KTH-CBH
SIMULATION METHODS IN MEDICAL ENGINEERING (CM2014)

LABORATORY REPORT
MOLECULAR SIMULATION LAB

GROUP 18

FEBRUARY 23, 2024

GROUP MEMBERS:
NAME: YIZHI DONG YIHAN XIAO

DATE PERFORMED:
FEB 9, 2024

Contents

1 Introduction

This lab is based on the offered script. It uses a simple algorithm to solve Newton's equations of motion for a set of 40 interacting particles. Some parameter of the script will be adjusted to see what influences the energy and dynamics of the particles.

2 Points to Address

2.1 Question a)

This question is about the concept of Verlet algorithm.

2.1.1

Derive the Verlet algorithm using the Taylor expansion for r(t+dt) and r(t-dt).
The Taylor expansion for r(t+dt) and r(t-dt) are:

$$r(t+\delta t) = r(t) + v(t)\delta t + a(t)\delta t^2/2 + b(t)t^3/6 + ...$$
$$r(t-\delta t) = r(t) - v(t)\delta t + a(t)\delta t^2/2 - b(t)t^3/6 + ...$$

From their approximation we can get:

$$r(t+\delta t) \approx r(t) + v(t)\delta t + a(t)\delta t^2/2 + b(t)t^3/6 \quad (1)$$
$$r(t-\delta t) \approx r(t) - v(t)\delta t + a(t)\delta t^2/2 - b(t)t^3/6 \quad (2)$$

From equation (1) + (2) we can get:

$$r(t+\delta t) = 2r(t) - r(t-\delta t) + a(t)\delta t^2$$

And this is the equation for the Verlet algorithm.

2.1.2

The expression for the velocity v(t) in Verlet algorithm:
From equation (1) - (2) we can get:

$$r(t+\delta t) - r(t-\delta t) = 2\delta t v(t)$$

So:

$$v(t) = \frac{r(t+\delta t) - r(t-\delta t)}{2\delta t}$$

2.1.3

Explain why the Verlet algorithm is not a self-starting algorithm.

$$r(t+\delta t) \approx r(t) + v(t)\delta t + \frac{1}{2}a(t)\delta t^2$$

Because with only the initial position $r(t)$ we can calculate the acceleration $a(t)$, but the initial speed $v(t)$ is unknown and need to be obtained somehow to calculate the next step.

2.1.4

Explain how we generate starting velocity in the script. PS: in the script a modified version of the Verlet algorithm is used.

According to the Maxwell–Boltzmann distribution, the velocity of the gas molecules in each direction follows a Gaussian distribution that can be described as:

$$f(v_x) = \sqrt{\frac{m}{2\pi k_b T}} e^{-\frac{mv_x^2}{2k_b T}}$$

The molecular dynamics simulation is run in a two-dimensional space, in order to generate the initial velocity, the script created two uniform random variables x_1 and x_2, which range from -1 to 1, and used a polar form Box-Muller transform to generate Gaussian random numbers.

$$z_1 = \sqrt{-2ln({x_1}^2 + {x_2}^2)} \frac{x_1}{\sqrt{{x_1}^2 + {x_2}^2}}$$

$$z_2 = \sqrt{-2ln({x_1}^2 + {x_2}^2)} \frac{x_2}{\sqrt{{x_1}^2 + {x_2}^2}}$$

Since the velocity in one dimension follows a Gaussian distribution with a standard deviation of $\sqrt{\frac{k_b T}{m}}$, so the velocities would be:

$$v_1 = z_1 \sigma = \sqrt{-2ln({x_1}^2 + {x_2}^2)} \frac{x_1}{\sqrt{{x_1}^2 + {x_2}^2}} \sqrt{\frac{k_b T}{m}}$$

$$v_2 = z_2 \sigma = \sqrt{-2ln({x_1}^2 + {x_2}^2)} \frac{x_2}{\sqrt{{x_1}^2 + {x_2}^2}} \sqrt{\frac{k_b T}{m}}$$

2.2 Question b)

The atoms interact with van der Waals interactions, which we model using the Lennard-Jones potential. As was explained before, the parameter ϵ is the well-depth. Perform a simulation using a Lennard-Jones potential with a depth of 1.5 kJ/mol (keep the initial σ value) and compare the results with the simulations you have already performed for Argon atoms. In particular:

2.2.1

The following figure is the potential energy curve in function of the distance.

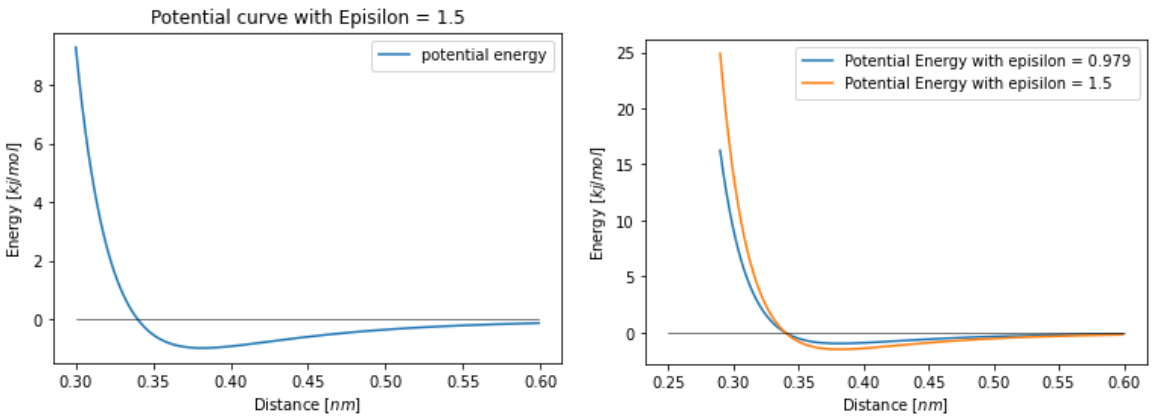

Figure 1: Potential energy curve in Function of Distance

According to the figures, the shape of the energy curve doesn't change much. In both curves, potential energy reaches 0 at the same point, and approach the lowest point at the similar place. And there's also plot with both curves in it.

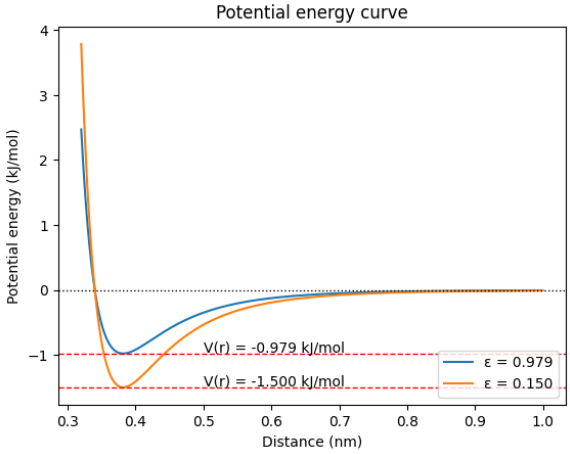

Figure 2: Potential energy curve with different epsilon values

For what has changed, except for the balance point (potential energy = 1), the absolute value of potential energy with the well-depth equaling 1.5 becomes larger. In the left side of the balance point, the curve with $\epsilon = 1.5$ is on top of the curve with $\epsilon = 0.979$, which means the repulsive force becomes larger. In the right side of the balance point, the curve with $\epsilon = 1.5$ is under the curve with $\epsilon = 0.979$, which means the attraction force becomes larger.

2.2.2

In this part, the changes in the energies derived from the Python simulations will be described. The analysis is supported by the following graph.

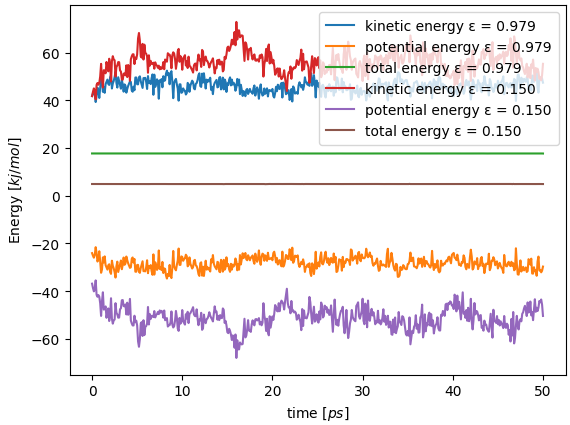

Figure 3: Overall Simulation results with different epsilon values

According to this plot, the kinetic energy doesn't change much, and the minor change may be caused by the random generated initial speed. For potential energy, the absolute value of the potential energy becomes larger in curve with $\epsilon = 1.5$, which is consistent with the conclusion on one particle derived in the previous part. Finally, Due to increase of the absolute value of the negative potential energy, the total energy of the system decreases, but still conserves.

2.3 Question c)

Total energy ($E_{tot} = E_{pot} + E_{kin}$) in the simulation should be conserved, so it is good practice to monitor this during the simulation. Describe what happens to the total energy if you use a time step larger than 0.01 ps. What is the largest time step that one can use to simulate the Argon atoms and keep energy conservation. Support your results with simulation outputs.

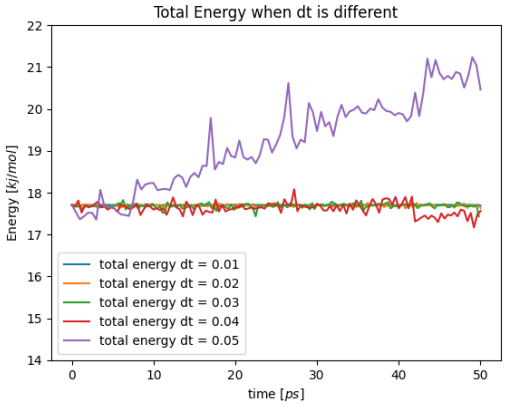

Figure 4: Total energy during the simulation with different time steps

It can be inferred from the plot that when $dt = 0.04$, ripples can be seen in the results, but the simulation could still continue. when $dt = 0.05$, the system became unstable, and the energy conservation is broken.

So some repeated simulations are done with different time steps. Each simulation lasts 50ps, and the time step is increased by 0.001ps in each iteration. Then the variance of the total energy during each iteration of simulation are compared.

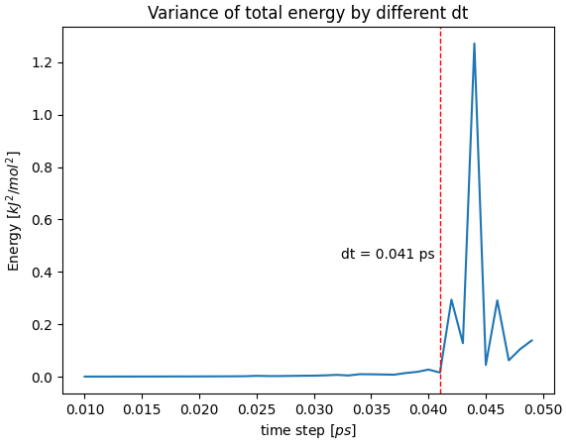

Figure 5: Variance of total energy by different time step

It can be seen from the plot that starting from $dt = 0.041$ the variance begin to rapidly increase, so the largest time step that one can use to simulate the Argon atoms and keep energy conservation is 0.04ps.

2.4 Question d)

Describe the effect of mass and temperature on the kinetic energy. To achieve this, perform simulations at higher and lower temperatures and different mass values. Tips: change only one value in each simulation.

According to the figures above, kinetic energy increased as temperature increased, and they almost had a linear relationship, and kinetic energy didn't vary much when mass changed.

Theoretically, average kinetic energy only relates to temperature, and has a linear relationship below, and the result of simulation confirmed it:

$$E_k = \frac{3}{2} k_B T$$

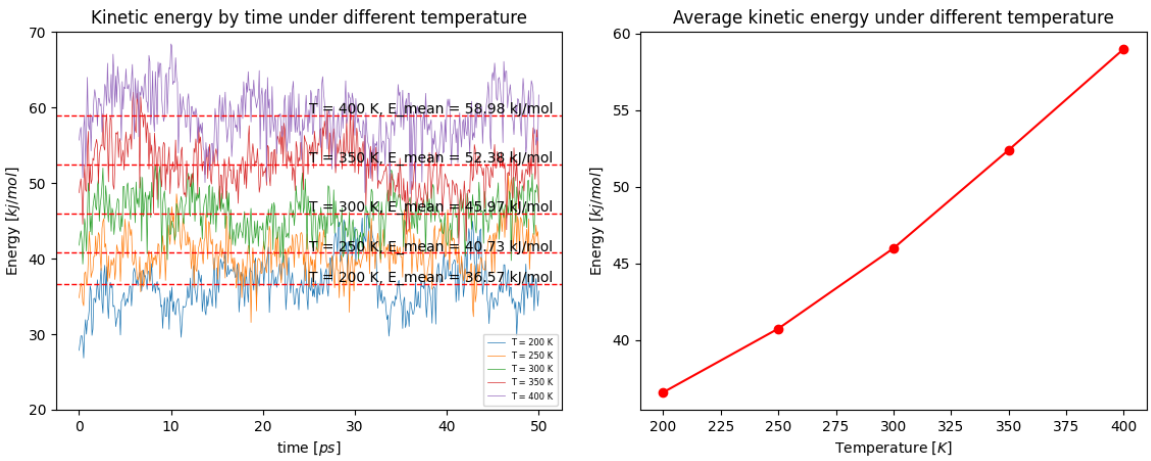

Figure 6: Result of simulation under different temperature

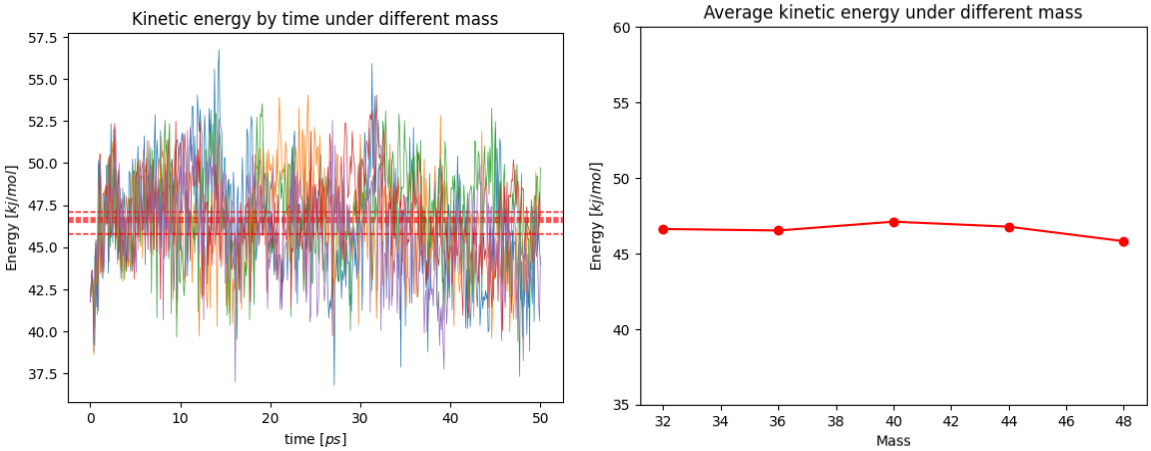

Figure 7: Result of simulation under different mass

2.5 Question e)

Finally, imagine that you want to simulate the conformation rearrangement of a small protein (a system of ¡100000 atoms) in detail. (the time scale of the process is nanoseconds):

2.5.1 Choice between Monte-Carlo or Molecular Dynamics

Molecular dynamics may be the better choice.

Monte-Carlo uses random sampling to build and explore the conformation of a protein, which samples potential energy's global minimum with high efficiency. However, the protein's realistic evolution by time is missed. So, it may be more suitable for studying the equilibrium properties.

Molecular Dynamics makes use of Newton's laws of motion. So the trajectories of atoms, as well as the velocity is calculated step by step over time. Because we want to simulate the conformation rearrangement in detail, this method is better at calculating dynamics. Also, it can provide information about the thermodynamics which is related to kinetics energy [3]. In conclusion, Molecular Dynamics is the better methods because the trajectories and velocity are calculated step by step, then the conformation rearrangement can be studies in detail.

2.5.2 Choice between atomistic or coarse grained model

An atomistic molecular model may be the better choice.

A coarse grained model simplifies the representation of a protein and set of interaction. By grouping several atoms as one particle, the computational cost is decreased, thus longer time scale is allowed. However, details of dynamics and conformation are lost. It is more suitable when studying the large-scale motions and interactions of proteins.

In an atomistic model, each atom of a protein is studied explicitly. It provides details of conformation and dynamics, and the trajectories and velocity of each particle is calculated separatedly [4]. So, to simulate the conformation rearrangement of a small protein in detail, an atomistic model may be better, because 100000 is not a large number, with the help of modern computer.

3 Conclusion

References

[1] Majdolhosseini, Maryam & Villa, Alessandra. *Molecular Simulation Computer Lab & Report Instructions*. 2024.

[2] Majdolhosseini, Maryam & Villa, Alessandra. *Computer lab- Molecular Simulation*.

[3] Shukla, R., Tripathi, T. (2020). Molecular Dynamics Simulation of Protein and Protein–Ligand Complexes. In: Singh, D.B. (eds) Computer-Aided Drug Design. Springer, Singapore. https://doi.org/10.1007/978-981-15-6815-2_7

[4] Kaynak BT, Krieger JM, Dudas B, Dahmani ZL, Costa MGS, Balog E, Scott AL, Doruker P, Perahia D and Bahar I (2022) Sampling of Protein Conformational Space Using Hybrid Simulations: A Critical Assessment of Recent Methods. Front. Mol. Biosci. 9:832847. doi: 10.3389/fmolb.2022.832847