Statistical mechanics handout 3

Explain what is meant by the term sample space

Explain what is meant by a probability mass

Explain what is meant by normalisation in the context of probability

Probability Imagine an experiment performed in the future. The outcome from this experiment, X, is something that we cannot possibly know in advance of doing the experiment. We might know at the very least however we know that X will be a one from amongst a set of mutually exclusive outcomes - the sample space Ω. We can then define a quantity known as the probability mass function, p_i. The probability mass for outcome i is defined as:

$$p_i = P(X = x_i)$$

Where here $P(X = x)$ is the probability that the probability that the outcome of our putative experiment, X, is equal to the the particular outcome, x_i. The probability mass function (vector) must be normalised i.e.:

$$\sum_{i=1}^{N} p_i = 1$$

where N is the number of outcomes the experiment could have.

Write the definition of the expectation of a function

Expectation The expectation of a function of a random variable $g(X)$ is defined as:

$$\langle g(X) \rangle = \sum_{i=1}^{N} g(x_i) p_i$$

State the axioms of information theory

Write an expression for the information contained in a probability distribution

Information We define a quantity I (the information) contained in a probability distribution by requiring that this quantity has the following properties (Khinchine)

- The information depends only on the probability distribution.
- The uniform distribution contains the minimum information.
- If we enhance the sample space with impossible events the information does not increase.
- Information is additive.

It is possible to show, starting from these axioms, that the information contained in a probability distribution that has N possible outcomes in the sample space, Ω, that have probabilities given by the vector $\mathbf{p}$, is equal to:

$$I(\mathbf{p}) = k \sum_{i=1}^{N} p_i \ln p_i$$

Summary

Any vector of probabilities must be normalised ($\sum_i p_i = 1$) and will have an information given by $I(\mathbf{p}) = k \sum_{i=1}^{N} p_i \ln p_i$. If each outcome of the experiment has a value b_i associated with it we can define an expectation for this quantity as $\langle B \rangle = \sum_i b_i p_i$

What function do you have to minimise to find the minimum in the function $I(\mathbf{p})$ subject to the constraint $C(\mathbf{p}) = K$?

Constrained Optimisation and Lagrange multipliers To find the optimum value of a function, $I(\mathbf{p})$, subject to some constraints, $C_j(\mathbf{p}) = K_j$ you introduce the set of Lagrange multipliers, $\{\lambda_j\}$ and look for the unconstrained optimum of the extended function:

$$F(\mathbf{p}, \{\lambda_j\}) = I(\mathbf{p}) + \sum_j \lambda_j (C_j(\mathbf{p}) - K_j)$$

In other words you try to find a point where:

$$\frac{\partial F(\mathbf{p}, \{\lambda_j\})}{\partial p_i} = 0 \qquad \forall \quad i \qquad \text{and} \qquad \frac{\partial F(\mathbf{p}, \{\lambda_j\})}{\partial \lambda_j} = 0 \qquad \forall \quad j$$

The basis of this technique is the recognition that, at a constrained optimum the gradients[a] of the constraint functions, $\{\nabla C_j(\mathbf{p})\}$, are either parallel or antiparallel to the gradient of the target function, $\nabla I(\mathbf{p})$.

[a] a vector that points along the direction in which the function is changing most rapidly

Explain the meaning of the term microstate and state the principle of equal a priori probability

Explain how one can calculate the entropy from the probability distribution

Axiom of equal a priori probabilities Any physical system can adopt one of a number of microstates. For a quantum system these are the various quantum levels. For a classical system these are the various combinations of velocity and positions of the atoms in the system. The principle of equal a priori probabilities assumes that any system in equilibrium is equally likely to be in any one of the accessible microstates. Consequently, to determine the probability of being in any microstate one has to minimise the information contained in the distribution. However, to make the analogy with classical thermodynamics clearer we will maximise minus the information, which we will assert is equal to the entropy:

$$S = -k_B \sum_i p_i \ln p_i$$

The sum here runs over the set of all microstates.

Explain what is meant by phase space

Write an expression that can be used to calculate an ensemble average

Phase space The set of all the possible microstates the system can adopt is known as phase space. Each of the microstates in phase space will have associated with it values for all the various extensive variables (volume, energy, number of atoms, magnetisation and so on). In addition to these properties each of the microstates will have a probability associated to it. We can thus determine the ensemble averages of the energy, volume and number of atoms using:

$$\langle E \rangle = \sum_i E_i p_i \qquad \langle V \rangle = \sum_i V_i p_i \qquad \langle N \rangle = \sum_i N_i p_i$$

where the sums runs over all the microstates in phase space and V_i, E_i, N_i and p_i are the volume, the energy, the number of atoms and the probability of microstate i.

Summary

Any physical system can adopt multiple **microstates**. These set of of all these microstates is referred to as **phase space.** There will be values of the extensive variables associated with each of these microstates as well as a probability of being in each of the microstates in the system.

Explain the constraints we have to incorporate when we find the probability distribution that maximises the entropy

Write the extended function that needs to optimised and the derivative of this function with respect to p_i.

How to determine the probability of being in a microstate The axiom of equal a priori probabilities tells us that we can find the probability of being in a state by maximising the entropy, $S = -k_B \sum_{i=1}^{N} p_i \ln p_i$. However, when doing this we must recognise that the problem to be solved here is a constrained optimisation. In particular, normalisation requires that we have:

$$\sum_i p_i = 1$$

Furthermore, if each of the microstates in phase space has some set of properties $\{b^{(j)}\}_i$ (a volume, energy, number of atoms, etc) we also assert that the ensemble average for each of these properties $\langle B^{(j)} \rangle$ must be finite. In other words:

$$\sum_i p_i b_i^{(j)} = \langle B^{(j)} \rangle$$

The extended function we generate using Lagrange's method of undetermined multipliers is thus:

$$I(\mathbf{p}, \lambda_0, \{\lambda^{(j)}\}) = -k_B \textstyle\sum_i p_i \ln p_i - k_B \lambda_0 \left(\sum_i p_i - 1\right) - k_B \sum_j \lambda_j \left(\sum_i p_i b_i^{(j)} - \langle B^{(j)} \rangle\right)$$

Differentiating this with respect to p_i we obtain:

$$\left(\frac{\partial I}{\partial p_i}\right) = -\ln p_i - \psi - \sum_j \lambda^{(j)} b_i^{(j)} \quad \rightarrow \quad p_i = e^{-\psi} e^{-\sum_j \lambda^{(j)} b_i^{(j)}}$$

where we define $\psi = \lambda_0 + 1$.

Write an expression for the generalised partition function

What is the partial derivative $\left(\frac{\partial \psi}{\partial \lambda^{(j)}}\right)$ equal to

The generalised partition function We can derive a value for the $e^{-\psi}$ that appeared in our final expression for the probability, p_i, of being in microstate i by remembering that our probability mass vector must be normalised:

$$\sum_i p_i = e^{-\psi} \sum_i e^{-\sum_j \lambda^{(j)} b_i^{(j)}} = 1 \quad \rightarrow \quad e^{\psi} = \sum_i e^{-\sum_j \lambda^{(j)} b_i^{(j)}}$$

This quantity, e^{ψ}, is called the generalised partition function. It is generally given the symbol Z and we can use it to rewrite the probability of being in a microstate as: $p_i = \frac{e^{-\sum_j \lambda^{(j)} b_i^{(j)}}}{Z}$ We can also determine $\langle B^{(j)} \rangle$ from the partition function by using:

$$\frac{\partial \psi}{\partial \lambda^{(j)}} = \frac{\partial}{\partial \lambda^{(j)}} \ln\left[\sum_i e^{-\sum_j \lambda^{(j)} b_i^{(j)}}\right] = \frac{-\sum_i b_i^{(j)} e^{-\sum_j \lambda^{(j)} b_i^{(j)}}}{Z} = -\langle B^{(j)} \rangle$$

Write an expression for the entropy in terms of ψ, $\{\lambda^{(j)}\}$ and $\{\langle B^{(j)} \rangle\}$

The entropy We next remember that the entropy can be calculated using $S = -k_B \sum_i p_i \ln p_i$. Inserting the formula for the probability that we derived above this is:

$$\frac{S}{k_B} = \sum_i p_i \left[\psi + \sum_j \lambda^{(j)} b_i^{(j)}\right] = \psi \sum_i p_i + \sum_j \lambda^{(j)} \sum_i p_i b_i^{(j)} = \psi + \lambda^{(j)} \langle B^{(j)} \rangle$$

Summary

We can find the probability of being in any microstate by performing a constrained optimisation using the method of Lagranges undetermined multipliers. This process leads us to a natural definition for the generalised partition function.

Write an expression for the differential $\mathrm{d}S$ as a linear combination of the differentials $\{\mathrm{d}\alpha_k\}$ and $\{\mathrm{d}\langle B^{(j)}\rangle\}$

Explain the meaning of the term generalised force

Changes in entropy We now suppose that the values of the set of properties (the $\{b^{(j)}\}_i$) for each of the microstates depend on some set of external parameters $\{\alpha_k\}$. We then ask if it is possible to write an expression for the differential, $\mathrm{d}S$, of the entropy change. Using the expression that we arrived at for the entropy in the previous part we can write:

$$\frac{\mathrm{d}S}{k_B} = \mathrm{d}\psi + \sum_j \lambda^{(j)} d\langle B^{(j)}\rangle + \sum_j \langle B^{(j)}\rangle \mathrm{d}\lambda^{(j)}$$

Furthermore, remembering that the value of the properties of each of the microstates depend on α and recalling our old friend from PDEs, we arrive at:

$$\mathrm{d}\psi = \sum_k \left(\frac{\partial\psi}{\partial\alpha_k}\right)\mathrm{d}\alpha_k + \sum_j \left(\frac{\partial\psi}{\partial\lambda^{(j)}}\right)\mathrm{d}\lambda^{(j)} = \sum_k \left(\frac{\partial\psi}{\partial\alpha_k}\right)\mathrm{d}\alpha - \sum_j \langle B^{(j)}\rangle \mathrm{d}\lambda^{(j)}$$

Inserting this result into the previous expression and calculating $\left(\frac{\partial\psi}{\partial\alpha_k}\right)$ using the definition of the generalised partition function we arrive at:

$$\frac{\mathrm{d}S}{k_B} = \sum_k \langle F_k\rangle \mathrm{d}\alpha_k + \sum_j \lambda^{(j)} \mathrm{d}\langle B^{(j)}\rangle \quad \text{where} \quad \frac{F_k}{k_B T} = -\sum_j \lambda^{(j)} \left\langle \frac{\partial b^{(j)}}{\partial \alpha_k} \right\rangle$$

F_k is a quantity known as a generalised force. Each of the terms inside the summation is an ensemble average of the derivative of the $b^{(j)}$ properties with respect to α_k.

Write an expression for the ensemble average of an extensive quantity

Write an expression for the ensemble average of an intensive quantity

Explain the connection between Lagrange multipliers and intensive quantities

Extensive and intensive quantities In statistical thermodynamics thermodynamic variables are either held fixed or they are calculated by taking ensemble averages over phase space. It is easiest to understand this process by thinking about the extensive variables. Each of the microstates has an associated value for the set of extensive variables - these are the set of $\{b^{(j)}\}_i$ values that we have discussed in previous sections. If we assert that the volume must equal V_1 then the system obvious cannot be in any microstate with a volume not equal to V_1. In other words the probability of being any state i with $V_i \neq V_1$ must be equal to zero.
Imagine we do not fix the energy. The consequence of this is that there is a finite probability of being in microstates of any energy as long as they have volume V_1. Obviously, however, the probabilities of being in high energy microstates will be considerably lower than those of being in low energy microstates. Our requirement that the ensemble average of the energy be finite introduces a Lagrange multiplier and as it turns out the value of this Lagrange multiplier is related to the value of the conjugate intensive thermodynamic variable. In our example with the energy it is related to the temperature. Furthermore, because a Lagrange multiplier is just a fixed number, the value of the corresponding intensive thermodynamic variable must be fixed.
Ensemble averages for non-constrained thermodynamic variables can be calculated using:

$$\text{Extensive:} \quad \langle B^{(j)}\rangle = \sum_i b_i^{(j)} p_j \qquad \text{Intensive:} \quad \left\langle \frac{\partial b^{(j)}}{\partial \alpha_k} \right\rangle = \sum_i \left(\frac{\partial b_i^{(j)}}{\partial \alpha_k} \right) p_j$$

In the above expression α_k is an extensive variable that is held fixed.

Summary

We calculate thermodynamic variables for equilibrated macro states by taking ensemble averages over all the microstates in phase space.