Unit 3: Canonical ensemble

15 Feb.

3.1 The thermal reservoir

3.1.1 Replicas and occupation numbers

While it is relatively easy to prevent particle exchange, for example by sealing gases inside airtight containers, it is not practical to forbid energy exchange as would be needed to fully isolate statistical systems. Any thermal insulator is imperfect, and even in the deepest reaches of space the system would still be bombarded by cosmic microwave radiation. In practice it is more convenient to work with physical systems that are characterized by their (intensive) temperatures rather than their (extensive) internal energies.

This leads us to define the **canonical ensemble** to be a statistical ensemble characterized by its fixed temperature T and conserved particle number N, with the temperature held fixed through contact with a **thermal reservoir**.

The second part of this definition connects the fixed temperature to the fundamental fact of energy conservation (the first law of thermodynamics). This is done by proposing that our system of interest Ω is in thermal contact with a much larger external system Ω_{res} — the thermal reservoir, sometimes called a "heat bath". The overall combined system $\Omega_{\text{tot}} = \Omega_{\text{res}} \otimes \Omega$ is governed by the micro-canonical ensemble, with conserved total energy $E_{\text{tot}} = E_{\text{res}} + E \approx E_{\text{res}}$, while the energy E of Ω is allowed to fluctuate. The key qualitative idea is that, in thermodynamic equilibrium, Ω has a negligible effect on the overall system. In particular, the temperature of that overall system—and therefore the temperature of Ω, by intensivity—is set by the reservoir and remains fixed even as E fluctuates. This effectively generalizes the setup we used to analyze heat exchange in the previous section, where we saw that thermal contact causes a net flow of energy from hotter systems to colder systems. When these systems are 'natural', this cools the hotter one by heating the colder one.

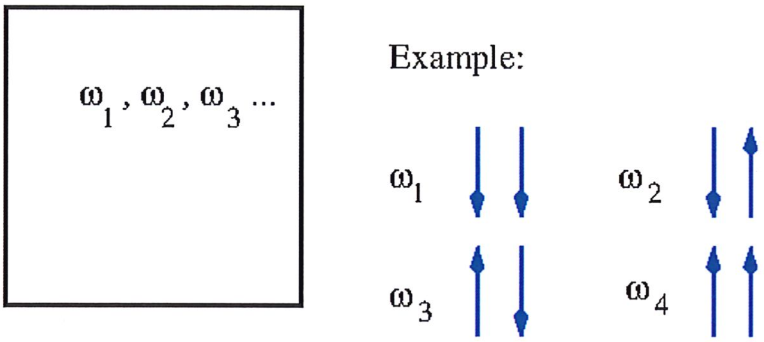
The mathematical implementation of this argument, as developed by Gibbs, proceeds by considering a well-motivated ansatz for the form of the thermal reservoir Ω_{res}. The goal, which will be useful to keep in mind as we go through the lengthy analysis, is to show that the specific form of Ω_{res} is ultimately irrelevant. This will allow us to work directly with the system of interest, Ω, independent of the details of the thermal reservoir that fixes its temperature.

Without further ado, we take Ω_{tot} to consist of many ($R \gg 1$) identical **replicas** of the system Ω that we're interested in. All of these replicas are in thermal contact with each other, and in thermodynamic equilibrium.[6] Choosing any one

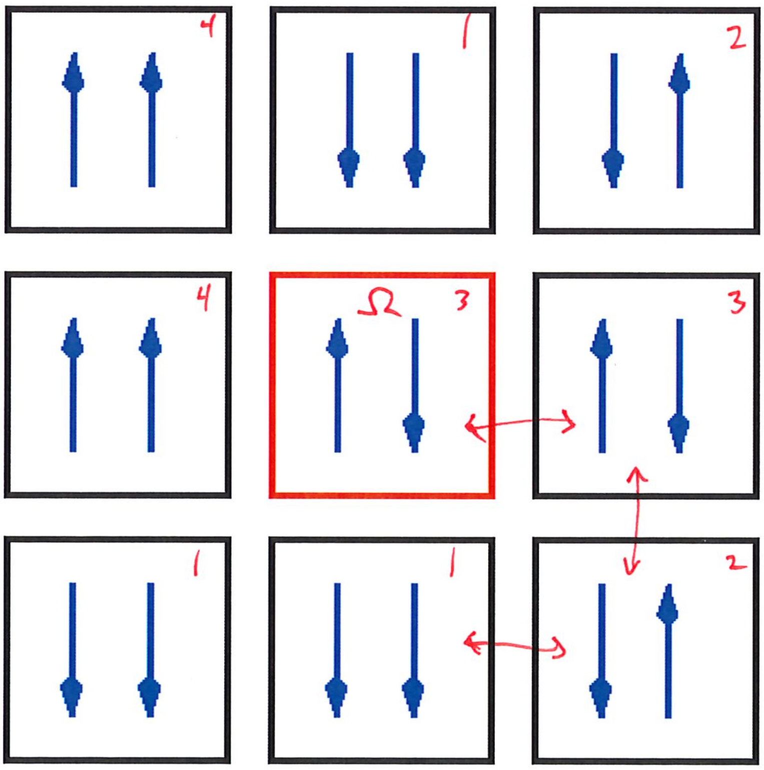
[6]The thermal contact between any two replicas can be indirect, mediated by a sequence of intermediate replicas. This transitivity of thermodynamic equilibrium is sometimes called the zeroth law of thermodynamics. It declares that if systems Ω_A & Ω_B are in thermodynamic equilibrium while systems Ω_B & Ω_C are in thermodynamic equilibrium, then Ω_A & Ω_C must also be in thermodynamic equilibrium.

of the replicas to be the system of interest, Ω, the other $R-1 \gg 1$ replicas collectively form the thermal reservoir Ω_{res}. Assuming we want to study reasonable systems Ω, this ansatz ensures that Ω_{res} is also reasonable, simply much larger.

An extremely small example of this setup is illustrated by the figures below, where the system of interest is just $N = 2$ spins. For now we assume the spins are *distinguishable*, so that $\downarrow\uparrow$ and $\uparrow\downarrow$ are both distinct micro-states. This means that each individual replica has the $M = 4$ micro-states ω_i defined below.



To form the overall system Ω_{tot} we now bring together the $R = 9$ replicas shown below. We draw boxes around each replica to remind us that they are allowed to exchange only energy with each other, while the $N = 2$ spins are conserved in each replica. We pick out one of these replicas (coloured red) to serve as the system Ω we will consider. The other 8 are the thermal reservoir Ω_{res} that fixes the temperature of Ω.



$W_1 = \downarrow\downarrow \quad n_1 = 3$

$W_2 = \downarrow\uparrow \quad n_2 = 2$

$W_3 = \uparrow\downarrow \quad n_3 = 2$

$W_4 = \uparrow\uparrow \quad n_4 = 2$

A convenient way to analyze the overall system of R replicas, Ω_{tot}, is to define the **occupation number** n_i to be the number of replicas that adopt the micro-state $\omega_i \in \Omega$ in any given micro-state of Ω_{tot}. The index $i \in \{1, 2, \cdots, M\}$ runs over all M micro-states of Ω. In the example above, three of the replicas have the micro-state $\omega_1 = \downarrow\downarrow$, meaning $n_1 = 3$. What are the occupation numbers $\{n_2, n_3, n_4\}$ for the other three ω_i in the figures above? Are all replicas are accounted for, $\sum_i n_i = R$?

$n = \{3, 2, 2, 2\}$

$\sum_i n_i = 9 = R$

Normalizing the occupation number by R gives us a well-defined *occupation probability*, $p_i = n_i/R$ with $\sum_i p_i = 1$. This p_i is the probability that if we choose a replica at random it will be in micro-state ω_i.

Now let us consider conservation of energy, which continues to apply to the total energy E_{tot} of the overall system Ω_{tot}. We assume that each replica's energy E_r is independent of all the other replicas. This is guaranteed for the non-interacting systems we will focus on until Unit 9, and also holds when interactions are allowed within each replica but not between different replicas. The thermal contact between replicas allows E_r to fluctuate subject to conservation of E_{tot}, but there are at most M possible values E_i it can have, corresponding to the M micro-states $\omega_i \in \Omega$. Some distinct micro-states $\omega_i \neq \omega_j$ may have the same energy $E_i = E_j$, which doesn't affect the analysis. This allows us to rearrange a sum over replicas into a sum over the micro-states of Ω:

$$E_{\text{tot}} = \sum_{r=1}^{R} E_r = \sum_{i=1}^{M} n_i E_i, \tag{27}$$

with the occupation number n_i counting how many times micro-state ω_i appears among the R replicas. We can assume that R and M are both finite, so we don't need to worry about rearranging infinite sums.

3.1.2 Partition function

Following Gibbs, we have already taken the thermal reservoir Ω_{res} to consist of $R-1$ replicas of the system of interest, Ω. The next step is to further simplify the mathematics by assuming that the overall R-replica system Ω_{tot} is fully specified by a fixed set of M occupation numbers $\{n_i\}$. This is equivalent to assuming that the occupation probabilities $\{p_i\}$ are constant in time, as a reflection of thermodynamic equilibrium. From Eq. 27, we see that this ensures conservation of the total energy E_{tot}, and we can apply the micro-canonical tools we developed in the previous unit. Recall our ultimate goal of showing that such details of the thermal reservoir are irrelevant to the system Ω.

Based on the conservation of E_{tot}, we want to determine the (intensive) temperature of Ω_{tot}, which fixes the temperature of the system of interest, Ω. According to our previous work, to do this we first need to compute the overall number of micro-states M_{tot} as a function of E_{tot}, from which we can derive the micro-canonical entropy and temperature since the system is in thermodynamic equilibrium. From the fixed occupation numbers n_i, we already know how many times each micro-state ω_i appears among the R replicas. To determine M_{tot} we just need to count how many possible ways there are of distributing the $\{n_i\}$ micro-states among the R replicas.

If we consider first the micro-state ω_1, the number of possible ways of distributing n_1 copies of this micro-states among the R replicas is just the binomial coefficient

$$\binom{R}{n_1} = \frac{R!}{n_1!\,(R-n_1)!}.$$

Moving on to ω_2, we need to keep in mind that n_1 replicas have already been assigned micro-state ω_1, so there are only $R-n_1$ replicas left to choose from. What is the resulting number of possible ways of distributing these n_2 micro-states?

$$\binom{R-n_1}{n_2} = \frac{(R-n_1)!}{n_2!\,(R-n_1-n_2)!} \qquad \binom{R-n_1-n_2}{n_3}, \quad \ldots$$

Repeating this process for all micro-states $\{\omega_1, \omega_2, \cdots, \omega_M\}$, and recalling that $(R - \sum_i n_i)! = 0! = 1$, you should obtain a product that 'telescopes' to

$$M_{\text{tot}} = \frac{R!}{n_1!\, n_2!\, \cdots\, n_M!}. \tag{28}$$

From this we can see that the order in which we assign micro-states to replicas is irrelevant, since integer multiplication is commutative.

Thanks to thermodynamic equilibrium, the entropy of the micro-canonical overall system Ω_{tot} is

$$S(E_{\text{tot}}) = \log M_{\text{tot}} = \log(R!) - \sum_{i=1}^{M} \log(n_i!),$$

where the dependence on E_{tot} enters through the occupation numbers via Eq. 27. With $R \gg 1$ and $n_i \gg 1$ for all $i = 1, \cdots, M$, we can approximate each of these logarithms using the first two terms in Stirling's formula,

$$\log(N!) = N \log N - N + \mathcal{O}(\log N) \approx N \log N - N \qquad \text{for } N \gg 1.$$

$N \sim 10^{23}$

$\log N \sim 50$

In order for *every* occupation number to be large, $n_i \gg 1$, the number of replicas must be much larger than the number of micro-states of Ω. As we have discussed before, the number of micro-states M is typically a very large number, so with

$M \sim e^N$

$R \gg M$ we are formally considering truly enormous thermal reservoirs! This enormity helps ensure that the detailed form of the reservoir will be irrelevant.

Applying the approximation above, what do you find for $S(E_{\text{tot}})$ in terms of R and n_i? What is the entropy in terms of the occupation probabilities $p_i = n_i/R$?

$$S(E_{\text{tot}}) = \log(R!) - \sum_{i=1}^{M} \log(n_i!) \approx R \log R - \cancel{R} - \sum_i \left(n_i \log n_i - \cancel{n_i}\right)$$

$$\approx R \log R - \sum_i n_i \log n_i$$

$$= \cancel{R \log R} - R \sum_i p_i \left(\cancel{\log R} + \log p_i\right)$$

$$= -R \sum_i p_i \log p_i$$

$\sum_i n_i = R$

$n_i = R p_i$

$\sum_i p_i = 1$

In your result, the dependence on E_{tot} now enters through the occupation probabilities p_i. In order to determine the temperature, we have to express $S(E_{\text{tot}})$ directly in terms of E_{tot}. We do this by applying our knowledge that thermodynamic equilibrium implies maximal entropy.

$\frac{1}{T} = \left.\frac{\partial S}{\partial E}\right|_N$

Following the same steps as in Section 2.2.3, we maximize the entropy, now with two Lagrange multipliers to account for two constraints on the occupation probabilities:

$$\sum_{i=1}^{M} p_i = 1 \qquad\qquad \sum_{i=1}^{M} n_i E_i = R \sum_{i=1}^{M} p_i E_i = E_{\text{tot}}.$$

Writing everything in terms of occupation probabilities, we therefore need to maximize the modified entropy

$$\overline{S} = -R \sum_{i=1}^{M} p_i \log p_i + \alpha \left(\sum_{i=1}^{M} p_i - 1\right) - \beta \left(R \sum_{i=1}^{M} p_i E_i - E_{\text{tot}}\right).$$

Here we've chosen the sign of β for later convenience. What is the occupation probability p_k that maximizes $\overline{S}$?

$$0 = \frac{\partial \overline{S}}{\partial p_k} = -R\left(\log p_k + 1\right) + \alpha - \beta R E_k$$

$$\log p_k = -1 + \frac{\alpha}{R} - \beta E_k$$

$$p_k = \exp\left[-\left(1 - \frac{\alpha}{R}\right) - \beta E_k\right] = \frac{\exp[-\beta E_k]}{\exp\left[1 - \frac{\alpha}{R}\right]} = \frac{1}{Z} e^{-\beta E_k}$$

By defining a new parameter Z in terms of α, you should find

$$p_k = \frac{1}{Z} e^{-\beta E_k}. \tag{29}$$

As before, we need to fix the parameters $\{Z, \beta\}$ by demanding that the two constraints above are satisfied. The first of these constraints is straightforward and produces an important result:

$$1 = \sum_{i=1}^{M} p_i = \frac{1}{Z} \sum_{i=1}^{M} e^{-\beta E_i} \qquad \Longrightarrow \qquad Z(\beta) = \sum_{i=1}^{M} e^{-\beta E_i}. \tag{30}$$

Equation 30 defines the canonical **partition function** $Z(\beta)$, a fundamental quantity in the canonical ensemble, from which many other derived quantities can be obtained.

$Z(\beta)$ still depends on the other as-yet-unknown parameter $\beta(E_{\text{tot}})$. Applying our second constraint, Eq. 27, relates β to E_{tot}:

$$E_{\text{tot}} = R \sum_{i=1}^{M} p_i E_i = \frac{R}{Z(\beta)} \sum_{i=1}^{M} E_i \, e^{-\beta E_i} = R \frac{\sum_{i=1}^{M} E_i \, e^{-\beta E_i}}{\sum_{i=1}^{M} e^{-\beta E_i}}. \tag{31}$$

This relation is a bit complicated, but will suffice for our goal of expressing the entropy in terms of E_{tot}. Inserting Eq. 29 for p_i into your earlier result for the entropy, what do you obtain upon applying Eqs. 30 and 31?

$$S(E_{\text{tot}}) = -R \sum_{i=1}^{M} p_i \log p_i = -R \sum_i p_i \log\left(\frac{1}{Z} e^{-\beta E_i}\right)$$

$$= R \log Z \underbrace{\sum_i p_i}_{1} + R\beta \sum_i p_i E_i$$

$$= R \log Z + \beta E_{tot}$$

15 Feb.

There is a pleasant simplification when we take the derivative to determine the temperature. Defining $\beta' \equiv \frac{\partial}{\partial E_{\text{tot}}} \beta(E_{\text{tot}})$, we have

$$\frac{1}{T} = \frac{\partial}{\partial E_{\text{tot}}} S(E_{\text{tot}}) = \frac{\partial}{\partial E_{\text{tot}}} \left[E_{\text{tot}} \beta + R \log Z(\beta)\right] = \beta + E_{\text{tot}} \beta' + R \frac{1}{Z} \frac{\partial Z(\beta)}{\partial \beta} \beta'.$$

Using Eq. 31 we can compute

$$\frac{1}{Z} \frac{\partial Z(\beta)}{\partial \beta} = \frac{1}{Z} \frac{\partial}{\partial \beta} \sum_{i=1}^{M} e^{-\beta E_i} = -\frac{1}{Z} \sum_{i=1}^{M} E_i \, e^{-\beta E_i} = -\sum_{i=1}^{M} p_i E_i = -\frac{E_{\text{tot}}}{R},$$