Exercise 07

The Boltzmann distribution

Deadline: Please, send your solutions to **luca.donati@fu-berlin.de** by **Friday, 25 June, 10.15 a.m.**

7.1 Free energy perturbation I (10 P)

Consider four states - A, B, C, and D - which have the following harmonic potential energy functions

$$\begin{aligned} V_A(x) &= 2 \cdot (x+2)^2 - 2 \\ V_B(x) &= 1.2 \cdot x^2 \\ V_C(x) &= 5 \cdot (x-1.5)^2 - 2 \\ V_D(x) &= 0.5 \cdot (x-3)^2 + 1 \end{aligned} \tag{1}$$

Assume that

$$\beta = \frac{1}{k_B T} = 1\,. \tag{2}$$

(a) Plot the Boltzmann distributions of these four states into a single graph. (2 P)

(b) Calculate all possible free energy differences between these four states using the free energy perturbation formula (3 P)

$$\Delta A_{AB} = A_B - A_A = -k_B T \ln \left\langle \exp\left(\frac{V_A(x) - V_B(x)}{k_B T}\right)\right\rangle_A . \tag{3}$$

(c) Verify that the the free energy differences are consistent with the fact that the free energy is a state function. (2 P)

(d) Order the states according to their free energies. Can you understand the ordering by considering the potential energy functions? (3P)

7.2 Principle Component Analysis (10 P)

We are going to realize a Principle Component Analysis (PCA) of the trajectory available on the website.

(a) Generate a scatter plot of the trajectory.

(b) Calculate the mean values $\langle x \rangle$, $\langle y \rangle$ and mark on the axes of the plot.

(c) Calculate the covariance matrix $C_{ij} = \langle (x_i - \langle x_i \rangle)(x_j - \langle x_j \rangle) \rangle$ with $i.j = 1, 2$. In MATLAB, you can compute the covariance matrix with the function C=cov(x,y).

(d) Calculate the eigenvalues and eigenvectors of the covariance matrix.

(e) Plot the eigenvectors into the scatter plot centered at ($\langle x \rangle$, $\langle y \rangle$).

(f) Project the trajectory onto the eigenvectors following this procedure:

   (a) Sort the eigenvector matrix according to the order of the eigenvalues. The first column has to be the eigenvector with the highest eigenvalue (this is the first principal component), the last column has to be the eigenvector with the lowest eigenvalue. You can sort the eigenvalues in descending order, using the function sort(lambda,'descend').

   (b) Transpose the new eigenvector matrix.

(c) Subtract the mean values $\langle x \rangle$, $\langle y \rangle$ from the trajectory and transpose the new dataset.

(d) Multiply the eigenvector matrix by the adjust dataset to get the new dataset respect to the principal components.

(e) Plot the new trajectory.

(g) Find the time points in which the projected trajectories have the largest and the lowest value. Mark the corresponding data points in the scatter plot.