Freie Universität Berlin
Fachbereich Physik
June 5th 2018

Prof. Dr. Roland Netz
Douwe Bonthuis
Jan Daldrop
Philip Loche

Statistical Physics and Thermodynamics (SS 2018)

Problem sheet 7

Hand in: Friday, June 15 during the lecture

http://www.physik.fu-berlin.de/en/einrichtungen/ag/ag-netz/lehre/

1 Ideal gas in the canonical ensemble (12 points)

The canonical partition function of an ideal gas is given by

$$Z = \frac{1}{N!}\left(\frac{V}{\lambda_t^3}\right)^N, \ \ \lambda_t = \frac{h}{\sqrt{2\pi m k_\mathrm{B} T}}. \tag{1}$$

a) Obtain the free energy $F(N,V,T)$ from the canonical partition function. **(1 point)**

b) Start from the definition $F = U - TS$ and use the first law of thermodynamics $dU = TdS - PdV$ to derive the differential relation

$$dF = -PdV - SdT. \tag{2}$$

(1 point)

c) Use your results from a) and b) to construct the entropy $S(N,V,T)$. **(2 points)**
Hint: Note the T-dependence of λ_t.

d) Obtain the caloric equation of state

$$U(N,V,T) = \frac{3}{2}Nk_\mathrm{B}T \tag{3}$$

from the relation $F = U - TS$ and by using your results from a) and c). **(1 point)**

e) Construct an expression for the entropy $S(U,V,N)$ from your previous results, and compare your result to the Sackur-Tetrode equation derived in the lecture

$$S(N,V,U) = -k_\mathrm{B}N\ln\left(\frac{N}{V}\right) + \frac{3}{2}k_\mathrm{B}N\ln\left(\frac{U}{N}\right) + k_\mathrm{B}N\left(\frac{3}{2}\ln\left(\frac{4\pi m}{3h^2}\right) + \frac{5}{2}\right). \tag{4}$$

(2 points)
Hint: Start by obtaining $S(N,V,T)$ and $T(N,V,U)$.

f) Invert eq. (4) to get an expression for $U(S,V,N)$. **(1 point)**

g) Calculate the temperature as a function of S,V,N by an appropriate derivative of $U(S,V,N)$. **(2 points)**

h) Invert this equation again to obtain a result for $S(N,V,T)$. Is it equal to your result from c)? **(1 point)**

i) As an application, calculate the heat capacity C_V of the ideal gas from the caloric equation of state eq. (3). **(1 point)**

2 Shannon entropy (8 points)

The Shannon entropy for a discrete random variable x_i is defined as

$$H = -\sum_i p_i \log p_i. \tag{5}$$

a) Consider a binary random variable with $x_i \in \{0, 1\}$. Calculate the Shannon entropy for

   (a) $p_0 = p_1 = 0.5$,

   (b) $p_0 = 0.9, \;\; p_1 = 0.1$,

   (c) $p_0 = 1, \;\; p_1 = 0$.

   Use your results to explain why the Shannon entropy is a measure for information content. **(1 point)**

Remark: If one replaces the logarithms in eq. (5) by base-2 logarithms, the Shannon entropy is a measure for the minimum number of bits required for a lossless compression of the output.

b) Show that $H \geq 0$. **(1 point)**

c) Derive the maximal value of H for a random variable N possible outcomes x_i. **(2 points)**
   Hint: Introduce a Lagrange multiplier to ensure $\sum_i p_i = 1$ when you maximize H.

d) Given the definitions

$$H_x = -\sum_i p(x_i) \log p(x_i), \tag{6}$$

$$H_y = -\sum_j p(y_j) \log p(y_j), \tag{7}$$

   and

$$H_{xy} = -\sum_{ij} p(x_i, y_j) \log p(x_i, y_j), \tag{8}$$

   show that $H_{xy} \leq H_x + H_y$, where the equality holds if and only if x_i and y_i are independent. **(2 points)**

A related quantity to the Shannon entropy is the so-called Kullback-Leibler divergence for two (discrete) probability distributions p, q. It is defined by

$$D(p||q) = \sum_i p_i \log(\frac{p_i}{q_i}). \tag{9}$$

e) Calculate the Kullback-Leibler divergence for the two cases

   (a) $p_i = q_i$,

   (b) $q_i = 1/N = \text{const.}$

   Use your results to argue why the Kullback-Leibler divergence is a measure for the information lost if one approximates a (true) probability distribution p by a (model) distribution q. **(1 point)**

The cross entropy of two distributions p and q is defined as

$$H_{\text{ce}}(p; q) = -\sum_i p_i \log q_i. \tag{10}$$

f) Express the cross entropy in terms of the Shannon entropy and the Kullback-Leibler divergence. **(1 point)**

Remark: The cross entropy and the Kullback-Leibler divergence play an important role as objective functions in machine learning.