Due date: November 26, 2024
Please submit a pdf with answers (add figures if appropriate) as well as the notebook used to generate them.

Principal component analysis (PCA) can be used to identify a low dimensional representation of data while keeping as much of the variance in the data as possible. As discussed in the lecture, PCA might not be suitable for the analysis of dynamical simulations data, in particular for the identification of reaction coordinates of rare events, as the direction of largest variance in the data and of the slowest motion can differ. To perform PCA, we will use the function provided in the notebook.

1 2D double well (6 pt)

As a first model potential, we will use a simple double well system in two dimensions:

$$V(x, y) = (x^2 - 1.5)^2 + 0.05y^2 \tag{1}$$

1.1 Potential visualization

Implement a function to compute the potential and its derivatives and visualize the potential in a range of $x \in [-3, 3]$ and $y \in [-10, 10]$ (for generating a grid of points you can use `numpy.meshgrid`). For this potential, the two stable states are defined as A for $x < -0.8$ and B for $x > 0.8$.

1.2 Trajectory data

We use Molecular Dynamics (MD) to simulate a particle in this potential. An example code for the MD algorithm is given in the jupyter notebook. The corresponding forces $f_x = -\partial V(x, y)/\partial x$ and $f_y = -\partial V(x, y)/\partial y$ and the kinetic energy $E_{\text{kin}} = \frac{1}{2}mv^2$ need to be added. Run a simulation in the potential with $\Delta t = 0.05$, $\beta = 3.0$, particle mass $m = 1.0$, and $\gamma = 1.0$ for a total of 10^6 steps. Initial velocities are drawn from the corresponding Maxwell-Boltzmann distribution and subsequently rescaled to match the target temperature. The temperature is related to the kinetic energy by $T = 2E_{\text{kin}}/N_f$ where N_f is the number of degrees of freedom. Print out the x- and y-position, the running average of β, and if the system is in state A or B every 10^5 steps. Save the particle position every 100 steps (10^4 slices in total).

Plot the x and y position as a function of time. How does each coordinate evolve as a function of time? Make a scatter plot of the trajectory data over the range $x \in [-6, 6]$ and $y \in [-6, 6]$. Describe the distribution of points in the xy-plane.

1.3 PCA

Fit a PCA model using the trajectory data. Report the first and second principal components and visualize their direction together with a scatter plot of the trajectory data in the xy-plane. Plot the trajectory data projected on the first principal component as a function of time. Can you identify the transition between state A and B in this projection? Explain your observation.

2 z-potential (4 pt)

As a second model potential, we use the z-potential:

$$\begin{aligned} V(x, y) = & \frac{x^4 + y^4}{20480} - 3e^{-0.01(x+5)^2 - 0.2(y+5)^2} - 3e^{-0.01(x-5)^2 - 0.2(y-5)^2} \\ & + \frac{5e^{-0.2(x+3(y-3))^2}}{1 + e^{-x-3}} + \frac{5e^{-0.2(x+3(y+3))^2}}{1 + e^{x-3}} + 3e^{-0.01(x^2+y^2)} \end{aligned} \quad (2)$$

defined over the range $x \in [-15, 15]$ and $y \in [-10, 10]$.

2.1 Potential visualization

Visualize the potential in a range of $x \in [-15, 15]$ and $y \in [-10, 10]$. For this potential, the two stable states are defined as A for $x < -4.0, y < -3.0$ and B for $x > 4.0, y > 3.0$.

2.2 Trajectory data

Use the Langevin MD integrator and run a simulation in the potential with $\Delta t = 0.05$, $\beta = 1.5$, and $\gamma = 1.0$ for a total of 10^7 steps. Print out the x- and y-position, the running average of β, and if the system is in state A or B every 10^6 steps. Save the trajectory every 1000 steps (10^4 slices in total).

Plot the x and y position as a function of time. How does each coordinate evolve as a function of time? Make a scatter plot of the trajectory data over the range $x \in [-15, 15]$ and $y \in [-10, 10]$. Describe the distribution of points in the xy-plane.

2.3 PCA

Fit a PCA model using the trajectory data. Report the first and second principal components and visualize their direction together with a scatter plot of the trajectory data in the xy-plane. Plot the trajectory data projected on the first principal component as a function of time. Can you identify the transition between state A and B in this projection? Explain your observation.