UNIVERSITY OF ABERDEEN

HONOURS DISSERTATION

CHRISTOPHER JAMES THOMSON
2012

THIS THESIS WAS SUBMITTED AS PART OF THE REQUIREMENT FOR THE MENG. DEGREE IN ENGINEERING.

School of Engineering 2012 Univ. of Aberdeen

ABSTRACT

The objective of this thesis is to determine a general technique for converting continuous potentials to equilivant discrete potentials. Discrete potentials have many desirable properties.

CONTENTS

Final word count: 5606

LIST OF FIGURES

LIST OF TABLES

NOMENCLATURE

Acronyms/Terminology

FCC Face Centered Cubic, 22

MD Molecular Dynamics, 9

RDF Radial Distribution Function, 30

SPEADMD Step Potentials for Equilbria and Discontinuous Molecular Dynamics, 15

Operators

∇ Gradient operator, 11

Variables/Constants

$\mathbf{a}$ Acceleration vector of a particle, 11

Φ_{ij} Intermolecular potential between particles i and j, 11

σ Characteristic length, taken to be particle diameter, 12

ε Characteristic energy, usually well depth, 12

F Force, 11

F_{ij} Intermolecular force between particle i and j, 11

F_i Force acting on particle i, 11

m Mass of a particle i, 11

ACKNOWLEDGEMENTS

I would like to dedicate this work to..

CHAPTER

ONE

INTRODUCTION

Process simulation packages have become an integral part of chemical engineering design. Central to these simulation software packages is the ability to calculate thermodynamic and transport properties of fluids quickly and accurately. Many modern processes rely on molecular scale effects... (Absorption, membrane technology (reverse osmosis), catalysis)....

To better understand these large scale systems, we need to improve our understanding of the smaller scales. Experiments are hard, can't hold a ruler up to a molecule, everything is too fast, too small to see. (X-ray crystallography). theory is hard, lots of molecules, can't solve even three molecules motion analytically (need cite). simulations are great, don't try to solve analytically. Since 50's (alder wainwright), computers are faster, modern sims are amazing.

At the heart of these simulations are models for the atoms and molecules involved. There are two classes of models, discrete and contiuous.

In chapter **??**, the something is discussed

CHAPTER

TWO

MOLECULAR MODELS

In this chapter, the dynamics and models used to represent molecular systems are discussed. First, the arguments for selecting classical mechanics over quantum mechanics as the underlying dynamics for the simulations are presented. Then, the two major classifications of the available classical models, discrete and continuous potentials, are discussed.

2.1 Classical Mechanics

2.1.1 Validity of Classical Mechanics

The underlying assumption behind many molecular dynamics simulations is that the particles move according to the laws of classical mechanics. Strictly, atoms and molecules should be treated using quantum mechanics, due to their size and speed. However molecular dynamics makes two assumptions that allow these quantum mechanical effects to be ignored.

The first is the Born-Oppenheimer Approximation which allows the motion of electrons and the nucleus to be treated separately. Since the nucleus is much larger than the electrons and hence less affected by quantum mechanics, it is treated as a classical particle. The electrons on the other hand are represented using a potential [1].

The second assumption is that any quantum mechanical effects should average out. Molecular dynamics is rarely interested in the motion of a single electron, it is more concerned with the statistical average over every particle.

These assumtions are usually valid unless very light atoms (such as hydrogen or helium) are being simulated, or the particles are vibrating at very high rates [2].

2.1.2 Newton's Second Law of Motion

The fundamental identity of classical mechnanics is Newton's Second Law of Motion (Eq. (2.1.1)). This equation allows the prediction of a particle's trajectory provided that an initial position and velocity is known; and the forces acting on that particle can be calculated for any position or velocity.

$$\mathbf{F} = m\mathbf{a} \tag{2.1.1}$$

If a force depends only on the position of a particle, it is known as a conservative force. Almost all forces considered in molecular dynamics are of this type because atoms or molecules do not lose energy due to friction or any other dissipative process.

Conservative molecular forces can be of one of two types. The first are body forces, which only depend on a single particle's absolute position, such as gravity. However gravity is usually disregarded in MD as atoms and molecules have such low masses that gravity has very little effect on them.

The second and most important type are intermolecular forces that depend on position relative to other particles. In general, the intermolecular forces are a function of all particle positions; However, usually the total force acting on a particle i is assumed to be a sum of the forces between i and every other particle j. This is known as the binary interaction assumption and the summation of forces is given by

$$\mathbf{F}_i = \sum_{j \neq i}^{N} \mathbf{F}_{ij} \tag{2.1.2}$$

where N is the total number of particles in the system. Force calculations are limited to pairs as this is simpler, but there are examples of n-body forces available in the literature (e.g. see Ref. [3]).

The intermolecular forces used in molecular dynamics are frequently described using a potential. Provided that the force is conservative, it can be calculated from its potential using the following equation

$$\mathbf{F}_{ij} = -\nabla \Phi_{ij} \tag{2.1.3}$$

where ∇ denotes the gradient of the potential, Φ_{ij}. Using this definition, it is clear to see that by defining an intermolecular potential, the intermolecular forces are also described. In the literature, intermolecular force fields are usually described using potentials and in the following sections the two types of potentials will be discussed.

2.2 Continuous Potentials

Models for the intermolecular interactions are typically defined using functions for the potential energies between molecules. A very popular potential used in molecular dynamics simulations is the Lennard-Jones potential [4] given by the following expression

$$\Phi(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right] \tag{2.2.1}$$

where ε is the depth of the energy well and σ is the distance where the potential between two particles is zero. This is illustrated in Fig. 2.1. Despite its simplicity, it gives comparable results to experimental values [5] for noble gases due to their monatomic nature.

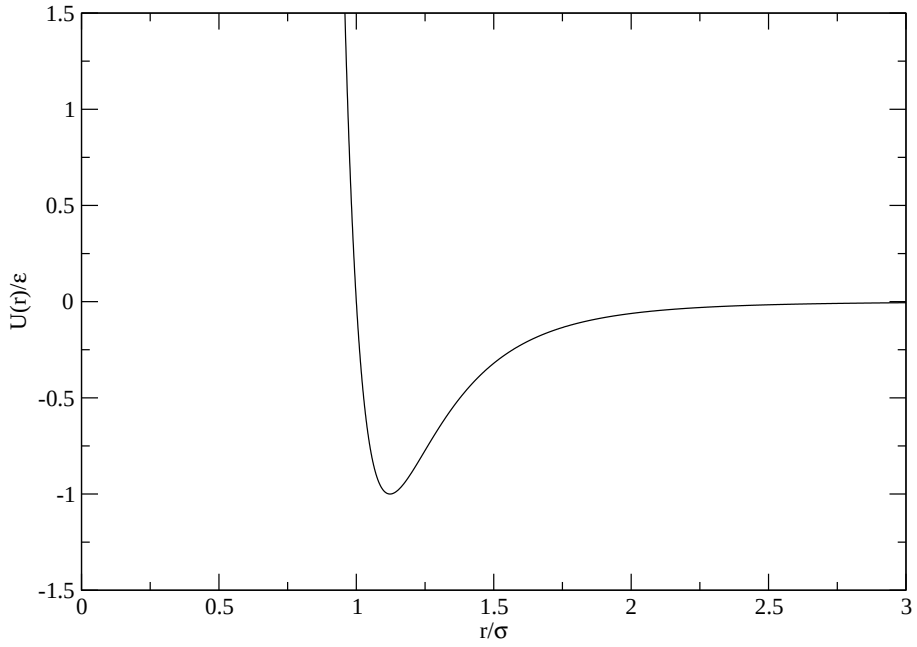

Figure 2.1: Plot of the Lennard-Jones potential

The power 12 term, $(\sigma/r)^{12}$, gives the potential a repulsive core caused by Pauli repulsion of overlapping electron shells. This repulsion is more accurately represented by an exponential function which forms the basis of the Buckingham potential [6]. However the Buckingham potential is not as popular as the Lennard-Jones potential because the expontential function is more computationally expensive to calculate [7].

The power 6 term, $(\sigma/r)^{6}$, in the Lennard-Jones potential represents the attractive Van der Waals forces. This attractive well and repulsive core can be seen in the force plot of the Lennard-Jones potential, Fig. 2.2.

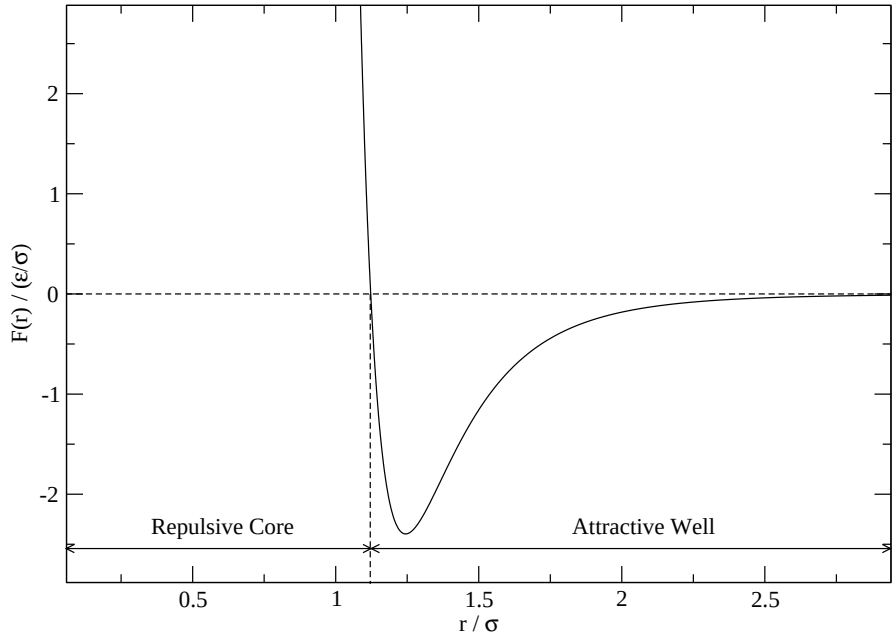

Figure 2.2: Plot of the force between a pair of Lennard-Jones particles separated by a distance r

$$\begin{aligned}\Phi = &\sum_{\text{bonds}} k_b(r-r_0)^2 + \sum_{UB} K_{UB}(S-S_0)^2 + \sum_{\text{angles}} k_\theta(\theta-\theta_0)^2 \\ &+ \sum_{\text{dihedrals}} k_\chi[1+\cos(n\chi-\delta)] + \sum_{\text{improper}} k_\psi(\psi-\psi_0)^2 \\ &+ \sum_{i=1}^{N-1}\sum_{j>i}^{N}\left\{4\varepsilon\left[\left(\frac{\sigma}{r_{ij}}\right)^{12} - \left(\frac{\sigma}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}}\right\}\end{aligned} \tag{2.2.2}$$

The Lennard-Jones potential can be included in part of more complex potentials such as the CHARMM (Chemistry at HARvard Molecular Mechanics) potential for simulating proteins (Eq. (2.2.2)) [8]. Here the Lennard-Jones potential is used to represent Van der Waals forces while Coulomb's Law $\left(\frac{q_i q_j}{r_{ij}}\right)$ is modelling longer range electrostatic interactions. The first five terms of Eq. (2.2.2) are contraints on bond movements (streching, Urey-Bradley bond vibration, angle bending, dihedral angle, and improper dihedral angle motion respectively, see Ref. [9] for descriptions of these bond movements).

2.3 Discrete Potentials

Discrete potentials differ from continuous potentials because they have discontinuities. The simplest discrete potential is that of the hard sphere (see Fig. 2.3).

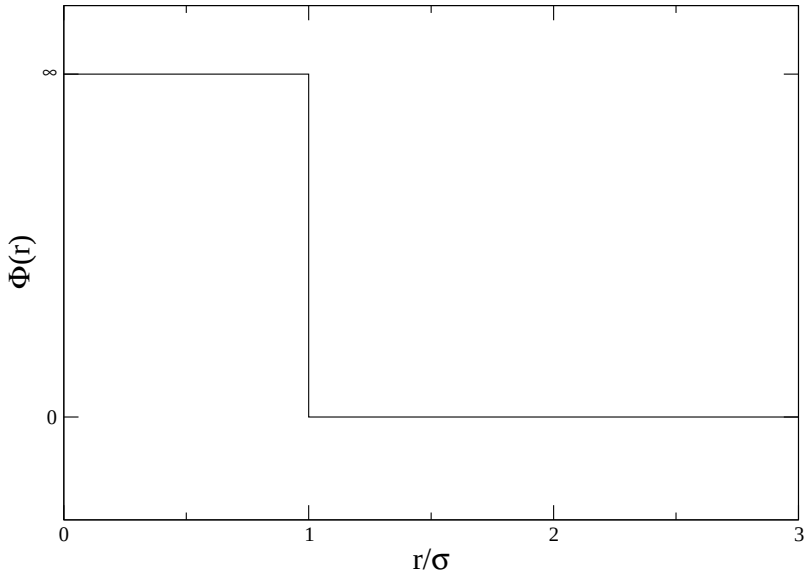

Figure 2.3: Plot of the hard sphere potential

Hard spheres were the first molecular models ever simulated [10] due to their relative simplicity. The potential for hard spheres is shown in equation (2.3.1), where σ is the diameter of the spheres.

$$\Phi = \begin{cases} \infty & \text{if } |\mathbf{r}_i - \mathbf{r}_j| < \sigma \\ 0 & \text{if } |\mathbf{r}_i - \mathbf{r}_j| > \sigma \end{cases} \tag{2.3.1}$$

By calculating the force between two hard spheres using equation (2.1.3) shows that hard spheres exert no force until the particles collide whereupon they experience an infinite repulsive force. This means that hard spheres cannot overlap as they move have to exceed an infinite force to do so.

The hard sphere potential can be elaborated upon by adding an attractive well outside the hard core (see Fig. 2.4. This square well potential takes the form of equation (2.3.2) [11], where λ is the outer radius of the core in terms of the hard core diameter σ.

$$\Phi = \begin{cases} \infty & \text{if } |\mathbf{r}_i - \mathbf{r}_j| < \sigma \\ -\varepsilon & \text{if } \sigma < |\mathbf{r}_i - \mathbf{r}_j| < \lambda\sigma \\ 0 & \text{if } |\mathbf{r}_i - \mathbf{r}_j| > \lambda\sigma \end{cases} \tag{2.3.2}$$

The square well potential is similar to the Lennard-Jones potential in that they have a steep repulsive core followed by an attractive well, however the two give quite different simulation results. In order to model molecular interactions more effectively, stepped potentials have to be used.

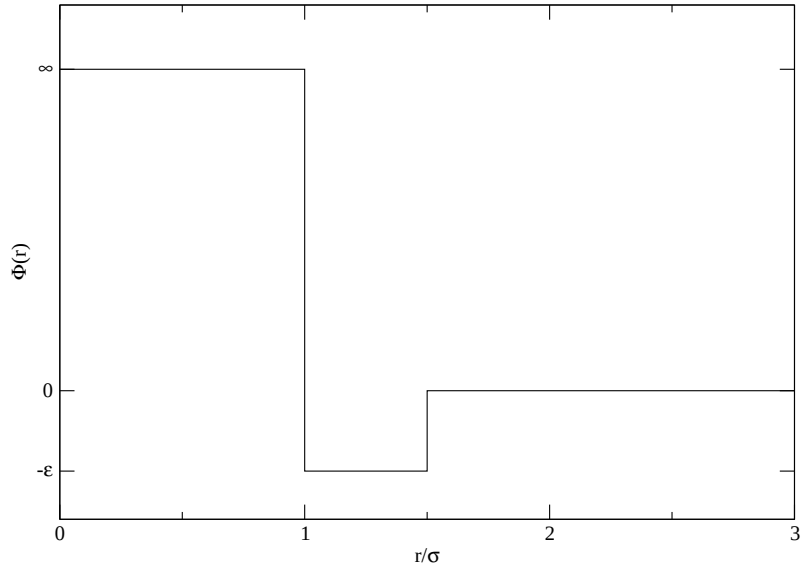

Figure 2.4: Plot of the square well potential with $\lambda = 1.5$

Stepped potentials are a combination of square wells and square shoulders to mimic the behaviour of a continuous potential. Many types of stepped potentials have been developed in the literature. Chapela *et al.* [12] created a stepped version of the Lennard-Jones potential shown in Figure 2.5.

The SPEADMD (Step Potentials for Equilibria and Discontinuous Molecular Dynamics) project has created stepped potentials that represent a range of hydrocarbons [13–15]. These potentials have been converted to equations of state that match experimental values relatively well.

Stepped potentials have also been used to model protein interactions using PRIME.

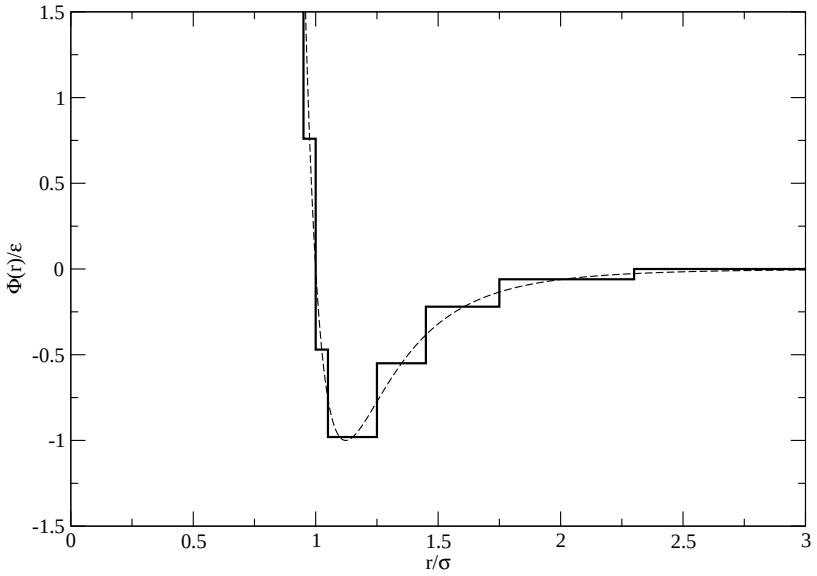

Figure 2.5: A stepped version (solid line) of the continuous Lennard-Jones potential (dashed line) created by Chapela *et al.*

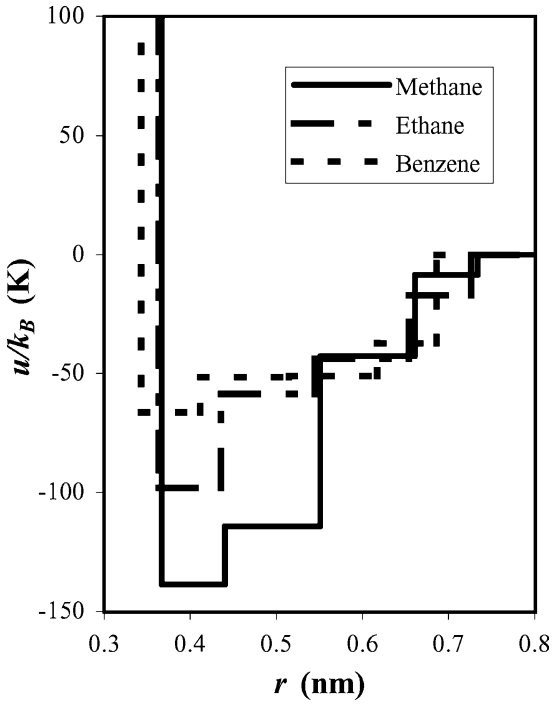

Figure 2.6: SPEADMD stepped potentials for methane, ethane and benzene [13].

CHAPTER

THREE

MOLECULAR DYNAMICS

3.1 Applying Classical Mechanics to Molecules

The most fundamental part of molecular dynamics is solving Newton's Second Law of Motion (Eq. (3.1.1)).

$$\mathbf{F} = m\frac{\partial^2 \mathbf{r}}{\partial t^2} \tag{3.1.1}$$

While the forces acting on the particles can be calculated from their intermolecular potentials, the future positions of the particles become the solution of a differential equation. The techniques used to solve this differential equation are different depending on whether the potential is continuous or discontinuous.

3.2 Force-Driven Simulation

3.2.1 Introduction

For continuous potentials it is not possible to solve for the particle's future position analytically. This because the force is not constant so the differential equation has an infinite number of derivatives. Therefore numerical methods for solving differential equations have to be used; these are known as integrators.

3.2.2 Integrators

The majority of numerical integrators are based on Taylor Series (Eq. 3.2.1).

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \frac{\partial \mathbf{r}(t)}{\partial t}(\Delta t) + \frac{1}{2}\frac{\partial^2 \mathbf{r}(t)}{\partial t^2}\Delta t^2 + \frac{1}{3!}\frac{\partial^3 \mathbf{r}(t)}{\partial t^3}\Delta t^3 + \frac{1}{4!}\frac{\partial^4 \mathbf{r}(t)}{\partial t^4}\Delta t^4 + ... \tag{3.2.1}$$

The simplest integrator is Euler's Method which is just the Taylor Series truncated after the acceleration term (Eq. 3.2.2).

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(\Delta t) + \frac{1}{2}\mathbf{a}\Delta t^2 \tag{3.2.2}$$

However this method suffers from large errors and is highly unstable (i.e. it amplifies any errors) [16] and is therefore rarely used. The Verlet Integrator [17] improves upon Euler's method by combining the forward timestep with a reverse timestep (Eq. (3.2.3a)). This method is actually third order as the third (and first) derivative is cancelled out during its derivation. The Verlet integrator does not include an equation to calculate the future velocity, as knowledge of the velocity is not required for the integrator. However it is often desired to know velocities to calculate physical properties so the central difference used by Verlet is often used to calculate velocities (Eq. (3.2.3b)).

$$\mathbf{r}(t+\Delta t) = 2\mathbf{r}(t) - \mathbf{r}(t-\Delta t) + \mathbf{a}(t)\Delta t^2 \tag{3.2.3a}$$

$$\mathbf{v}(t+\Delta t) = \frac{\mathbf{r}(t+\Delta t) - \mathbf{r}(t-\Delta t)}{2\Delta t} \tag{3.2.3b}$$

Integrators suffer from two key failings that cause a systematic gain of energy known as "energy drift". Firstly, integrators are based on infinite Taylor series which cannot be fully implemented and have to be truncated after a certain number of terms; this introduces an error. Secondly, integrators are unable to predict values of forces that have discontinuities in them, such as discrete potentials or discontinuities introduced by truncating potentials to improve simulator speed. There are a couple of types of integrators that try and reduce these problems.

The first method to improve the traditional integrator is the predictor-corrector integrator. These use a truncated Taylor series to calculated a predicted value for the future position and higher order time derivatives. The force is then calculated at this predicited position, then the difference between the predicted acceleration and the corrected acceleration calculated from the force is used to correct the position and time derivatives.

The most popular predictor-corrector integrator is that of Gear [18] using the 5th order algorithm. The predicted value for the n^{th} time derivative is shown in Eq. (3.2.4). By defining $\Delta\mathbf{a} = \mathbf{a}^C - \mathbf{a}^P$, where the superscripts C and P denote the corrected and predicted values respectively, the corrected time derivatives can be calculated using Eq. (3.2.5) with coefficients from Eq. (3.2.6).

$$\frac{\partial^n}{\partial t^n}\mathbf{r}^P(t+\Delta t) = \sum_{k=i}^{5} \frac{1}{k!}\frac{\partial^n}{\partial t^n}\mathbf{r}(t)\Delta t^k \tag{3.2.4}$$

$$\frac{\partial^n}{\partial t^n}\mathbf{r}^{\,C}(t+\Delta t) = \frac{\partial^n}{\partial t^n}\mathbf{r}^{\,P}(t+\Delta t) + \frac{c_i}{\Delta t^i}\left(\frac{\Delta t^2}{2}\Delta\mathbf{a}\right) \tag{3.2.5}$$

$$c_0 = \frac{3}{16},\ \ c_1 = \frac{251}{360},\ \ c_2 = 1,\ \ c_3 = \frac{11}{18},\ \ c_4 = \frac{1}{6},\ \ c_5 = \frac{1}{60} \tag{3.2.6}$$

Gear's algorithm, while more accurate at short timesteps than Verlet's integrator [16], loses precision at long timesteps and is computationally expensive.

The other method used to try and mitigate the failings of other types of integrator is the sympletic integrator. These have the useful property in that they, on average, conserve energy [19]. The most common symplectic integrator used in MD is the Velocity Verlet Integrator [20] shown in (3.2.7).

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \frac{1}{2}\mathbf{a}(t)\Delta t^2 \tag{3.2.7a}$$

$$\mathbf{v}(t+\Delta t) = \mathbf{v}(t) + \frac{\mathbf{a}(t)+\mathbf{a}(t+\Delta t)}{2}\Delta t \tag{3.2.7b}$$

The popularity of the Velocity Verlet is due to its computational simplicity and its accuracy and stability at relatively long timesteps. It can even be expanded [21] to maintain its accuracy and stability at very long timesteps at a small extra computational cost. However the Velocity Verlet cannot be used in systems that do not conserve energy, i.e. systems with dissipative forces.

3.3 Event-driven Simulation

3.3.1 Introduction

The methods for integrating continuous potentials rely on calculating the forces acting on the particles, however this is problematic when working with discrete potentials. The gradient (and hence the force) is either infinite at the steps or zero between them.

Therefore a new technique for simulating these discrete potentials is needed. There is no force between the steps therefore the acceleration (and higher order time derivatives) are zero and the Taylor series (equation (3.2.1)) can be reduced to equation (3.3.1).

$$\mathbf{r}(t+\Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t \tag{3.3.1}$$

This allows the prediction of when particles reach a step, and using the conservation of energy and momentum, the post collision velocities of the particles can be calculated (this is expanded upon in section 3.3.3). This method is known as event-driven molecular dynamics and it is an analytical solution of Newton's Second Law of Motion.

3.3.2 Collision Time Prediction

The prediction of the future positions of the particles using Eq. (3.3.1) requires the time before the next discontinuity is reached by any pair of particles. This section outlines how these collision times are calculated.

For hard sphere simulations there are two conditions that must be satisfied in order for a collision to occur. Firstly, the particles must be moving towards each other and secondly, the particles must pass close enough to each other to collide. These conditions can be expressed mathematically in equations (3.3.2a) and (3.3.2b) respectively [16].

$$\mathbf{v}_{ij} \cdot \mathbf{r}_{ij} < 0 \tag{3.3.2a}$$

$$(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})^2 - v_{ij}^2(r_{ij}^2 - \sigma^2) \geq 0 \tag{3.3.2b}$$

The time to collision can then be calculated using the quadratic in equation (3.3.3). While there are two solutions, only the earliest collision (the negative root) needs to be considered. The second root gives the time when the particles leave after passing through each other which, for hard sphere, cannot happen.

$$\Delta t = \frac{(-\mathbf{v}_{ij} \cdot \mathbf{r}_{ij}) \pm \sqrt{(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})^2 - v_{ij}^2(r_{ij}^2 - \sigma^2)}}{v_{ij}^2} \tag{3.3.3}$$

Since $\mathbf{v}_{ij} \cdot \mathbf{r}_{ij}$ must be negative for the collision, there is the possibility of catastrophic cancellation [22], if $(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})^2 \gg v_{ij}^2(r_{ij}^2 - \sigma^2)$. Therefore it is advisable to use the positive root from the alternate form of the quadratic equation given in equation (3.3.4) [23].

$$\Delta t = \frac{r_{ij}^2 - \sigma^2}{(-\mathbf{v}_{ij} \cdot \mathbf{r}_{ij}) \mp \sqrt{(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})^2 - v_{ij}^2(r_{ij}^2 - \sigma^2)}} \tag{3.3.4}$$

When considering stepped potentials many of the same principles apply, except there are now two possible "collisions". The first, when two particles enter a step is treated identically to hard spheres. The other event, when the particles leave the step, is calculated using the second, later root of the quadratic. In order to prevent loss of numerical precision, the leaving time should be calculated using equation (3.3.5).

$$\Delta t = \begin{cases} \frac{r_{ij}^2 - \sigma^2}{(-\mathbf{v}_{ij} \cdot \mathbf{r}_{ij}) - \sqrt{(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})^2 - v_{ij}^2(r_{ij}^2 - \sigma^2)}}, & \text{if } \mathbf{v}_{ij} \cdot \mathbf{r}_{ij} > 0 \\ \\ \frac{(-\mathbf{v}_{ij} \cdot \mathbf{r}_{ij}) + \sqrt{(\mathbf{v}_{ij} \cdot \mathbf{r}_{ij})^2 - v_{ij}^2(r_{ij}^2 - \sigma^2)}}{v_{ij}^2}, & \text{if } \mathbf{v}_{ij} \cdot \mathbf{r}_{ij} < 0 \end{cases} \tag{3.3.5}$$

3.3.3 Collision Dynamics

Once the time of the next collision is known, the particles can be moved to their new locations, using Eq. (3.3.1). However, in order to predict future collisions the post-collision velocities of the colliding particles must also be calculated.

The simplest collision between two particles is an elastic bounce, where the velocities and just exchanged along the separation vector between the two particles. The change in velocity during the collision for particles i and j is shown in Eq. (3.3.6).

$$\Delta \mathbf{v}_i = -(\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij})\hat{\mathbf{r}}_{ij} \tag{3.3.6a}$$

$$\Delta \mathbf{v}_j = (\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij})\hat{\mathbf{r}}_{ij} \tag{3.3.6b}$$

For stepped potential system, the collision dynamics are more complex. When two particles collide they must pay an energy "cost" to proceed through the step. This energy cost $\Delta\Phi$ is the difference in the energy of the current step and the step the particles are going into, and is shown in Eq. (3.3.7).

$$\Delta\Phi = U_{\text{next step}} - U_{\text{current step}} \tag{3.3.7}$$

If the kinect energy of the particles is insufficient, the pair bounce off the step and the post-collision velocities are calculated using equation (3.3.6). However, if the particles can pay this cost i.e. the inequality (3.3.8) is true, then the particles can enter the step.

$$\frac{1}{4}m(\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij})^2 > \Delta\Phi \tag{3.3.8}$$

The change in the velocities of particles i and j after going through a step are shown in equation (3.3.9) where A is given in equation (3.3.10), the derivation of these equations is given in Appendix A. If the particles are entering a step, the positive root of A is used, whereas if the particles are leaving a step it is the negative root that should be used.

$$\Delta \mathbf{v}_i = \frac{A}{m}\hat{\mathbf{r}}_{ij} \tag{3.3.9a}$$

$$\Delta \mathbf{v}_j = -\frac{A}{m}\hat{\mathbf{r}}_{ij} \tag{3.3.9b}$$

$$A = -\frac{m}{2}\left((\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij}) \pm \sqrt{(\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij})^2 - \frac{4}{m}\Delta\Phi}\right) \tag{3.3.10}$$

It can be noticed that Eq. (3.3.6) can be derived from Eqs. (3.3.9) and (3.3.10) by setting $\Delta\Phi = 0$.

3.4 Force-driven Simulators

In order to compare the effectiveness of step potentials a set of comparison results from a continuous potential is needed. The method to simulate a continuous potential is given in this section.

Force-driven (or time driven) simulators are the most popular method of simulating particles due to their relative simplicity and ability to handle continuous potentials. Simulators of this kind were pioneered by Rahman [5] who predicted physical properties of liquid argon using a Lennard-Jones potential with reasonable accuracy.

The distinguishing feature between force-driven and event-driven simulators is the way in which they move through time. During force-based simulations particles' positions and velocities are calculated at uniform intervals of time, Δt using the forces acting on the particles. These newly calculated values are then used to predict the next set of particle positions. This is then repeated over the desired simulation time.

The general algorithm for a force driven simulator is as follows:

1. Initialisation
2. Calculate particles' future positions
3. Calculate the forces acting on the particles
4. Calculate the future velocities of particles
5. Run thermostat (if enabled)
6. Measure properties
7. Repeat steps 2-6 for the desired number of iterations

3.4.1 Initialisation

The particles are initialised in a Face Centered Cubic (FCC) structure. The use of the FCC lattice is common when simulating Lennard-Jones potentials as the first force-driven simulation [5] was carried out using liquid Argon which crystalises to a FCC lattice.

Particle velocities are assigned randomly from a Gaussian distribution with a mean, $\mu = 0$, and a standard deviation, $\sigma = \sqrt{T^*}$, where T^* is the desired reduced temperature. The velocities are then rescaled to ensure there is no net shift in linear momentum in any direction by applying (3.4.1) in each orthongonal direction.

$$v_i^{new} = v_i^{old} - \frac{1}{N}\sum_{i}^{N} v_i^{old} \tag{3.4.1}$$

In this disseration the Velocity Verlet integrator is used as all forces are conservative.

3.4.2 Periodic Boundary Conditions

When simulating a system it is necessary to have a boundary to prevent the particles from moving away from each other into infinity. However solid walls have a large effect on the properties of a system so an alternative method is needed. Since the earliest MD simulations [24] have used periodic boundary conditions to solve this problem.

The concept behind periodic boundary conditions is to create a pseudoinfinite system made up of tessellated images of the simulated system (as shown in figure 3.1). When a particle leaves the primary cell, its image enters from the opposite side. This ensures mass is conserved in each cell. This, however, brings the problem of particles being able to interact with multiple images of another particle, or even with themselves. The minimum image criterion is used to ensure particles only interact with the closest image of another particle.

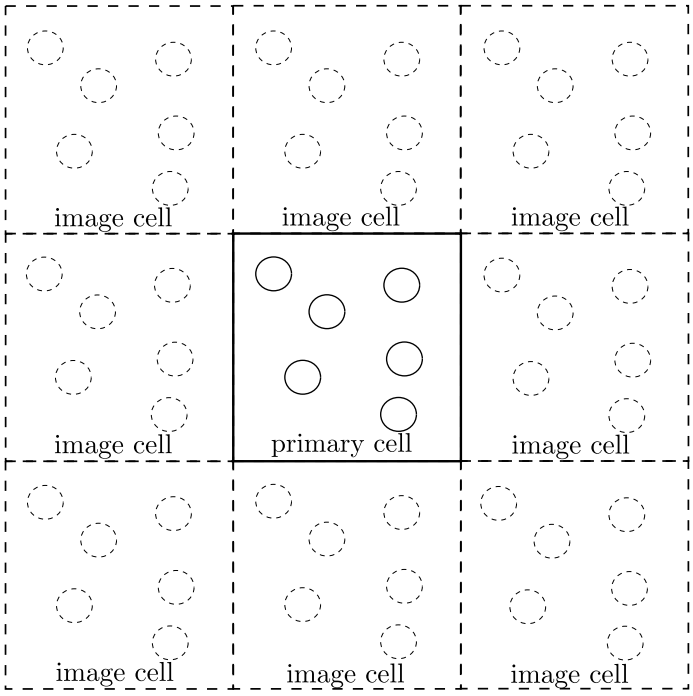

Figure 3.1: Figure showing 2D periodic boundary conditions. Only the nearest 8 images (dashed) to the simulated system (solid) are shown.

3.4.3 Thermostat

All simulations in this dissertation used the caniocal NVT ensemble, i.e. the number of particles, N; the volume of the system, V and the temperature, T were kept constant during the simulation. While keeping the number of particles and system size constant is simple, controlling the temperature is more complex.

The simplest method for controlling the temperature is to rescale the particle velocities to match the desired temperature using equation (3.4.2).

$$\mathbf{v}_{\text{new}} = \mathbf{v}_{\text{old}} \sqrt{\frac{T_{\text{desired}}}{T_{\text{current}}}} \tag{3.4.2}$$

However this does not allow energy fluctuations that should exist in a NVT system, therefore an Andersen thermostat[25] is used. An Andersen thermostat works by colliding a random particle with a ghost particle at the desired temperature. In this force-driven simulation, this is achieved by reassigning the velocities of 5% of particles from a Gaussian distribution at the correct temperature (similar to section 3.4.1).

3.4.4 Optimisation

Due to the time-consuming nature of computer simulations there are a number of techniques used to speed up simulations. Since the calculation of the forces on the particles is the most time-consuming part of force-driven simulations [2], almost all optimising techniques focus on this aspect.

The first technique to improve simulation speed is to truncate the potential. Since continuous potentials tend to zero as particles get further away from each other, significant time can be saved by selecting a cut-off radius at which the potential is taken to be zero. The form of a truncated potential is shown in equation (3.4.3). For the Lennard-Jones potential a cut-off radius of 3σ is used as $\Phi(3\sigma) = -0.00548\varepsilon$ and $F(3\sigma) = -0.0109\varepsilon/\sigma$ are both approximately 0.5% of the minimum values.

$$\Phi(r) = \begin{cases} \Phi(r) & \text{if } r \leq r_{\text{cut-off}} \\ 0 & \text{if } r > r_{\text{cut-off}} \end{cases} \tag{3.4.3}$$

While truncating the potential prevents the calculation of the forces it still requires the computation of the distance between particles. These extraneous calculations can be eliminated by using a neighbour list.

There are two main types of neighbour list used in molcular dynamics simulations. The first is the use of Verlet lists [17], and this is the type used in this disseration for the force-driven simulation. A Verlet list is a list of all the particles within a certain radius of a particle. If this list was updated every timestep, this would be no improvement on the original method, but by making the Verlet radius larger than the truncation radius, these list can be used for multiple timesteps. Haile [16] recommends using a Verlet radius of 3.3σ, and updating the list every 10 timesteps, and this is what was done in this dissertation.

Another method is to use cell-linked lists [23], this method involves dividing the system into a grid and only the particles in the same cell or a neighbouring cell are taken into account. This method can however be inefficient as the length of each cell must be at least the truncation cut-off radius wide to prevent particles being missed out. However

this means a volume of $27r^3_{\text{cut-off}}$ (as in 3 dimensions each cell has 26 neighbours) are considered but only particles within the spherical volume $\frac{4}{3}\pi r^3_{\text{cut-off}}$ should be checked; this means the volume checked is over six times larger than it needs to be.

Mattson and Rice [26] improve this by reducing the length of each cell to less than the cut-off radius, but this results in more neighbouring cells having to be checked i.e. for a cell length of $0.5r_{\text{cut-off}}$ the nearest 124 (a $5 \times 5 \times 5$ grid) cells are considered neighbours. This means there is a compromise as smaller cells mean that less volume is checked, but also means the cell lists are made obsolete quicker and therefore have to generated more frequently.

3.5 Event-Driven Simulators

Though force-driven simulators are more popular the first MD simulation was done using an event-driven simulator by Alder and Wainwright [10]. Event-driven simulators differ from force-driven simulators in that they move through time by jumping from the point when a discontinuity occurs (known as a collision) to the next discontinuity. These collisions are taken to be instantaneous and only one can occur at any particular time.

The general algorithm for a event driven simulator shown below.

1. Initialisation
2. Find the next event
3. Process the event
4. Update the event list
5. Measure properties
6. Repeat steps 2-5 for the desired number of events

CHAPTER

FOUR

MEASURING THERMODYNAMIC PROPERTIES

4.1 Introduction

Molecular dynamics is a useful tool to predict macroscopic properties of particles systems. Many of the identities and methods to measure these properties are derived in statistical mechanics. However, even when the system is at equilibrium, these properties fluctuate around a mean (see Figure 4.1) therefore it is common to take time averages of these values. These time averages are denoted with angle brackets $\langle\ \rangle$, and the time average of a property, A, is shown in (4.1.1) [16].

$$\langle A \rangle = \lim_{t \to \infty} \frac{1}{t} \int_{t_0}^{t_0+t} A(\tau) d\tau \tag{4.1.1}$$

This time average can be calculated precisely in event-driven simulations for several properties such as pressure, kinetic energy and potential energy. These only change at collisions and therefore are constant for the time between the collision and can be calculated by (4.1.2).

$$\langle A \rangle = \frac{1}{t} \sum_{t_0}^{t_o+t} A(\tau) \Delta\tau \tag{4.1.2}$$

In force-driven simulators all properties change continuously and hence time averages cannot be calculated precisely, however approximations can be made. If properties are measured every uniform period of time, the time average can be approximated by equation (4.1.3), where M is the number of measurements taken.

$$\langle A \rangle = \frac{1}{M} \sum_{1}^{M} A(\tau) \tag{4.1.3}$$

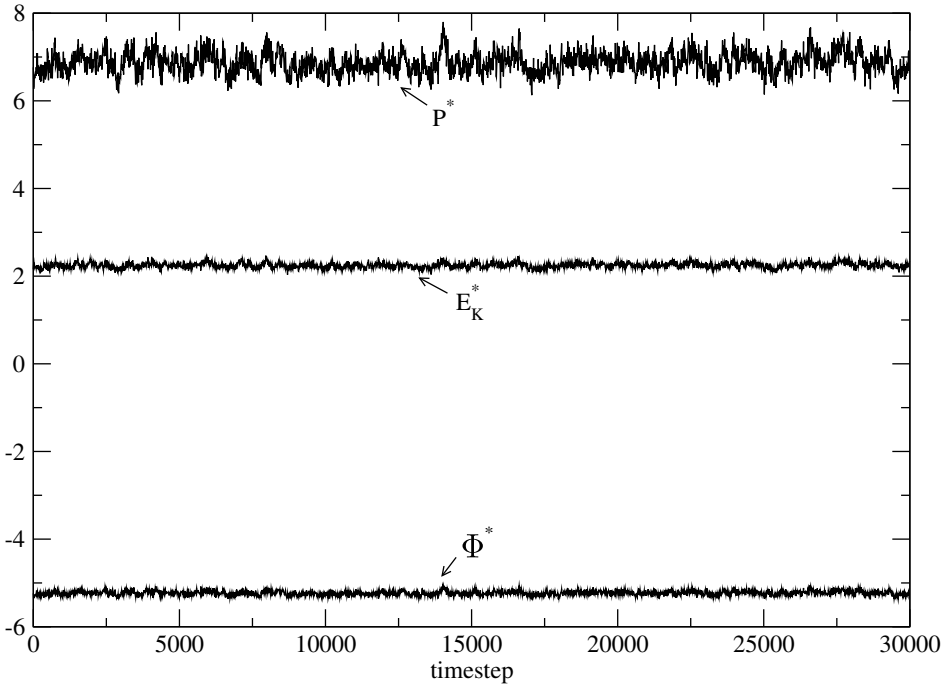

Figure 4.1: Plot showing flucuation of pressure ($P^* = P\sigma^3/\varepsilon$), kinetic energy per particle ($E_K^* = E_K/N\varepsilon$) and potential energy per particle ($\Phi^* = U/N\varepsilon$). Results are from a force-driven simulation involving 864 particles, at a density $\rho^* = 0.9$ and temperature $\langle T^* \rangle = 1.497$. Values were collected every 10 timesteps where each timestep was $\Delta t^* = 0.005$.

4.2 Units

In molecular dynamics simulations, properties are frequently measured in dimensionless forms [16]. These "reduced units" are usually denoted with an asterisk. In order to achieve this a number of fundamental dimensions are needed: a characteristic length σ, a characteristic energy ε, and the mass of one particle m. In the case of the Lennard-Jones potential, the characteristic length and energy are taken as: the distance of the root, and the depth of the attractive well respectively. A table of reduced forms are given in table 4.1.

4.3 Energy

Perhaps one of the most important properties to measure in MD simulations is the total internal energy of the system. For isolated systems, i.e. systems where mass or energy cannot enter or leave the system, this internal energy is the sum of kinetic and potential

Table 4.1: Table of reduced forms of various quanities used in this dissertation [16]

Quantity	Reduced forms
Density	$\rho^* = N\sigma^3/V$
Energy	$U^* = U/\varepsilon$
Force	$F^* = F\sigma/\varepsilon$
Length	$r^* = r/\sigma$
Pressure	$P^* = P\sigma^3/\varepsilon$
Temperature	$T^* = kT/\varepsilon$
Time	$t^* = t/(\sigma\sqrt{m/\varepsilon})$
Velocity	$v^* = v\sqrt{m/\varepsilon}$

energy (equation (4.3.1)).

$$U = E_K + \Phi \tag{4.3.1}$$

The total kinetic energy in the system is the sum of the kinetic energy of each particle, as shown in equation (4.3.2).

$$E_K = \sum_{i}^{N} m v_i^2 \tag{4.3.2}$$

The potential energy of the system is the sum of the potential energy between every pair of particles (for a pairwise potential), and is shown in equation (4.3.3).

$$\Phi = \sum_{i\ <\ j} \sum \Phi(r_{ij}) \tag{4.3.3}$$

Event-driven simulators strictly conserve energy, therefore the kinetic and potential energy can be measured at the beginning of the simulation and then updated whenever either changes e.g. when a collision occurs.

4.4 Temperature

The velocity distrubution of particles is given by the Maxwell distribution [16], shown in equation (4.4.1), where k_B is the Boltzman Constant.

$$f(v_x)dv_x = \sqrt{\frac{m}{2\pi k_B T}} e^{-\frac{m v_x^2}{2k_B T}} \tag{4.4.1}$$

This is the form of a Gaussian distribution and it can be shown [27] that the mean square velocity in any direction is as shown in equation (4.4.2).

$$\bar{v_x^2} = \frac{k_B T}{m} \tag{4.4.2}$$

Making the assumption that the velocity distribution is the same in each direction, the temperature can be expressed as equation (4.4.3), by taking the average temperature in each direction.

$$T^* = k_B T = \frac{mv^2}{3N} = \frac{2}{3N} E_K \tag{4.4.3}$$

This allows the calculation of the temperature from the kinetic energy of the system.

4.5 Pressure

The pressure in a molecular dynamics simulation is calculated using the virial equation of state (equation (4.5.1)) [27].

$$\frac{PV}{Nk_B T} = 1 + B_2\rho + B_3\rho^2 + B_4\rho^3 + ... \tag{4.5.1}$$

The coefficients, $B_2, B_3, B_4, ...$ are known as the second, third, fourth, etc. virial coeffcients. Values for these virial coefficients are available in the literature for common potentials such as hard spheres [28] or Lennard-Jones [29]. Physically these coefficients represent the contribution to the pressure by two, three, four, etc. particles interacting, and since most potentials are pairwise the pressure contribution is truncated at the second virial coefficient, the form of which in three dimensions is given in equation (4.5.2) [30].

$$B_2 = -2\pi \int_0^\infty \left(e^{\mathcal{U}(r)/kT} - 1\right) r^2 dr \tag{4.5.2}$$

However these virial coefficients do not apply to molecular dynamics simulations that use periodic boundary conditions [16], therefore an alternative method is required.

Using kinetic theory and measuring momentum flux during the simulation an expression for the second virial coefficient can be created for both simulation methods. In force driven simulations, the second virial can be calculated using equation (4.5.3) [16].

$$B_2\rho = \frac{1}{3NkT} \left\langle \sum_{i} \sum_{<j} \mathbf{F}_{ij} \cdot \mathbf{r}_{ij} \right\rangle \tag{4.5.3}$$

Calculating the pressure in event-driven simulators is more complex due to the lack of forces in the simulation, however by keeping track of the momentum flux at each collision an average pressure can be calculated using equation (4.5.4)[31]. Here N_{coll} is the total number of collisions during the time t.

$$B_2\rho = \frac{m}{3}\frac{N_{\text{coll}}}{Nt}\langle \mathbf{r}_{ij} \cdot \Delta\mathbf{v}_i \rangle_{\text{coll}} \tag{4.5.4}$$

4.6 Radial Distrution Function

One of the most important measurements in molecular dynamics is the radial distribution function (RDF) . The RDF provides information concerning the arrangement of the particles, and can be determined experimentally [32]. This means that the RDF can be used to determine the agreement of MD simulations and experimental results.

Statistically, the RDF is the ratio of the probability of finding a particle a distance r from another particle to the expected probability if the particles were randomly distributed [33]. Since potentials usually have a hard core there is very little chance of finding a particle within the radius of another particle, therefore the RDF is zero very close to the particle. However as the radial distance tends to infinity the probability of finding a particle tends to the randomly distributed probability so the RDF tends to one, which can be seen in Fig. 4.2.

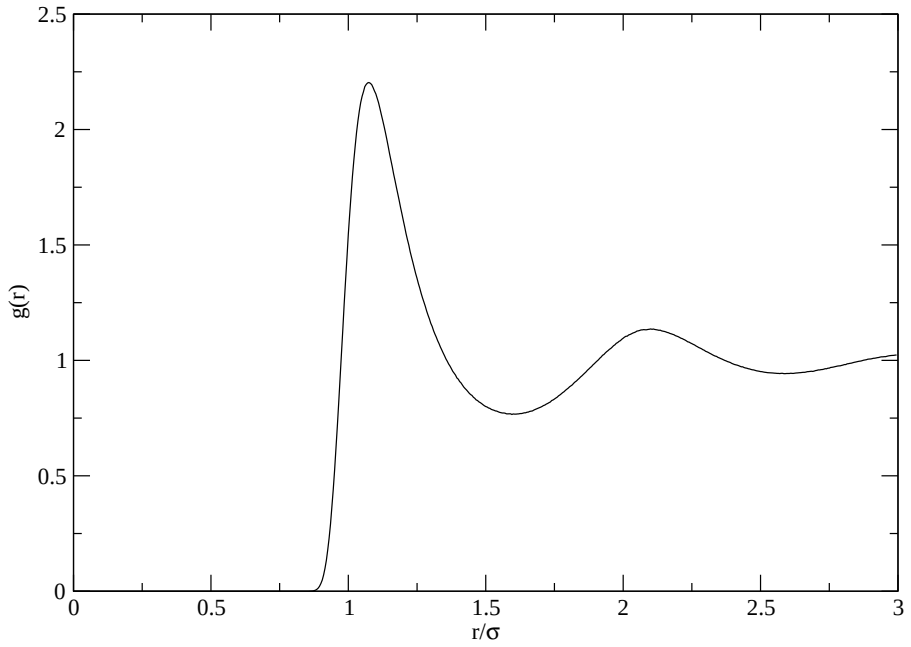

Figure 4.2: Plot of the radial distribution function, $g(r)$, for a continuous Lennard-Jones potential. Results taken from a force-driven simulation at $\rho^* = 0.7$ and $T^* = 1.5$

The RDF is measured in MD simulations by splitting the radial distance into a number of bins. The number of pairs of particles in a particular bin is then counted. The value of

the RDF for bin n can be calculated using the following equation:

$$g(n) = \frac{N_n}{NV_{\text{bin}}\rho} \tag{4.6.1}$$

where V_{bin} is the volume of the bin, and N_n is the number of particle pairs in bin n. This measurement illustrates another definition of the RDF. The radial distribution function is the ratio between the density at a specific distance from another particle to the average density of the system.

Since the RDF contains information about the potential, simulations with discrete potentials produce radial distribution functions with discontinuities in them. Figure 4.3 shows the RDF produced when using the stepped potential shown in Fig. 2.5.

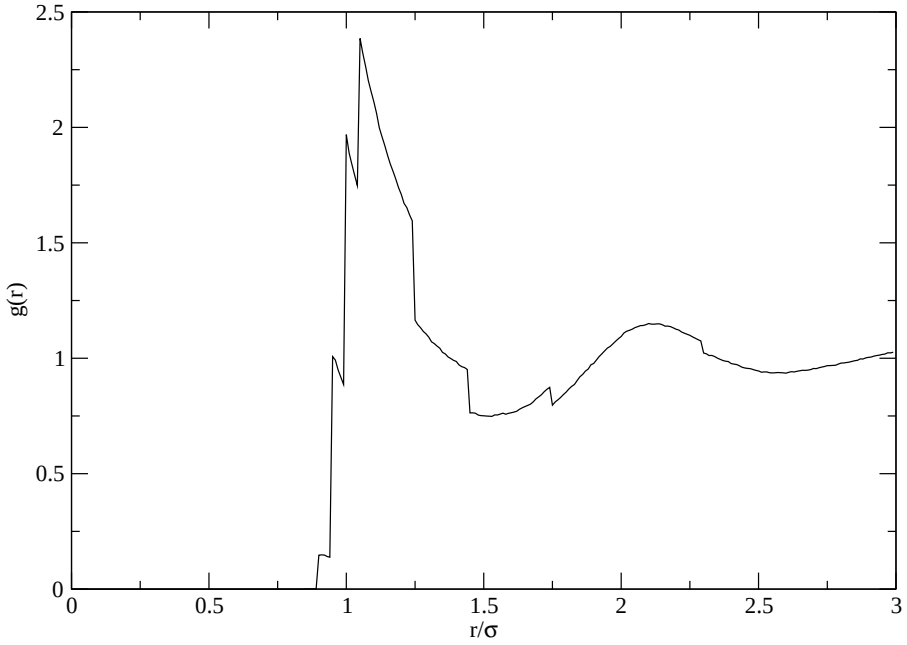

Figure 4.3: Plot of the radial distribution function, $g(r)$, for a discrete stepped potential. Results taken from a event-driven simulation at $\rho^* = 0.7$ and $T^* = 1.5$

This discontinuous RDF can be converted to a continuous function by using the indirect correlation function, $y(r)$, [12]. The indirect correlation function can be defined for both continuous and discontinuous RDF and potentials using:

$$y(r) = g_d(r)e^{\beta\Phi_d(r)} = g_c(r)e^{\beta\Phi_c(r)} \tag{4.6.2}$$

where $\beta = 1/k_BT$ and the subscripts c and d denote the continuous and discontinous functions respectively. The continous form of the RDF can then be calculated using

Eq. (4.6.3). Figure 4.4 shows the continuous form of the discontinuous RDF shown in Fig. 4.3.

$$g_c(r) = g_d(r)e^{[\Phi_d(r)-\Phi_c(r)]\beta} \tag{4.6.3}$$

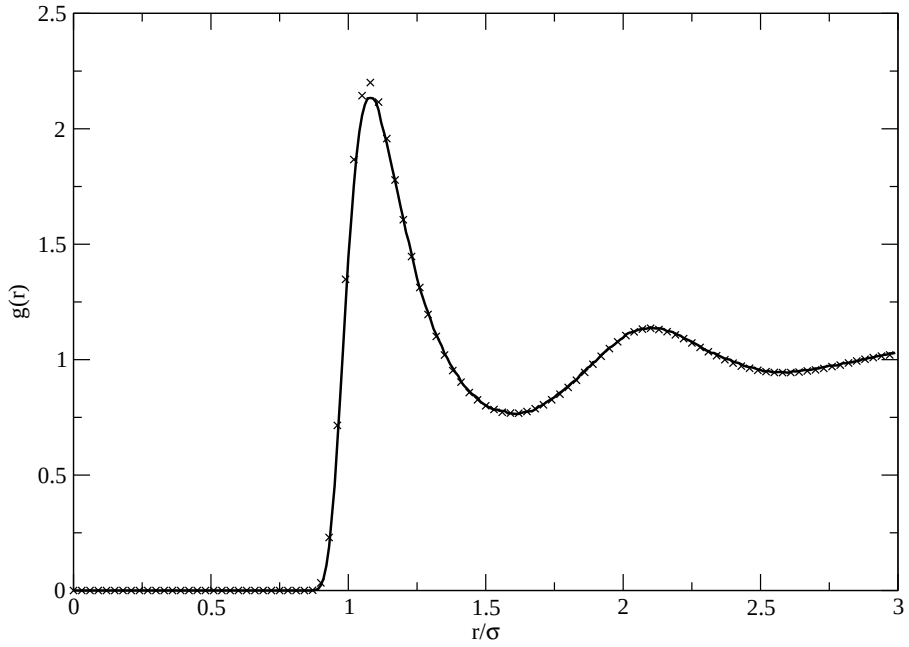

Figure 4.4: The continuous form of the radial distribution function of a discrete potential is shown with the solid line. The crosses show the RDF obtained from the continuous Lennard-Jones potential. Both were obtained at $\rho^* = 0.7$ and $T^* = 1.5$

4.7 Measuring Properties using the Radial Distribution Function

Another reason why the radial distribution functions is so important is that it can be used to calculate other thermodynamic properties. Any pair function can be related to the RDF [33], but the two most important are the potential energy and the pressure, which can be related to the RDF using Eqs. (4.7.1) and (4.7.2) respectively.

$$\Phi = 2\pi\rho N \int_0^\infty \Phi(r)g(r)r^2 dr \tag{4.7.1}$$

$$P = \rho k_B T + \frac{2\pi}{3}\rho^2 \int_0^\infty rF(r)g(r)r^2 dr \tag{4.7.2}$$

However, in MD simulations these properties are usually measured using the methods discussed in Sections 4.3 and 4.5 as these are usually more accurate.

4.8 Long-Range Corrections

In Section 3.4.4 the truncation of the potential to increase simulator performance was discussed. While particles far apart exert little force between each other, the contributions to the thermodynamic properties can still be significant. Therefore it is common in MD simulations to account for the truncated potential by adding on a long-range correction.

These long range corrections can be calculated using Eqs. (4.7.1) and (4.7.2). Since the RDF tends to one as separation distance tends to infinity, $g(r)$ is assumed to be one in these corrections. Similarly since the potential, $\Phi(r)$, and the force, $F(r)$, tend to zero as r increases, these are assumed to be zero. Making these assumptions and by splitting the integrals Eqs. (4.7.1) and (4.7.2) can be integrated to give equations:

$$\Phi = \Phi_{\text{MD}} + \frac{8\pi\rho}{3r_c^3}\left(\frac{1}{3r_c^6} - 1\right) \tag{4.8.1}$$

$$P = P_{\text{MD}} + \frac{16\pi\rho^2}{3r_c^3}\left(\frac{2}{3r_c^6} - 1\right) \tag{4.8.2}$$

where r_c is the truncation radius of the potential, and the subscript MD denotes the property measurement from the MD simulation. These give the long-range corrections for the potential energy and the pressure respectively.

CHAPTER

FIVE

FROM CONTINUOUS TO DISCONTINUOUS

State the case why we want to run discontinuous systems, the advantages (fast, extremely stable-> infinitely hard potentials, and disadvantages (underdeveloped set of potentials in the literature, complex algorithm). These disadvantages can be overcome by writing a good general EDMD program and finding a a general method to convert the hard work in soft potentials to stepped potentials.

CHAPTER

SIX

RESULTS

6.1 Benchmarking

6.1.1 Introduction

After a MD simulator has been created it is necessary to compare its results with those generated by others, to verify that the simulator works correctly.

6.1.2 Force based code verusus NIST and ESpReSSo

6.1.3 Event Driven

Event driven codes are significantly more complex than time-stepping codes, so we need more tests

HARD SPHERES v LEO

Stepped Potential of Chapela

6.2 Chapela's dumb stepping and very good stepping versus force based.

Show that dumb stepping doesn't work, show how good chapela's results actually are when he tries. Proves this is possible.

6.3 Stepping in probability versus action

Hard cores dominate the freezing behaviour (see Alder and Wainwrights famous paper) and therefore for high density pressure etc. but the stepping in equal probability doesnt do this.

Table 6.1: Comparison of results obtained by the event-driven simulator with literature values. t_{avg} is the average time between collisions, $\langle \hat{\mathbf{r}} \cdot \Delta \mathbf{v} \rangle_{coll}$ is the average momentum transfer per collision, and D is the coefficient of diffusion.

ρ	t_{avg}		$\langle \hat{\mathbf{r}} \cdot \Delta \mathbf{v} \rangle_{coll}$		D	
	Simulator	Lue	Simulator	Lue	Simulator	Lue
0.3	0.3052	0.3052	1.775	1.772	0.53	0.55
0.4	0.1944	0.1942	1.776	1.773	0.341	0.359
0.5	0.13024	0.13031	1.774	1.7724	0.247	0.247
0.6	0.08966	0.08968	1.771	1.7721	0.169	0.173
0.7	0.0625	0.0625	1.773	1.776	0.114	0.113
0.8	0.04365	0.0436	1.772	1.772	0.064	0.065
0.9	0.03029	0.03024	1.773	1.772	0.033	0.0327

6.4 Hard core position

Try setting the inner step to infinite energy

Use barker henderson, an old attempt to make hard sphere match everything else. (too far out).

talk about probability of finding a particle in the core. Talk about sigma try out 3, 4, and 5.

6.5 Temperature comparisons

Show that chapelas solution gets worse faster than ours.

6.5.1 Event-Driven Simulator

The event-driven simulator was first tested running a hard sphere simulation before testing the more complex stepped potentials. A single 'step' with a energy requirement sufficiently large such that no particle could enter it. The simulation was run once at a range of densities using 864 particles at a reduced temperature of $T^* = 1$ for 5 million collisions, the results were compared with those of Lue [31] in table 6.1. The agreement between results is good and lies within statistical uncertainty. The largest discrepancies are in the values for the coefficient of diffusion at low densities which is probably due to Lue's values were obtained after 10 million collisions.

The simulator was then benchmarked using a step potential. The results were compared with Chapela et al [12] using their 'Case 6' steps. The simulation was run for 1.5 million collisions using 864 particles. Each simulation was run ten times and the mean values and standard deviations are given in table 6.2

Table 6.2: Comparison of results obtained by the event-driven simulator with literature values using stepped potentials. Numbers in parenthesis indicate the uncertainty in the final digit.

ρ	$\langle T \rangle$		$\langle U \rangle$		$\langle P \rangle$	
	Simulator	Chapela et al	Simulator	Chapela et al	Simulator	Chapela etal
0.85	0.719(3)	0.72	-6.04(7)	-5.80	-0.5(4)	0.54
0.85	1.339(8)	1.34	-5.130(9)	-5.14	4.08(4)	4.08
0.85	2.35(1)	2.35	-4.24(2)	-4.20	8.78(9)	8.86
0.85	3.37(2)	3.37	-3.48(2)	-3.49	12.90(9)	13.00
0.85	4.59(1)	4.60	-2.67(1)	-2.68	17.31(8)	13.43
0.75	0.811(2)	0.81	-5.095(3)	-5.08	-0.20(2)	-0.24
0.75	1.309(9)	1.31	-4.67(1)	-4.63	1.81(5)	1.84
0.75	2.49(1)	2.49	-3.88(1)	-3.82	5.80(4)	5.95
0.75	3.59(2)	3.59	-3.26(1)	-3.22	9.03(7)	9.20
0.65	1.309(8)	1.31	-4.081(8)	-4.06	0.80(3)	0.81
0.65	2.61(1)	2.61	-3.42(1)	-3.41	3.86(5)	3.89
0.65	3.79(1)	3.79	-2.926(9)	-2.94	6.34(7)	6.33

6.6 References

[1] Jasper, A. W.; Nangia, S.; Zhu, C.; Truhlar, D. G. Non-Born - Oppenheimer Molecular Dynamics. *Acc. Chem. Res.* **2006**, *39*, 101–108, DOI: 10.1021/ar040206v.

[2] Frenkel, D.; Smit, B., *Understanding Molecular Simulations*, 2nd Edition; Academic Press: 2002, DOI: 10.1016/B978-012267351-1/50006-7.

[3] Tersoff, J. New empirical approach for the structure and energy of covalent systems. *Phys. Rev. B* **1988**, *37*, 6991–6999, DOI: 10.1103/PhysRevB.37.6991.

[4] Lennard-Jones, J. The Determination of Molecular Fields I & II. *Proc. R. Soc. Lon. Ser-A* **1924**, *106A*, 441–477, DOI: 10.1098/rspa.1924.0081.

[5] Rahman, A Correlations in the Motion of Atoms in Liquid Argon. *Phys. Rev.* **1964**, *136*, A405–A411, DOI: 10.1103/PhysRev.136.A405.

[6] Buckingham, R. A. The Classical Equation of State of Gaseous Helium, Neon and Argon. *Proc. R. Soc. Lon. Ser-A* **1938**, *168*, 264–283, DOI: 10.1098/rspa.1938.0173.

[7] White, D. N. J. A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 517–521, DOI: 10.1023/A:1007911511862.

[8] MacKerell, A. D. et al. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of Proteins. *J. Phys. Chem. B* **1998**, *102*, 3586–3616, DOI: 10.1021/jp973084f.

[9] Spoel, D. van der; Lindahl, E.; Hess, B.; Buuren, A. R van; Apol, E.; Meulenhoff, P.; Tieleman, D. P.; Sijbers, A. L. T. M.; Feenstra, K. A.; Drunen, R. van; Berendsen, H. J. C. Gromacs User Manual version 4.5.4., www.gromacs.org.

[10] Alder, B. J.; Wainwright, T. E. Phase Transition for a Hard Sphere System. *J. Chem. Phys.* **1957**, *27*, 1208–1209, DOI: 10.1063/1.1743957.

[11] Barker, J. A.; Henderson, D. Perturbation Theory and Equation of State for Fluids: The Square-Well Potential. *J. Chem. Phys.* **1967**, *47*, 2856–2861, DOI: 10.1063/1.1712308.

[12] Chapela, G.; Scriven, L. E.; Davis, H. T. Molecular dynamics for discontinuous potential. IV. Lennard-Jonesium. *J. Chem. Phys.* **1989**, *91*, 4307, DOI: 10.1063/1.456811.

[13] Elliot Jr., J. R. Optimized step potential models for n-alkanes and benzene. *Fluid Phase Equilib.* **2002**, *194-197*, 161–168, DOI: 10.1016/S0378-3812(01)00664-1.

[14] Unlu, O.; Gray, N. H.; Gerek, Z. N.; Elliott, J. R. Transferable Step Potentials for the Straight-Chain Alkanes, Alkenes, Alkynes, Ethers, and Alcohols. *Ind. Eng. Chem. Res.* **2004**, *43*, 1788–1793, DOI: 10.1021/ie034036m.

[15] Vahid, A.; Elliott Jr., J. R. Transferable intermolecular potentials for carboxylic acids and their phase behavior. *AIChE Journal* **2010**, *56*, 485–505, DOI: 10.1002/aic.11966.

[16] Haile, J. M., *Molecular Dynamics Simulation: Elementary Methods*; Wiley-Interscience: 1997.

[17] Verlet, L. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.* **1967**, *159*, 98–103, DOI: 10.1103/PhysRev.159.98.

[18] Gear, C. W., *Numerical Initial Value Problems in Ordinary Differential Equations*; Forsythe, G., Ed.; Prentice-Hall: 1971.

[19] Hairer, E.; Lubich, C.; Wanner, G. Geometric numerical integration illustrated by the Stormer-Verlet method. *Acta Numerica* **2003**, *12*, 399–450, DOI: 10.1017/S0962492902000144.

[20] Swope, W. C.; Andersen, H. C.; Berens, P. B.; Wilson, K. R. A computer simulation method for the calculation of equilibrium contants for the formation of physical clusters of molecultes: Application to small water clusters. *J. Chem. Phys.* **1982**, *76*, 637–649, DOI: 10.1063/1.442716.

[21] Khakimov, Z. New integrator for molecular dynamics simulations. *Comput. Phys. Commun.* **2002**, *147*, 733–736, DOI: 10.1016/S0010-4655(02)00387-9.

[22] Goldberg, D. What every computer scientist should know about floating-point arithmetic. *ACM Comput. Surv.* **1991**, *23*, 5–48, DOI: 10.1145/103162.103163.

[23] Pöschel, T.; Schwager, T., *Computational Granular Dynamics: Models and Algorithms*; Springer: 2005.

[24] Alder, B. J.; Wainwright, T. E. Studies in Molecular Dynamics. I. General Method. *J. Chem. Phys.* **1959**, *31*, 459–466, DOI: 10.1063/1.1730376.

[25] Andersen, H. C. Molecular dynamics simulations at constant pressure and/or temperature. *J. Chem. Phys.* **1980**, *72*, 2384–2393, DOI: 10.1063/1.439486.

[26] Mattson, W.; Rice, B. M. Near-neighbour calculations using a modified cell-linked list method. *Comput. Phys. Commun.* **1999**, *119*, 135–148, DOI: 10.1016/S0010-4655(98)00203-3.

[27] Landau, L.; Lifshitz, E., *Statistical Mechanics*, 2nd; Course of Theoretical Physics, Vol. 5; Pergamon: Oxford, 1968.

[28] Labík, S.; Kolafa, J.; Malijevský, A. Virial coefficients of hard spheres and hard disks up to the ninth. *Phys. Rev. E* **2005**, *71*, 021105, DOI: 10.1103/PhysRevE.71.021105.

[29] Schultz, A. J.; Kofke, D. A. Sixth, seventh and eighth virial coefficients of the Lennard-Jones model. *Mol. Phys.* **2009**, *1007*, 2309–2318, DOI: 10.1080/00268970903267053.

[30] Smith, J. M.; Van Ness, H. C.; Abbott, M., *Introduction ot Chemical Engineering Thermodynamics*, 7th; McGraw-Hill: New York, 2005.

[31] Lue, L. Collision statistics, thermodynamics, and transport coeffiecients of hard hyperspheres in three, four, and five dimensions. *J. Chem. Phys.* **2005**, *122*, 044513, DOI: 10.1063/1.1834498.

[32] Yarnell, J. L.; Katz, M. J.; Wenzel, R. G. Structure Factor and Radial Distribution Function for Liquid Argon at 85 °K. *Phys. Rev. A* **1973**, *7*, 2130–2144, DOI: 10.1103/PhysRevA.7.2130.

[33] Allen, M.; Tildesley, D. J., *Computer Simulation of Liquids*; Oxford University Press: 1987.

APPENDIX

A

DERIVATION OF COLLISION DYNAMICS FOR STEPPED POTENTIALS

Considering a collision between particles i and j, each with mass, m with a step energy difference of $\Delta\Phi$, the conservation of momentum is shown in equation (A.0.1). Here the prime indicates post-collision values.

$$m\mathbf{v}_i + m\mathbf{v}_j = m\mathbf{v}_i' + m\mathbf{v}_j' \tag{A.0.1}$$

The momentum change of each particle must occur along the separation vector between the two particles, which can be expressed by equation (A.0.2), where A is an arbitary coefficient.

$$m\mathbf{v}_i - m\mathbf{v}_i' = -(m\mathbf{v}_j - m\mathbf{v}_j') = -A\hat{\mathbf{r}}_{ij} \tag{A.0.2}$$

Energy must also be conserved in the system so equation (A.0.3) must also apply. This can be rewritten to equations (A.0.4) and (A.0.5)

$$\frac{1}{2}mv_i^2 + \frac{1}{2}mv_j^2 = \frac{1}{2}mv_i'^2 + \frac{1}{2}mv_j'^2 + \Delta\Phi \tag{A.0.3}$$

$$v_i^2 - v_i'^2 + v_j^2 - v_j'^2 - \frac{2}{m}\Delta\Phi = 0 \tag{A.0.4}$$

$$(\mathbf{v}_i - \mathbf{v}_i') \cdot (\mathbf{v}_i + \mathbf{v}_i') + (\mathbf{v}_j - \mathbf{v}_j') \cdot (\mathbf{v}_j + \mathbf{v}_j') - \frac{2}{m}\Delta\Phi = 0 \tag{A.0.5}$$

Equation (A.0.2) can now be substituted into (A.0.5) to give equation (A.0.6).

$$\frac{A}{m}\hat{\mathbf{r}}_{ij}(\mathbf{v}_j - \mathbf{v}_i + \mathbf{v}_j' - \mathbf{v}_i') - \frac{2}{m}\Delta\Phi = 0 \tag{A.0.6}$$

Equation (A.0.2) and the definition of the separation velocity vector ($\mathbf{v}_{ij} = \mathbf{v}_i - \mathbf{v}_j$) can be substituted into equation (A.0.6) to give (A.0.7).

$$-\frac{A^2}{m} - A\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij} - \Delta\Phi = 0 \tag{A.0.7}$$

This is a quadratic equation in terms of A therefore it's roots must be given by equation (A.0.8).

$$A = -\frac{m}{2}\left((\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij}) \pm \sqrt{(\mathbf{v}_{ij} \cdot \hat{\mathbf{r}}_{ij})^2 - \frac{4}{m}\Delta\Phi}\right) \tag{A.0.8}$$

From equation (A.0.2), the change in velocity of each particle is given in equations (A.0.9)

$$\Delta\mathbf{v}_i = \frac{A}{m}\hat{\mathbf{r}}_{ij} \tag{A.0.9a}$$

$$\Delta\mathbf{v}_j = -\frac{A}{m}\hat{\mathbf{r}}_{ij} \tag{A.0.9b}$$