

# Worksheet 1: Properties and Fitting of Atomistic Water Models

Frank Uhlig, Jens Smiatek, Johannes Zeman, Christian Holm

Institute for Computational Physics, University of Stuttgart

April 18, 2017

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Short Questions - Short Answers (5 points)</b>	<b>2</b>
<b>3</b>	<b>Polarization energy in bulk water (5 points)</b>	<b>3</b>
<b>4</b>	<b>Atomistic Water Simulations with GROMACS (10 points)</b>	<b>4</b>
4.1	Visualizing the System . . . . .	4
4.2	Running the Simulations . . . . .	5
4.3	Analysis . . . . .	6

## General Remarks

- Deadline for the report is **Monday, 1st May 2017, 12:00 noon**
- In this worksheet, you can achieve a maximum of 20 points.
- The report should be written as though it would be read by a fellow student who attends the lecture, but doesn't do the tutorials.
- To hand in your report, send it to your tutor via email.
  - Frank (fuhlig@icp.uni-stuttgart.de)
- Please attach the report to the email. For the report itself, please use the PDF format (we will *not* accept MS Word doc/docx files!). Include graphs and images into the report.

- The report should be 5–10 pages long. We recommend using L<sup>A</sup>T<sub>E</sub>X. A good template for a report is available online.
- The worksheets are to be solved in **groups of two or three** people.

## 1 Introduction

In this worksheet, you will first tackle several theoretical tasks related to water models and their construction. In the second part, you will create your own water model using the open-source Molecular Dynamics software package GROMACS. In the remainder of the worksheet, you will run various analyses on the simulation data and investigate the properties of your “optimal” water model.

All files required for this tutorial can be found in the archive `templates.tar.gz` which can be downloaded from the lecture’s homepage.

## 2 Short Questions - Short Answers (5 points)

Have a look at [http://www.edinformatics.com/interactive\\_molecules/water.htm](http://www.edinformatics.com/interactive_molecules/water.htm).

<b>Task</b>	(5 points)
Answer the following questions:	
<ul style="list-style-type: none"> <li>• Why is a water molecule polar?</li> <li>• What is a hydrogen bond?</li> <li>• What are the values for the H-O-H angle in water and the typical distance of a hydrogen bond (<i>not</i> O-H bond)?</li> <li>• What are the main differences between various atomistic water models?</li> <li>• What is the energy of a SPC water molecule, at its equilibrium geometry, in gas phase.</li> </ul>	

### Hints

- The above website provides some nice applets to play with. Anyway: Don’t trust the results of the measurements you can make there!

### 3 Polarization energy in bulk water (5 points)

Common, empirical molecular force fields are either parametrized against accurate quantum mechanical calculations or experimental data. In the latter case, one samples the parameter space in a given force-field expression until satisfactory agreement at a (or various) state point(s) is achieved. Often used experimental data is the heat of vaporization, which provides information about the strength of association in the bulk liquid compared to the gas phase.

<b>Task</b>	(2 points)
<ul style="list-style-type: none"><li>• Derive an explicit formula for the heat of vaporization <math>H_{\text{vap}}</math> and simplify it with respect to the volume-work term. Assume that the gas is ideal and that there is no change in geometry or vibrations when going from liquid to gaseous state.</li><li>• Identify the terms to be calculated in a molecular dynamics simulation, for a theoretical estimate of <math>H_{\text{vap}}</math>.</li></ul>	

One often neglected term in empirical force fields is the electronic polarizability. It accounts for the distortion of the electronic density of a molecule due the electric field of the surrounding molecules in a liquid, and vice versa. Therefore, induced multipole moments need to be calculated self-consistently at each step of the simulation. To avoid this additional computational effort, most force fields account for the electronic polarizability implicitly. In these force fields the multipole moments are altered compared to their gas phase values to value which represent, sort of, average moments in the liquid phase.

We will focus on the leading term, the induced dipole moment, and the corresponding dipole polarizability. A correction to the heat of vaporization, due to the neglect of polarizabilities, can be derived from the respective dipole polarizabilities, and average induced dipole moments. The correction accounts for the energy needed to distort the electronic cloud. To calculate such a correction, we will assume, within the so-called Drude oscillator model, that the inducible dipole is simply modelled as a pair of charges connected by a harmonic spring.

<b>Task</b>	(3 points)
<ul style="list-style-type: none"><li>• Derive a formula for the dipole polarizability of the Drude oscillator in terms of its charge and force constant.</li><li>• What is the polarization energy in the Drude model in terms of induced dipole moment and polarizability? Include this term in your final expression for the heat of vaporization.</li></ul>	

## 4 Atomistic Water Simulations with GROMACS (10 points)

In this exercise, you will perform several simulations to optimize a given water model. During the original optimization procedure of this model, someone forgot to include the polarization energy correction discussed in sec. 3. You will re-optimize the water model taking this correction into account. The optimization procedure focusses on the partial-charge distribution of the water model. You will run several simulations using different parameter sets and based on the heat of vaporization and other data optimize the initially given water model.

All simulations are performed with GROMACS. GROMACS is a freely available Molecular Dynamics software package available at [www.gromacs.org](http://www.gromacs.org). You can either use the computers in the ICP CIP pool or install GROMACS on your own computer. The simulations focus on improving the SPC model.

Download the archive `templates.tar.gz` from the lecture website. After unpacking, you will find different files and directories in the included directory:

- `gen.sh`: (incomplete) bash script to help setup calculations
- `setup/`: setup files for simulation
  - `conf.gro`: pre-equilibrated water structure with 216 solvent molecules
  - `grompp.mpd`: parameters for the simulation
  - `topol.top`: topology file for the water simulations
  - `index.ndx`: file needed for the analyses
  - `tut.itp`: topology file of unoptimized water model
- `reference_data/`: reference data
  - `rdf_oo.dat`: experimental oxygen-oxygen radial pair-distribution function
  - `pol_cor.dat`: numerical values of polarization energy correction

### 4.1 Visualizing the System

You can visualize the system using `vmd` (Visual Molecular Dynamics):

```
$> vmd conf.gro
```

In the `vmd` console, you can use the following command to show the simulation box:

```
> pbc box
```

## Hint

- By default, all molecules are represented by lines only. To get a nicer looking visualization, use the menu in the “VMD Main” window and click on *Graphics > Representations...*. In the “Graphical Representations” dialog that will pop up, use the *Drawing Method* drop-down menu and select *CPK*.
- You can also save or load a “visualization state” in the drop-down menu under “File”, that will contain information on the trajectory files you loaded and the graphical representations you used to visualize them. This allows you to easily load and visualize data according to your preferences.

## 4.2 Running the Simulations

It is a good idea to run the different simulations in separate directories. Create directories for each partial charge set that you want to investigate. Copy the files from the archive into each of these directories and change them accordingly. The archive also has an incomplete script to help you setup your calculations. Some minor fixes are needed before you can actually use it.

First, we will start a simulation using the unoptimized TUT water model. In order to have the required GROMACS commands available in your terminal, execute the following command:

```
$> source /group/sm/2016/tutorial_02/gromacs-4.6.7/bin/GMXRC
```

To prepare the first simulation, change to a directory you just created (e.g., `tut`) and call the GROMACS preprocessor `grompp`:

```
$> cd tut
$> grompp
```

If `grompp` completed without errors, the simulation can be started with

```
$> mdrun -v
```

The system will then be simulated for 500 ps with a time step of 2 fs at 300 K.

<b>Task</b>	(3 points)
Perform one simulation for each of the water models.	

## Hints

- To tell GROMACS about the different charge sets you have to change the charges in the `tut.itp` file
- Use the commands `grompp` and `mdrun` for each simulation as described above.

### 4.3 Analysis

In the following tasks, you will analyze the simulation data in various ways. Luckily, GROMACS comes with a variety of easy-to-use analysis tools.

#### Radial Distribution Function

The radial distribution functions (RDFs) give a first hint on the local structure of the systems. To compute the RDFs, you can use the command

```
$> g_rdf -n index
```

**Task** (1 point)

- Compute the RDFs of the systems you have simulated.
- Compare the RDFs of the different water models with each other and the given reference and interpret the results (peaks, distances between peaks, differences between water models,...)

#### Hint

- You can use the program `xmgrace` to plot the output `rdf.svg` files.

#### Hydrogen Bond Analysis

A crucial feature of water is the pronounced effect of hydrogen bonds between the oxygen and hydrogen atoms. The occurrence of these bonds is largely responsible for many important properties of water. We can determine the number of hydrogen bonds within the simulated system via

```
$> echo -e "0\n0" | g_hbond
```

**Task** (1 point)

- Run the hydrogen bond analysis for the different water models.
- Calculate the average number of hydrogen bonds per water molecule. What is the meaning of donors and acceptors?

## Mean Square Displacement

In three-dimensional space, the diffusion coefficient  $D$  can be calculated from the mean square displacement (MSD) function

$$\langle \Delta r^2(t) \rangle = 6Dt \quad (1)$$

for large enough  $t$ . You can calculate the MSD using GROMACS with

```
$> g_msd -n index
```

### Task

(2 points)

- Compute the MSD for each of the different water models.
- Have a look at the output files `msd.xvg` for the different water models and compare them. What are the differences? Can you identify the diffusional ( $\langle \Delta r^2(t) \rangle \propto t$ ) and the ballistic regime?
- Determine the fitting range that GROMACS uses and check if it is a sensible choice for your analysis. If not, adapt the options to the `g_msd` command.

## Heat of vaporization

The heat of vaporization  $H_{\text{vap}}$  provides information on the binding strength in the liquid phase compared to the gas phase. It is commonly expressed in units of kJ/mol. The necessary energies can be extracted from the simulations using GROMACS:

```
$> g_energy
```

Further input then needs to be supplied on the console to determine which properties to calculate. For automatization you can also input the latter data input using a command similar to this:

```
$> echo <num1> [num2,...] | g_energy
```

where `num1`, `num2`, ... are integers identifying quantities you want to calculate (careful, they depend on the settings of your simulation).

### Task

(3 points)

- Compute the heat of vaporization for your model systems, including polarization energy correction. Determine the optimal water model from this data (experimental value:  $H_{\text{vap}} = 44$  kJ/mol).
- Compare your choice to what you would have determined as optimal water model from diffusivity and structural properties.
- What other properties have been affected during the parametrization and what physical effects have been neglected in the simulation.