at the University of Freiburg

Master Thesis

February 21, 2021

Author:

Wilkin Wöhler

Supervision:

Tanja Schilling

Abstract

A raw writing of hard sphere nucleation, a simulation to measure quantiteis, and a analysis of data generate by means of the simulation. A test citation: [1]

Contents

1 Theory **1**

1.1 Hard sphere system . . . 1
1.2 metastable fluid/ phase diagram . . . 1
1.3 Classical nucleation theory :/ . . . 2
1.4 Memory including approaches . . . 3
1.5 Computer Precision . . . 4
1.6 Comparsion to Real world experiments . . . 4

2 Simulation details **5**

2.1 Algorithm and Simulation details . . . 5
2.1.1 Event driven molecular dynamics (EDMD) . . . 6
2.1.2 Details of the Implementation . . . 7
2.2 Probe of simulation code . . . 7
2.2.1 Diffusive behaviour . . . 7
2.2.2 Radial distribution function . . . 8
2.3 Estimate of required resources . . . 8
2.3.1 Calculation time estimates . . . 8
2.3.2 File sizes estimates . . . 8
2.4 Produced Data . . . 8
2.4.1 Equilibration steps . . . 8
2.4.2 Initial density . . . 8
2.5 Possible extensions . . . 8
2.5.1 Varying radius . . . 8
2.5.2 Varying mass . . . 8
2.5.3 Muliprocessing . . . 8

3 Data Analysis **9**

3.1 Diffusion of the lqiuid . . . 9
3.2 Diffusion of the metastable liquid . . . 9
3.3 Cluster growth . . . 9
3.4 Tensor of Gyration properties . . . 9

3.5 ACF largest cluster? 9
3.6 Nuclection time dilemma 9
3.7 Induction time by exponential distribution 9
3.8 Nucleation rate comparison 10
3.9 Memory Kernels 10

4 Conclusion - Summary **11**
4.1 Conclusion 11

5 Appendix **12**
.1 A 12

List of Figures

List of Tables

1 Theory

1.1 Hard sphere system

The Hard Sphere system is the simplest model of a fluid including interactions between the single particles. Its well known potential between particles i and j reads:

$$V(r_{ij}) = \infty \cdot \Theta(\sigma - r_{ij}) \tag{1.1.1}$$

In this equation r_{ij} indicates the distance between the two particles, σ is the diameter of the Hard Spheres and Θ is the Heavyside function.
While the ideal gas model without pair interactions already makes it possible to derive famous equations as $pV = NkT$, it does not include phase transitions yet. But these can be observed when granting the particles to take up space. Because it is the simplest model and it is well feasible for computer simulations the Hard Sphere system is very well suited to study basic properties of phase transitions.

Compared to experiments where similar systems can also be realized, general properties of the system at hand can be varied very precisely without much effort and position data of the single particles can be extracted easily as well, because they naturally are required for the simulation.

On the downside computer simulations are much more constraint in their size, but with todays computational possibilites system of the order of 1 million particles become tractable, and such computer simulations become a powerfull tool to study also phase transitions in simple systems.

The beginning of such simulations actually dates back to the beginning of electronic computer technology (cite Alder and Wainwright 1959). Since then more algorithms to increase efficency have been elaborated, and technology advanced giving today the possibility of studying large systems.

1.2 metastable fluid/ phase diagram

The equation of state for the simple Hard Sphere system has various approximations, (cite an overview paper), The probably most common approximation due to its simplicity is the Caranhan-Sterling ap-

proximation:

$$Z = \frac{1 + \eta + \eta^2 - \eta^3}{(1 - \eta)^3} \tag{1.2.1}$$

(cite https://aip.scitation.org/doi/10.1063/1.1672048) It approximates the compressibility factor Z depending on the packing fraction η for the Hard Sphere fluid.

For the stable brach after nucleation a common approximation is given by the Alamrza equation of state(cite https://aip.scitation.org/doi/full/10.1063/1.3133328).

$$\frac{p(v - v_0)}{k_B T} = 3 - 1.807846y + 11.56350y^2 + 141.6y^3 - 2609.26y^4 + 19328.09y^5 \tag{1.2.2}$$

where p is the pressure, v is the volume per particle $v_0 = \sigma^3/\sqrt{2}$ is the volume per particle at close packing, including the diameter of the spheres σ, and $y = p\sigma^3/(k_B T)$, where k_B is the Boltzman constant and T is the temperature of the crystal.

From these two equations of state we can draw the phase diagram: (include grapic of phase diagram.)

The chemical potential difference between the two equations of state can be calculate from the difference between the two equations of state.

Eventhough not further discussed in this thesis it might be said that for polydisperse radii the phase diagram becomes even richer as show for example in https://journals.aps.org/prl/pdf/10.1103/PhysRevLett.91.068301

1.3 Classical nucleation theory :/

Classical nucleation theory (CNT) has been proposed look who did this the first time and since then multiple times modified to describe various types of systems. Eventhough its predictions often deviate far from experimental results. look if tanja cited someone else here. Either way it can provide a reference to know what to expect roughly and also it can be checked by the simulation data.

CNT assumes that a spherical crystalite may form in the liquid with properties of the bulk crystal while the fluid remains with the properties of the bulk liquid. The difference in the free energy landscape is given by a surface and a volume term. The first arises from the surface tension γ between the fluid bulk and solid bulk phases. The second comes by the difference in chemical potential $\Delta\mu$. The whole expression reads:

$$\beta\Delta G(R) = 4\pi R\gamma - \frac{4}{3}\pi R^3 \rho \Delta\mu \tag{1.3.1}$$

Where ρ is the particle density of the solid phase.

The free energy barrier can be sketched like this (maybe sketch it for som mu and gamma).

As can be seen it has a maximum at R_{crit} the critical radius. If the radius of a cluster surpasses the critical radius, it is likely to continue to grow until it incorporates all available fluid. It is given by:

$$R_{crit} = \frac{2\gamma}{\rho\Delta\mu} \tag{1.3.2}$$

Furthermore the height of this barrier can be calculated to be

$$\beta\Delta G(R_{crit}) = \frac{16\pi\gamma^3}{3\rho^2(\Delta\mu)^2} \tag{1.3.3}$$

In the classical picture now we look at an activated process which has an rate given by the Arhenius law:

$$k = c\exp(-\frac{k_B T}{\Delta E}) \tag{1.3.4}$$

$$\Leftrightarrow \quad k = c\exp -\beta\Delta G(R_{crit}) \tag{1.3.5}$$

As the constant c is not further defined the absolute nucleation rate in this picture is not predicted but only set into realtion to ther rates. Check carefully in how far this is correct!
Given the equation of state for the liquid and fluid phase from section 1.2, and taking a literature value for γ cite the refernce to which the value corresponds we can calculate R_{crit} for various densities.
Plot of r_{crit} for different densities

1.4 Memory including approaches

CNT assumes a Markovian system, which means that it diffuses through a free energy landscape without granting it memory. But it has been shown by Kuhnbold ... cite Anjas Lennard-Jones System that for a Lennard-Jones system memory effects can not be neglected if an acurate desciption is desired. Such it must be assumed that also for the Hard Sphere system memory effects may be present.

To analyze such memory effects a framework by Hugues Meyer(cite Hugues) has been elaborated. In it a memory kernel for coarse-grained observables by means of projection operators is defined together with an efficent algorithm to calculate such.
With the derived memory kernel an equation of motion for the observable can be dervied, ressmbling mostly the Generalized Langevin equationbu is called non-stationary Generalized Langevin Equation because the memory kernel can depend not only on $K(t-\tau)$ but instead on two times $K(t,\tau)$:

$$\frac{dA_t}{dt} = \omega(t)A_t + \int_0^t d\tau K(\tau, t)A_\tau + \eta_t \quad , \tag{1.4.1}$$

For the Markovian process the memory kernel is approximately given by a Dirac delta functional $\delta(\tau - t)$, in which case the Langevin equation is recovered.

1.5 Computer Precision

The precision of computer certainly influences the outcome of simulations. (talk with fabian if he found any papers reagarding varying results fromvarying precision over time)
And it should always be kept in mind that the simulation only approximates the real world. Even smallest variations in the last digits of positions, changes the simulation radical after a certain number of steps. This comes by the fact that we face a complex system with chaotic behaviour, which means that even small variations will grow exponentally until the system has nothing in common anymore. Look if you can visulaize such an behaviour by As it can be seen after x timesteps the two system have radicallly changed.

1.6 Comparsion to Real world experiments

Since the 1950s ? people have systhesized hard sphere like systems in the lab. Today a whole zoo of systems is known. All of these system have in common, that the hard spheres are in a bath of a fluid, which surrounds them. This fluids density and optical refractive index shhould match the density and optical refractive index of the hard spheres to prevent segmentation and to enable optical measurements of the system. maybe elaborate a little more on why each of the two is necessary

The absence of the bath in simple hard sphere simulations is probably the largest difference to the hard sphere systems in the laboratory. It has been argued that the difference can be circumvented by normalizing with a diffusion length or time, but a discussion on the possibility of hydrondynamic effects changing the behaviour of the lab system compared to simulations is ongoing at the moment.

For simulations it is much harder to include the bath becuase it introudces a multiple number of particles, making it very slow to simulate large systems.

2 Simulation details

During the coarse of the master thesis an event driven molecular dynamic (EDMD) simulation code has been elaborated. The choice to use the EDMD approach is taken because interst in the actual dynamics of the system were desired. This means that simulations probing the phase space of the system instead of the dynamics, like Monte Carlo (MC) simulation schemes, are not suited.
Furthermore the discontinous potential of the hard spheres is an obstacle not easy to face in regular molecular dynamics (MD) schemes, where the Newtonian equation of motion for the particles is numerically integrated.
The EDMD approach on the other site actually requires these discontinuities as will be discussed in the following sections, together with some details of the program.

2.1 Algorithm and Simulation details

In this section we will highlight the main differences to regular MD simulations, as they are the main tool to otherwise probe the dynamics of the system. Furthermore we will stick to the hard sphere example when discussing the EDMD simulations, but it can be kept in mind that the EDMD approach also allows to simulate particles with other potentials as long as the potentials are only containing step functions.

The decisive difference between EDMD simulations and regular MD schemesis that, instead of evaluating all pair and external forces on each particle and then evolving the whole system to the next time step, EDMD simulations do not have a predefined time step, but the system is evolved from one event to the next one. An event in this context is defined as the time where the next collision in the whole system takes place.

The event prediction algorithm is follows closely the approach proposed by Bannerman et. al [**Bannerman2014**] which will be discussed in the next section.

2.1.1 Event driven molecular dynamics (EDMD)

For the prediction of events in EDMD simulations an overlap function $f_{ij}(t)$ between particles i and j is defined, where the squared quantities are used merely because they are easily accessible.

$$f_{ij}(t) \coloneqq |\vec{r}_j(t) - \vec{r}_i(t)|^2 - \sigma^2 \tag{2.1.1}$$

$$\begin{aligned} &\text{with} \quad \vec{r}_i(t) = \vec{r}_i(t_0) + (t - t_0)\ \vec{v}_i(t_0)\,, \\ &\qquad \Delta t \coloneqq t - t_0\,, \\ &\qquad \vec{v}_{ij}(t) \coloneqq \vec{v}_j(t) - \vec{v}_i(t)\,, \\ &\qquad \vec{r}_{ij}(t) \coloneqq \vec{r}_j(t) - \vec{r}_i(t)\,, \\ &\Leftrightarrow \quad \vec{r}_{ij}(t) = \vec{r}_{ij}(t_0) + \Delta t\ \vec{v}_{ij}(t_0) \end{aligned} \tag{2.1.2}$$

$$f(t) = (\vec{r}_{ij}(t_0) + \Delta t\ \vec{v}_{ij}(t_0))^2 - \sigma^2 \tag{2.1.3}$$

$$f(t) = |\vec{r}_{ij}(t_0)|^2 + \Delta t^2\ |\vec{v}_{ij}(t_0)|^2 - 2\Delta t\ \vec{r}_{ij}(t_0) \cdot \vec{v}_{ij}(t_0) - \sigma^2 \tag{2.1.4}$$

The overlap function has the property that it is negative for two particles being closer than their diameter, 0 for at collision and positive if neither overlapping nor touching. The calculation of the next collison thus is to calculate the roots of eq. 2.1.4.

Solving for Δt with $|\vec{r}_{ij}(t_0)|^2 \coloneqq rr$, $|\vec{v}_{ij}(t_0)|^2 \coloneqq vv$ and $\vec{r}_{ij}(t_0) \cdot \vec{v}_{ij}(t_0) \coloneqq rv$ is rather trivial:

$$0 = rr + vv\ \Delta t^2 - 2rv\ \Delta t - \sigma^2 \tag{2.1.5}$$

$$\Leftrightarrow \quad 0 = \Delta t^2 - \frac{2rv}{vv}\ \Delta t + \frac{rr - \sigma^2}{vv} \tag{2.1.6}$$

$$\Leftrightarrow\ \Delta t = -\frac{rv}{vv} \pm \sqrt{\left(\frac{rv}{vv}\right)^2 - \frac{rr - \sigma^2}{vv}} \tag{2.1.7}$$

But a caveat when executing on a floating point machine is present as can be seen when considering which solution is of interest. As for a possible collision it is necessary that the two particles move towards each other we can conclude that the scalar product is required to be positive $rv > 0$, because otherwise the particles are already moving away from each other.

Also the quadratic formula has two solutions, corresponding to the entry and the exit of the overlap. Because the entry has to be prior to the exit, we further conclude that interest lies on the smaller solution that is:

$$\Delta t = \frac{-rv + \sqrt{(rv)^2 - vv(rr - \sigma^2)}}{vv} \tag{2.1.8}$$

Now for the case where the distance of the spheres is already close to the diameter of the spheres we find $(rv)^2 >> (rr - \sigma^2)$, which results in a cancelation of two large numbers leaving a small number. Floating point number operations are inherently bad suited because they tend to large inaccuracy in this case. Rewriting eq. 2.1.8 by making use of the third binomial formula look if this is fine to write. leads to:

$$\Delta t = \frac{(rr - \sigma^2)}{-rv - \sqrt{(rv)^2 - vv(rr - \sigma^2)}} \tag{2.1.9}$$

Comparably this does not contain a cancelation of the type seen before and such is better suited for the use in a computer simulation. cite Goldberg '91

For this purpose all possible collisions of the particles have to be pre calculated and then stored in a future event list (FEL), which then is sorted by the time of occurance. From the FEL it can be read when the next collision between which two particles will take place. The collisions outcome itself can be analytically calculated equation of collision where the two particles are assigned new momenta. Collision between more than two particles is close to excluded by treating the hard spheres as actual hard spheres.

After every collision new events have to be calculated for collision now possible for the two particles. These have to be sorted into the FEL, and a new cycle beginns.

Because some events may have been calculated for the previous ballistic trajectory of a particle which collided in the mean time, it is necessary to establish an interaction count for the particles. Every time an event occurs it is incremented by one. Each event now has

2.1.2 Details of the Implementation

Add Details of for example FEL, and backupevent handling, double time precision, reset sim

2.2 Probe of simulation code

To probe we have to measure known quantities

2.2.1 Diffusive behaviour

Show the diffusive behavior of at least the fluid

2.2.2 Radial distribution function

Show a RDF of the fluid , if possibible with the theoretical cervus-pevick approximation

2.3 Estimate of required resources

2.3.1 Calculation time estimates

GIve some profiling numbers of the simulation Also conclude that missing q6q6 O(N^2), broke the walltime.

2.3.2 File sizes estimates

Show the estimate on the file size

2.4 Produced Data

Overview of produced data with visualized snapshot?

2.4.1 Equilibration steps

Show dependence of equilibration steps on simulation

2.4.2 Initial density

Show dependence of initial density on simulation

2.5 Possible extensions

Highlit possible extenisions which migth be doable to implement.

2.5.1 Varying radius

Vary the radius of the spheres. Rewquirements and thought on that.

2.5.2 Varying mass

Same as with radius, highlight the points in the code requiring a change.

2.5.3 Muliprocessing

Idea on Parrallelizing the code. Probably more advanved, but for larger systems (Really large files) doable. (Actually RAM might become a serious problem).
Besenrein, Samstag 6.3. mittag, 13 h

3 Data Analysis

This is the analysis part

3.1 Diffusion of the lqiuid

This contain analysis of diffusion in the liquid to prove the simulations accuray

3.2 Diffusion of the metastable liquid

This contains Diffusion constants to normalize the rates

3.3 Cluster growth

Cluster growth depending on density

3.4 Tensor of Gyration properties

Well only swamp here, but it can be shown to conclude the swamp.

3.5 ACF largest cluster?

Just in case anything can be seen here

3.6 Nuclection time dilemma

Evaluation of induction time. Prolem with accuracy and precision. Compare methods.

3.7 Induction time by exponential distribution

Obtain exp assumption and best estimator

3.8 Nucleation rate comparison

All Nucleation rates that can be found.-> mayhap ask Hajo.

3.9 Memory Kernels

Memory kernels of systems at various densities. Depends strongly on what is found here

4 Conclusion - Summary

4.1 Conclusion

5 Appendix

.1 A

Bibliography

[1]D. M. Heyes, M. J. Cass, J. G. Powles, and W. A. Evans, "Self-diffusion coefficient of the hard-sphere fluid: System size dependence and empirical correlations", Journal of Physical Chemistry B **111**, 1455–1464 (2007).