Lecture-15: The Boltzmann Distribution

1 The Boltzmann Distribution

The fundamental purpose of statistical physics is to understand how microscopic interactions of particles (atoms, molecules, etc.) can lead to macroscopic phenomena. It is unreasonable to try to calculate how each and every particle is behaving. Instead, we use probability and statistics to model the behavior of a large group of particles as a whole.

Definition 1.1 (Configurations space). The state/configuration of an N particle system is denoted by $X : \Omega \to \mathcal{X} \triangleq \mathcal{Z}^N$, where the state of each particle $i \in [N]$ is represented by the random variable $X_i : \Omega \to \mathcal{Z}$.

Remark 1. We will limit ourselves to configuration spaces $\mathcal{Z}$ which are (a) finite sets, or (b) smooth, compact, finite dimensional manifolds.

Definition 1.2 (Observable). Any observable is map from configuration space $\mathcal{X}$ to $\mathbb{R}$ and denoted by $\mathcal{O} : \mathcal{Z}^N \to \mathbb{R}$ such that for any configuration $x \in \mathcal{X}^N$ we have the observable $\mathcal{O}(x)$.

Remark 2. Observables can, at least in principle, be measured through an experiment. In contrast, the configuration of a system usually cannot be measured.

Definition 1.3 (Energy). **Energy function** $E : \mathcal{X} \to \mathbb{R}$ is a special observable, whose exact form depends on the level of interaction of the particles.

Example 1.4 (Interacting particle system energy). We consider three different examples of energy function for an N-particle system.

Non-interacting. The total energy of the system depends only on the energies of the individual particles, and has the form $E(x) = \sum_{i=1}^{N} E_i(x_i)$.

Two-body interaction. The total energy of the system depends on the interaction between pairs of particles, and has the form $E(x) = \sum_{i=1}^{N} E_i(x_i) + \alpha \sum_{\substack{i,j=1 \\ i<j}}^{N} E_{ij}(x_i, x_j)$.

k-body interaction. The total energy of the system depends on the interaction among a set of k particles, and has a term of the form $\sum_{F \subseteq [N]:|F|=k} E_F(x_F)$, where $x_F \triangleq \{x_i : i \in F\}$. However, in real physical systems, interactions above the two or three body levels are rarely considered.

Definition 1.5. Given a finite configuration space $\mathcal{X}$ and the energy function $E : \mathcal{X} \to \mathbb{R}$, the **partition function** is defined as $Z(\beta) \triangleq \sum_{x \in \mathcal{X}} e^{-\beta E(x)}$. Here, the quantity β is the inverse temperature, i.e. $\beta = 1/T$.

Remark 3. In physics, β is actually defined as $1/(k_B \mathrm{T})$, where k_B is the Boltzmann constant. Here, we simply define $k_B = 1$.

Remark 4. The partition function is so called because it describes how particles are distributed or *partitioned* into different energy levels of a system.

Definition 1.6. Given a finite configuration space $\mathcal{X}$ and energy function $E : \mathcal{X} \to \mathbb{R}$, the **Boltzmann distribution** is the probability distribution for the system to be in state $x \in \mathcal{X}$, given by $P\{X = x\} = \mu_\beta(x) \triangleq \frac{e^{-\beta E(x)}}{Z(\beta)}$, where $Z(\beta)$ is the partition function such that $\sum_{x \in \mathcal{X}} \mu_\beta(x) = 1$.

Definition 1.7. The **Boltzmann average** of an observable $\mathcal{O} : \mathcal{X} \to \mathbb{R}$ is defined as its expectation under the Boltzmann distribution, and is denoted by

$$\langle \mathcal{O} \rangle \triangleq \sum_{x \in \mathcal{X}} \mathcal{O}(x) \mu_\beta(x) = \frac{1}{Z(\beta)} \sum_{x \in \mathcal{X}} e^{-\beta E(x)} \mathcal{O}(x).$$

Example 1.8 (Independent coin tosses). Let us consider the case of n independent coin tosses with outcomes denoted by $X : \Omega \to \mathcal{X} \triangleq \{H, T\}^n$, where the coin bias is $p \triangleq P\{X_i = H\}$ for all $i \in [n]$. For a configuration x in the configuration space $\mathcal{X} \triangleq \{H, T\}^N$, we define the fraction of number of heads in n coin tosses as $A_H(x) \triangleq \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{x_i = H\}}$. It follows that

$$P\{X = x\} = p^{nA_H(x)} (1-p)^{n - nA_H(x)} = e^{-\beta E(x)},$$

where the energy $E : \mathcal{X} \to \mathbb{R}$ is defined as $E(x) \triangleq -\frac{n}{\beta}\Big[A_H(x) \ln p + (1 - A_H(x)) \ln(1-p)\Big]$ for all configurations $x \in \mathcal{X}$. From the definition of entropy and KL distance, we observe that

$$E(x) = \frac{n}{\beta}\Big[D(A_H(x) \| p) + H(A_H(x))\Big].$$

Example 1.9 (Spin-1/2 system). One of the intrinsic properties of elementary particles is spin, denoted by σ. An **Ising spin** takes values in $\sigma \in \mathcal{Z} = \{+1, -1\}$. The energy of the particle in spin state σ in a magnetic field B is given by $E(\sigma) = -B\sigma$. Since there are only two states, the partition function is

$$Z(\beta) = \sum_{\sigma \in \mathcal{Z}} e^{-\beta E(\sigma)} = e^{-\beta B} + e^{\beta B}.$$

The probability of the particle being in spin state σ is given by

$$\mu_\beta(\sigma) = \frac{e^{-\beta E(\sigma)}}{Z(\beta)} = \frac{e^{\beta B \sigma}}{e^{-\beta B} + e^{\beta B}}.$$

The average value of the spin, called the **magnetization** is

$$\langle \sigma \rangle = \sum_{\sigma \in \mathcal{Z}} \sigma \mu_\beta(\sigma) = \frac{e^{\beta B} - e^{-\beta B}}{e^{-\beta B} + e^{\beta B}} = \tanh(\beta B).$$

High temperatures, $T \gg |B|$ **or** $\beta \to 0$. We have $e^{\beta B} \approx e^{-\beta B}$, so $\langle \sigma \rangle \approx 0$.

Low temperatures, $\beta \to 0$. We have two cases. If $B > 0$, then $e^{\beta B} \gg e^{-\beta B}$, so $\langle \sigma \rangle \approx 1$. If $B < 0$, then $e^{\beta B} \ll e^{-\beta B}$, so $\langle \sigma \rangle \approx -1$. That is, we have $\langle \sigma \rangle \approx \text{sign}(B)$.

Example 1.10 (Pott's spin). Another spin variable is the **Potts spin** with q states, which takes values in $\mathcal{X} = [q] = \{1, 2, \ldots, q\}$. In this case, the energy of a particle with Potts spin σ, when it is placed in a magnetic field of intensity B pointing in direction r, is given by

$$E(\sigma) = -B \mathbb{1}_{\{\sigma = r\}}.$$

The partition function is given by

$$Z(\beta) = e^{\beta B} + \underbrace{e^0 + e^0 + \ldots + e^0}_{q-1 \text{ times}} = e^{\beta B} + (q-1).$$

The Boltzmann distribution is computed to be

$$\mu_\beta(\sigma) = \frac{e^{-\beta E(\sigma)}}{Z(\beta)} = \frac{e^{\beta B \mathbb{1}_{\{\sigma = r\}}}}{e^{\beta B} + (q-1)}.$$

The Boltzmann average of energy is given by

$$\langle E \rangle = -B\frac{e^{\beta B}}{e^{\beta B} + (q-1)} + (\text{all other terms are } 0) = \frac{-Be^{\beta B}}{e^{\beta B} + (q-1)}.$$

Example 1.11 (Particles in a bottle). Let us consider the case of a particle in a closed container (bottle), which is placed in a gravitational field. The space of configurations is the three-dimensional volume of the bottle, which can be represented as $\mathcal{X} = \text{BOTTLE} \subset \mathbb{R}^3$.

The (potential) energy of a particle in a gravitational field is $E(x) = mgh(x) = wh(x)$, where $h(x)$ corresponds to the height of the configuration x. Here, x is not to be confused with the x-coordinate; it is simply a representation of the state of the particle.

In this case, the partition function is difficult to calculate, because $\mathcal{X}$ is an infinite set. However, we can still compute some related quantities in a straightforward fashion. Letting the partition function be $Z(\beta)$, we see that the Boltzmann density is

$$\mu_\beta(x) = \frac{e^{-\beta wh(x)}}{Z(\beta)}, \qquad \mu_\beta(y) = \frac{e^{-\beta wh(y)}}{Z(\beta)}, \qquad \frac{\mu_\beta(x)}{\mu_\beta(y)} = e^{\beta w(h(y)-h(x))}.$$

Let us fix $y = 0$. We get $\frac{\mu_\beta(x)}{\mu_\beta(0)} = e^{-\beta wh(x)}$. The above equation tells us the greater the height of the particle, the smaller the probability of the particle being in that state. This makes physical sense, because we expect most of the particles to be at low heights due to gravity. Note that this result was obtained without any explicit knowledge of the partition function.

2 Temperature Limits

Let us now see that happens to the Boltzmann distribution at high and low temperatures. The probability of the particle being in state x according to the Boltzmann distribution is

$$\mu_\beta(x) = \frac{e^{-\beta E(x)}}{\sum_{x\in\mathcal{X}} e^{-\beta E(x)}}.$$

2.1 High temperature limits

At high temperatures $\beta \to 0$, and therefore we have $e^{-\beta E(x)} \to 1$ for all values of energy. The partition function is simply $Z(\beta) = \sum_{x\in\mathcal{X}} e^{-\beta E(x)} = |\mathcal{X}|$. The probability of the particle being in any state x is

$$\lim_{\beta\to 0} \mu_\beta(x) = \frac{1}{|\mathcal{X}|},$$

which is a uniform distribution, i.e. in the high temperature limit, we expect every state to become equally probable. Note that this is in accordance with the Ising spin model. At high temperatures, the spin of the particles become randomly orientated, so the average spin (magnetization) will be 0.

2.2 Low temperature limits

The case of low temperature is more interesting. We need the following definitions to understand the low temperature limit.

Definition 2.1. The **ground state energy** of the system is the lowest energy attainable by it, denoted by $E_0 = \inf\{E(x) : x \in \mathcal{X}\}$. A state corresponding to the ground state energy is called a **ground state**, and the set of ground states is given by $\mathcal{X}_0 = \{x \in \mathcal{X} : E(x) = E_0\}$.

In terms of the ground state $\mathcal{X}_0$, we can write the Boltzmann distribution as

$$\mu_\beta(x) = \frac{e^{-\beta(E(x)-E_0)}}{|\mathcal{X}_0| + \sum_{x\in\mathcal{X}\setminus\mathcal{X}_0} e^{-\beta(E(x)-E_0)}}$$

Note that since $E(x) > E_0$ for all $x \in \mathcal{X} \setminus \mathcal{X}_0$, the second term in the denominator goes to 0 as $\beta \to \infty$. The term in the numerator is 1 if $E(x) = E_0$ and it is 0 otherwise. Therefore, the probability distribution can be written as

$$\mu_\beta(x) = \frac{1}{|\mathcal{X}_0|}\mathbb{1}_{\{x\in\mathcal{X}_0\}}.$$

So, at low temperatures, all the particles will settle to the ground state, and all the other energy levels are left unpopulated. We will be primarily interested in the low temperature limit of systems, as it provides insight into the "low energy structure" of the system. At high temperatures, the states are just random, which does not give us much information.

3 Thermodynamic Potentials

In the study of the Boltzmann distribution, it is often useful to study a few functions depending on the system which summarize the system and provide useful information. Such functions are called **thermodynamic potentials** and are written as functions of the inverse temperature β. The following are the commonly used thermodynamic potentials.

Definition 3.1. For an interacting particle system, we respectively define **free energy**, **free entropy**, **internal energy**, and **cannonical entropy** as

$$F(\beta) \triangleq -\frac{1}{\beta}\ln Z(\beta), \qquad \Phi(\beta) \triangleq -\beta F(\beta) = \ln Z(\beta), \qquad U(\beta) \triangleq \frac{\partial}{\partial\beta}\left(-\Phi(\beta)\right), \qquad S(\beta) \triangleq \beta^2\frac{\partial F(\beta)}{\partial\beta}. \tag{1}$$

Remark 5. Using these definitions, we derive some relations between the potentials which will be useful later.

Lemma 3.2. *Following statements are true for thermodynamic potentials.*

1. *Relation between free energy, internal energy, and canonical entropy:* $F(\beta) = U(\beta) - \frac{1}{\beta}S(\beta) = -\frac{1}{\beta}\Phi(\beta)$.
2. *Internal energy is the expected energy under the Boltzmann distribution, i.e.* $U(\beta) = \langle E\rangle$.
3. *Canonical entropy is the Shannon entropy of the Boltzmann distribution, i.e.* $S(\beta) = H(\mu_\beta)$.
4. *The free entropy function* $\Phi(\beta)$ *is convex in inverse temperature* β.

Proof. We recall the definitions for all thermodynamic potentials.

1. From the definition of free energy and free entropy, we can write

$$U(\beta) = \frac{\partial}{\partial\beta}\left(\beta F(\beta)\right) = F(\beta) + \beta\frac{\partial F(\beta)}{\partial\beta} = F(\beta) + \frac{1}{\beta}S(\beta).$$

2. From the definition of free entropy, we can write

$$U(\beta) = \frac{\partial}{\partial\beta}\left(-\Phi(\beta)\right) = \frac{\partial}{\partial\beta}(-\ln Z(\beta)) = -\frac{1}{Z(\beta)}\frac{\partial Z(\beta)}{\partial\beta}.$$

For a finite configuration space $\mathcal{X}$, we can exchange the partial derivative and the summation, to write

$$U(\beta) = -\frac{1}{Z(\beta)}\frac{\partial}{\partial\beta}\sum_{x\in\mathcal{X}} e^{-\beta E(x)} = -\frac{1}{Z(\beta)}\sum_{x\in\mathcal{X}} e^{-\beta E(x)}(-E(x)) = \sum_{x\in\mathcal{X}}\mu_\beta(x)E(x) = \langle E\rangle\,.$$

3. From the result in first two parts and the definition of free entropy, we can write the canonical entropy as

$$S(\beta) = \Phi(\beta) + \beta U(\beta) = \ln Z(\beta) + \sum_{x \in \mathcal{X}} \beta E(x) \mu_\beta(x) = \sum_{x \in \mathcal{X}} \mu_\beta(x)(\beta E(x) + \ln Z(\beta)).$$

   Now note that $\mu_\beta(x) = \frac{e^{-\beta E(x)}}{Z(\beta)}$. Taking natural log on both sides, we have $\ln \mu_\beta(x) = -\beta E(x) - \ln Z(\beta)$. Therefore, we recognize that $S(\beta) = -\sum_{x \in \mathcal{X}} \mu_\beta \ln \mu_\beta(x) = H(\mu_\beta)$.

4. We can write the second partial derivative of free entropy as first partial derivative of the negative free energy, and hence

$$\frac{\partial^2}{\partial \beta^2} \Phi(\beta) = -\frac{\partial}{\partial \beta} U(\beta) = -\frac{\partial}{\partial \beta} \left(\frac{\sum_{x \in \mathcal{X}} e^{-\beta E(x)} E(x)}{Z(\beta)} \right) = \sum_{x \in \mathcal{X}} \mu_\beta(x) E(x)^2 + \sum_{x \in \mathcal{X}} \mu_\beta(x) E(x) \frac{1}{Z(\beta)} \frac{\partial Z(\beta)}{\partial \beta}.$$

   We recall that $-\frac{\partial \ln Z(\beta)}{\partial \beta} = \langle E \rangle$, to obtain

$$\frac{\partial^2}{\partial \beta^2} (\Phi(\beta)) = -\frac{\partial^2}{\partial \beta^2} (\beta F(\beta)) = \langle E^2 \rangle - \langle E \rangle^2 .$$

   This is auto-covariance of energy E under the Boltzmann distribution μ_β, and is always positive. This implies the convexity of free entropy as a function of inverse temperature.

□

4 Temperature limits of thermodynamic potentials

We wish to study the behavior of the thermodynamic potentials in the limiting cases of high temperature ($\beta \to 0$) and low temperature ($\beta \to \infty$).

4.1 High temperature limit

In the high temperature limit, we take $\beta \to 0$, and the Boltzmann distribution is approximately the uniform distribution over the set of all states in the state space $\mathcal{X}$. Taking the Taylor series expansion around $\beta = 0$, we get that

$$S(\beta) = H(\mu_\beta) = \ln |\mathcal{X}| + \Theta(\beta),$$

where $|\mathcal{X}|$ denotes the cardinality of the state space $\mathcal{X}$. The internal energy $U(\beta)$ is given by

$$U(\beta) = \langle E \rangle_0 + \Theta(\beta),$$

where $\langle E \rangle_0 = \frac{1}{|\mathcal{X}|} \sum_{x \in \mathcal{X}} E(x)$ is the internal energy at $\beta = 0$. From the above equations, it can be seen that

$$F(\beta) = \langle E \rangle_0 - \frac{1}{\beta} \ln |\mathcal{X}| + \Theta(\beta).$$

4.2 Low temperature limit

In the low temperature limit, we take $\beta \to \infty$ and the Boltzmann distribution is approximately the uniform distribution over the set of ground states with the lowest energy. Let E_0 be the ground state energy and $\mathcal{X}_0$ be the ground state. For a finite configuration space $\mathcal{X}$, we have

$$E_0 \triangleq \min \{E(x) : x \in \mathcal{X}\}, \qquad \mathcal{X}_0 \triangleq \{x \in \mathcal{X} : E(x) = E_0\}.$$

Definition 4.1. We define the energy difference from the ground state energy as $\Delta E(x) \triangleq E(x) - E_0$ for all configurations $x \in \mathcal{X}$. The **energy gap** ΔE is defined to be the difference between the ground state energy level and the next lowest energy level, i.e. $\Delta E \triangleq \min \{\Delta E(y) : y \in \mathcal{X} \setminus \mathcal{X}_0\}$. Let $\mathcal{X}_1 \triangleq \{x \in \mathcal{X} : \Delta E(x) = \Delta E\}$ be the set of states with the energy level $E_0 + \Delta E$.

Proposition 4.2. *In the low temperature limit of $\beta \to \infty$, we have*

$$E_0 + \Theta(e^{-\beta\Delta E}), \qquad S(\beta) = \ln|\mathcal{X}_0| + \Theta(e^{-\beta\Delta E}), \qquad F(\beta) = E_0 - \frac{1}{\beta}\ln|\mathcal{X}_0| + \Theta(e^{-\beta\Delta E}).$$

Proof. We write the Boltzmann distribution in terms of ground state $\mathcal{X}_0$, the ground state energy E_0, and the energy difference $\Delta E(x)$ as

$$\mu_\beta(x) = \frac{e^{-\beta\Delta E(x)}}{|\mathcal{X}_0| + \sum_{x\in\mathcal{X}\setminus\mathcal{X}_0} e^{-\beta\Delta E(x)}}.$$

From the above, it can be seen that for $x \notin \mathcal{X}_0$, $\mu_\beta(x)$ is of the order of $e^{-\beta\Delta E}$ for large β.

1. We can write the internal energy for large β, in terms of $\mathcal{X}_0, \mathcal{X}_1$, and energy gap ΔE as

$$U(\beta) = E_0 + \Delta E \sum_{x\in\mathcal{X}_1} \mu_\beta(x) + \sum_{x\in\mathcal{X}\setminus(\mathcal{X}_0\cup\mathcal{X}_1)} \Delta E(x)\mu_\beta(x).$$

2. Recall that $\ln|\mathcal{X}_0|$ is the entropy of the uniform distribution over the ground state. Hence, we can write the canonical entropy for large β in a similar fashion

$$S(\beta) = \ln|\mathcal{X}_0| + \Theta(e^{-\beta\Delta E}).$$

3. Combining the expression for internal energy and canonical entropy, we can write the free energy for large β as

$$F(\beta) = E_0 - \frac{1}{\beta}\ln|\mathcal{X}_0| + \Theta(e^{-\beta\Delta E}).$$

□

Remark 6. In many settings, the high temperature case is not very interesting since the system is almost uniformly random. We consider a few example systems and study the thermodynamic potentials in the general case and in the low temperature limit where $\beta \to \infty$.

Example 4.3 (Two level system). Consider a system consisting of only two states $\mathcal{Z} \triangleq \{0, 1\}$, with the energy function given by $\epsilon_0 \triangleq E(0), \epsilon_1 \triangleq E(1)$. Without loss of any generality, we can assume $\epsilon_0 < \epsilon_1$, and consequently define the energy gap as $\Delta \triangleq \epsilon_1 - \epsilon_0$. The partition function can be computed to be

$$Z(\beta) = \sum_{x\in\mathcal{Z}} e^{-\beta E(x)} = e^{-\beta\epsilon_0} + e^{-\beta\epsilon_1} = e^{-\beta\epsilon_0}(1 + e^{-\beta\Delta}),$$

and the Boltzmann distribution is given by

$$\mu_\beta = \Big(\frac{1}{1+e^{-\beta\Delta}}, \frac{e^{-\beta\Delta}}{1+e^{-\beta\Delta}}\Big).$$

We can compute the free energy as

$$F(\beta) = -\frac{1}{\beta}\ln Z(\beta) = -\frac{1}{\beta}\ln\left(e^{-\beta\epsilon_0}(1+e^{-\beta\Delta})\right) = \epsilon_0 - \frac{1}{\beta}\ln(1+e^{-\beta\Delta}).$$

The internal energy can be computed to be

$$U(\beta) = \langle E\rangle = \frac{\epsilon_0 + (\epsilon_0+\Delta)e^{-\beta\Delta}}{1+e^{-\beta\Delta}} = \epsilon_0 + \Delta\frac{e^{-\beta\Delta}}{1+e^{-\beta\Delta}}.$$

We can either compute the canonical entropy directly as $S(\beta) = H(\mu_\beta)$ or compute it to be

$$S(\beta) = \beta(U(\beta) - F(\beta)) = \beta\Delta\frac{e^{-\beta\Delta}}{1+e^{-\beta\Delta}} + \ln(1+e^{-\beta\Delta}).$$

From the above equations, we see that, at low temperatures (for high values of β),

$$U(\beta) = \epsilon_0 + \Theta(e^{-\beta\Delta}), \qquad S(\beta) = \ln(1) + \Theta(e^{-\beta\Delta}),$$

which is intuitive, since ϵ_0 is the energy of the ground state and $\ln(1) = 0$ is the entropy of the uniform distribution over the ground state since it consists of only one state, and the extra terms are in the order of $e^{-\beta\Delta}$, which was what was expected.

Example 4.4 (Cylindrical bottle). Consider a single particle in a cylindrical bottle with configuration space $\mathcal{X} = B \times [0, d]$, where the base of the cylinder B is of arbitrary shape with area $|B|$. Let the energy function be the potential energy, $E(x) = wh(x)$, where $h(x)$ is the height of x from the base of the cylinder and w is a scalar. We can then compute the partition function for the continuous Boltzmann distribution as

$$Z(\beta) = \int_{x\in\mathcal{X}} e^{-\beta E(x)} dx = |B| \int_0^d e^{-\beta wu} du = \frac{|B|}{\beta w}\left(1 - e^{-\beta wd}\right).$$

The free energy is then given by

$$F(\beta) = -\frac{1}{\beta}\ln Z(\beta) = -\frac{1}{\beta}\ln|B| - \frac{1}{\beta}\ln\left(\frac{1 - e^{-\beta wd}}{\beta w}\right).$$

The internal energy is given by

$$U(\beta) = \langle E\rangle = \int_{x\in\mathcal{X}} \frac{E(x)e^{-\beta E(x)}}{Z(\beta)} du = \frac{\beta w}{(1 - e^{-\beta wd})}\int_0^d wue^{-\beta wu} du = \frac{1}{\beta} - \frac{wd}{e^{\beta wd} - 1}.$$

From the above, we can determine the canonical entropy as

$$S(\beta) = \beta(U(\beta) - F(\beta)) = 1 - \frac{\beta wd}{e^{\beta wd} - 1} + \ln|B| + \ln\left(\frac{1 - e^{-\beta wd}}{\beta w}\right) = \ln\left(\frac{|B|\, e}{\beta w}\right) - \frac{\beta wd}{e^{\beta wd} - 1}.$$

In the low temperature limit, we see that the second term vanishes. Comparing it with the intuition that at low temperatures, the entropy is the logarithm of the size of the ground state, we see that in the low temperature limit, the particle in the cylinder occupies a volume of order $\frac{|B|e}{\beta w}$. Since the area of the base is $|B|$, the cylinder is occupied up to a height of order $\frac{e}{\beta w}$ from the base.

Example 4.5 (Spherical bottle). Consider the system in which a particle is confined to a sphere $\mathcal{X}$ of radius R centred at $(R, 0, 0)$ in the cartesian coordinate, where the configuration space is given by

$$\mathcal{X} = \{(x_1, x_2, x_3) : (x_1 - R)^2 + x_2^2 + x_3^2 \leqslant R^2\}.$$

Let the energy function be the potential energy, proportional to the height from the (x_2, x_3) plane, i.e,

$$E(x) = wx_1.$$

Then, we can write the partition function as

$$Z(\beta) = \int_{\mathcal{X}} e^{-\beta wx_1} dx_1 dx_2 dx_3 = \int_0^{2R} dx_1 e^{-\beta wx_1} \int_{S_{x_1}} dx_2 dx_3,$$

where S_{x_1} is the subset of $\mathcal{X}$ at height x_1 from the (x_2, x_3) plane. Clearly, S_{x_1} is a circle of radius $\sqrt{R^2 - (x_1 - R)^2}$ and hence its area is $\pi(R^2 - (x_1 - R)^2) = \pi(2Rx_1 - x_1^2)$. Therefore, we can compute the partition function to be

$$Z(\beta) = \int_0^{2R} \pi(2Rx_1 - x_1^2)e^{-\beta w x_1} dx_1 = \frac{2\pi R}{(\beta w)^2}\left[e^{-2\beta w R}\left(1 + \frac{1}{\beta w R}\right) + 1 - \frac{1}{\beta w R}\right]. \tag{2}$$

Now, we can calculate the free energy in terms of partition function as

$$F(\beta) = -\frac{1}{\beta}\ln Z(\beta) = -\frac{1}{\beta}\ln\left(\frac{2\pi R}{(\beta w)^2}\right) - \frac{1}{\beta}\ln\left(\frac{e^{-2\beta w R}(\beta w R + 1) + \beta w R - 1}{\beta w R}\right),$$

and internal energy as the expected energy under the Boltzmann distribution as

$$U(\beta) = \langle E \rangle = \frac{1}{Z(\beta)}\int_{x \in \mathcal{X}} w x_1 e^{-\beta w x_1} dx = \frac{w\pi}{Z(\beta)}\int_0^{2R} u(2Ru - u^2)e^{-\beta w u} du.$$

We can explicitly compute the integration to obtain

$$U(\beta) = \frac{w\pi}{Z(\beta)}\left[\frac{4R^2}{(\beta w)^2}e^{-2\beta w R} + \frac{8R}{(\beta w)^3}e^{-2\beta w R} + \frac{6}{(\beta w)^4}e^{-2\beta w R} + \frac{4R\beta w - 6}{(\beta w)^4}\right].$$

From substituting equation (2) for the partition function in above, we obtain in the low temperature limit (for β large)

$$U(\beta) \approx \left(\frac{4\pi R\beta w^2}{(\beta w)^4} - \frac{6w\pi}{(\beta w)^4}\right)\frac{1}{Z(\beta)} = \left(\frac{4\pi R\beta w^2}{(\beta w)^4} - \frac{6w\pi}{(\beta w)^4}\right)\frac{\beta w^2}{2\pi R}\left(\frac{1}{1 - \frac{1}{R\beta w}}\right).$$

Using the expansion $\frac{1}{1-x} = 1 + x + x^2 + ...$, we obtain

$$U(\beta) = w\left(\frac{2}{\beta w} + \frac{3}{(\beta w)^2 R}\right)\left(1 + \frac{1}{\beta w R} + \Theta\left(\frac{1}{\beta^2}\right)\right) = \frac{2}{\beta} + \Theta\left(\frac{1}{\beta^2}\right).$$

Therefore, the canonical entropy in the low temperature limit is given by

$$S(\beta) = \beta(U(\beta) - F(\beta)) = \beta\frac{2}{\beta} + \beta\Theta\left(\frac{1}{\beta^2}\right) + \ln\left(\frac{2\pi R}{(\beta w)^2}\right) = \ln\left(\frac{2\pi e^2 R}{\beta^2 w^2}\right) + \Theta\left(\frac{1}{\beta}\right).$$

In a low temperature state, since the distribution is uniform over the ground state and the entropy is the logarithm of the size of the ground state, we see intuitively that the particle occupies a volume of the order of $\left(\frac{2\pi R e^2}{\beta^2 w^2}\right)$ at the bottom of the sphere.