

Worksheet 3: Molecular Dynamics 2 and Observables

Olaf Lenz Fatemeh Tabatabaei Marcello Sega

December 1, 2012

Institute for Computational Physics, University of Stuttgart

Contents

1	General Remarks	1
2	Introduction	2
3	Saving and Restarting the Simulation	3
4	Simple Observables	3
5	Equilibration	4
6	Setting up and Warming up the System	6
7	Molecular Dynamics at a Desired Temperature	7
8	Advanced Observables	8

1 General Remarks

- Deadline for the report is **Thursday, 13th December 2012, 10:00**.
- On this worksheet, you can achieve a maximum of 20 points.
- The report should be written as though it would be read by a fellow student who attends to the lecture, but does not do the tutorials.
- To hand in your report, send it to your tutor via email
 - Olaf (olenz@icp.uni-stuttgart.de; Thursday, 14:00 – 15:30)

– Jens (smiatek@icp.uni-stuttgart.de; Friday, 8:00 – 9:30)

- Please attach the report to the email. For the report itself, please use the PDF format (we will *not* accept MS Word DOC files!). Include graphs and images into the report.
- If the task is to write a program, please attach the source code of the program, so that we can test it ourselves.
- The report should be 5–10 pages long. We recommend to use L^AT_EX. A good template for a report is available.
- The worksheets are to be solved in groups of two or three people.

2 Introduction

On this worksheet, you will continue to do the Molecular Dynamics simulations of the Lennard-Jones system that you have started on the last worksheet. While on the last worksheet, the focus was on implementing an MD simulation of a LJ system, on this worksheet you will learn how to set up and analyze a system by measuring some observables and their mean values.

All files that are required for this tutorial can be found in the archive `templates.tar.gz` that can be downloaded from the lecture's homepage.

You will start with a program that is very similar to the final program of the last worksheet, *i.e.* a LJ simulation with Verlet and cell lists, written in C and Python. Again, to compile the C-part of the program, use the following command:

```
> python setup.py build_ext --inplace
```

As the simulation runs in this work will typically take a few minutes, it is split into two programs. The Python program `ljsim.py` will perform the simulation itself and measure different observables. Compared to the program from the last worksheet, it was extended by two features. First of all, it does not compute the energy and write out the coordinates to the VTF file in every timestep, but only every `measurement_stride` time steps. This saves both memory as well as computation time. Furthermore, at the end of a simulation run, the program will write the measured energies into a file `ljsim.dat`. To do that, it uses the Python module `pickle`, that can be used to write almost any Python object into a file and read it from there.

The Python program `ljanalyze.py` can then read this data file and plot the data. Like this, it is not necessary to repeat the costly simulation over and over again, but you can plot the results in various ways by just modifying `ljanalyze.py`.

Throughout this worksheet, you will be asked to perform simulations of a system with the following parameters: $N = 1000$ LJ particles at a density of $\rho = 0.316$ with a time step $\Delta t = 0.01$. Initially, the particles are set up on a regular cubic lattice and given small random velocities.

3 Saving and Restarting the Simulation

Unfortunately, it is usually not clear beforehand how long a simulation actually needs to be run. Furthermore, as you will see in the following tasks, it might be necessary to turn on or off specific measurements or algorithms after some simulation time has elapsed.

Therefore, it is useful to be able to stop a simulation after some time and restart it from the state, possibly with other algorithms activated.

Task	(? points)
<ul style="list-style-type: none">• Study the programs <code>ljsim.py</code> and <code>ljanalyze.py</code>. Focus on those parts of the program that you do not know yet, in particular reading data from and writing data to a file.• Extend the program <code>ljsim.py</code> such that you can restart the simulation from the point where it left off in the last run, and simulate the system for some more time.	

Hints

- Think about which variables need to be saved in a file to be able to restart the simulation (*i.e.* the *state* of the system) and which variables should/can stay in the Python program itself.
- To plan the program and to establish a workflow of how to perform a whole simulation, it might be a good idea to first read the rest of the worksheet.
- You can store the current state of the simulation in the same data file where the energies are stored, and also restart the simulation from that file. In fact, the measured energies should probably be part of the state, as you want to append the newly measured energies to the end of the old energies.

4 Simple Observables

Performing a simulation is futile if we do not at least also measure some observables of the simulated system, so that we can watch how they change over time and compute their average values.

First of all, we want to measure the energy components of the system, *i.e.* the *kinetic energy* E_{kin} and the *potential energy* E_{pot} . The program already computes the total energy, which is just a sum of both of these components. In a microcanonical ensemble, the total energy should remain mostly constant, but the energy components may vary.

Another interesting observable is the temperature T_m of the system. It can be measured via the equipartition theorem, which states that the kinetic energy per degree of freedom

is $1/2 k_B T_m$. As point-like particles have 3 translational degrees of freedom each, the temperature can be computed as follows:

$$\frac{1}{2} k_B T_m = \frac{E_{\text{kin}}}{3N} \quad (1)$$

Finally, we want to compute the *pressure* P of the system. From statistical mechanics it can be shown that the pressure can be computed from two contributions. One of them is the ideal gas contribution, the other one comes from the interactions. Without giving details, we just state the formula:

$$P = \frac{1}{V} \left(\frac{1}{3} \sum_{i=0}^N m \mathbf{v}^2 + \sum_{i=0, j < i}^N \mathbf{F}_{ij} \cdot \mathbf{r}_{ij} \right). \quad (2)$$

where V is the volume of the system and \mathbf{F}_{ij} is the force between particle i and j .

Task

(? points)

- Extend `ljsim.py` such that the energy components are measured and returned from the function `compute_energy()`.
- Add a function `compute_pressure()` that computes the pressure of the system.
- Did you implement the function in the Python- or in the C-part of the program? Explain why you did it in the one and not the other part!
- Add a function `compute_temperature()` that computes the temperature of the system.
- Did you implement the function in the Python- or in the C-part of the program? Explain why you did it in the one and not the other part!
- All of the observables should be measured whenever the energy was measured in the original program.
- The time series of the various observables should be written to the data file.

5 Equilibration

Imagine you want to measure the density of water at ambient conditions and you have two samples: one stored in a freezer and one on a shelf. If you take the one from the freezer and do the measurement immediately, the result will differ from the one on the shelf. If you leave the frozen sample on a table for some time and let it come to

equilibrium with the surroundings, the result will be the same as for the one from the shelf. In a simulation we are in a similar situation. Usually we do not know at the beginning what the system looks like under the simulation conditions. On the contrary, we are often performing the simulation to answer this question. We prepare the system in an arbitrary initial configuration and let it equilibrate with a virtual heat bath.

Equilibration is manifested in a time-drift of all observables. On the contrary, once the equilibrium state is reached, *all* observables fluctuate around a constant value. Some observables relax to equilibrium quickly while others may take much longer. How long the equilibration takes depends on system properties as well as on how far from equilibrium it is. A system cannot be considered as equilibrated if there exists one single observable which exhibits a time-drift. Note that it is even possible that an observable exhibits a non-monotonic behavior, and that the direction of the time-drift reverses after some time. When you want to compute statistical averages of the observables, you should do so only after the system is equilibrated.

Unfortunately, there is no recipe to determine whether a system has equilibrated or not. Instead, it is necessary to have a thorough look at the plots, and probably also to visualize the system to see what is going on. When the measured quantities do not drift anymore, most probably the system is equilibrated, but this is no guarantee! To decide if we can start collecting data, we have to make sure that the time over which the observables are fluctuating around a constant value exceeds the estimated duration of equilibration considerably. We may also need to apply some physical knowledge of the simulated system.

Often, plots of the raw data of the observables are fluctuating very strongly so that it is hard to see whether the observables still display a time-drift on a scale that is smaller than the fluctuations themselves. To make such a drift visible, one can smooth the curve of the observable. The simplest way to do so are *running averages*. If you have a time series of N measurements of the observable O_i , the running average time series \hat{O}_i is defined by:

$$\hat{O}_i = \frac{1}{M} \sum_{i-M/2}^{i+M/2} O_i \quad (3)$$

where M is the *window size* of the running average, *i.e.* the number of points over which the average is performed. The larger the window size is chosen, the more smoothed is the function.

Task

(? points)

- Let the simulation run for 1000 time units and plot the different observables over time.
- Implement a function `compute_running_average(O,M)` that computes the running average, where `O` is a Numpy-Array containing the time series of the observable and `M` is the window size of the running average. The function should return a NumPy array of the time series of the running average.
- Plot the running averages of the pressure, temperature and energy components for window sizes of 10 and 100.
- Explain the behavior of the observables. What is happening in the system?
- At what time would you feel confident that the system is equilibrated? Provide plots or images that support your claim.

Hints

- Visualize the system with VMD to understand what is happening.
- To simplify plotting of the running average, the function `compute_running_average` should return an array of the same size as the input array. The values of \hat{O}_i for $0 < i < \frac{M}{2}$ and $N - \frac{m}{2} < i < N$ can be set to `NaN`.

6 Setting up and Warming up the System

At the moment, `ljsim.py` sets up the different LJ particles on a cubic lattice. This avoids that the particles overlap and overly large energies are produced. However, this is a very artificial state of the system and only works in the particular case of a LJ system where all particles have the same size. Therefore, we want to learn how this can be done cleanly.

In general, it is very difficult to set up an arbitrary system that does not produce these overlaps. Instead, in the most cases it is sufficient to set up a system with overlapping particles in an arbitrary fashion, *e.g.* by randomly placing the particles, and then use a *warmup* to get rid of the overlaps.

There are different possibilities how to do a warmup. For example, one can try to minimize the energy via a gradient descent algorithm (or any other optimization algorithm), or one can slowly let the particles grow until they have reached their final size. Another frequently used procedure works like this: During the warmup phase, one does a “normal” simulation of the system but limits the forces acting on a single particle to

a maximal value (*force capping*). Like this, the overlapping particles feel a force that pushes them apart, but these forces cannot become overly large.

Of course, once there are no very large forces left, the warming up procedure has to be turned off, as it modifies the physics of the system.

Task

(? points)

- Extend `ljsim.py` such that the components of the force acting on a particle can be limited to a maximal (absolute) value of 20.
- It should be possible to turn on or off force capping via a constant (or maybe even a command line parameter?), so that you can later measure the properties of the system without force capping.
- When you measure observables during the warmup, their values will be bogus, so turn off all measurements while force capping is active.
- However, you still need to know whether or not the warmup is finished and there are no overly large forces anymore. One possibility to solve this problem is to run the warmup not for a specified time, but simply until no force is capped anymore.
- Did you implement the force capping in the Python- or in the C-part of the program? Explain why you did it in the one and not the other part!
- Change the setup procedure of the simulation such that the particles are created at random positions.
- Test whether the new setup with warmup works as intended.

7 Molecular Dynamics at a Desired Temperature

Because we are using the energy-conserving velocity verlet algorithm and we keep a constant number of particles in a box of constant volume, we are simulating the microcanonical NVE ensemble by construction. However, experiments are typically performed at a given constant temperature. To be able to compare simulations to experimental data, one would prefer to simulate the system at a predefined temperature, rather than at a predefined total energy, *i.e.* to simulate a canonical NVT ensemble.

Therefore, we need to be able to keep the temperature constant, *i.e.* we need an extension to the MD algorithm that mimicks a heat bath that the particles are coupled to. Such an algorithm is called a *thermostat*. There are a number of different thermostat algorithms with different advantages and disadvantages.

The simplest thermostat is the *velocity rescaling* thermostat, which also bases on the equipartition theorem (see Eq. (1)). The idea of the velocity rescaling algorithm is simply to multiply all velocities by a factor such as to achieve the desired temperature.

One important problem of the velocity rescaling thermostat is that it destroys the Maxwell-Boltzmann distribution of velocities in the system. This means that although a simulation keeps the temperature constant, it does not really produce the NVT ensemble. On the next worksheet, you will learn about more advanced thermostats that actually simulate the canonical NVT ensemble.

Here, we will first use the velocity rescaling thermostat to drive the system to the desired temperature, but then turn off the thermostat and measure the properties of the system in NVE ensemble with an average temperature close to the desired temperature. Note that the equilibration might still not be finished even when the temperature is stable.

Task

(? points)

- Derive how to compute the factor with which to multiply the velocities to achieve the desired temperature T from equation 1.
- Extend `ljsim.py` such that the velocities are rescaled to yield the desired temperature directly after the observables are measured. It should be possible to turn on or off the thermostat via a constant (or maybe even a command line parameter?), so that you can later measure the properties of the system without the thermostat.
- Did you implement the velocity scaling in the Python- or in the C-part of the program? Explain why you did it in the one and not the other part!
- Run the simulation with the standard parameters (see above) at the desired temperatures $T \in \{0.3, 1.0, 2.0\}$ and plot the evolution of the various observables. Run the simulation long enough until you observe no more time drift in any of the observables.
- Explain the observed behavior.

8 Advanced Observables

An observable containing information about the structure of the system is the *radial distribution function* (RDF), which is defined by

$$g(r) = \frac{1}{\rho^2 4\pi r^2 dr} \sum_{ij} \langle \delta(r - |r_{ij}|) \rangle \quad (4)$$

The RDF describes the probability to find a particle at a particular distance r to another particle. In the ideal gas, $g(r) = 1$ for all values of r . It can be used to compute all

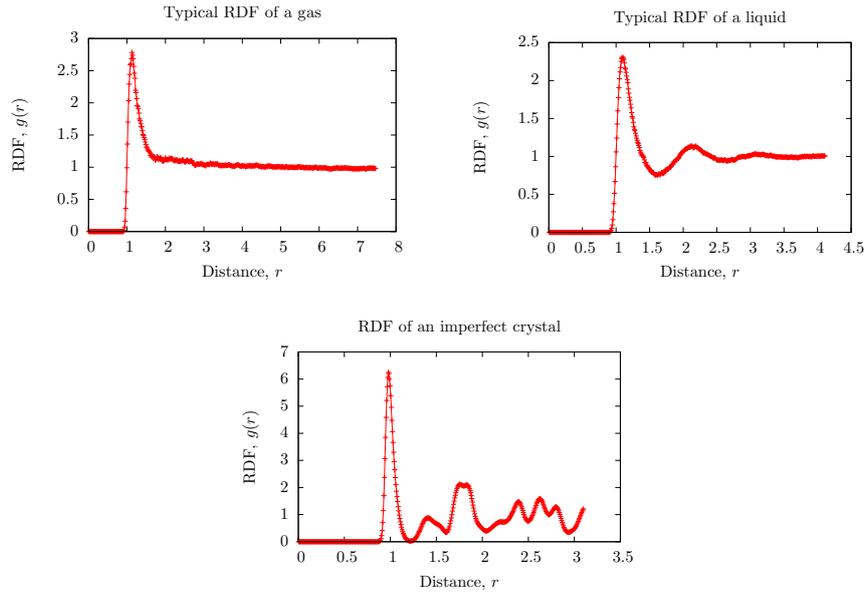


Figure 1: Examples of radial distribution functions obtained from simulations.

kinds of thermodynamic observables. Figure 1 displays plots of the typical RDF in the different phases. For small distances, the RDF is 0, which corresponds to the fact that the particles cannot overlap. At larger distances, the RDF approaches 1, which means that the probability to find a particle in that distance corresponds to the overall density of the system.

... to be completed ...