

# Fluid simulations using the Lennard Jones Potential.

## Intro

Molecular dynamics (MD) is a method of simulating systems of atoms or molecules moving and interacting with one another. These simulations can show how a system changes with time, not only in terms of the positions and velocities of the particles but also resulting bulk properties of the system such as pressure, temperature, potential energy, and density. Certain state variables can also be fixed within a tolerance. This enables researchers to explore the state of a system at very small spatial and temporal scales with comparable ease in terms of time and cost compared to practical methods. The downsides of MD are that approximations must be made for the simulations to be feasible; time taken to run a simulation is one consideration as is the fact that the Schrodinger equation can't be solved for multi-electron systems.

A specific example of the use of molecular dynamics is drug discovery. Many drugs function as competitive inhibitors to proteins. Exploring the exact shape of an active site in a protein and how this can interact with different possible drugs is a very powerful tool for developing a drug with the correct affinity for the application. The exact methodology does rely on some practical techniques. The structure of the protein can be determined using x-ray crystallography, however the structure of proteins in a crystal tends to differ from their structure in solution. A MD simulation can be carried out using the crystal structure of the protein in a box surrounded by water as a starting point. This structure can then be allowed to equilibrate to give a better idea of the structure of the protein and its active site in solution. One approximation made in these sorts of studies is to model most of the bonds in the protein using a simple model such as the Lennard Jones potential. The active site of the protein can then be modelled using quantum mechanics as this tends to be the focus of studies.

Another use of molecular dynamics is to simulate drug delivery. Drug delivery tends to be optimised by trials where different doses are used on test subjects to optimise the drug delivery. Here simulations can be used to explore the transport rate of drugs across tissues. This is useful for optimising drug delivery to give the correct dose in the correct location depending on the properties of the drug and the size of the person and could help to make better guesses on the correct doses during trials which would help cut out the need for test patients and improve certainty of the right dose in practice.

## Thermodynamic ensemble

## Verlet Algorithm

The simulations used below make use of the Lennard Jones Potential to simulate the potential and forces experienced by the atoms and the Verlet Algorithm to determine the position of the particles at the next time step.

The position  $X$  of an atom at time  $t$  can be represented as:  $X(t)$

Taylor expansions of  $x(t + \delta t)$  and  $x(t - \delta t)$  up to the fourth order  $O(\delta t^4)$  can be used to derive the Verlet Algorithm like so:

Taylor expansion of  $x(t + \delta t)$  about  $t$ :

$$x(t + \delta t) \approx x(t) + \frac{\dot{x}(t)}{1!}(\delta t) + \frac{\ddot{x}(t)}{2!}(\delta t)^2 + \frac{x'''(t)}{3!}(\delta t)^3$$

Taylor expansion of  $x(t - \delta t)$  about  $t$ :

$$x(t - \delta t) \approx x(t) - \frac{\dot{x}(t)}{1!}(\delta t) + \frac{\ddot{x}(t)}{2!}(\delta t)^2 - \frac{x'''(t)}{3!}(\delta t)^3$$

Adding the two expansions together and cancelling out the terms gives:

$$x(t - \delta t) + x(t + \delta t) \approx 2x(t) + \ddot{x}(t)(\delta t)^2$$

Substituting acceleration  $\ddot{x}(t)$  with  $\frac{F(t)}{m}$  (this comes from  $F = ma$ ) and rearranging gives:

$$x(t + \delta t) = 2x(t) - x(t - \delta t) + \frac{F(t)}{m}(\delta t)^2$$

Above is the classical Verlet algorithm. There is another algorithm known as the Velocity-Verlet Algorithm which uses particle velocities and forces as opposed to positions and forces as above. An advantage of the Velocity-Verlet Algorithm would be higher accuracy because force has a more direct effect on velocity (single integral) as opposed to a double integral to find the change in position due to a force, thus using velocity would reduce the uncertainty for each time step; the force is constant between adjacent timesteps. Further, using velocity in each time step would allow for more direct calculation of the state of the system because the velocity of each particle pertains to the kinetic energy in the system and the temperature. Using positions would require an extra calculation step to find the average velocity between two timesteps for each particle.

## Lennard Jones Potential

The Lennard Jones potential is the equation used to calculate the  $F$  term in the Verlet Algorithm above.

$$U(r) = 4\epsilon \left( \frac{\sigma^{12}}{r^{12}} - \frac{\sigma^6}{r^6} \right)$$

### *Lennard Jones potential*

The Lennard Jones potential describes the potential energy of interaction of two particles moved together from infinity, hence the potential at infinity is zero. Taking the differential of the potential with respect to space also gives a force. The  $\frac{1}{r^6}$  term describes the attractive term due electron – nuclear attraction and dominates at large distances  $\frac{1}{r^6} > \frac{1}{r^{12}}, r > 1$ . The  $\frac{1}{r^{12}}$  term is repulsive and becomes very large at small distances due to nuclear repulsion.

The potential energy  $U(r)$  is zero when  $r = \sigma$  hence  $U(\sigma) = 0$ .

$$\frac{dU(r)}{dr} = F(r) = 4\epsilon \left( \frac{12\sigma^{12}}{r^{13}} - \frac{6\sigma^6}{r^7} \right)$$

Substituting  $\sigma$  into  $F(r)$  gives  $F(\sigma) = \frac{24\epsilon}{\sigma}$ . There is a very large repulsive force at zero potential, which makes sense because there was an attractive force all the way from infinity which would have resulted in a negative potential beforehand.

Setting  $F(r) = 0$  and solving for  $r$  gives  $F\left(2^{\frac{1}{6}}\sigma\right) = 0$ . One could say the equilibrium distance is at  $r = 2^{\frac{1}{6}}\sigma$ .

Substituting the equilibrium distance into  $U(r)$  and simplifying gives a well depth of  $\epsilon$ .  $U\left(2^{\frac{1}{6}}\sigma\right) = -\epsilon$ .

Evaluating the integrals:

$$\int_{2\sigma}^{\infty} U(r) dr = -0.6035$$

$$\int_{2.5\sigma}^{\infty} U(r) dr = -0.2030$$

$$\int_{3\sigma}^{\infty} U(r) dr = -0.0821$$

This shows that after a certain point even with the particles quite close, the potential becomes negligible hence most simulations will have a cut off around  $3\sigma$  because considering the interaction of further particles increases computation time with little change to the result.

Along with a cut off, simulations are run in a box where if a particle exits the box, it enters the box at the opposite / inverse of the position it left. If a particle is at position (0.5, 0.5, 0.5) in a box from (0, 0, 0) to (1, 1, 1) and moves along the vector (0.7, 0.6, 0.2) the particle will end up in the position (0.2, 0.1, 0.7).

### Reduced units:

Real units are scaled to reduced units which tend to have values around the order of magnitude  $10^0$  which makes inputting data, collection, and analysis easier.

Some examples of reduced units:

Distance  $r^* = \frac{r}{\sigma}$  'distance in units of  $\sigma$ ', energy  $E^* = \frac{E}{\epsilon}$ , temperature  $T^* = \frac{k_B T}{\epsilon}$ . Using these given units, other useful units in reduced form can be derived using unit analysis:

$$\epsilon = \frac{1}{2} m^* v^2, v^2 = \frac{\sigma^2}{t^2}, \text{ hence } t^* \text{ has units } \sqrt{\frac{m^*}{\epsilon}} \times \sigma.$$

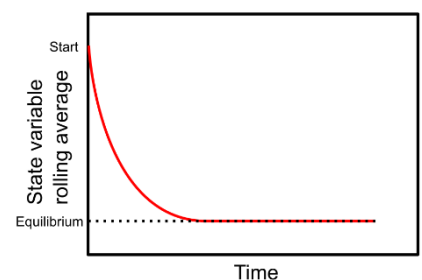
$$p = \frac{F}{A}, F^* \text{ has units } \frac{\epsilon}{\sigma} \text{ and } A \text{ has units } \sigma^2 \text{ hence pressure has units } \frac{\epsilon}{\sigma^3}.$$

$$\text{Given } \sigma = 0.34 \text{ nm}, \frac{\epsilon}{k_B} = 120 \text{ K}, r^* = 3.2, T^* = 1.5, r = 1.088 \text{ nm}, T = 120 T^* = 180 \text{ K}$$

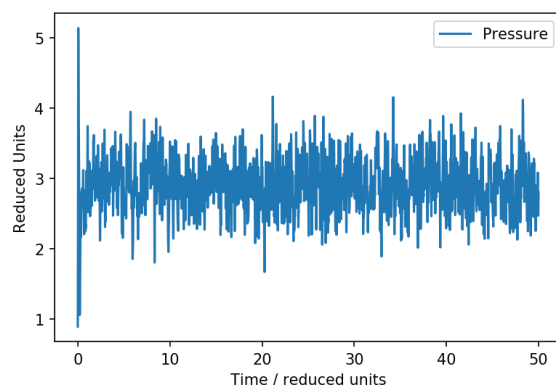
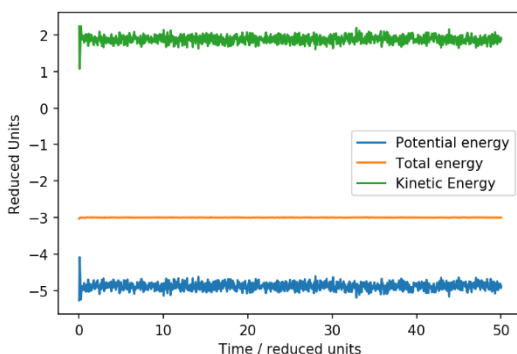
$$k_b \times 120 = 1.38 \times 10^{-23} \times 120 = 1.656 \times 10^{-21} \text{ J} = 0.01 \text{ eV}$$

### Equilibration.

Reaching equilibrium or Steady State for a molecular dynamics simulation is when the simulation starts from a point away from equilibrium and converges on a constant rolling average for state variables like pressure and potential energy. The system does not start at equilibrium because it is arranged like a crystal which is unrealistic for a fluid thus the state variables will converge on a steady state. It is important to let the system reach equilibrium before sampling so that your results are indicative of a fluid system as opposed to the starting positions which tends to be a crystal.

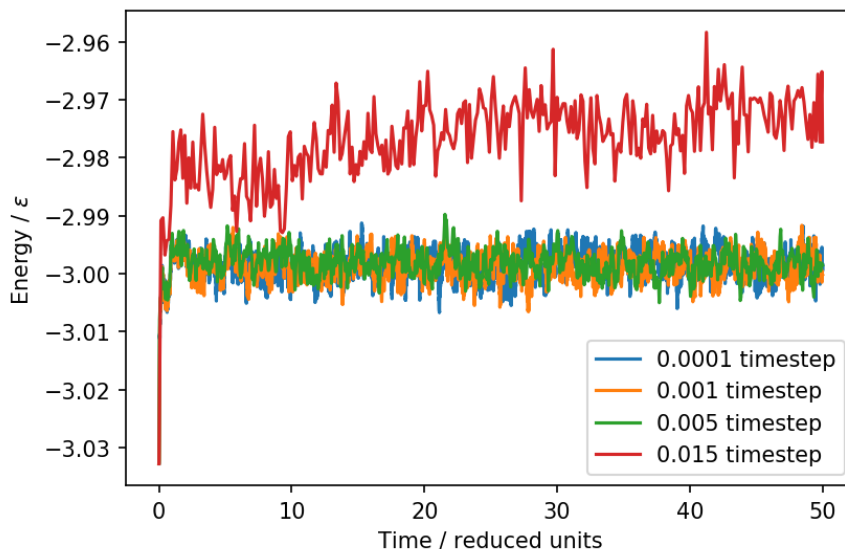
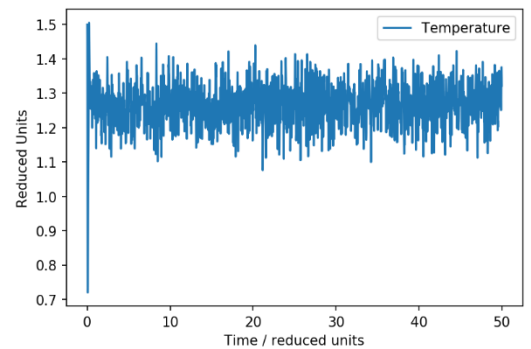


### Equilibration plots:



These plots reach equilibrium because you can clearly see them very quickly converge on a single value represented by a horizontal line of best fit. Interestingly, they appear to oscillate about the equilibrium position so determining when the system has equilibrated would be better done by considering the variance of points from the rolling average.

## Testing timesteps

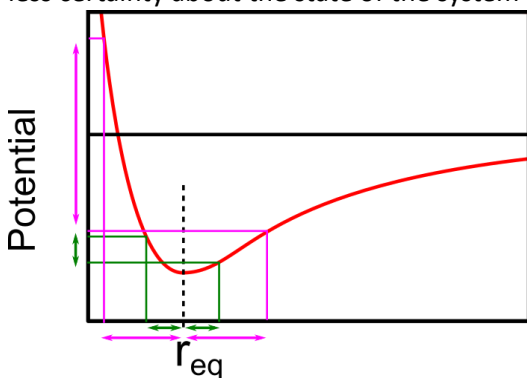


Here there are 2 factors to consider: Quality of the produced data and the time taken to run.

In terms of data quality, more data points are always better because it allows you to make more accurate conclusions due to the higher certainty afforded by the agreement of more data points - better average, smaller standard error.

Between the timesteps, the smallest timestep does not give a significant benefit over the 0.001 timestep as you can see from the graph, they have a similar grouping of points and map onto one another quite well.

The 0.005 timestep is visibly a downgrade simply because the larger timestep means fewer data points and therefore less certainty about the state of the system at a given point.



The highest timestep 0.015 appears to have caused the system to have a higher average total energy. I believe this is because the larger timestep also means the atoms will make larger movements from one step to another and in this case, they can jump closer together than higher timesteps without experiencing a large repulsive force and therefore on average the system has closer atoms experiencing larger repulsion and larger potential energy.

The reason this increases the total energy instead of reducing the total energy is because the Lennard - Jones potential is asymmetric with the left (repulsive) side of the well having a steeper gradient:

This exaggerated plot (figure 3) shows that a larger jump from the equilibrium due to a larger timestep (assuming the same velocity in both cases) results in a larger gap between the repulsive and attractive potentials and therefore an increase in the system's potential energy at larger distances from equilibrium.

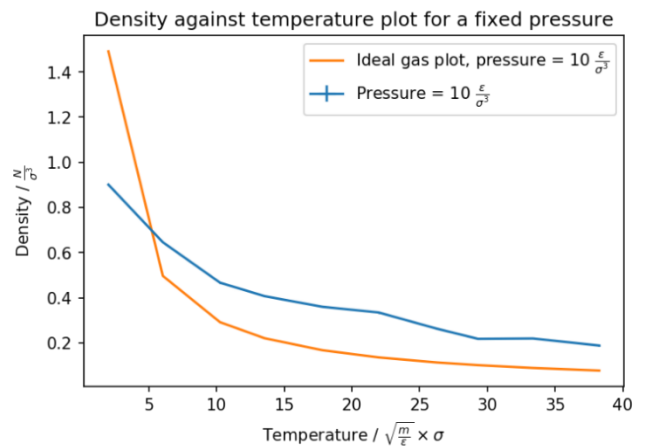
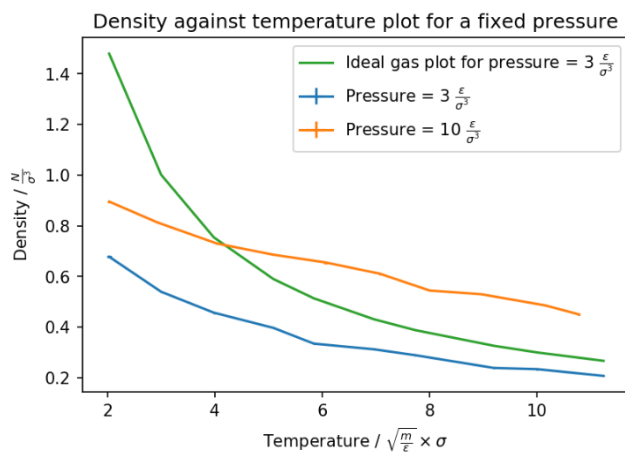
In terms of times taken:

Timestep	Runtime / s
0.0001	77

0.001	7.9
0.005	1.7
0.01	0.84

One can see that runtime scales linearly with the size of the timestep. As discussed above, 0.0001 was not significantly better than 0.001 and 0.005 was noticeably worse than 0.001 therefore the timestep I will be using moving forwards is 0.001 because 8 seconds isn't too long to run (77s is a bit too much) and gives good quality results that have similar grouping and follow the smaller equilibration plateau as the smaller timestep.

## Simulating specific conditions:



Deviation from the ideal gas law increases with pressure and decreases with temperature.

The increase in deviation with pressure is due to this fact: The ideal gas law assumes no potential energy / no interaction of the particles in the gas; therefore this assumes that the particles very close together with no repulsion. The Lennard Jones potential is well approximated by the ideal gas law at large distances because the Lennard Jones potential predicts a very small amount of interaction at large distances, tending towards zero. At higher temperatures particles are spread out more and therefore interact to a lesser extent hence the good approximation. However, at small distances the particles experience significant changes in potential and repel therefore they don't get as close together at low temperatures as the ideal gas law predicts.

Increasing pressure forces the particles closer together. The closer the particles the greater their interaction / greater forces experienced. Larger forces correspond to larger deviations from the ideal gas law which assumes zero potential / zero interaction between particles in the gas. Further, the ideal gas law assumes that particles in a fluid can get infinitely close due to assuming no interaction hence much higher densities are predicted by the ideal gas law at high temperatures compared with the simulations. Lower pressures correspond with lower densities where particles are further apart and therefore interact to a lesser extent therefore making the ideal gas law a better approximation at lower pressures.

I expect my simulation to be in better agreement with the Van der Waals equation because this equation contains a correction factor  $ap^2$ ,  $p = \left(\frac{n}{V}\right)^2$  which considers intermolecular interactions and becomes large for high densities where the particles are close together like in the above simulations. This term contains the density  $p$ , which will have to get smaller to account for higher temperatures and so this would fit the graph above where the density is lower than predicted by the ideal gas equation.

$$p = a\left(\frac{n}{V}\right)^2(V - nb) = nRT$$

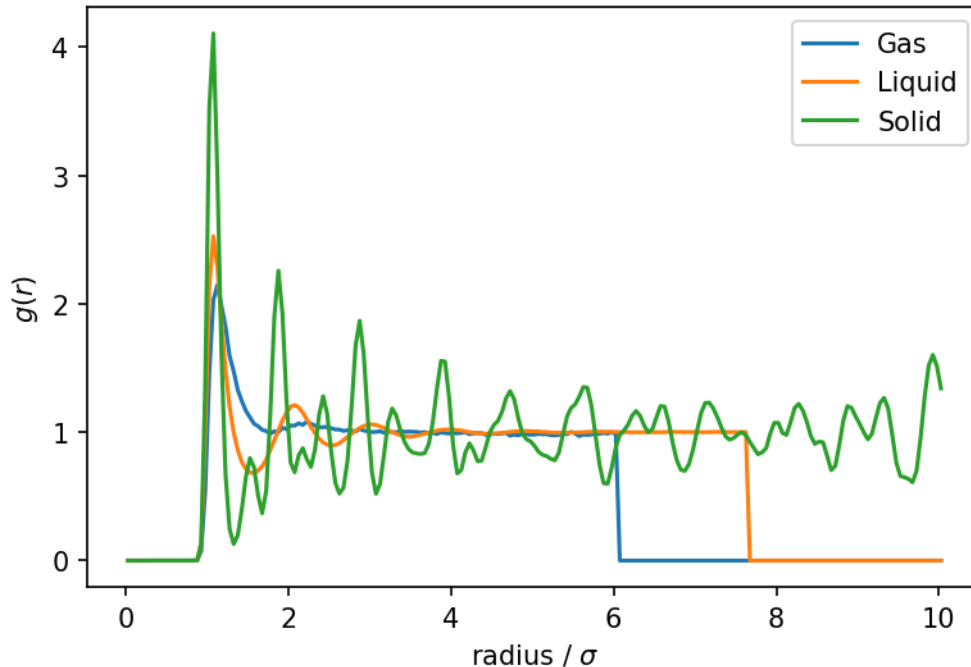
*Van der Waals equation.*

The correction factor  $nb$  accounts for the finite size of molecules. This term would be very small for the above single-atom simulations where the interatomic interactions account for the size of the atoms. This term is useful for

fluids made of larger many atom molecules where significant intermolecular interactions can be at distances smaller than the size of the molecule itself, which occupies a volume, subtracting from the total volume of the system.

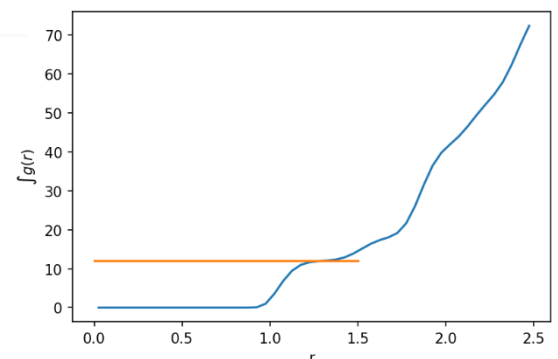
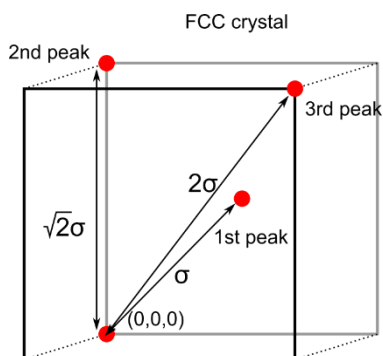
## RDF

The RDF is the probability of finding a particle at distance  $r$  from another target particle<sup>1</sup>. The first peak in the RDF corresponds to the coordination number of your target atom. A larger value of the RDF corresponds to a larger coordination number.



The RDF for the solid has sharp well-defined peaks with a long range. This is because atoms oscillate about fixed lattice points in a solid. Hence at long distances there are still well-defined spots where there will be an atom. This oscillation causes the broadening of the peaks at higher temperatures.

For the liquid there is a sharp peak and a few broad peaks at a short range, but the RDF becomes flat / structureless at long range. This is because the atoms in liquids exhibit bonding but have enough kinetic energy that they are not



fixed to specific lattice points. This means that there is an equal probability of finding particles at every point in space far away from your target particle aka the 'bulk'. The first sharp peak is due to the first coordination sphere. There is a lower chance of finding particles than in the bulk just outside the first coordination sphere due to repulsion of atoms from the first coordination sphere. The first peak is lower than that of the solid because liquids tend to be less dense and more spread out. Gasses are not structured, and the particles are highly mobile therefore you only see one peak for the initial coordination sphere with the RDF quickly becoming flat as there is equal probability of finding a particle everywhere.

Given that the distance of the first peak is  $\sigma$  the lattice spacing is  $\sqrt{3}\sigma$  by just applying Pythagoras' theorem. One could also use  $\frac{1}{d_{hkl}^2} = \frac{1}{h^2 + k^2 + l^2}$  and rearrange for  $d$  knowing that  $h=k=l = \sigma$ .

Extrapolation to the y axis from the first minimum (low gradient in integral) in figure 6 shows that the coordination number is 12. By taking the integral to the first minimum, the coordination number of the other two phases was found. The liquid has a coordination number of 12 and the gas has a coordination number of 6.

## Conclusion

The Lennard Jones potential models systems of atoms well because atoms are spherical, do not possess dipoles, don't and don't change shape. To model water, you would need to account for the covalent bonding in the water molecules; you would not want your atoms separating. One would also need to account for the rotational forces experienced by the molecules changing their orientation. One would need to account for the dipole – dipole interactions of the water using a dipole – dipole potential which considers the relative angle of each water molecule. You would also need to account for the hydrogen bonding exhibited which will act at close range with a specific bond distance and angle and will therefore have its own potential function.

The cut off used in the above calculations has been  $2.5 \sigma$  as written in the input files. The advantages a larger cut off is more accurate calculations due to considering the interaction of more of the particles with each other particle. This has diminishing returns because particles that are further apart exhibit weaker interaction (usually exponentially). Further, increasing the cut off increases computation costs exponentially since you are effectively increasing the size of a sphere containing all the relevant particles for each other particle. The volume of a sphere is  $\frac{4}{3}\pi r^3$  so increasing the cut off can significantly increase the number of particles considered in each timestep. Therefore, it is important to find a balance between cut off and computation time due to the diminishing returns and large increase in computation time.

If  $\epsilon$  was very small, then a liquid phase could not be observed.

Finite size effects are effects due to the size of the repeating simulation box used. In the above simulations the finite size of the experiment had a drastic effect on the results because the system being modelled would be homogeneous in a similar way on the macro scale with only one type of atom. The cut off for the potential calculation was exactly half the side length of the box so no particle would be seeing more than one copy of every other particle anyway. There were some issues with the finite size of the simulations such as the distance that the RDF could be determined however the RDF devolved to the bulk for both fluids way before the cut off due to repetition so RDF could be accurately determined. For systems that are mixtures of atoms or contain larger molecules then the box should be increased in size to accommodate for the larger unit size of particles to avoid repetition within the cut off for interaction.

Using SHAKE or RATTLE to fix bond length in systems with large molecules would be very useful because it would allow you to increase your timestep without your bonds spontaneously breaking. As discussed above, larger timesteps allows for atoms to jump further up potential cliffs than would otherwise be feasible. This could lead to undesired bond breaking in covalent molecules if bonds aren't fixed. Having a larger timestep with larger molecules would greatly help with simulation time and would allow for larger systems with big molecules to be simulated, allowing diffusion to be modelled for example.

## Bibliography.

1. Molecular Simulation/Radial Distribution Functions - Wikibooks, open books for an open world. Accessed December 1, 2021. [https://en.wikibooks.org/wiki/Molecular\\_Simulation/Radial\\_Distribution\\_Functions](https://en.wikibooks.org/wiki/Molecular_Simulation/Radial_Distribution_Functions)