

## CHAPTER 3

### MOLECULAR DYNAMICS

Molecular dynamics is a computational technique regularly used in condensed matter physics, chemistry, and related fields. It consists of the temporal evolution of a system of  $N$  particles interacting with each other by means of determined rules. In classical molecular dynamics, the evolution of a system is based on Newton's 2<sup>nd</sup> law of motion,  $\vec{F} = m\vec{a}$ , and the force is the gradient of a certain potential which is a function of all the particle coordinates. However, this method manipulates entirely at the microscopic level. Statistical interpretation is necessary for connection between the microscopic occurrence calculated by simulations and the macroscopic properties measured by means of experiment.

#### ***3.1 Statistical Mechanics Principles***

##### **3.1.1 Ensembles**

The essence of an ensemble of any system is the collection of microscopic states of the system in conjunction with a number of constraints. The microscopic states can be described in many ways. However, in this dissertation, only phase space will be considered. A point in the phase space represents an instantaneous state of the system. There are two famous ensembles in statistical mechanics, the microcanonical ensemble and canonical ensemble.

###### ***3.1.1.1 Microcanonical ensemble***

This is the ensemble of the system with the following constraints:

1. No particles are permitted to enter or leave the system (isolated system).

2. Size of the system is fixed.
3. Total energy of the system is constant.

The fundamental statistical postulate asserts that in an equilibrium the system, with the energy lying between  $E$  and  $E+\delta E$ , has equal probability,  $P(E)$ , to be found in any one of accessible states.

### 3.1.1.2 Canonical ensemble

This ensemble is that of a system with the following constraints.

1. No particles are permitted to enter or leave the system.
2. Size of the system is fixed.
3. The temperature of the system is kept constant.

To keep the temperature constant, it is necessary to bring the canonical system into contact with a heat reservoir. The system + reservoir is treated with a microcanonical ensemble. Defining  $\Omega(E)$  as the number of microscopic states of any system at energies between  $E$  and  $E + \delta E$ , the entropy of the system is defined as

$$S = k_B \ln \Omega(E), \quad (3.1)$$

and the temperature is defined by

$$\frac{1}{T} = \beta = \frac{\partial S}{\partial E} = k_B \left( \frac{\partial}{\partial E} \ln \Omega \right). \quad (3.2)$$

The probability to find the system in state  $i$  with energy  $E_i$  depends on the number of states available to the reservoir, and is given by between

$$P(E_i) = \frac{\exp(-\beta E_i)}{\sum_I \exp(-\beta E_i)}. \quad (3.3)$$

### 3.1.2 System properties

A macroscopic property system  $G$  can be observed via an average value over all states in the ensemble,

$$G_{observed} = \langle G_{state} \rangle, \quad (3.4)$$

or can be derived from those average values. In the molecular dynamics method, the system evolves from state to state over time. If we write  $\Gamma(t)$  to represent a point in phase space at time  $t$ , the instantaneous value can be written as  $G(\Gamma(t))$  and the average value along the observation time is [Lucus 1991]

$$\langle G \rangle = \langle G(\Gamma(t)) \rangle = \frac{1}{T_{observed}} \int_t^{t+T_{observed}} G(\Gamma(t)) dt. \quad (3.5)$$

Among the interesting macroscopic properties are the kinetic and potential energy. The kinetic energy of the system (K.E.) can be calculated by summing up the kinetic energy of individual particles:

$$\langle K.E. \rangle = \sum_i \frac{1}{2} m_i \langle v_i^2 \rangle. \quad (3.6)$$

The potential energy is also easy to calculate by summing up the interaction energy of all particles in the system.

Another interesting property is the temperature of the system. It can be calculated by the equipartition theorem,

$$T = \frac{2}{3} \frac{\langle K.E. \rangle}{N k_B}, \quad (3.7)$$

and the pressure can be calculated by

$$P = \frac{1}{3} \frac{N}{V} m \langle v^2 \rangle, \quad (3.8)$$

where  $N$  is number of particles in the system of interest of volume  $V$ .

### 3.2 Molecular Dynamics Procedure

The procedure used in this simulation can roughly be separated into 3 steps. The first one is the initialization, reading data from storage devices and preparing necessary constants. Initial velocities and positions of all particles in the system are assigned. In the next step, motion of the particles is carried out. Forces acting on each particle, due to interactions with its neighbors, are calculated. Succeeding positions and velocities are then calculated, and this operation is repeated for a large number of timesteps. The last step of the simulation is to calculate system properties and store all positions and velocities for the future investigations.

#### 3.2.1 The predictor-corrector algorithm

The finite difference approach is a standard method for solving systems of differential equations. A basic concept, called the predictor-corrector algorithm, supposes that we know the positions, velocities, etc. at time  $t$ . Then we can estimate positions, velocities, etc. at time  $t + \delta t$  from a Taylor series expansion around  $t$

$$\begin{aligned} \bar{r}^p(t + \delta t) &= \bar{r}(t) + \delta t \bar{v}(t) + \frac{1}{2} \delta t^2 \bar{a}(t) + \frac{1}{6} \delta t^3 \bar{b}(t) + \dots \\ \bar{v}^p(t + \delta t) &= \bar{v}(t) + \delta t \bar{a}(t) + \frac{1}{2} \delta t^2 \bar{b}(t) + \dots \\ \bar{a}^p(t + \delta t) &= \bar{a}(t) + \delta t \bar{b}(t) + \dots \\ \bar{b}^p(t + \delta t) &= \bar{b}(t) + \dots, \end{aligned} \quad (3.9)$$

where  $\bar{r}(t)$ ,  $\bar{v}(t)$ ,  $\bar{a}(t)$ , and  $\bar{b}(t)$  represent positions, velocities, accelerations and the first derivative of accelerations, respectively, at time  $t$ . The superscript  $p$  marks a *predicted* value. However, in the present step, these formulas cannot generate sufficiently precise trajectories as time evolves because the value of  $\bar{b}(t)$  is not given yet

and is thus initially set to zero. One way to improve the result is to determine  $\bar{b}(t)$  from  $\bar{a}(t)$  and  $\bar{a}(t+\delta t)$  using

$$\bar{b}(t) \approx \frac{\bar{a}(t+\delta t) - \bar{a}(t)}{\delta t}. \quad (3.10)$$

In the next step, the corrector step, the new  $\bar{b}(t)$  is used to recalculate the new set of  $\bar{r}(t+\delta t)$ ,  $\bar{v}(t+\delta t)$ ,  $\bar{a}(t+\delta t)$ , and  $\bar{b}(t+\delta t)$  using formulas (3.9). The corrector step may be iterated. Nevertheless, in our simulations only one corrector step was performed.

Another way to obtain a better trajectory is the method introduced by Gear [1966; 1971]. The corrected accelerations,  $\bar{a}^c(t+\delta t)$ , at predicted positions,  $\bar{r}^p(t+\delta t)$ , are calculated and brought to compare with predicted accelerations,  $\bar{a}^p(t+\delta t)$ , yielding the estimated size of the error in the prediction step:

$$\Delta\bar{a}(t+\delta t) = \bar{a}^c(t+\delta t) - \bar{a}^p(t+\delta t). \quad (3.11)$$

The error size and results from the predictor step are used in the correction step, e.g., by

$$\begin{aligned} \bar{r}^c(t+\delta t) &= \bar{r}^p(t+\delta t) + c_0 \Delta\bar{a}(t+\delta t) \\ \bar{v}^c(t+\delta t) &= \bar{v}^p(t+\delta t) + c_1 \Delta\bar{a}(t+\delta t) \\ \bar{a}^c(t+\delta t) &= \bar{a}^p(t+\delta t) + c_2 \Delta\bar{a}(t+\delta t) \\ \bar{b}^c(t+\delta t) &= \bar{b}^p(t+\delta t) + c_3 \Delta\bar{a}(t+\delta t). \end{aligned} \quad (3.12)$$

The values of corrector coefficients depend on the choice of process used in the calculation, such as the order of finite difference being used in the algorithm.

### 3.2.2 Boundary conditions

The number of particles in computer simulations is usually limited by the efficiency of the computer. Only a few thousand particles, compared to  $6 \times 10^{23}$  per mole of a real system, can be simulated in a tiny container. It can be supposed that a

large fraction of particles lies near the container walls. Those particles will experience quite different interactions from particles in the bulk. Nevertheless, this problem can be overcome by introducing periodic boundary conditions to simulate infinite systems [Born and Karman 1912]. The container, usually of cubic shape, is replicated throughout space to form an infinite lattice. The particles in the original container and its periodic images will progress in the same way in the course of simulations. When a particle leaves the container by crossing a boundary, there will be an image of that particle entering the container on the opposite boundary.

### 3.2.3 Cut-off limit

Ideally, the potential energy should be computed from an infinite summation of all interactions. However, in the case of short-range interactions, rapidly dropping to zero, we can neglect the interactions beyond a distance called the cut-off limit. Normally, it is never greater than half of the periodic box-length to include only the nearest images of distinguishable particles in interaction calculations.

### 3.2.4 Long-range interactions

Interactions over distances far beyond half of the periodic box length are called long-range interactions. Sometimes they cannot be ignored, and they make computer simulations more difficult. Two standard techniques for this problem are Ewald summation [Ewald 1921] and reaction field correction [Onsager 1936].

The Ewald summations handle all particles with their images as they are in a solid crystallized system. In case of coulombic interaction the potential energy of the system,  $V$ , can be written down as

$$V = \frac{1}{2} \sum_{\vec{n}} \left( \sum_{i,j} Z_i Z_j \frac{1}{|\vec{r}_{ij} + \vec{n}L|} \right), \quad (3.13)$$

where  $i, j$  represent all particles in the periodic box, and  $\bar{n}$  is a vector of integers representing the periodic images. This summation conditionally converges to a known limit for numerous sorts of long-range potentials.

The reaction field method, introduced without the assumption of periodicity, treats all particles beyond the cut-off sphere as forming a continuum with a given dielectric constant. Therefore, any charge lying inside the cut-off sphere will polarize the continuum and create a reaction field at the center.

### 3.2.5 Shifted and shifted-force potentials

The truncation of the potential at a cut-off introduces some difficulties in defining a consistent potential and force, since the potential function,  $V(r_{ij})$ , contains a discontinuity at  $r_{ij} = r_{cutoff}$ . The problem may be avoided by shifting the potential function by an amount  $V_c = V(r_c)$ . However, the force at  $r_{ij} = r_{cutoff}$  is still discontinuous. For conservation of energy, the shift-force potential method [Streett, Tildesley, and Saville 1978] is introduced, in which

$$V_{shift-force}(r_{ij}) = \begin{cases} V(r_{ij}) - V_c - V'(r_c)(r_{ij} - r_c) & r_{ij} \leq r_c \\ 0 & r_{ij} > r_c \end{cases} \quad (3.14)$$

### 3.2.6 Calculation of macroscopic properties

#### 3.2.6.1 Structural properties

Liquid structure cannot be written as an exact arrangement in space as in the case of crystalline solids. It is usually be described in terms of comparative arrangements of atoms with other atoms. One essential measure is the radial distribution function (RDF), represented by  $g_{ij}(r)$ . This function presents the probability per unit volume of finding particle  $j$  at distance  $r$  from particle  $i$ . The RDF is normalized to be 1 for a completely random distribution. Another quantity related to

the RDF is the running integration number, represented by  $n_{ij}(r)$ , easily calculated from

$$n_{ij}(r) = \rho \int_0^r g_{ij}(r) 4\pi r^2 dr, \quad (3.15)$$

where

$$\rho = \text{number density} = N/V.$$

An average coordination number is defined as the value of  $n_{ij}(r)$  at the first minimum of  $g_{ij}(r)$  following the first peak. Any particles at distances less than this first minimum belong to the first solvation shell.

### 3.2.6.2 Dynamical properties

The most common way to present dynamical properties of the system is to formulate via the time-correlation function,

$$C_{AB}(t) = \frac{\langle A(0)B(t) \rangle}{\langle A(0)B(0) \rangle}, \quad (3.16)$$

where  $\langle A(0)B(t) \rangle$  can be calculated using the averaging formula,

$$\langle A(0)B(t) \rangle = \frac{1}{t_{\max}} \sum_{t_0=1}^{t_{\max}} A(t_0)B(t_0+t). \quad (3.17)$$

Another meaningful dynamical property is the time-autocorrelation function, defined as

$$C_{AA}(t) = \frac{\langle A(0)A(t) \rangle}{\langle A(0)^2 \rangle}, \quad (3.18)$$

and yet another is the self-diffusion coefficient,  $D$ , which is related to the velocity autocorrelation function by the Green-Kubo relation [Frenkel and Smit 1996],

$$D = \frac{1}{3} \int C_{vv}(t) dt. \quad (3.19)$$



The elementary Stokes-Einstein theory of diffusion of liquid [Frenkel 1946] shows that the relation of the self-diffusion coefficient,  $D$ , and viscosity,  $\nu$ , is

$$D = \frac{1}{6\pi a} \frac{k_B T}{\nu}, \quad (3.20)$$

where  $a$  is the radius of fluid particle,  $k_B$  is Boltzmann's constant, and  $T$  is the absolute temperature.