\mu}\\ &= -kT\log\mathcal{Z} \end{aligned}\tag{1.99}$$

and hence

$$\mathcal{Z} = e^{-\beta\Phi} \tag{1.100}$$

so $\mathcal{Z}$ and Φ in the grand canonical ensemble are related in the same way as Z and F in the canonical ensemble.

Φ is clearly an extensive quantity, so if we rescale the size of the system by a factor λ we obtain

$$\Phi(T, \lambda V, \mu) = \lambda\Phi(T, V, \mu) \tag{1.101}$$

where we used the fact that V is extensive but T, μ are intensive. Now differentiate this expression w.r.t. λ and set $\lambda = 1$. The result is

$$V \left(\frac{\partial \Phi}{\partial V} \right)_{T,\mu} (T, V, \mu) = \Phi(T, V, \mu) \tag{1.102}$$

But from (1.98) the LHS is $-pV$ so we have the useful relation[5]

$$\Phi(T, V, \mu) = -p(T, \mu)V \tag{1.103}$$

where on the RHS we've used the fact that p is intensive so $p(T, V, \mu)$ must be independent of V. Substituting this back into (1.98) gives the *Gibbs-Duheim relation*:

$$dp = \frac{S}{V} dT + \frac{N}{V} d\mu \tag{1.104}$$

1.17 Summary of ensembles

We have defined three different ensembles: microcanonical, canonical and grand canonical. Their different properties are summarized in the table. The final column states which quantity is extremized in equilibrium.

ensemble	system	independent variables	thermodynamic potential
microcanonical	isolated	E, V, N	S maximized
canonical	in contact with heat reservoir	T, V, N	F minimized
grand canonical	in contact with heat and particle reservoir	T, V, μ	Φ minimized

---

[5] Note that this argument does not work for $F(T, V, N)$: extensivity implies only that F must have the form $Vf(T, N/V)$ for some function f.