3_Statistical_Physics_Definitions

May 30, 2024

3 Statistical Physics Definitions

The field of statitical physics uses approaches of statistics and probability theory to address physical problems. It considers large populations and derives expressions for the ensemble (or the macrostate) of a system from the microscopic states in the system.

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.special import factorial as fac

%matplotlib inline
%config InlineBackend.figure_format = 'retina'

plt.rcParams.update({'font.size': 12,
                     'axes.titlesize': 18,
                     'axes.labelsize': 16,
                     'axes.labelpad': 14,
                     'lines.linewidth': 1,
                     'lines.markersize': 10,
                     'xtick.labelsize' : 16,
                     'ytick.labelsize' : 16,
                     'xtick.top' : True,
                     'xtick.direction' : 'in',
                     'ytick.right' : True,
                     'ytick.direction' : 'in',})
```

3.1 Entropy

3.1.1 Entropy Definition by Boltzmann

The term of entropy becomes very important in that context. It measures the number of different ways a system can be rearranged to yield the same macrostate. It is, thus, an indicator for the microscopic degeneracy of a macrostate. In this context the definition of entropy by **Boltzmann** is well known, i.e.,

$$S = k_\mathrm{B} \ln(W) \tag{1}$$

where W is the number of microstates corresponding to a system's macrostate with and energy E.

Here $k_{\mathrm{B}} = 1.38064852 \times 10^{-23}\,\mathrm{m^2\,kg\,s^{-2}\,K^{-1}}$ is the Boltzmann constant. Below are two examples of how to use the formula for the calculation of the entropy.

Example: Entropy of an N letter word

SMP PMS MSP MPS SPM

SSS PPP MMM MMP MMS

$$\vdots$$

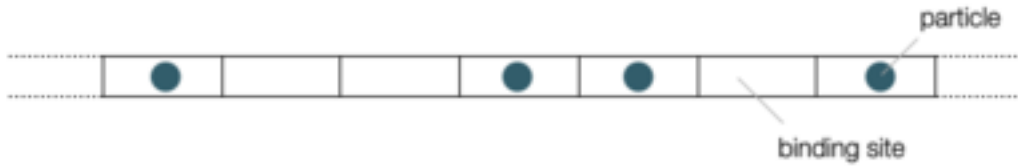Consider the number of states W of a word with N letters of M different characters. The letter can be arranged in

$$W = M^N \tag{2}$$

different ways such that the entropy is given by

$$S = N\,k_{\mathrm{B}}\ln(M) = k_{\mathrm{B}}\sum_1^N \ln(M) \tag{3}$$

which just tells that entropy is an additive quantity.

Example: Arrangement of molecules along a chain

particle

binding site

We consider a linear molecules (perhaps a DNA) that has N binding sites for, e.g., proteins. N_{p} sites are occupied with a protein where the binding energy is equal for each site. The number of different ways in which the N_{p} proteins can be arranged on the N sites is given by the binomial coefficient

$$W(N_{\mathrm{p}}; N) = \frac{N!}{N_{\mathrm{p}}!(N - N_{\mathrm{p}})!}. \tag{4}$$

Therefore, the entropy is given by

$$S = k_{\mathrm{B}}\ln\left(\frac{N!}{N_{\mathrm{p}}!(N - N_{\mathrm{p}})!}\right) \tag{5}$$

which can be further simplified using the identity

$$\ln(N!) = \sum_{n=1}^{N} \ln(n) \tag{6}$$

and the *Stirling approximation*

$$\sum_{n=1}^{N} \ln(n) \approx \int_{1}^{N} \ln(x)\,\mathrm{d}x \approx N\ln(N) - N. \tag{7}$$

This, finally, leads to

$$S = -k_{\mathrm{B}}N[c\ln(c) + (1-c)\ln(1-c)] \tag{8}$$

with $c = N_{\mathrm{p}}/N$ being the mean occupation of each site or the probability to find a state occupied.

Below you just find some Python code calculating the entropy as a function of "concentration" using the Stirling approximation and the original formula. You also recognize there, that the Stirling formula is not yet very good, since $N = 100$.

```
[3]: epsilon=1e-3
     c=np.linspace(epsilon,1-epsilon,100);N=100
     Np=np.arange(1,N)

     def S(x):
         return(np.log(fac(N)/(fac(x)*fac(N-x))))
```

```
[7]: fig=plt.figure()
     plt.ion()
     plt.plot(c,-(c*np.log(c)+(1-c)*np.log(1-c)),'r-',label="Stirling")
     plt.plot(Np/N,S(Np)/N,'k--',label="Original")
     plt.xlabel("$c$")
     plt.ylabel(r"$S$ [$k_\mathrm{B}$]")
     plt.title("Accuracy of the Stirling Formula")
     plt.tight_layout()
     plt.legend()
     plt.savefig("img/stirling.pdf")
     plt.close(fig)
```

Accuracy of the Stirling Formula
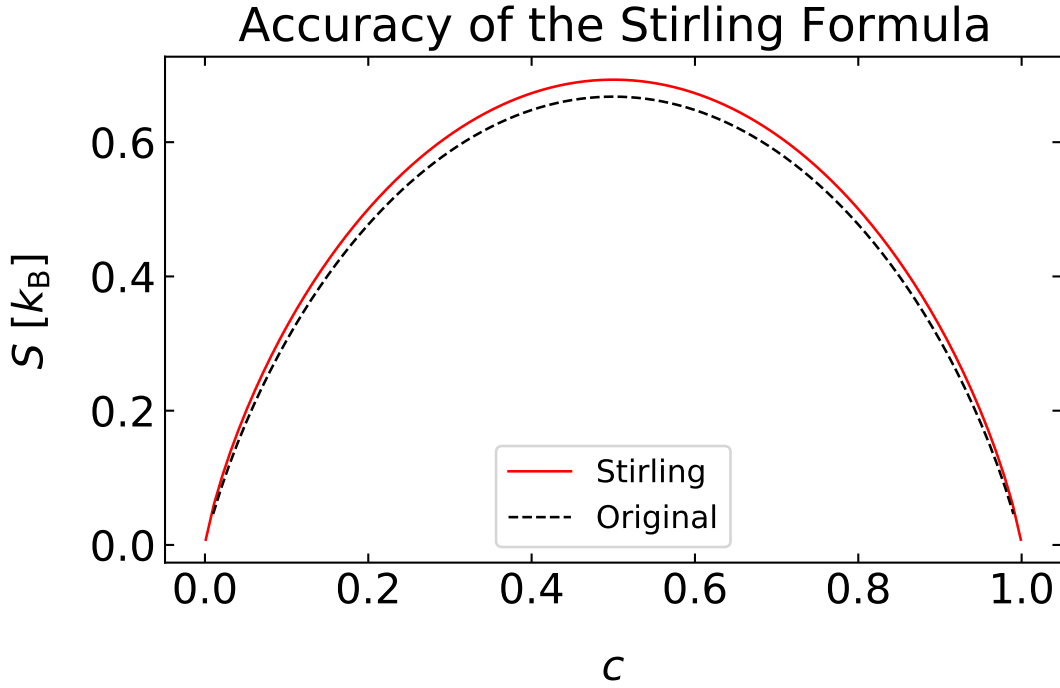
3.1.2 Shannon Entropy

A different access to entropy comes from the field of information theory and has been devised by Claude Shannon. Information theory is trying to mathematically assess the information content of a measurement facing uncertainty. It will turn out further below that this alternative description results in the Boltzmann distribution and effectively amounts to making a best guess about the probability distribution given some limited knowledge about the system such as the average energy.

The Shannon entropy is defined by

$$S\left(p_1, p_2, ..., p_N\right) = S\left(\{p_i\}\right) = -\sum_{i=1}^{N} p_i \ln p_i \tag{9}$$

and relates to its thermodynamic version, the Gibbs entropy

$$S\left(\{p_i\}\right) = -k_\mathrm{B} \sum_{i=1}^{N} p_i \ln p_i, \tag{10}$$

where p_i is the probability of the ith microstate (or outcome). The example below will show, that if only the normalization of the probability is known, maximization of the Shannon entropy will directly lead to an equal probability of events (uniform distribution). Later, we see that similar calculations can be done to yield the Boltzmann distribution.

Example: Uniform distribution

To figure out that the Shannon entropy is indeed delivering some useful measure, we will have a look at a measurement which has N outcomes (e.g., rolling a dice). We, of course, know that in

this case all numbers of the dice have equal probability, but we can test this by maximizing the Shannon entropy as required by our thermodynamic considerations earlier.

To do so, we use the technique of Lagrange multipliers, which allows us to set a constraint while maximizing the entropy. This contraint is for this example, that

$$\sum_i^N p_i = 1 \tag{11}$$

i.e., that the probability is normalized to 1. With this constraint we maximize the entropy by adding an additional term with the contrain multilied by the Lagrange multiplier λ:

$$S' = -\sum_i p_i \ln p_i - \lambda \left(\sum_i p_i - 1 \right). \tag{12}$$

We see, that if the probability is normalized to 1, we do not change the entropy. Our procedure is to find that set of probabilities p_i which maximizes this augmented entropy function.

The derivative of the augmented entropy with respect to λ yields the normalization condition, i.e.,

$$\frac{\partial S'}{\partial \lambda} = -\left(\sum_i p_i - 1 \right) \stackrel{!}{=} 0. \tag{13}$$

Differentiation with respect to the probabilities yields

$$\frac{\partial S'}{\partial p_i} = -\ln p_i - 1 - \lambda \stackrel{!}{=} 0 \tag{14}$$

which directly gives

$$p_i = e^{-1-\lambda}. \tag{15}$$

Together with the normalization condition we therefore obtain

$$\sum_{i=1}^N e^{-1-\lambda} = 1 \tag{16}$$

and since the exponent does not depend on i we find

$$e^{-1-\lambda} = \frac{1}{N} \tag{17}$$

or

$$p_i = \frac{1}{N} \tag{18}$$

which is the expected equal probability of finding one of the N outcomes.

3.2 Boltzmann Distribution

Our previous consideration of the state functions has shown, that thermal equibrium is associated with a minimum in free energy. As the free energy consists of internal energy U (or enthalpy H) and an entropic term $(-TS)$, we may understand this minimization as a competition between the minimization of the internal energy and a maximization of the entropy (since it is $-TS$). The figure below illustrates this competition for a gas in the gravity field.
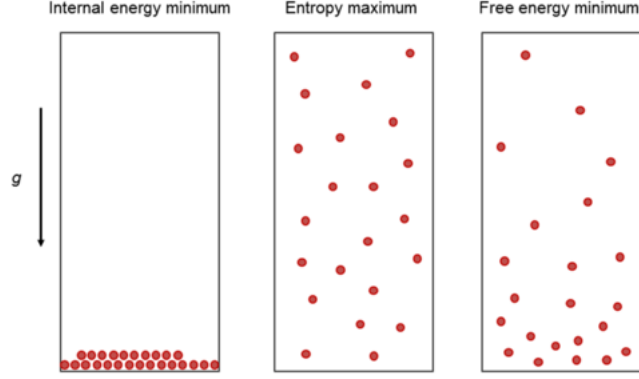
The internal energy minimization yields just a condensed layer at the bottom of the container, while the entropy maximization will try to spread the particles evenly (middle picture). The compromise of both at finite temperature is given by the **barometric height formula**, i.e.,

$$p(z) = p_0 \exp\left(-\frac{mgz}{k_{\mathrm{B}}T}\right),$$ (19)

where $p(z)$ is the probability to find a particle at height z, m is the mass of a particle, g is the gravitational acceleration and p_0 is a normalization constant. The result actually gives a hint at some very fundamental distribution, which always provides the free energy minimum in thermal equilibirum. This distribution is the Boltzmann distribution.

The Boltzmann distribution is an approach of statistical physics to describe a thermodynamic system in equilibrium. The idea is hereby to deliver probability distributions for the probability of all different microstates. Key distinguishing feature of different microstates is their energy E_i (that was so far neglected in the examples above), where i indicates the ith microstate.

The Boltzmann distribution tells us precisely the probability of finding a given microstate with energy E_i: If a particle is in equilibrium with its environment then the probability of finding the particle in state i with energy E_i is

$$p(E_i) = \frac{1}{Z} \exp\left(-\frac{E_i}{k_{\mathrm{B}}T}\right).$$ (20)

The normalization factor $1/Z$ contains the so-called partition function Z:

$$Z = \sum_i \exp\left(-\frac{E_i}{k_{\mathrm{B}}T}\right) = \mathrm{const.}$$ (21)

It ensures that the total probablity to find a system in any of the states is

$$\sum_i p(E_i) = 1. \tag{22}$$

3.2.1 Mean Energy

The Boltzmann distribution is useful to calculate also expectation values, for example, of the total energy of the system (the mean energy $\langle E \rangle$).

The mean is defined by:

$$\langle E \rangle = \frac{1}{Z} \sum_{i=1}^{N} E_i \exp\left(-\frac{E_i}{k_B T}\right). \tag{23}$$

Abbrevating $\beta = (k_B T)^{-1}$ we find

$$\langle E \rangle = \frac{1}{Z} \sum_{i=1}^{N} \left(-\frac{\partial}{\partial \beta} \exp\left(-\beta E_i\right)\right), \tag{24}$$

where the sum is nothing else than the derivative of the partition function

$$\langle E \rangle = -\frac{1}{Z} \frac{\partial}{\partial \beta} Z \tag{25}$$

or just

$$\langle E \rangle = -\frac{\partial}{\partial \beta} \ln(Z). \tag{26}$$

3.2.2 Free Energy

Employ the Gibbs entropy $S = -k_B \sum_{i=1}^{N} p_i \ln p_i$ to find a relation between the free energy F (or G) and the partition function. Inserting the probability $p_i = Z^{-1} \exp(-\beta E_i)$ and doing some transformations yields

$$S = k_B (\ln(Z) + \beta < E >). \tag{27}$$

Using

$$F = U - TS \tag{28}$$

for the free energy, we can insert the above result for the entropy and obtain

$$F = -k_B T \ln(Z) \tag{29}$$

(or G in the same way).

Note that this is the total energy and not the mean energy (internal, enthalpy or free energy) of the states in the system. The partition function thus allows us to calculate the free energy.

3.2.3 Deriving the Boltzmann Distribution

There are a number of ways to derive the Boltzmann distribution. We will have a quick look at a classical derivation of the Boltzmann distribution for a closed system, e.g., a system which is in contact with a reservoir as depicted below.
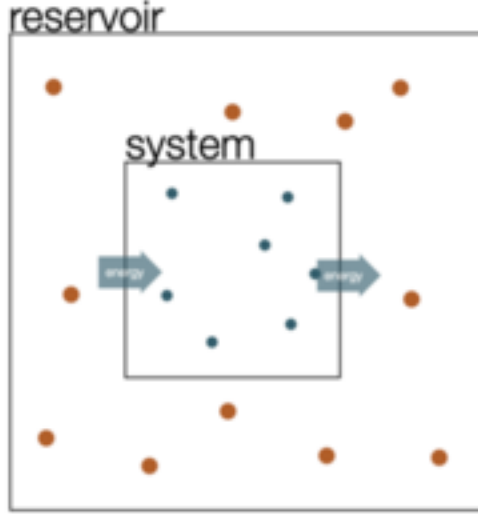
System (index s) and reservoir (index r) have total energy $E_{\text{tot}} = E_{\text{r}} + E_{\text{s}}$. We assert now, that the probability to find the system in a specific microstate $p(E_{\text{s}}^{(i)})$ with the energy $E_{\text{s}}^{(i)}$ is directly proportional to the number of states available to the reservoir, when the system is in that state. The ratio of the probabilities of two states is then equal to the ratio of the number of states of the reservoir, i.e.,

$$\frac{p(E_{\text{s}}^{(1)})}{p(E_{\text{s}}^{(2)})} = \frac{W_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(1)})}{W_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(2)})}. \tag{30}$$

Here, the function $W_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(1)})$ is the number of states available to the reservoir, when the system is having the energy $E_{\text{s}}^{(1)}$.

We can now rewrite the above equation in terms of the entropy using $W(S(E)) = \exp(S(E)/k_{\text{B}})$ such that

$$\frac{W_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(1)})}{W_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(2)})} = \frac{\exp(S_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(1)})/k_{\text{B}})}{\exp(S_{\text{r}}(E_{\text{tot}} - E_{\text{s}}^{(2)})/k_{\text{B}})}. \tag{31}$$

We may now expand the entropy to first order

$$S_{\text{r}}(E_{\text{tot}} - E_{\text{s}}) \approx S_{\text{r}}(E_{\text{tot}}) - \frac{\partial S_{\text{r}}}{\partial E} E_{\text{s}} \tag{32}$$

considering that E_s is only very tiny as compared to the total energy of the reservoir. Using the thermodynamic identity that

$$\frac{\partial S_r}{\partial E}\Big|_{V,N} = \frac{1}{T} \tag{33}$$

we finally find

$$\frac{p(E_s^{(1)})}{p(E_s^{(2)})} = \frac{\exp(-E_s^{(1)}/k_B T)}{\exp(-E_s^{(2)}/k_B T)} \tag{34}$$

which corresponds to the ratio of two Boltzmann distributions

$$p(E_s^{(i)}) = \frac{1}{Z} \exp\left(-\frac{E_s^{(i)}}{k_B T}\right), \tag{35}$$

where Z is the previously mentioned normalization factor, which is called *partition function.*

Example: Barometric height formula

We have mentioned already the barometric height formula giving the probability of finding a particle at a height z. To derive that, we consider a constant gravitational force $F = -mg$ along the z-direction such that the potential energy is given by $E = mgz$ assuming that $E = 0$ at $z = 0$.

The probability for finding the particle at position z is therefore

$$p(z) = \frac{1}{Z} \exp\left(-\frac{mgz}{k_B T}\right) = \frac{1}{<z>} \exp\left(-\frac{z}{<z>}\right). \tag{36}$$

Normalization provides the value of the partition function

$$Z = \int_0^\infty \exp\left(-\frac{mgz}{k_B T}\right) dz = \frac{k_B T}{mg}. \tag{37}$$

We may further calculate the mean height, which is the sedimentation length in sedimentation problems:

$$<z> = \frac{1}{Z} \int_0^\infty h \exp\left(-\frac{mgz}{k_B T}\right) dz = \frac{k_B T}{mg} \tag{38}$$

and the mean energy

$$<E> = -\frac{\partial}{\partial \beta} \ln(Z) = +\frac{\partial}{\partial \beta} \ln(\beta mg) = \frac{1}{\beta} = k_B T = mg <z>. \tag{39}$$

Example: Boltzmann distribution is the maximum entropy distribution in which the average energy is prescribed as a constraint

We can also obtain the Boltzmann distribution from the Shannon entropy by constraining the Shannon entropy. With only the normalization as a constraint in entropy maximization, we obtained equally likely microsates. If we now constrain the mean energy $\langle E \rangle$ of the system, we obtain a distribution which maximizes the entropy under this condition. The mean energy is given by

$$\langle E \rangle = \sum_i E_i p_i, \tag{40}$$

such that we can add another constraint to our augmented entropy. We now have a Lagrange multiplier λ for the normalization of the probability and a second one β which is multiplied by the energy constraint.

$$S' = -\sum_i p_i \ln p_i - \lambda \left(\sum_i p_i - 1 \right) - \beta \left(\sum_i p_i E_i - \langle E \rangle \right). \tag{41}$$

Taking the derivative

$$\frac{\partial S'}{\partial p_i} = -\ln p_i - 1 - \lambda - \beta E_i \overset{!}{=} 0 \tag{42}$$

results in

$$p_i = \mathrm{e}^{-1-\lambda-\beta E_i} \tag{43}$$

and together with the normalization condition $\sum p_i = 1$ finally

$$\mathrm{e}^{-1-\lambda} = \frac{1}{\sum_i \mathrm{e}^{-\beta E_i}}, \tag{44}$$

where we already recognized that we can replace the prefactor $\mathrm{e}^{-1-\lambda}$ by $1/Z$ with Z being the partition function

$$Z = \sum_i \mathrm{e}^{-\beta E_i}. \tag{45}$$

Overall this, therefore, leads to the Boltzmann distribution

$$p_i = \frac{\mathrm{e}^{-\beta E_i}}{\sum_i \mathrm{e}^{-\beta E_i}} \tag{46}$$

which is quite interesting. We have just fixed the mean energy of the system and maximized the entropy. The Boltzmann distribution is therefore the probability distribution which maximizes the entropy under as little as possible additional information (just the mean energy).

The only thing that is missing in the above formula is an expression for the value of β, the Lagrange multiplier. This can be obtained when knowning the mean energy. In thermal euilibrium, this mean energy can be obtained from the equipartition theorem.