Project 3: Linked list and thermostat

Outi Haapanen, Lotta Gustavsson, Markus Rauhalahti, Gaoyue Bi

May 4, 2017

1 Introduction

In this project we implemented a linked cell -algorithm and Langevin thermostat as a python program `molmodp3.py`. The program may be ran by typing `python3 molmod3.py` on a Linux or a Mac. The input parameters may be changed by editing the program code.

2 Linked cell algorithm

Lists are used to calculate the non-bonded interactions between particles. The trivial way to check pairwise interacting particles is to make a list of all pairwise interacting particles and test all possible particle pairs by running the whole list through – the method scales as N^2. This method gets time- and energy-consuming as the number of particles increases. To improve the efficiency, cell-linked lists are used in simulations to calculate the forces between unconnected particles within a cut-off radius. Cell-linked lists scale linearly which makes them a lot more efficient.

The implementation of linked cell algorithm is made by dividing the simulation box (L_x,L_y,L_z) into subboxes. The size of a subbox is greater than or equal to $\frac{L}{\lfloor \frac{L}{r_c} \rfloor}$ where r_c is the cutoff radius. Each box has 26 neighbours but because of periodic boundary conditions only 13 of them must be considered. A list of box pairs is made and stored in an array list – 14 different varieties of box pairs. In the beginning, all possible pre-calculations are done and box-paired linked lists are created. After initialization during every timestep, the absolute position and the subbox of each particle is stored. Linked lists of particles in subboxes are created and box-paired linked lists are run through. The forces on each particle are calculated and added to the total force of the particle if the distance of particles are less than the cut-off radius.

3 Canonical ensemble (NVT) and Langevin thermostat

Canonical ensemble is a statistical ensemble that represents the possible states in a system that is in thermal equilibrium at constant temperature. In canonical ensemble, the amount of substance, volume and temperature are conserved and it is also called "constant temperature molecular dynamics" (CTMD). The energy of exothermic and endothermic processes occurring in this system are in interaction and thus exchanged with a thermostat, heat bath etc. Since the system can exchange energy with the thermostat, the states of the system will differ in total energy. Introducing a thermostat in molecular dynamics simulations requires an algorithm that is able to interact with the particles' kinetic energy. Various thermostats have been developed to do this, for example Nose-Hoover, Behrensen and Langevin.

Langevin dynamics/thermostat is one method of controlling the temperature in molecular dynamics simulation. Obtaining canonical ensemble distribution is a sum of right thermostat and its parameters, timesteps and system size.

The basis of general molecular dynamics simulations is Newton's second law, but Langevian dynamics includes collisional and frictional parameters in the equation. The Langevian dynamics equation thus becomes

$$m\frac{d^2\mathbf{r}}{dt^2} = \mathbf{F} - m\gamma\frac{d\mathbf{r}}{dt} + \mathbf{R} \tag{1}$$

where γ is collisional frequency ($1/\gamma$ is often called velocity relaxation time, which is the time taken for a particle to "forget" its initial velocity) [1, p. 373] [2]. It is to be noted that when the

collisional frequency is in the limit of zero and no random force is applied, equation (1) becomes the equation from the Newton's second law [3, p. 389]. This relates to the situation where no thermostat is applied to the system. Langevin equation can also be stated as

$$m\ddot{\mathbf{r}} = -\Delta U(\mathbf{r}) - \Gamma\dot{\mathbf{r}} + \xi(t) \tag{2}$$

where $\xi(t)$ is Gaussian-correlated noise source that mimics random impulses and $-\Gamma\dot{\mathbf{r}}$ is the diffusion factor. Usually, the random force, $\xi(t)$ is considered to be uncorrelated with the particle forces and velocities and the forces acting on the particles [3, p. 389]. Therefore, the random force has the following properties

$$\langle \xi(t) \rangle = 0 \tag{3}$$
$$\langle \xi(t)\xi(t') \rangle = 2A\delta(t - t') \tag{4}$$

where A is a constant.

4 Testing the code

4.1 Linked cell list

The linked cell algorithm was implemented as derived in the Allen and Tildesley book [4]. The implementation differs by the use of the native list data structure in python instead of using linked lists. Both lists are extendable, but the python lists are in this case not called by calling for a head of a list.

The test are ran using the following parameters as outputted by the program

```
Molecular Modeling, project 3
Parameters:
   N = 50
   Steps = 100
   dt = 0.001
   L = 5
   cutoff = 1
LJ parameters:
   sigma = 0.001
   epsilon = 0.01
Linked cell parameters:
   Number of cells = 125
   Cell length = 1.0
```

by varying the particle number N. The scaling of the program is shown in 4.2.

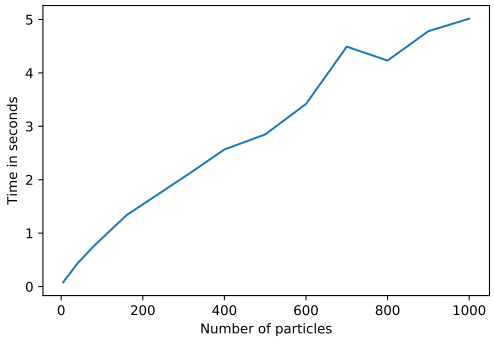

Figure 1: Scaling behaviour using the linked cell algorithm

The program obtains a near linear scaling. The inconsistencies are perhaps due other calculations performed on the computer simultaneously.

N	Total	Iterations	Forces	List generation	I/O
50	0.490445569	0.26786484	0.135453798	0.045567791	0.040893397
500	2.93300215	1.674962974	0.469114637	0.435651919	0.347906669
5000	20.852120938	12.466260794	0.57809586	4.250284018	3.514218079

Table 1: Time consumption in seconds of different parts of the program

The time consumption of the different parts are shown for 50, 500, and 5000 particles in table 4.1

The time taken for I/O and list generation makes sense as they scale linearly. The I/O part is at least optimized by limiting the frequency the program spits out the rsults.

The scaling of forces is insensible as the time taken changes very little upon a large increase of the system size. This indicates that there are errors in the algorithm.

4.2 Langevin thermostat

The Langevin thermostat was implemented by substracting the momentum term and the term with γ and the temperature from the force.

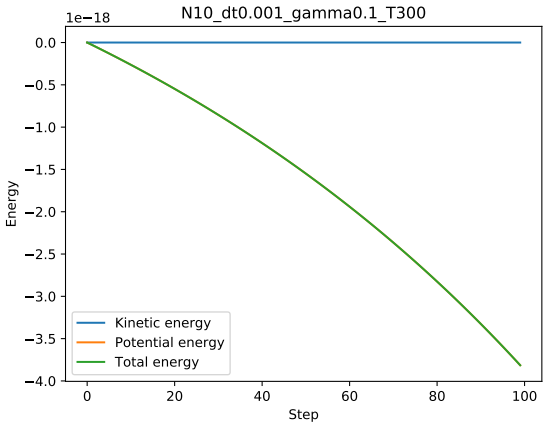

Figure 2: dt = 0.001

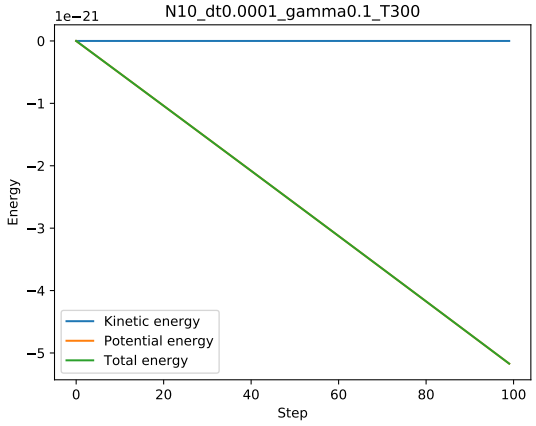

Figure 3: dt = 0.0001

As can be seen, the energies are not stable at all. This is not affected by the timestep, but is due to faulty implementations.

5 Conclusions

A linked cell algorithm and the Langevin thermostat were implemented. The code is buggy. The scaling is somewhat linear thanks to the linked cell algorithm, whereas the Langevin thermostat does not seem to work. The program was nonetheless written. The clarity of the implementation presents the understanding of the algorithm, which through some debugging could have given correct results.

References

[1] Alan Hinchcliffe. *Molecular Modelling for Beginners.* Wiley, 2003.

[2] M. J Abraham, D. van der Spoel, E. Lindahl, B. Hess, and the GROMACS development team. GROMACS User Manual, version 5.0.7. 2017.

[3] Andrew R. Leach. *Molecular Modelling Principles and Applications.* Pearson Education Limited, 2010.

[4] Mike P Allen and Dominic J Tildesley. *Computer simulation of liquids.* Oxford university press, 1989.