Luis Vollmers | **Dr. Maria Reif** | **Prof. Martin Zacharias**

Molecular Dynamics (SS2025)
Exercise 3

Temperature Coupling

Report Tasks

1. Include the commented thermostat-`MDP` file of your choice.
2. For all six production runs, plot T, E_{kin}, E_{tot} and E_{pot} vs simulation time and describe them. Particularly, find differences and similarities between the thermostats.
3. Calculate and compare the C_{v} of each production run with the literature value and interpret the results.
4. Visualize the polymer trajectory with `vmd` and describe any unusual behaviour. You may include screenshots of representative conformations.
5. Compare the `MDP` file from the current simulation with last week's `MDP` file and speculate which alterations might have caused the behaviour at hand.

Figure 1: Cartoon Depiction of the Flying-Ice-Cube effect.

Comparing Different Temperature Coupling Algorithms

In the exercise folder, further simulation input files can be found apart from the polymer files. These files comprise two different water models and three different temperature coupling algorithms and they can thus be used to perform six different simulation sequences in total. In the context of an MD simulation, water model means the collection of force field terms describing the true atomistic behaviour of real water up to a certain approximation. Similar to the polymer model of the last exercise sheet. Going by the book, one simulation sequence consists of an energy minimisation, an equilibration and a production run. The goal is to compare the specific heat capacity C_v for all different water models and thermostats with the real world measurement from the literature.
There are 6 different MDP files for this part of the exercise: `eq_nve.mdp`, `prod_nve.mdp`, `eq_berendsen.mdp`, `prod_berendsen.mdp`, `eq_nhc.mdp` and `prod_nhc.mdp`. They can be categorised into equilibration MDP files and production MDP files. The simulations that utilize the equilibration MDP files feature less timesteps. The purpose is to heat up the system and let it settle into its thermodynamic equilibrium without skewing the production data by factoring in this non-equilibrium process. Eventually, the production MDP files and their respective simulations serve the purpose of data collection. The measurements of interest are equilibrium properties and their accuracy increases the longer the simulations run. Therefore, each production simulation takes quite some time to finish and only the production output files will be used for calculating C_v and other properties.

Additionally, the six MDP files can be categorised by the thermostat they are using. The three temperature-coupling algorithms used here are: `None` (NVE-Ensemble), `Berendsen` and `Nose-Hoover-Chain`. They all perform very differently in generating the correct canonical ensemble, which is thermodynamically connected to the heat capacity. The NVE simulation is supposed to perform worst, since it resembles the wrong ensemble alltogether.
To get an idea of how GROMACS implements the different thermostats and therefore the different ensembles, it is recommended to read through the MDP files and compare them, e.g. with `vimdiff <file1> <file2>`. Furthermore, for an advanced understanding of the MDP files, it is helpful to use the GROMACS manual[1] to add a comment behind each line that explains their meaning. Comments can be inserted into an MDP file with `; <Comment>`. After commenting, a line could look like the following:

```
intergrator = md ; Leap-Frog algorithm for time integration
```

GROMACS will ignore everything behind the semi-colon. Include one completely commented MDP file of your choice in the report.

[1] https://manual.gromacs.org/documentation/2018/user-guide/mdp-options.html

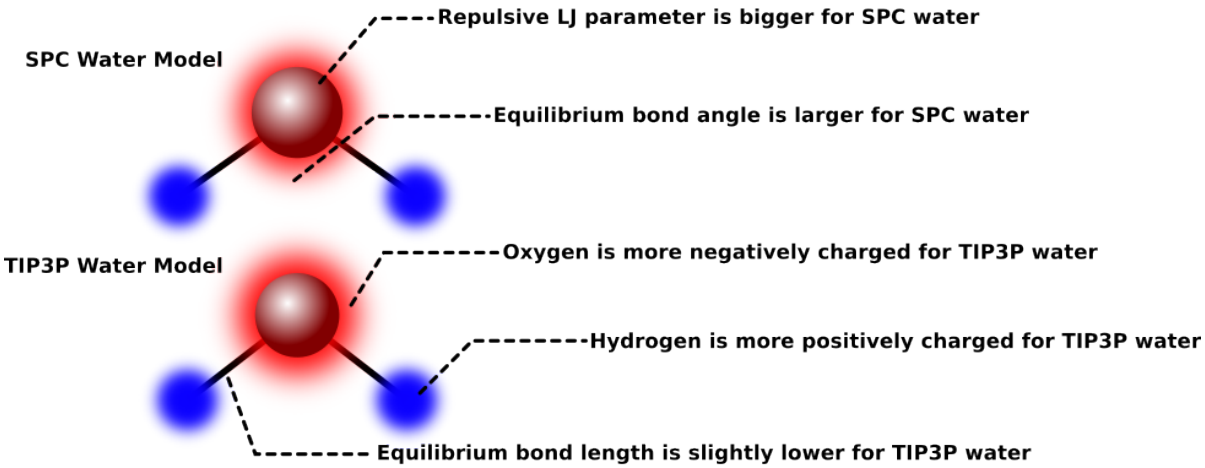

Figure 2: TIP3P and SPC water in comparison. The bond lengths, angles, charges and Lennard-Jones parameters are true to scale and depicted by the different features of each molecule. The charge is depicted by the blue and red blurs and the repulsive Lennard-Jones parameter is depicted by the red bubble. Even though the differences are barely visible, the impact on the accuracy is immense and highlights the importance of precise parameterisation in MD. The exact parameters are listed elsewhere (see below).

For a more consistent comparison, C_v shall be calculated for two distinct water models. In this exercise, the `SPC` and the `TIP3P` water model will be used. The differences between these water models can be investigated by reading their `*.itp` files or by inspecting the wikipedia webpage[2]. Although they behave quite different, both of them are so called three-point-water-models, containing three interaction sites only, representing the three atoms of the water molecule. There are also four-point-water-models and five-point-water-models that benefit from a better charge distribution, but are computationally more expensive.
The `MDP` files do not contain information about the water model used in the simulation sequence. In fact, this information is stored in the `GRO` and `TOP` file which are called `spcbox.gro`, `spc.top`, `tip3pbox.gro` and `tip3p.top`. The differences of these files can be investigated easily by using the aforementioned command called `vimdiff`. As a reminder, the topology contains information about the forcefield used, i.e. the parameters that distinguish `TIP3P` and `SPC` water from each other and the number of molecules present in the system. The `GRO` files contain the box dimensions, the atom coordinates and, if given, their velocities.

The following procedure describes how to conduct the simulation sequences. The first step is to minimise the water boxes. This is done using `gmx grompp` followed by `gmx mdrun`, e.g. for `SPC` water:

```
gmx grompp -f em.mdp -c spcbox.gro -p spc.top -o em_spc.tpr
gmx mdrun -v -deffnm em_spc
```

Notice that the water boxes only need to be minimised once, not for every temperature coupling. Once the minimisation is ready, conduct the equilibration in a similar matter, e.g. for the Nose-Hoover-Chain thermostat:

```
gmx grompp -f eq_nhc.mdp -c em_spc.gro -p spc.top -o eq_nhc_spc.tpr
gmx mdrun -v -deffnm eq_nhc_spc
```

The given `flags` always have their command-specific meaning. E.g. `-f` needs to be followed by the (correct) `MDP` file for the `gmx grompp` command. The flag `-v` prompts the `gmx mdrun`

[2] `https://en.wikipedia.org/wiki/Water_model`

to be more transparent about it's procedures. For detailed information, consult the command line reference[3].
Finally, an example for the execution of a production run can be seen below:

```
gmx grompp -f prod_nhc.mdp -c eq_nhc_spc.gro -p spc.top -o prod_nhc_spc.tpr
gmx mdrun -v -deffnm prod_nhc6_spc
```

All consecutive runs, covering all water models and temperature coupling algorithms respectively, shall be executed accordingly.

Once a production run finishes, the data collection can begin. For this purpose the command `gmx energy` is recommended. Each simulation produces an `EDR` file that stores energy information over the course of the simulation run. `gmx energy` uses this file to calculate all sorts of properties. The properties of interest in this exercise are the temperature T, kinetic energy E_{kin}, total energy E_{tot} and potential energy E_{pot}. Furthermore, `gmx energy` can directly calculate the heat capacity C_{v} if the flags `fluct_props` and `-nmol` are specified. The complete command could look like the following:

```
gmx energy -f prod_nhc_spc.edr -s prod_nhc_spc.tpr -o nhc_spc.xvg -fluct_props -nmol <num>
```

Upon command execution the energetic properties to be processed by the program will be defined via user input. They can either be selected by a number or by writing out the complete name of that property. Include the plots of the four energetic properties (saved to `nhc_spc.xvg` for the example above) in your report for each of the production runs and describe differences in the plot appearance between the temperature coupling algorithms.
Additionally, compare the six values calculated for C_{v} with the literature value of $74.4\,\frac{\mathrm{J}}{\mathrm{mol\,K}}$ (for water at 300 K and 1 bar)[4] and report the relative deviation. If all options were set correctly, `gmx energy` prints the values to the terminal automatically, but only to the terminal.
Herein, C_{v} is used as a measure for the quality of the canonical ensemble generated by the temperature coupling algorithm. Based on that, propose a best-practice thermostat. `GROMACS` calculates C_{v} from the total energy fluctuations $\mathrm{RMSD}(E_{\mathrm{tot}})$ the specified molecule number N_{mol}, the ideal gas constant R and the temperature T[5].

$$C_{\mathrm{v}} = \frac{(\mathrm{RMSD}(E_{\mathrm{tot}}) \cdot 1000)^2 \cdot N_{\mathrm{mol}}}{RT^2} \tag{1}$$

The Flying-Icecube-Effect with a Polymer Model

In the prior exercies, a polymer was simulated with different force field terms. Instead of simulating this polymer again yourself, there is a pre-made trajectory contained in the online repository that shows some strange artifacts. The three input files used to generate the MD run, `polymer.gro`, `polymer*.top`, and `polymer.mdp`, should be downloaded together with the output files. Upon comparison of the `polymer*.top` with the corresponding file from the prior exercise, it is revealed that all force field terms are activated in the current run. However, visualizing the trajectory reveals a very strange, unnatural behavior, which entered scientific literature under the name *Flying-Icecube-Effect*. Some snapshots of the trajectory should be featured in the report, together with a description of the behavior of the polymer during the simulation. Some instructions are listed below as a reminder on how to visualize trajectories nicely in VMD.

[3] `https://manual.gromacs.org/documentation/2018/user-guide/cmdline.html`
[4] Engineering ToolBox, (2004). Water - Specific Heat vs. Temperature. [online] Available at: `https://www.engineeringtoolbox.com/specific-heat-capacity-water-d_660.html` [2023-05-17]
[5] M. P. Allen and D. J. Tildesley, Computer Simulation of Liquids, Clarendon, Oxford, 1987 under *Heat Capacity*.

```
vmd polymer.gro polymer_non_bonded.xtc
```

`Graphics` ⇒ `Representations` ⇒ window opens on Tab `Draw Style` ⇒ `Drawing Method` ⇒ `Licorice` (⇒ optional: adjust Bond Radius)

Furthermore, a white background is better suited for inclusion in scientific reports.

`Graphics` ⇒ `Colors` ⇒ `Display` ⇒ `Background` ⇒ `white`

Given the odd behaviour of the polymer, it might come in handy to display several frames at once to hint towards the dynamics of the system.

`Graphics` ⇒ `Representations` ⇒ switch to Tab `Trajectory` ⇒ change the option for `Draw Multiple Frames` ⇒ E.g.: `1000:1005`

The unnatural behavior is even easier to observe after increasing the `Step` value from 1 to 3. Obviously, since the last week's simulation did not exhibit any oddities, there must be a difference in the input files of both runs, and indeed, these differences can be found in the `polymer.mdp`. Inspect the current and last week's `polymer.mdp` files for differences and argue which change caused this behavior and how it was circumvented in last week's tutorial. The command `vimdiff` is recommended for a fast and error-proof comparison. Include a brief assessment of whether this section's findings confirm your proposal for a best-practice thermostat or not.