

# Worksheet 2: Diffusion Processes and Properties of Atomistic Water Models

## Solutions

Johannes Zeman

May 13, 2016

Institute for Computational Physics, University of Stuttgart

### Contents

<b>1</b>	<b>Remarks</b>	<b>1</b>
<b>2</b>	<b>Short Questions - Short Answers (5 points)</b>	<b>2</b>
<b>3</b>	<b>Diffusion Processes - Ballistic and Diffusional Regime (7 points)</b>	<b>3</b>
<b>4</b>	<b>Atomistic Water Simulations with GROMACS (8 points)</b>	<b>8</b>
4.1	Visualizing the System . . . . .	8
4.2	Running the Simulations . . . . .	8
4.3	Analysis . . . . .	10
	<b>Bibliography</b>	<b>14</b>

### 1 Remarks

The Solutions provided here show possible approaches to solve the questions from the corresponding worksheet and may not be exhaustive.

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

<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 difference between the SPC and the SPC/E water model?</li></ul>	



---

### Solution

- Water is a molecule consisting of two hydrogen atoms covalently bound to an oxygen atom. Since oxygen has a higher electronegativity, the probability of presence for electrons involved in the covalent bonds is shifted towards the oxygen atom. This leads to a negative partial charge at the oxygen and positive partial charges at the hydrogen atoms, effectively making the molecule polar.
- A hydrogen bond is an intra- or intermolecular electrostatic attractive interaction between polar groups of molecules and/or polar molecules. It can occur whenever hydrogen atoms are bound to highly electronegative elements such as nitrogen, oxygen, or fluorine. Hydrogen bonds are not to be confused with covalent bonds or van-der-Waals interactions, they are in fact relatively strong dipole-dipole attractions. Their interaction energy lies between those of van-der-Waals interactions (lower) and covalent bonds (higher).
- H-O-H angle in a water molecule:  $\approx 104.5^\circ$   
Hydrogen bond distance:  $1.2 - 3.0 \text{ \AA}$  (according to donor-acceptor distances categorized in [1] with an assumed donor-hydrogen bond length of  $1 \text{ \AA}$ )
- The main differences lie in the number of interaction sites, (non-)flexibility, and the incorporation of polarization effects.
- In contrast to SPC, the SPC/E model incorporates a correction term in the potential energy taking average polarization effects into account.

### 3 Diffusion Processes - Ballistic and Diffusional Regime (7 points)

The Langevin equation is an effective description of particle diffusion in an *implicit* solvent. The interaction of particles with the surrounding solvent is replaced by their coupling to a virtual heat bath in the form of a dissipative and a stochastic force. The corresponding one-dimensional differential equation is given by

$$m \frac{dv(t)}{dt} = -\gamma v + \eta \quad (1)$$

with  $\gamma = m\zeta$  and  $\eta dt = \Gamma dW(t)$ . The first term on the right hand side of the Langevin equation describes the dissipative force whereas the second term incorporates the stochastic force in terms of Gaussian white noise generated by a Wiener process.

**Task** (3 points)  
Solve the Langevin equation to obtain  $\langle v^2(t) \rangle$ . Have a look at the hints below!

#### Hints

- First, obtain a general solution  $v(t)$  of equation (1). The general solution of an equation of the form  $\frac{dy(x)}{dx} + p(x)y(x) = q(x)$  is given by  $y(x) = \frac{1}{u(x)} \left( \int_0^x u(x')q(x')dx' + C \right)$  with  $u(x) = \exp\left(\int p(x)dx\right)$ .
- Subsequently, obtain a particular solution: Determine the value of  $C$  by requiring  $v(0) = v_0$  as an initial condition.
- The first and second moments of a Wiener process are given by  $\langle dW(t) \rangle = 0$  and  $\langle dW(t')dW(t'') \rangle = dt'$ . This means that an integral  $\int \langle dW(t) \rangle$  will vanish and a double integral of the form  $\int_0^t \int_0^t f(t')f(t'')\langle dW(t')dW(t'') \rangle$  will simplify to  $\int_0^t f^2(t')dt'$ . Use this to obtain an expression for  $\langle v^2(t) \rangle = \langle v(t)v(t) \rangle$ .

---

#### Solution

Langevin equation:

$$\frac{dv(t)}{dt} + \frac{\gamma}{m}v = \frac{1}{m}\eta$$

The general solution of a non-homogeneous first-order ODE of the form

$$\frac{dy(x)}{dx} + p(x)y(x) = q(x)$$

is given by

$$y(x) = \frac{1}{u(x)} \left( \int_0^x u(x')q(x')dx' + C \right) \text{ with } u(x) = \exp\left(\int p(x)dx\right) .$$

Here:

$$x = t, \quad y = v, \quad p = \frac{\gamma}{m} = \zeta, \quad q = \frac{\eta}{m}, \quad \text{and} \quad u = e^{\zeta t}$$

and thus the general solution of equation (1) is

$$v(t) = e^{-\zeta t} \frac{1}{m} \int_0^t e^{\zeta t'} \eta dt' + C e^{-\zeta t}.$$

The condition  $v(0) = v_0$  yields  $C = v_0$  and the particular solution is therefore

$$v(t) = v_0 e^{-\zeta t} + e^{-\zeta t} \frac{1}{m} \int_0^t e^{\zeta t'} \eta dt'.$$

To obtain  $\langle v^2(t) \rangle$ , we need to compute the expected value of  $v(t)v(t')$  applying  $\eta dt = \Gamma dW(t)$ :

$$\begin{aligned} \langle v^2(t) \rangle &= \langle v_0^2 \rangle e^{-2\zeta t} + \frac{2\Gamma \langle v_0 \rangle}{m} e^{-2\zeta t} \int_0^t e^{\zeta t'} \langle dW(t') \rangle \\ &+ e^{-2\zeta t} \frac{\Gamma^2}{m^2} \int_0^t \int_0^t e^{\zeta t'} e^{\zeta t''} \langle dW(t') dW(t'') \rangle \end{aligned}$$

With  $\langle dW(t') \rangle = 0$  the second addend on the right hand side vanishes and with  $\langle dW(t') dW(t'') \rangle = dt'$  the integral in the third addend simplifies so that we obtain

$$\begin{aligned} \langle v^2(t) \rangle &= \langle v_0^2 \rangle e^{-2\zeta t} + e^{-2\zeta t} \frac{\Gamma^2}{m^2} \int_0^t e^{2\zeta t'} dt' \\ &= \langle v_0^2 \rangle e^{-2\zeta t} + \frac{\Gamma^2}{2\zeta m^2} (1 - e^{-2\zeta t}), \end{aligned}$$

which is the solution we've been looking for.

---

In a canonical ensemble, particle velocities are Maxwell-Boltzmann distributed:

$$P(v) = 4\pi \left( \frac{m}{2\pi k_B T} \right)^{\frac{3}{2}} v^2 e^{-\frac{m}{2k_B T} v^2} \quad (2)$$

Thus, it follows for the expectation value of the kinetic energy per particle that

$$\langle E_{\text{kin}} \rangle = \frac{1}{2} m \langle v^2 \rangle = \int_0^\infty \frac{1}{2} m v^2 P(v) dv = \frac{1}{2} k_B T, \quad (3)$$

which is nothing but the equipartition theorem for one dimension.

**Task** (1 point)  
 Give an expression for  $\Gamma$  as a function of temperature in the limit of long time scales ( $t \rightarrow \infty$ ).

### Hints

- Use equation (3) together with the previously calculated expectation value  $\langle v^2(t) \rangle$ . Think about which terms vanish for  $t \rightarrow \infty$ !

### Solution

Previously calculated solution for  $\langle v^2(t) \rangle$ :

$$\langle v^2(t) \rangle = \langle v_0^2 \rangle e^{-2\zeta t} + \frac{\Gamma^2}{2\zeta m^2} (1 - e^{-2\zeta t})$$

At long time scales,  $e^{-2\zeta t}$  vanishes so that

$$\lim_{t \rightarrow \infty} (\langle v^2(t) \rangle) = \frac{\Gamma^2}{2\zeta m^2}.$$

Together with the equipartition theorem in one dimension ( $\langle v^2 \rangle = \frac{k_B T}{m}$ ) and  $\gamma = \zeta m$  this yields

$$\Gamma = \pm \sqrt{2\gamma k_B T}$$

Solving the Langevin equation (1) for the displacement  $x(t) = \frac{dv(t)}{dt}$  assuming  $x(0) = x_0 = 0$  allows to obtain the following expression for the mean square displacement:

$$\langle x^2(t) \rangle = \frac{\langle v_0^2 \rangle}{\zeta^2} (1 - e^{-\zeta t})^2 + \frac{\Gamma^2}{m^2 \zeta^2} t - \frac{\Gamma^2}{2m^2 \zeta^3} (3 - 4e^{-\zeta t} + e^{-2\zeta t}) \quad (4)$$

For very short time scales, the displacement  $x(t)$  can be approximated by the free particle term

$$\lim_{t \rightarrow 0} (x(t)) = v_0 t \quad (5)$$

**Task**

(2 points)

- Use equation (5) together with the result of the equipartition theorem (3) to obtain an expression of  $\langle x^2(t) \rangle$  for  $t \rightarrow 0$  (ballistic regime).
- Use equation (4) in the limit of long time scales ( $t \rightarrow \infty$ ) together with your result for  $\Gamma$  from the previous task to obtain an expression of  $\langle x^2(t) \rangle$  for  $t \rightarrow \infty$  (diffusional regime).
- How does  $\langle x^2(t) \rangle$  scale with  $t$  in the respective cases?

**Solution**

- Ballistic regime:

$$\lim_{t \rightarrow 0} (\langle x^2(t) \rangle) = \langle v_0^2 \rangle t^2 = \frac{k_B T}{m} t^2$$

- Diffusional regime:

$$\lim_{t \rightarrow \infty} (\langle x^2(t) \rangle) = \frac{\Gamma^2}{\gamma^2} t = \frac{2k_B T}{\gamma} t$$

- According to the results above, the MSD scales quadratically with  $t$  in the ballistic regime and linearly in the diffusional regime.

The diffusion coefficient  $D$  is connected to the mean square displacement  $\langle x^2(t) \rangle$  through the relation

$$\langle x^2(t) \rangle = 2n_dDt \quad (6)$$

with  $n_d$  being the number of dimensions (here  $n_d = 1$ ).

<b>Task</b>	(1 point)
<ul style="list-style-type: none"><li>• Provide an expression for <math>D</math> using the result for <math>\langle x^2(t) \rangle</math> at long time scales from the previous task.</li><li>• Compare your solution with the Stokes-Einstein equation. How is the drag coefficient <math>\gamma</math> related to solvent viscosity?</li></ul>	

---

### Solution

- Diffusion constant from diffusional regime of MSD:

$$D = \frac{k_B T}{\gamma}$$

- The Stokes-Einstein equation reads

$$D = \frac{k_B T}{2n_d \pi \eta a}$$

with  $n_d$  being the number of dimensions,  $\eta$  the viscosity of the surrounding medium and  $a$  the hydrodynamic radius of the particle of interest.

The drag coefficient  $\gamma = 2\pi n_d \eta a$  is thus directly proportional to solvent viscosity.

---

## 4 Atomistic Water Simulations with GROMACS (8 points)

In this exercise, you will perform several simulations with different water models using GROMACS to learn about their properties. GROMACS is a freely available Molecular Dynamics software package ([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 the SPC, SPC/E and TIP3P water models.

### 4.1 Visualizing the System

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

- `conf.gro`: pre-equilibrated water structure with 216 solvent molecules
- `grompp.mpd`: parameters for the simulation
- `topol.top`: topology file for the water simulations including a link to the force field parameters
- `index.ndx`: file needed for the analyses

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*.

### 4.2 Running the Simulations

For the different simulations, it is a good idea to run them in separate directories. Create three new directories named `spc`, `spce`, and `tip3p`. Copy the files from the archive into each of these directories.

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

```
$> source /group/sm/2015/gromacs-4.6.5/bin/GMXRC
```

In order to prepare the first simulation, change to the `spc` directory and call the GROMACS preprocessor `grompp`:

```
$> cd spc
$> 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.

**Task**

(3 points)

Perform one simulation for each of the water models.

**Hints**

- To tell GROMACS which water model to use, open the `topol.top` files in the different directories and change the line  
`#include "spc.itp"`  
to  
`#include "spce.itp"`  
or  
`#include "tip3p.itp"`  
respectively.
- Use the commands `grompp` and `mdrun` for each simulation as described above.

**Solution**

The simulations can be run with

```
$> mkdir spc spce tip3p
$> cp conf.gro index.ndx grompp.mdp topol.top spc
$> cp conf.gro index.ndx grompp.mdp topol.top spce
$> mv conf.gro index.ndx grompp.mdp topol.top tip3p
$> sed -i s/spc.itp/spce.itp/g spce/topol.top
$> sed -i s/spc.itp/tip3p.itp/g spce/topol.top
$> source /group/sm/2015/gromacs-4.6.5/bin/GMXRC
$> cd spc && grompp && mdrun
$> cd ../spce && grompp && mdrun
$> cd ../tip3p && grompp && mdrun
```

### 4.3 Analysis

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

#### 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

(2 points)

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

#### Hint

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

#### Solution

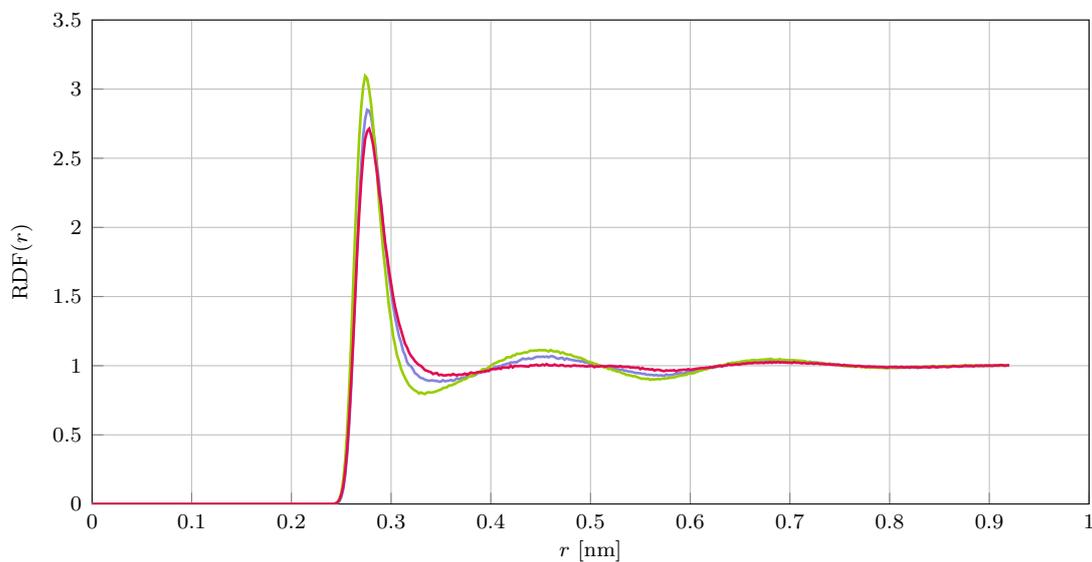


Figure 1: Radial distribution functions of SPC (—), SPC/E (—) and TP3P (—) water at 300 K

All RDFs show a first peak at roughly 0.28 nm. The SPC and SPC/E models exhibit a second peak at around 0.45 nm, while the TIP3P model barely shows any peaks apart from the first one. A third peak at 0.68 nm is barely visible for all models. The first peaks differ in magnitude, with TIP3P showing the smallest, and SPC/E the largest value. The lower first peak and the almost vanishing second and third peaks in the TIP3P model suggest that the angle of  $104.52^\circ$  leads to a less structured liquid compared to the other models using a tetragonal angle of  $109.47^\circ$ . The higher peaks and the slight shift of the SPC/E model's RDF to smaller distances likely occur due to its additional polarization energy term leading to a more pronounced structure.

---

### 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 mainly reliable 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?

---

### Solution

The following table shows the average number of hydrogen bonds and the average number of hydrogen bonds per molecule in water obtained from simulations of pure water using different water models:

Model	H-bonds	Molecules	H-bonds/molecule
TIP3P	361.517	216	3.35
SPC	374.177	216	3.46
SPC/E	388.834	216	3.60

**Table 1:** Average number of hydrogen bonds in water simulated with the TIP3P, SPC, and SPC/E model.

Comparison of table 1 with figure 1 suggests that a more structured liquid leads to a higher number of hydrogen bonds in the system. Since hydrogen bonds are shared between two molecules, the average number of hydrogen bonds per molecule has been computed as

$$\text{\#H-bonds/molecule} = 2 \frac{\text{\#H-bonds}}{\text{\#molecules}} \quad (7)$$

---

### Mean Square Displacement

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

$$\langle \Delta r^2(t) \rangle = 6Dt \quad (8)$$

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

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

#### Task

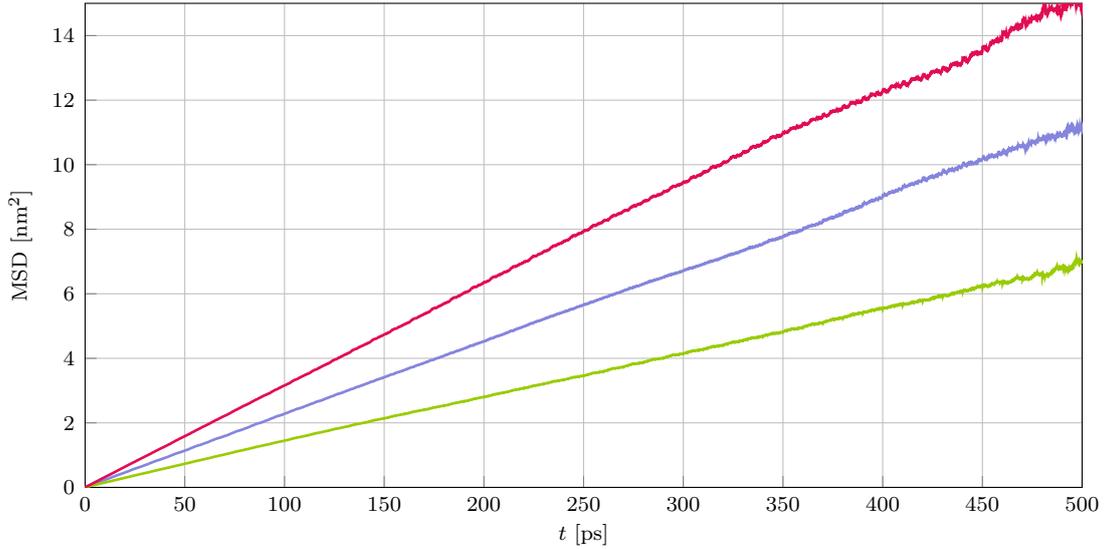
(2 points)

- 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?
- Does the form of the ballistic regime look as you would expect it from the second last task in section 3? If not, try to explain why!

---

### Solution

The following graph shows the MSDs of the oxygen atoms in the system for different water models:



**Figure 2:** Diffusional regime of the mean square displacements of the SPC (—), SPC/E (—) and TP3P (—) water’s oxygen atoms at 300 K. Note the linear behavior.

As expected for the diffusional regime, the MSDs show a linear behavior for long simulation times. The increasing noise towards the right hand side of the graph is due to the increasing statistical inaccuracy of the autocorrelations from which the MSDs are computed.

The slopes of the curves correspond to the respective self-diffusion coefficients  $D$  of the different water models which are listed in the following table:

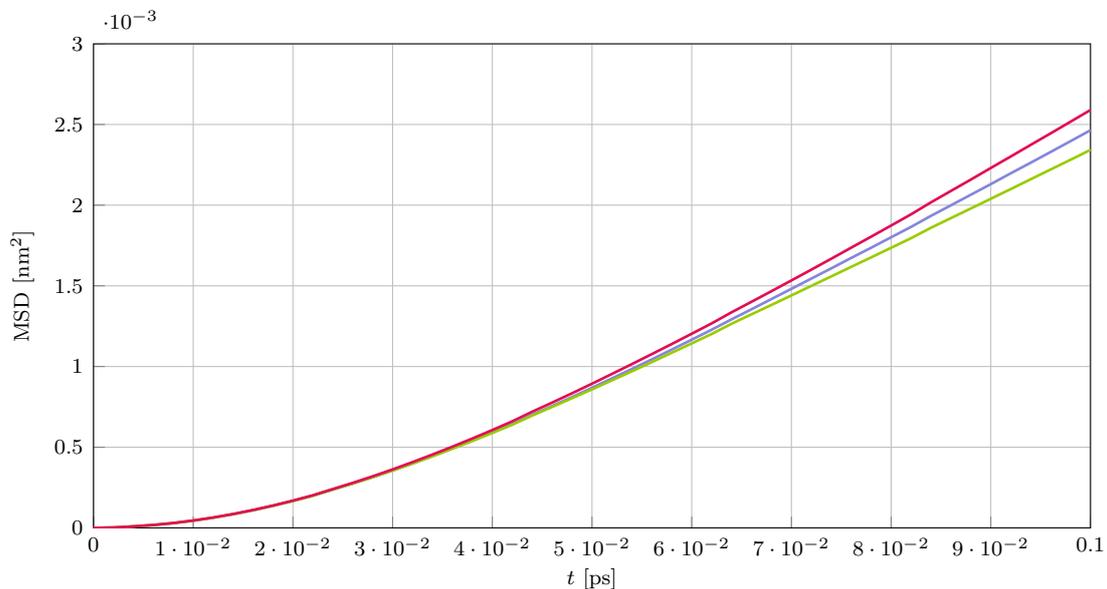
Model	$D$ [ $10^{-5}$ cm <sup>2</sup> /s]
TIP3P	$5.0503 \pm 0.6689$
SPC	$3.7258 \pm 0.0401$
SPC/E	$2.2690 \pm 0.0250$

**Table 2:** Self-diffusion coefficients of water simulated with the TIP3P, SPC, and SPC/E models.

The real self-diffusion coefficient of water is approximately  $2.4 \cdot 10^{-5}$  cm<sup>2</sup>/s at 300 K (interpolated from [2]), so the SPC/E model reproduces this property best.

The ballistic regime cannot be identified in figure 2. In fact, even zooming in does not really help here, since system trajectories were stored only every 50 time steps, corresponding to storage intervals of 0.1 ps. In order to be able to see the ballistic

regime, the parameter `nstxtcout` in the file `grompp.mdp` has to be set to a value of 1, causing system trajectories to be stored on a per-time-step basis. The resulting data is shown in the graph below.



**Figure 3:** Ballistic regime of the MSDs of the SPC (—), SPC/E (—) and TP3P (—) water's oxygen atoms at 300 K. Note the quadratic behavior.

As expected, the graphs exhibit a quadratic behavior.

---

## Bibliography

- [1] G. A. Jeffrey. *An introduction to hydrogen bonding*, volume 12. Oxford university press New York, 1997.
- [2] Manfred Holz, Stefan R. Heil, and Antonio Sacco. Temperature-dependent self-diffusion coefficients of water and six selected molecular liquids for calibration in accurate 1H NMR PFG measurements. *Phys. Chem. Chem. Phys.*, 2:4740–4742, 2000.