Assignment 1

- **All questions are of equal value.**
- **Important:** This assignment should be returned using GitHub. If you don't have an account, create a GitHub account, and in GitHub create a directory called Chem9664B. Make that directory private and share it with me. My username at GitHub is mejk (you need that for sharing). This assignment also assumes that you have now access to the command line terminal (Linux, WSL in Windows or Mac terminal), and have python and Jupyter Lab installed. Installation help is in the online course notes.
- **All of the tasks below should be in one single Jupyter notebook (ipynb file). Include the .gro file below in the directory with the notebook. Use comments/markdown!**

Due: Friday January 27, 2023 at midnight.

The data in the file *five_spc_waters.gro* has 5 water molecules in the Gromacs *.gro* format is fixed meaning that the positions of the columns are fixed. A gro file (as detailed in the Gromacs manual) contains the following information:

Line 1: Title. Afree format string. Optional time in ps after `'t='`
Line 2: The number of atoms
Line 3 and onwards until the second to last line: Information for each atom. Each line has

- residue number (5 positions, integer)
- residue name (5 characters)
- atom name (5 characters)
- atom number (5 positions, integer)
- position (in nm, x y z in 3 columns, each 8 positions with 3 decimal places)

In addition velocity can be included, but that is optional (and we don't have it in this file):

- velocity (in nm/ps (or km/s), x y z in 3 columns, each 8 positions with 4 decimal places)

This format is fixed, ie. all columns are in a fixed position.

The last line: has the box vectors, that is, the size of the simulation box.

Note that in the SPC model, each of the water molecules has an oxygen (OW) and two hydrogens (HW1 and HW2). *The residue number* shows the numbering of the water molecules and *atom number* is the numbering for each individual atom.

1. Write a small python program (in the Jupyter Lab environment) that reads the file from your

computer and as an output summarizes: a) the number of atoms, b) the size of the simulation box, c) the volume of the simulation box and c) the different atoms types. Remember that the output should be informative and, for example, tell what the types are.

2.Write a short python program that prints out the maximum, minimum and average values of the *x*-, *y*- and *z*-coordinates. Remember to the units!

3. Write a small python program that plots the position of each of the *atoms* in the *x*-*y* plane. Use a different colour for the oxygens and hydrogens. Use the box vectors to to draw a box (limit the coordinates) to the values given by them. Hint: look at the coordinate values and thing how the box vectors should be used. Remember the units on the *x* and *y* axes!

4. The *.gro* format is specific to Gromacs. One of the common (ASCII) format is the so-called .xyz format (names are give in the form *file_name.xyz*). The format is specified as follows:

> **First line:** total number of atoms (optional, but should be used as it is informative)
>
> **Second line:** molecule name or comment (optional, but should be used as it is informative)
>
> **Line 3 and onwards:** element symbol or atomic number, *x*, *y*, and *z*-coordinates, separated by spaces. It is a good idea to use the same number of decimals.
>
> **Note:** The gromacs *.gro* format has the box dimensions, but the xyz format doesn't contain it.

Write a python code that uses the information from the .gro file and produces (=writes) the corresponding *.xyz* file.

Important: When we write our own molecular dynamics program, we will output the data in the *xyz* format.