Soft Matter Simulation

2012

Thomas Rodgers

Contents

List of Figures

List of Tables

Nomenclature

Roman

$\mathbf{a}$	Acceleration	$\mathrm{m\,s^{-2}}$
A	Hamaker constant	
A	Helmholtz energy	$\mathrm{J\,mol^{-1}}$
a	Repulsion parameter	–
a_0	Effective cross sectional area per head group	$\mathrm{nm^2}$
C	Harmonic spring constant	$\mathrm{m^2\,kg\,s^{-2}}$
C_p	Constant pressure heat capacity	$\mathrm{J\,mol^{-1}\,K^{-1}}$
C_v	Constant volume heat capacity	$\mathrm{J\,mol^{-1}\,K^{-1}}$
DP	Degree of polymerisation	–
E	Energy	$\mathrm{J\,mol^{-1}}$
$\mathbf{F}$	Force	N
$\mathbf{F}^C$	Conservative force	N
$\mathbf{F}^{C,B}$	Bonded conservative force	N
$\mathbf{F}^{C,NB}$	Non-bonded conservative force	N
$\mathbf{F}^D$	Dissipative force	N
$\mathbf{F}^R$	Random force force	N
G	Gibbs free energy	$\mathrm{J\,mol^{-1}}$
$g(r)$	Radial distribution function	–
H	Enthalpy	$\mathrm{J\,mol^{-1}}$
K	Equilibrium constant	–
k^{AB}	Harmonic spring constant	$\mathrm{J\,mol^{-1}\,m^{-2}}$
k^{ABC}	Bending spring constant	$\mathrm{J\,mol^{-1}}$
k^B	Improper spring constant	$\mathrm{J\,mol^{-1}}$
k_B	Boltzmann constant	$1.38 \times 10^{-23}\ \mathrm{m^2\,kg\,s^{-2}\,K}$
KE	Kinetic energy	J
l_c	Length of the hydrophobic chain	nm
m	Mass	kg
$\bar{M}_n$	Number average molar mass	–
$\bar{M}_w$	Weight average molar mass	–
$\bar{M}_z$	Polydispersity index	–
N	Number of particles	–
n	Number density	–
N_0	Aggregation number	–
n_c	Total number of carbons per hydrocarbon chain	–
n_{Me}	Total number of methyl groups per hydrocarbon chain	–
N_s	Packing parameter	–
$\mathbf{p}$	Momentum	$\mathrm{kg\,m\,s^{-1}}$

P	Pressure	Pa
Q	Charge	C
Q	Partition function	–
$\mathbf{r}$	Position	
$\mathbf{r}_{ij}$	Distance between beads i and j	m
R	Ideal gas constant	$8.314\,\mathrm{J\,mol^{-1}\,K^{-1}}$
r	Colloids radius	m
r_0	Harmonic spring equlibrium length	m
R_g^2	Radius of Gyration	m^2
r_c	Interaction cut-off	m
S	Entropy	$\mathrm{J\,mol^{-1}\,K^{-1}}$
T	Temperature	K
t	Time	s
T_g	Glass transition temperature	K
T_m	Melting temperature	K
$\mathbf{u}$	Velocity	$\mathrm{m\,s^{-1}}$
U	Internal energy	$\mathrm{J\,mol^{-1}}$
$\mathbf{V}$	Potential energy	J
$\mathbf{v}$	Velocity	$\mathrm{m\,s^{-1}}$
$\mathbf{v}_{ij}$	Velocity difference between beads i and j	$\mathrm{m\,s^{-1}}$
V	Volume	m^3
v	Volume of hydrophobic core	nm^3
V_n	Torsional constant	$\mathrm{J\,mol^{-1}}$
W	Probability density function	–
x	Distance between colloids	m
Z	Partition function	–
Greek		
α	Improper angle	–
$(\chi N)_{eff}$	Effective Flory-Huggins parameter	–
ϵ	Lennard-Jones well depth	$\mathrm{J\,mol^{-1}}$
γ	Friction coefficient	–
κ^{-1}	Dimensionless compressibility	–
κ_T	Isothermal compressibility	–
Λ	Thermal de Broglie wavelength	m
λ	Temperature scaling factor	–
μ	Chemical potential	$\mathrm{J\,mol^{-1}}$
μ	Pressure scaling factor	–
ϕ	Volume fraction	–
φ	Improper angle	–
Ψ	Transition probability	–
ρ	Density	$\mathrm{kg\,m^{-3}}$
σ	Lennard-Jones distance	m
σ	noise amplitude	–
θ	Bond angle	–
θ	Nematic angle	r
ξ	Gaussian random number	
Acronyms		
CMC	Critical micelle concentration	

Chapter 1

Review of Soft Matter Systems

Contents

1.1 Introduction

Soft matter comprises a variety of mater in different physical states that are easily deformed by thermal stresses or thermal fluctuations, Figure 1.1. All soft matter materials share an important common feature, this is that the predominant physical behaviors occur at an energy scale comparable with room temperature thermal energy. Soft matter systems include liquids, colloids, polymers, foams, gels, granular materials, and a number of biological materials.

Soft matter systems often have interesting behaviours that are difficult to predict as they arise directly from its atomic or molecular constituents. This is because soft matter often self-organises into mesoscopic physical structures; these are much larger than the microscopic scale (the arrangement of atoms and molecules), and much smaller than the macroscopic (overall) scale of the material. The properties and interactions of these mesoscopic structures determine the macroscopic behaviour of the material. For example, the bubbles that comprise a foam are mesoscopic because they individually consist of a vast number of molecules, and yet the foam itself consists of a great number of these bubbles, and the overall mechanical stiffness of the foam emerges from the combined interactions of the bubbles, Figure 1.2. By way of contrast, in hard condensed matter physics it is often possible to predict the overall behaviour of a material because the molecules are organised into a crystalline lattice with no changes in the pattern at any mesoscopic scale.

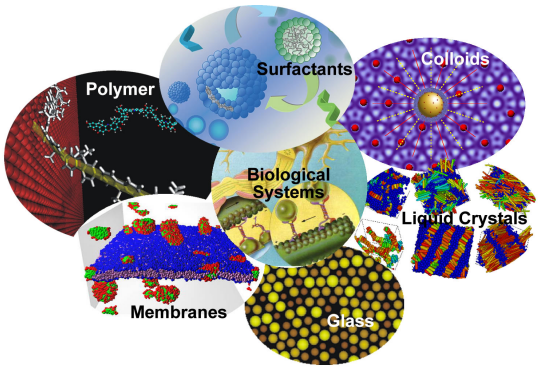


Figure 1.1: Examples of soft matter systems.

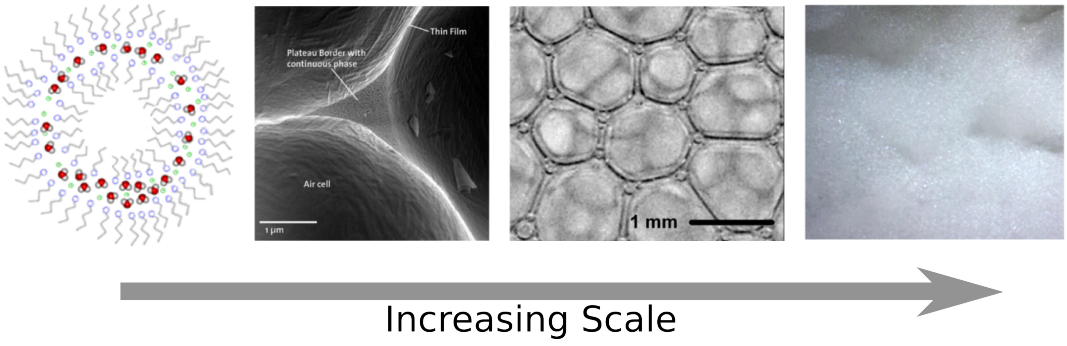


Figure 1.2: Range of scales in foam, from the molecular arrangement (left) to the macroscale structure (right).

Soft materials are important in a wide range of technological applications. They may appear as structural and packaging materials, foams and adhesives, detergents and cosmetics, paints, food additives, lubricants and fuel additives, rubber in tires, etc. In addition, a number of biological materials (blood, muscle, milk, yogurt) are classifiable as soft matter. Liquid crystals, another category of soft matter, exhibit a responsivity to electric fields that make them very important as materials in display devices (LCDs). In spite of the various forms of these materials, many of their properties have common physicochemical origins, such as a large number of internal degrees of freedom, weak interactions between structural elements, and a delicate balance between entropic and enthalpic contributions to the free energy. These properties lead to large thermal fluctuations, a wide variety of

forms, sensitivity of equilibrium structures to external conditions, macroscopic softness, and metastable states.

1.2 Colloids

A colloid is a substance microscopically dispersed evenly throughout another substance. A colloidal system consists of two separate phases: a dispersed phase and a continuous phase in which the colloid is dispersed. A colloidal system may be solid, liquid, or gas[1]. The dispersed-phase particles typically have a diameter of between approximately 1 and 1000 nanometers; however, liquid-liquid systems can have larger dispersed-phase droplets (especially if stabilised). The dispersed-phase particles or droplets are affected largely by the surface chemistry present in the colloid.

1.2.1 Classification

Because the size of the dispersed phase may be difficult to measure, and because colloids have the appearance of solutions, colloids are sometimes identified and characterized by their physico-chemical and transport properties. Colloids can be classified as in Figure 1.1.

Table 1.1: Classification of colloid systems.

Medium		Dispersed Medium		
		Gas	Liquid	Solid
Continuous Medium	Gas	**None** (all gases are mutually miscible)	**Liquid aerosol** fog, mist, hair spray	**Solid aerosol** smoke, cloud, inhalers
	Liquid	**Foam** shaving cream, head on beer	**Emulsion** milk, mayonnaise, shampoo	**Sol** paint, ink, blood
	Solid	**Solid foam** aerogel, pumice, styrofoam	**Gel** jelly, agar	**Solid sol** pearl

1.2.2 Interactions Between Colloid Particles

There are several forces that play an important role in the interaction of colloid particles:

van der Waals forces There is an attractive force between any pair of atoms or molecules, even when the atoms are uncharged or have no dipole moment. The origin of this force is quantum mechanical arising from the interaction between fluctuating dipoles in each of the atoms. The van der Waals potential actually depends on the

[1] Emulsion is often used interchangeably with colloid, but an emulsion specifically refers a liquid-liquid system

shape of the colloid particles, equation 1.2.1, where A is the Hamaker constant.

$$V(x) = -\begin{cases} \dfrac{Ar}{12x} & \text{Spherical particles of size } r \\ \dfrac{A}{12\pi x^2} & \text{Plate-like particles} \end{cases} \tag{1.2.1}$$

Electrostatic interaction Colloidal particles often carry an electrical charge and therefore attract or repel each other. The charge of both the continuous and the dispersed phase, as well as the mobility of the phases are factors affecting this interaction. Dissolved ions interact with the colloids and modify the nature of the electrostatic interactions. In particular, the electrostatic interactions are screened by dissolved ions.

Entropic forces According to the second law of thermodynamics, a system progresses to a state in which entropy is maximized. This can result in effective forces even between hard spheres.

Excluded volume repulsion It is impossible for two hard particles to overlap each other. This means that there is a finite size of the system.

Steric forces Polymer-covered surfaces or solutions containing non-adsorbing polymer can modulate interparticle forces, producing an additional steric repulsive force (which is predominantly entropic in origin) or an attractive depletion force between them.

1.3 Polymers

A polymer is a large molecule (macromolecule) composed of repeating structural units. These sub-units are typically connected by covalent chemical bonds. Although the term polymer is sometimes taken to refer to plastics, it actually encompasses a large class of compounds comprising both natural and synthetic materials with a wide variety of properties. Because of the extraordinary range of properties of polymeric materials they are widely used. Their roles range from synthetic plastics and elastomers to natural biopolymers such as nucleic acids and proteins that are essential for life. Natural polymeric materials such as shellac, amber, wool, silk and natural rubber have been used for centuries. A variety of other natural polymers exist, such as cellulose, which is the main constituent of wood and paper.

Most commonly, the continuously linked backbone of a polymer used for the preparation of plastics consists mainly of carbon atoms. A simple example is polyethene, whose repeating unit is based on ethene monomer. However, other structures do exist; for example, elements such as silicon form familiar materials such as silicones, examples being Silly Putty and waterproof plumbing sealant. Oxygen is also commonly present in polymer backbones, such as those of polyethylene glycol, polysaccharides (in glycosidic bonds), and DNA (in phosphodiester bonds).

1.3.1 Polymer Properties

Polymer properties are broadly divided into several classes based on the scale at which the property is defined as well as upon its physical basis. The most basic property of a polymer is the identity of its constituent monomers. A second set of properties, known as

microstructure, essentially describe the arrangement of these monomers within the polymer at the scale of a single chain. These basic structural properties play a major role in determining bulk physical properties of the polymer, which describe how the polymer behaves as a continuous macroscopic material. Chemical properties, at the nano-scale, describe how the chains interact through various physical forces. At the macro-scale, they describe how the bulk polymer interacts with other chemicals and solvents. The physical properties of a polymer are strongly dependent on the size or length of the polymer chain. For example, as chain length is increased, melting and boiling temperatures increase quickly. This is a result of the increase in chain interactions such as Van der Waals attractions and entanglements that come with increased chain length. Some of the main parameters used are:

Degree of Polymerisation is given by equation 1.3.1.

$$\text{DP} = \frac{\text{Molar Mass of Polymer}}{\text{Molar Mass of Repeat Unit}} \tag{1.3.1}$$

Number Average Molar Mass is given by equation 1.3.2.

$$\bar{M}_n = \frac{\sum_i N_i M_i}{\sum_i N_i} \tag{1.3.2}$$

Weight Average Molar Mass is given by equation 1.3.3.

$$\bar{M}_w = \frac{\sum_i N_i M_i^2}{\sum_i N_i M_i} \tag{1.3.3}$$

Polydispersity Index is given by equation 1.3.4 and for polydispersed polymers $n < w < z$ while as uniform polymers have $n = w = z$.

$$\text{PDI} = \bar{M}_z = \frac{\bar{M}_w}{\bar{M}_n} \tag{1.3.4}$$

Radius of Gyration describes the dimensions of a polymer chain as in equation 1.3.5, and is the average distance from the center of mass of the chain to the chain itself.

$$R_g^2 = \frac{1}{N} \sum_{k=1}^{N} \left(\mathbf{r}_k - \mathbf{r}_{\text{mean}}\right)^2 \tag{1.3.5}$$

1.3.2 Phase behaviour

Crystallinity

When applied to polymers, the term crystalline has a somewhat ambiguous usage. A synthetic polymer may be loosely described as crystalline if it contains regions of three-dimensional ordering on atomic (rather than macromolecular) length scales, usually arising from intramolecular folding and/or stacking of adjacent chains. Synthetic polymers may consist of both crystalline and amorphous regions; the degree of crystallinity may be expressed in terms of a weight fraction or volume fraction of crystalline material. Few synthetic polymers are entirely crystalline. Polymers with a degree of crystallinity approaching zero or one will tend to be transparent, while polymers with intermediate degrees of crystallinity will tend to be opaque due to light scattering by crystalline or glassy regions.

Melting point

The term melting point, when applied to polymers, is not a solid-liquid phase transition but a transition from a crystalline or semi-crystalline phase to a solid amorphous phase, abbreviated as T_m. Some polymers will actually decompose at high temperatures rather than melt.

Glass transition temperature

A parameter of particular interest in synthetic polymer manufacturing is the glass transition temperature (T_g), which describes the temperature at which amorphous polymers undergo a transition from a rubbery, viscous amorphous liquid, to a brittle, glassy amorphous solid. The glass transition temperature may be engineered by altering the degree of branching or crosslinking in the polymer or by the addition of plasticizer.

1.3.3 Polymer Mixtures

In general, polymeric mixtures are far less miscible than mixtures of small molecule materials. This effect results from the fact that the driving force for mixing is usually entropy, not interaction energy (enthalpic). In other words, miscible materials usually form a solution not because their interaction with each other is more favorable than their self-interaction, but because of an increase in entropy and hence free energy associated with increasing the amount of volume available to each component. This increase in entropy scales with the number of particles (or moles) being mixed. Since polymeric molecules are much larger and hence generally have much higher specific volumes than small molecules, the number of molecules involved in a polymeric mixture is far smaller than the number in a small molecule mixture of equal volume.

The energetics of mixing, on the other hand, is comparable on a per volume basis for polymeric and small molecule mixtures. This tends to increase the free energy of mixing for polymer solutions and thus make solvation less favorable. Thus, concentrated solutions of polymers are far rarer than those of small molecules. Furthermore, the phase behaviour of polymer solutions and mixtures is more complex than that of small molecule mixtures. Whereas most small molecule solutions exhibit only an upper critical solution temperature phase transition, at which phase separation occurs with cooling, polymer mixtures commonly exhibit a lower critical solution temperature phase transition, at which phase separation occurs with heating.

Block co-polymers (e.g. AAAAABBBBBAAAAA-) can also self-assemble into separated phase structures such that the different types of monomer units are separated from each other similar to surfactant liquid crystals, Figure 1.3.

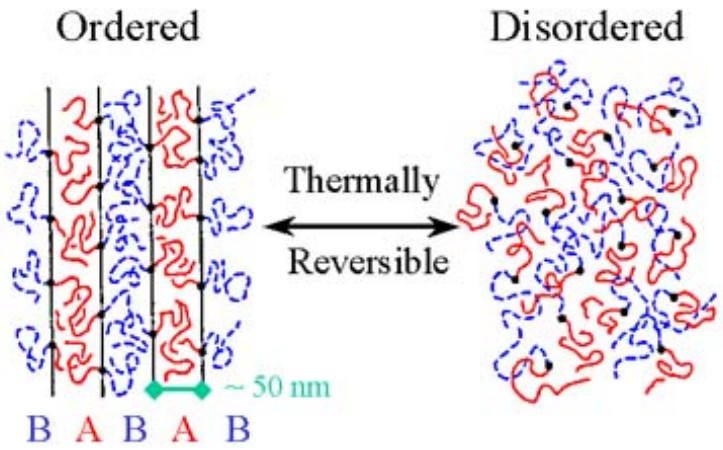


Figure 1.3: Block co-polymer ordering into lamellar phase on cooling.

1.4 Surfactants

The term surfactant comes from a catenation of surface active agent first coined by Antara Products in 1950 [1]. Surfactants have been used for many years in a wide range of everyday applications such as cleaning products, food, pharmaceuticals and paints to name but a few [2]. Surfactants even occur in nature forming biological structural systems [3]. Their functionality is due to their molecular structure in that they have parts soluble in a specific fluid (lyophilic) and parts insoluble in that same fluid (lyophobic).

In the case of water, as a solvent, they have a hydrophilic head group, which is water soluble, attached to a hydrophobic tail group, which is insoluble in water. This structure drives the formation of self-assembled aggregates by trying to keep the lyophobic parts away from the fluid they "dislike". A simplified surfactant molecule is shown by Figure 1.4. Most surfactants are chain-like molecules, although many colloidal particles can display amphiphilic behaviour.

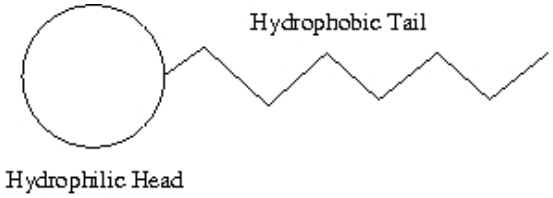


Figure 1.4: Schematic of simplified surfactant molecule.

1.4.1 Surfactant Classification

The hydrophobic tails of surfactants are generally limited to hydrocarbon, perfluorocarbon, and polydimethylsiloxane chains [3]. The hydrophilic head group is usually shorter and bulkier than the tail and its polarity is used to classify the surfactant. Surfactants are classified into four groups:

Anionic surfactants have a head group that is negatively charged with a positively charged counter-ion.

Cationic surfactants have a head group that is positively charged with a negatively charged counter-ion.

Zwitterionic surfactants have a head group that contain both a positively and negatively charged ion; these often comprise of a base connected to an acid.

Non-ionic surfactants have no overall charge and no ions, the hydrophilic part of the molecule is produced from polar species, commonly poly(ethylene oxide) groups are used.

These classifications accommodate the main types of surfactants, but are not exclusive: a surfactant can have a hydrocarbon tail, a poly(ethylene oxide) section, and a charged head group.

1.4.2 Surfactant Self-Assembly

As previously mentioned, when in solution surfactant molecules have a tendency to self assemble into different structures depending on the concentration, temperature, and surfactant structure [4]. At very low concentrations, surfactant molecules are solubilized as unimers. In a dilute solution the surfactants will self-assemble into aggregates such as micelles and cylinders. The form that gives the minimum free energy for the given set of conditions will be the optimum aggregate. The free energy can be said to be mainly made up of three terms:

Hydrophobic contribution which is favourable due to the hydrocarbon chains sequestering themselves within the interior of the aggregates. This reduces the enthalpy as it reduces the hydrocarbon-solvent repulsive interaction energy. This is the main driving force for surfactant self-assembly.

Surface term reflecting the resistance of the head groups packing together and their attraction the fluid. This increases the enthalpy as it increases the head-head repulsive interaction energy. This is the main force that limits micellar size.

Packing term reflecting the exclusion of fluid from the aggregate interior and the limited geometric structures available. This decreases the entropy as the free chains with many degree of freedom are packed together.

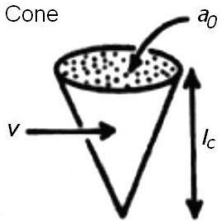
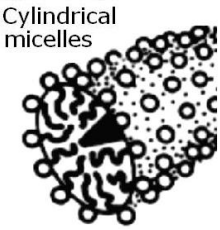
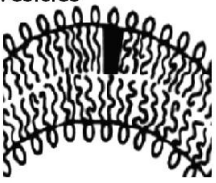
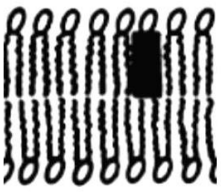
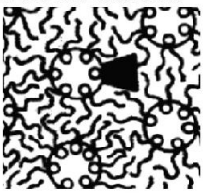
For these dilute systems the shape of the aggregates can be predicted using a packing parameter [5], surface curvature [6], or molecular simulations. The packing parameter, N_s, can be given by equation 1.4.1, where v is the volume of the hydrophobic chain, l_c is the length of the hydrophobic chains and, a_0 is the effective area per head group.

$$N_s = \frac{v}{l_c a_0} \tag{1.4.1}$$

The value of the packing parameter relates to the structure of the aggregate as shown in Table 1.2. This is only an approximation; each surfactant will behave differently depending on the conditions it is subjected to, and concentration of the surfactant itself, therefore a full phase diagram should be used to determine the aggregate structures if possible. Mixtures of the aggregates can exist, as well as intermediate structures between those shown in Table 1.2 [5].

Tanford (1972) [8] used the data of Reiss-Husson and Luzzati (1964) [9], generated by X-ray scattering, to produce relations for the values of the hydrocarbon core volume

Table 1.2: Shapes of aggregates formed from surfactants with different packing parameters [7]

Surfactant Type	Packing Parameter	Packing Shape	Aggregate
Single-chained surfactants with large headgroup areas	$\leq \frac{1}{3}$		
Single-chained surfactants with small headgroup areas	$\sim \frac{1}{2}$		
Double-chained surfactants with large head-group areas, fluid chains	$\frac{1}{2} - 1$		
Double-chained surfactants with small head-group areas, frozen chains	~ 1		
Double-chained surfactants with small head-group areas, poly unsaturated chains	> 1		

and the length of the hydrocarbon chains, equation 1.4.2, where n_c is the total number of carbons per hydrocarbon chain.

$$\begin{aligned} v &= (0.0274 + 0.0269 n_c)\,\mathrm{nm}^3 \text{ per hydrocarbon chain} \\ l_c &= (0.15 + 0.1265 n_c)\,\mathrm{nm} \text{ per hydrocarbon chain} \end{aligned} \tag{1.4.2}$$

The numbers in these relationships can be explained in terms of the molecular properties of the groups with in the hydrocarbon chains. The volume of a $-\mathrm{CH_2}-$ group is $0.0294\,\mathrm{nm}^3$, i.e. approximately the second number in the relationship for v, and the volume of a $-\mathrm{CH_3}$ group is about twice this value [10]: the difference in the two volumes thus gives rise to the first number in the relationship for v. The van der Waals radius of

a terminal methyl group is $0.21\,\mathrm{nm}$ and a carbon-carbon bond length is $0.154\,\mathrm{nm}$ [11]: this explains the numbers in the relationship for l_c. The bond length taken in the *trans* configuration gives an atom spacing of $0.133\,\mathrm{nm}$, which is approximately the 0.1265, and the van der Waals radius minus half this length is equal to 0.143, which is approximately 0.15. [1]

These relationships can be substituted into equation 1.4.1 to allow calculation of the packing parameter. The value to be calculated is the effective area per head group, a_0. This value is affected by the type of surfactant used, the material the surfactant is dissolved in, and the temperature [4].

The number of monomers required to form a spherical type micelle is termed an aggregation number, N_0, and it varies from typically 50 to 100, depending on the surfactant type, aggregate shape, temperature, and concentration. However, for a given surfactant under given conditions this number is fixed, therefore these micelles can be treated as mono-dispersed structures. Cylindrical micelles do not have such well defined aggregation numbers as the cylinders can grow to varying lengths, producing a poly-dispersed system [4].

1.4.3 Critical Micelle Concentration

Self-assemblies only form when it is thermodynamically favourable for them to do so; as low concentrations of surfactant then they can exist as unimers. As the concentration increases the aggregates will form; the concentration at which the micelles first form is called the critical micelle concentration, CMC. The chemical structure of the surfactant, the temperature, and any co-solutes are all known to influence the CMC [2].

Many physical properties of surfactant solutions undergo a sudden change at the CMC. These changes allow the value of the CMC to be determined for surfactant solutions. Figure 1.5 shows the general trends for some physical properties around the CMC.

Several methods have been used in the literature to measure the CMC of different surfactants in an aqueous solution, including surface tension ([13]), conductivity measurements ([14]), and fluorescence intensity ([15]).

1.4.4 Krafft Point

As stated, micelles will only form above the CMC, but they will also only form above a certain temperature. At low temperatures, the surfactant has low solubility and the concentration dissolved may be below the CMC. The point at which the dissolved concentration at the given temperature is equal to the CMC is called the Krafft point (or Krafft temperature); every surfactant has a characteristic Krafft point. The temperature dependence of the surfactant solubility in the region of the Krafft point is shown in Figure 1.6. Below the Krafft point, the solubility of the surfactants is less than the CMC and hydrated surfactant crystals will formed.

[1] Fennell Evans and Wennerström (1999) [4] give slightly modified versions of these relationships, where n_{Me} is the number of methyl groups. These contain values that appear to have been rounded from equation 1.4.2,

$$v = 0.027\,(n_c + n_{Me})\,\mathrm{nm}^3 \text{ per hydrocarbon chain}$$
$$l_c = (0.15 + 0.127 n_c)\,\mathrm{nm} \text{ per hydrocarbon chain.}$$

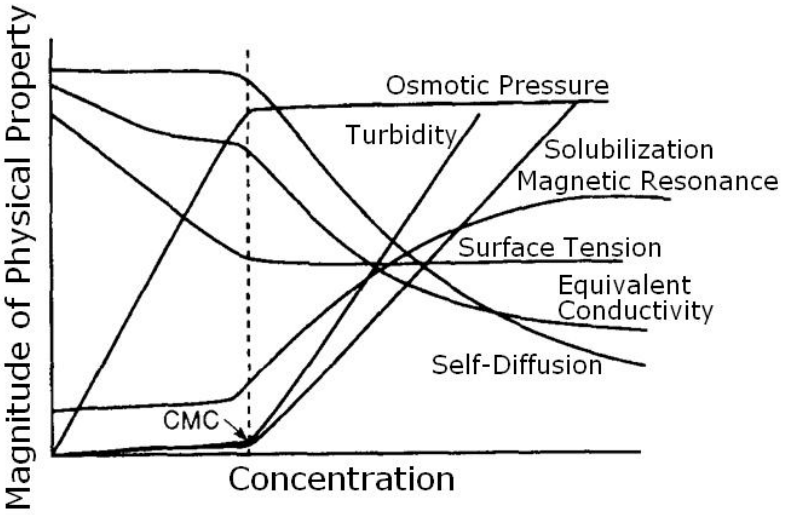


Figure 1.5: Effect of surfactant concentration on physical properties around the CMC [12].

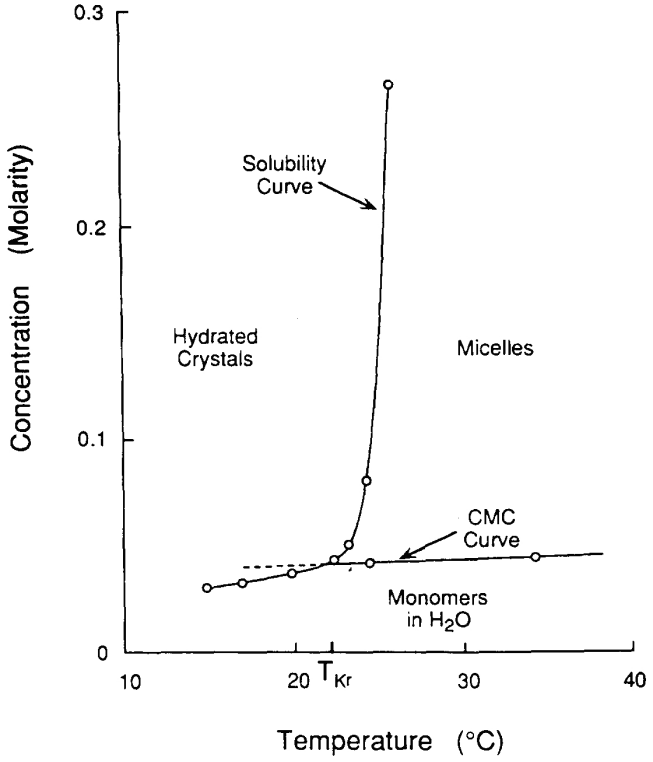


Figure 1.6: Representation of Krafft temperature.

1.5 Micellization

Micelles are the simplest and most thoroughly characterised self-organising structures. When surfactants associate into micelles, they form a liquid-like aggregate. No obvious mechanism leads to a specific aggregation number, as seem with micelles, therefore it is natural to describe the association in terms of a stepwise addition of a monomer, S to the aggregate, S_{n-1}, as in,

$$S + S_{n-1} \rightleftharpoons S_n.$$

If we assume that interactions between aggregates are negligible, then the equilibrium of the stepwise addition can be given by equation 1.5.1.

$$K_n = \frac{[S_n]}{[S]\,[S_{n-1}]} \tag{1.5.1}$$

In situations that involve aggregation numbers of order 100, technically we need a very large number of equilibrium constants. Therefore, we obtain more useful values and understanding with some simplified models. The three main models used are:

The **isodesmic model** assumes that the K_n is independent of n. This model does well to describe some dyes, but does not capture the cooperative nature associated with amphiphilic aggregation.

The **phase separation model** approximates aggregation as a phase separation process in which the activity of the monomer remains constant above the CMC. It is useful to capture the start mechanism of aggregation, but not the stop.

The **closed-association model** assumes that one aggregation number, N, dominates and relates the free energy of micellization to the measured CMC. This model captures both the start and the stop cooperative features associated with aggregation processes.

1.5.1 Isodesmic Model

The isodesmic model assumes that K_n is independent of n. In this case it can be shown that regardless of either the total concentration or of K that $[S]\,K < 1$. The aggregate distribution function can be given by equation 1.5.2, which decays exponentially with $[S_1] > [S_2] > [S_3]$.

$$f\,(n) = \frac{[S_n]}{\sum_{n=1}^{\infty} [S_n]} \tag{1.5.2}$$

The concentration of each aggregate size can be determined from knowledge of the monomer concentration and the equilibrium constant, equation 1.5.3, with the free energy of each monomer addition given by equation 1.5.4.

$$[S_n] = K^{n-1}\,[S] \tag{1.5.3}$$

$$\Delta G = -RT \ln K \tag{1.5.4}$$

The isodesmic model describes the association of some dyes in aqueous solution quite well [, dyemodel] but is less successful as a description of the formation of micelles as it does not show an abrupt onset in a narrow concentration range that typifies micelle formation. This means that this model cannot predict a CMC.

1.5.2 Phase Separation Model

Micelle formation has several features in common with the formation of a separate liquid phase. In terms of the equilibrium described by equation 1.5.1 the phase separation model assumes that aggregates with large n dominate all others except monomers. This assumption implies strong cooperativity because, once aggregation begins it becomes more and

more favourable to add another monomer until a large aggregation number is reached. The surfactant possesses a chemical potential, $\mu\,(\text{agg})$, in the aggregate and a different chemical potential, $\mu\,(\text{sol})$, in the solvent. When equation 1.5.5 is true, the monomers and aggregates coexist in equilibrium and $[S]$ is the CMC (neglecting other oligomers).

$$\mu\,(\text{agg}) = \mu\,(\text{sol}) + RT \ln [S] \tag{1.5.5}$$

The standard free energy of micelle formation, ΔG_{mic}, represents the standard free energy difference between a monomer in the micelle and the standard chemical potential in dilute solution, equation 1.5.6.

$$\Delta G_{mic} = \mu\,(\text{agg}) - \mu\,(\text{sol}) = RT \ln [\text{CMC}] \tag{1.5.6}$$

This method gives a very useful approximation for obtaining the free energy of micelle formation, but does not capture all the essential features of micelle formation. Although it describes the start mechanism it does not describe the stop mechanism.

1.5.3 Closed-Association Model

If it is assumed that one aggregation number, N, dominates for the aggregate sizes, and there only exists these aggregates and monomers, then the association can be simplified to,

$$N\,S \rightleftharpoons S_N,$$

producing the equilibrium constant as equation 1.5.7.

$$K_N = \frac{[S_N]}{[S]^N} \tag{1.5.7}$$

The total surfactant concentration can be expressed as equation 1.5.8.

$$[S]_T = N\,[S_N] + [S] = NK_N\,[S]^N + [S] \tag{1.5.8}$$

Experiments identify the CMC as the concentration at which surfactant preferentially starts to enter the aggregate. A good measure of this concentration point is where an added monomer is as likely to enter the aggregate as to remain in solution, i.e. equation 1.5.9.

$$\left.\frac{\partial\, N\,[S_N]}{\partial\,[S]_T}\right|_{CMC} = \left.\frac{\partial\,[S]}{\partial\,[S]_T}\right|_{CMC} = 0.5 \tag{1.5.9}$$

This means that the concentration of the surfactant in solution at the CMC in terms of the aggregate size and the association constant, equation 1.5.10, can be solved.

$$\begin{aligned} \left.\frac{\partial\,[S]_T}{\partial\,[S]}\right|_{CMC} &= 2 \\ [S]_{CMC}^{N-1} &= \left(N^2 K_N\right)^{-1} \end{aligned} \tag{1.5.10}$$

As the CMC refers to the total surfactant concentration equation 1.5.10 can be combined with equation 1.5.8 to give the CMC value, equation 1.5.11.

$$CMC = [S]_{CMC}\left(1 + N^{-1}\right) = \left(N^2 K_N\right)^{-1/(N-1)}\left(1 + N^{-1}\right) \tag{1.5.11}$$

From equation 1.5.11 it can be seen that the amount of micellized surfactant at the CMC is $[S]/N$, which becomes more and more negligible as the value of N increases.

From equation 1.5.8 a solution for the fraction of surfactant entering the aggregates on addition of surfactant can be calculated, equation 1.5.12.

$$\frac{\partial N\left[S_N\right]}{\partial\left[S\right]_T} = \frac{N^2\left[S_N\right]}{\left(\frac{\left[S_N\right]}{K_N}\right)^{1/N} + N^2\left[S_N\right]} \tag{1.5.12}$$

Figure 1.7 shows the change in the fraction of surfactant entering the aggregates for different values of the aggregation number, N. The larger the values of N the more abruptly the derivative $\partial N\left[S_N\right]/\partial\left[S\right]_T$ changes from low concentration to higher concentration. This means that for high values of N the CMC is a more easily defined point, with a discontinuity in the derivative at the CMC when $N \to \infty$.

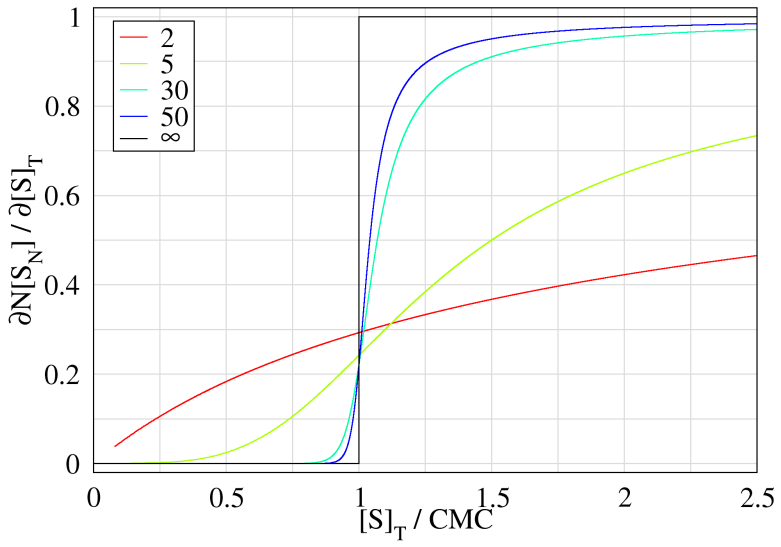


Figure 1.7: Variation of the fraction of added surfactant entering micelles for different aggregation numbers, N.

As in the isodesmic model the free energy for the aggregate formation per mole of aggregates can be given by equation 1.5.13.

$$\Delta G = -RT \ln K_n = -RT \ln\left[S_N\right] + NRT \ln\left[S\right] \tag{1.5.13}$$

1.6 Liquid Crystals

Many soft matter systems have liquid crystal phases. These phases, as suggested by the name, are phases that are part way between a crystalline and a liquid state, exhibiting properties of both. Crystals are characterised by having long range positional and directional order in all three spacial dimensions; in a liquid this is not the case and molecules can simply diffuse freely and have no internal order, except for some loose correlation between neighbours. Crystals can support stress without significant deformation, while liquids deform and flow. Liquid crystals exhibit some of the positional and directional order of the crystals, but have molecular motion, usually anisotropic, and their properties under shear can vary between very little deformation to free flowing.

Liquid crystals can be thermotropic or lyotropic:

Thermotropic liquid crystals are liquid crystals that do not require a solvent (although a solvent can be present); the order-disorder phase transition is purely due to a

temperature change. These are usually organic compounds with a rigid polyaromatic section connected to a flexible alkyl chain. The liquid crystal order is often determined by anisotropic nature of the molecule and anisotropic intermolecular attraction that thus arises. It is the rigid polyaromatic section that will determine the liquid crystal structure: these can be classed into two broad groups: calamitic (long, narrow molecules) and discotic (disc-shaped molecules). These tend to form nematic, smectic and, columnar liquid crystals.

Lyotropic liquid crystals are formed on the dissolution of amphiphilic molecules in a solvent with much smaller molecules. This means that lyotropic liquid crystal phase transitions are driven not just by temperature but also by solute concentration. These are often produced from polymers that align together and are repulsed by the solvent or from surfactant micellar structures (as individual surfactant molecules are generally not large or rigid enough to form liquid crystals in the absence of solvent). Surfactant micelles can form lyotropic liquid crystal phases, which tend to form at higher surfactant concentrations. There are six classes of surfactant liquid crystal phases: lamellar, hexagonal, cubic, nematic, gel (all of which are well documented), and intermediates (which are not) [3].

1.6.1 Thermotropic Liquid Crystal Phases

Nematic Phase

One of the most common liquid crystal phases is the nematic, this is also the simplest liquid crystal phase. In a nematic phase, the calamitic or rod-shaped molecules have no positional order, but they self-align to have long-range directional order with their long axes roughly parallel [16], Figure 1.8. Thus, the molecules are free to flow and their center of mass positions are randomly distributed as in a liquid, but still maintain their long-range directional order. Most nematics are uniaxial: they have one axis that is longer and preferred, with the other two being equivalent (can be approximated as cylinders or rods). However, some liquid crystals are biaxial nematics, meaning that in addition to orienting their long axis, they also orient along a secondary axis [17]. Nematics have fluidity similar to that of ordinary (isotropic) liquids but they can be easily aligned by an external magnetic or electric field. Aligned nematics have the optical properties of uniaxial crystals and this makes them extremely useful in liquid crystal displays (LCD) [18].

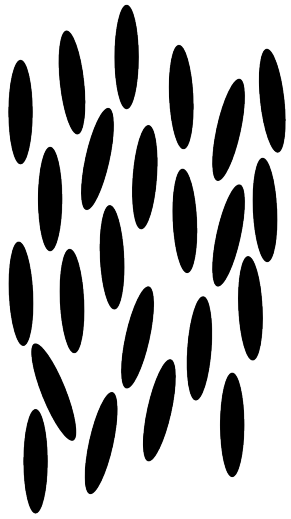


Figure 1.8: Schematic representation of a nematic phase liquid crystal.

As nematic liquid crystals are composed of rod-like molecules with the long axes of neighboring molecules aligned approximately to one another a dimensionless unit vector n called the director, is introduced to represent the direction of preferred orientation of molecules in the neighborhood of any point. Because there is no physical polarity along the director axis, n and $-n$ are fully equivalent [19]. The local nematic director, which is also the local optical axis, is given by the spatial and temporal average of the long molecular axes. A second rank symmetric traceless tensor order parameter is used to describe the orientational order of a nematic liquid crystal, although a scalar order parameter is

usually sufficient to describe uniaxial nematic liquid crystals. To make this quantitative, an orientational order parameter is usually defined based on the average of the second Legendre polynomial, equation 1.6.1.

$$S = \langle P_2(\cos\theta) \rangle = \left\langle \frac{3\cos^2\theta - 1}{2} \right\rangle \tag{1.6.1}$$

θ is the angle between the liquid crystal molecular axis and the local director (which is the 'preferred direction' in a volume element of a liquid crystal sample, also representing its local optical axis). The brackets denote both a temporal and spatial average. This definition is convenient, since for a completely random and isotropic sample, $S = 0$, whereas for a perfectly aligned sample $S = 1$. For a typical liquid crystal sample, S is on the order of 0.3 to 0.8, and generally decreases as the temperature is raised. A sharp drop of the order parameter to 0 is observed when the system undergoes a phase transition from an liquid crystal phase into the isotropic phase [20]. The order parameter can be measured experimentally in a number of ways, e.g. diamagnetism, birefringence, Raman scattering, or NMR [21].

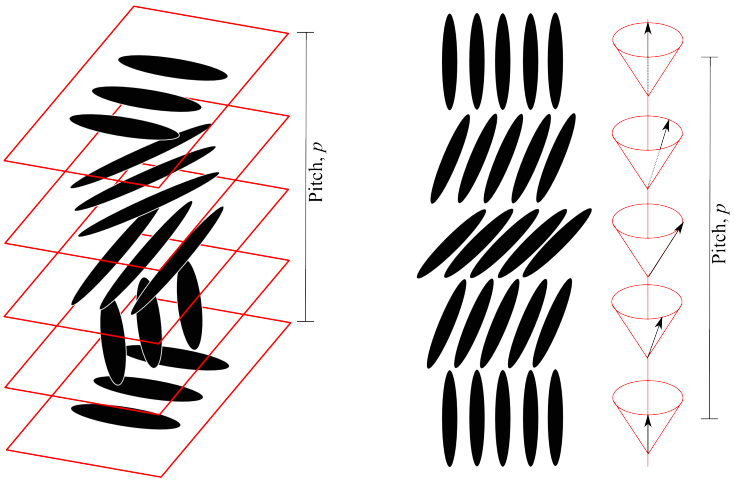
Smectic Phase

The smectic phase differs from the nematic phase in the fact that as well as the molecules being aligned they are also organized into layers. As the smectic phases, which are found at lower temperatures than the nematic, form well-defined layers, these can slide over one another. There are many different smectic phases, all characterized by different types and degrees of positional and orientational order [19], Table 1.3.

Chiral phases

The chiral nematic phase, N*, (also called the cholesteric phase) and the smectic C* phase exhibit chirality (handedness). This phase is often called the cholesteric phase because it was first observed for cholesterol derivatives. Only chiral molecules (i.e., those that have no internal planes of symmetry) can give rise to such a phase. This phase exhibits a twisting of the molecules perpendicular to the director, with the molecular axis parallel to the director. The finite twist angle between adjacent molecules is due to their asymmetric packing, which results in longer-range chiral order, Figure 1.9(a). In the smectic C* phase, the molecules have positional ordering in a layered structure (as in the other smectic phases), with the molecules tilted by a finite angle with respect to the layer normal, Figure 1.9(b).

The chiral pitch, p, refers to the distance over which the liquid crystal molecules undergo a full $360°$ twist (but note that the structure of the chiral nematic phase repeats itself every half-pitch, since in this phase directors at $0°$ and $\pm 180°$ are equivalent). The pitch typically changes when the temperature is altered or when other molecules are added to the liquid crystal host (an achiral liquid crystal host material will form a chiral phase if doped with a chiral material), allowing the pitch of a given material to be tuned accordingly. In some liquid crystal systems, the pitch is of the same order as the wavelength of visible light. This causes these systems to exhibit unique optical properties, such as Bragg reflection and low-threshold laser emission [22], and these properties are exploited in a number of optical applications [23].

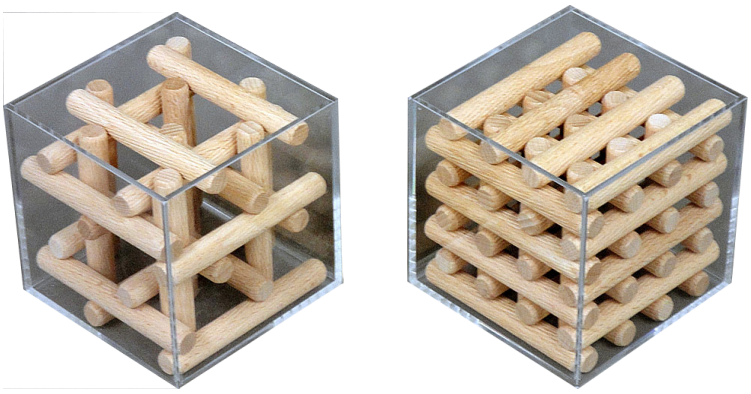


(a) (b)

Figure 1.9: Schematic representation of the (a) chiral nematic and the (b) smectic C* phases.

Blue phases

Blue phases are liquid crystal phases that appear in the temperature range between a chiral nematic phase and an isotropic liquid phase. Blue phases have a regular three-dimensional cubic structure of defects with lattice periods of several hundred nanometers, and thus they exhibit selective Bragg reflections in the wavelength range of visible light corresponding to the cubic lattice, Figure 1.10.



(a) (b)

Figure 1.10: Schematic representation of the blue phase (a) type I and (b) type II.

Although blue phases are of interest for fast light modulators or tunable photonic crystals, they exist in a very narrow temperature range, usually less than a few kelvin. Recently the stabilization of blue phases over a temperature range of more than 60 K including room temperature (260-326 K) has been demonstrated [24].

Discotic phases

Disk-shaped liquid crystal molecules can orient themselves in a layer-like fashion known as the discotic nematic phase, which can be chiral id the disks are chiral. The disks can also pack into ordered columns (or stacks), called a discotic columnar [25]. The columns

themselves may be organized into rectangular or hexagonal arrays, Figure 1.11. In the simplest case the short axes of the molecules lie parallel to the axis of the column and the columns are randomly distributed in space. More complicated discotic phases exist, where the short molecular axes lie at an angle to the column and translational order exists between the columns, analogous to the more complicated smectic phases.

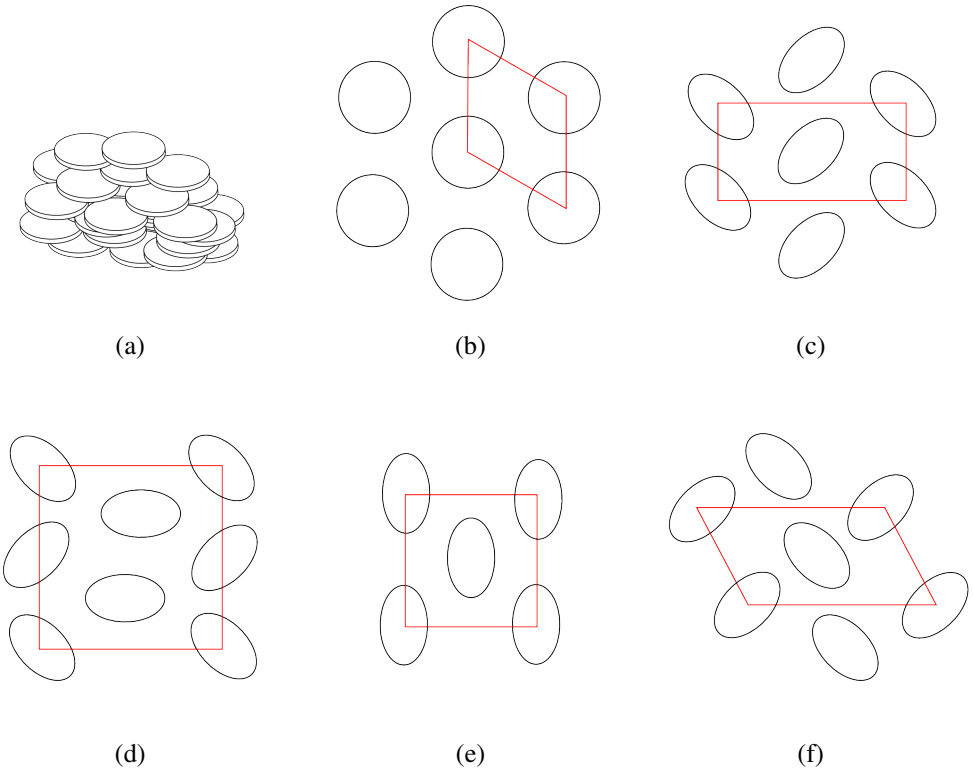


Figure 1.11: Schematic representation of the discotic phases (a) nematic N_D, (b) hexagonal D_{ho}, (c) rectangular $D_{rd(P2_{1/a})}$, (d) rectangular $D_{rd(P2/a)}$, (e) rectangular $D_{rd(2C/m)}$, and (f) oblique D_{ob}.

Other Thermotropic Phases

Bent-core molecules can form 'banana' phase liquid crystals [26]. Some of these phases are chiral although the molecules forming them are achiral. The chirallity is caused by the bend in the molecules being a fixed angle and direction.

Some high molecular mass polymers, liquid crystalline polymers, can also form liquid crystal phases [25]. These fall into two categories depending on where the mesogenic part of the molecule is located. If the mesogenic unit is contained within the main polymer chain then it is termed a main chain liquid crystal polymer. If it is attached to a side chain of the polymer then it is termed a side chain liquid crystal polymer. As well as depending on the nature of the mesogenic core, the mesophases formed by these materials are dependent on the flexibility of the polymer backbone and the side chains.

Closely related to liquid crystalline polymers are dendritic liquid crystals [27, 28]. These molecules consist of a central core with the mesogenic units attached to flexible spacers that radiate out from the core.

1.6.2 Lyotropic Liquid Crystal Phases

Lamellar Phase

The most common surfactant liquid crystal structure is the lamellar phase. The lamellar phase is built up of bilayers of surfactant separated by solvent, Figure 1.12. The thickness of the bilayer generally varies from about 1 to 1.9 times the all-*trans* alkyl chain length, while the thickness of the water layers varies over a much larger range of around 8 to 200 Å [3]. Depending on the surfactant, the bilayer can range from being stiff and planar to being very flexible. The lamellar phase does not normally flow under gravity but can have a relatively low apparent viscosity. The lamellar phase is also birefringent and can be easily identified from its optical properties, i.e. oily streaks and Maltese crosses [29]. X-ray diffraction studies show sharp reflections in the ratio $1 : 1/2 : 1/3 \ldots$ due to the repeating nature of the structure with the repeat spacings being the sum of the water and the alkyl chain layers.

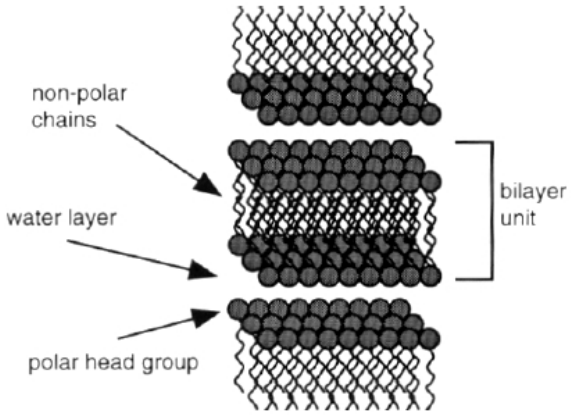


Figure 1.12: Schematic representation of a lamellar phase liquid crystal.

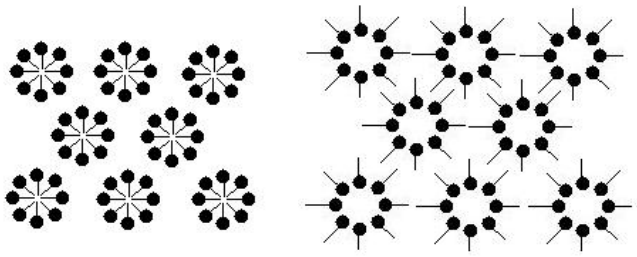
The lamellar phase is generally considered to play a central role in the evolution of the other liquid crystal phases, as their geometric features are relatively small changes to the lamellar structure [30].

Hexagonal Phases

The next most common liquid crystal type is the hexagonal phase, which itself consists of two types: "normal hexagonal" and "reversed hexagonal". The normal phase is water continuous, while the reversed phase is alkyl chain continuous. Both these phases consist of structures similar to cylindrical-type micelles but closely packed together, Figure 1.13. Both hexagonal phases usually have a relatively high apparent viscosity and do not flow easily. X-ray diffraction studies of both types show sharp reflections in the ratio $1 : 1/\sqrt{3} : 1/\sqrt{4} : 1/\sqrt{7} : 1/\sqrt{12} \ldots$ due to the nature of the cylinders.

Cubic Phases

Cubic phase liquid crystals encompass the broadest range of liquid crystals as they can be based around one of several cubic lattices (primitive, face-centered, or body-centered) and can be either made up of small micelles or three-dimensional bicontinuous aggregates, both in the normal and reversed arrangements. It is still partially unclear which types of cubic structures can occur for the different aggregates [31]. The two different classes of the cubic phase (micellar and bicontinuous) can be distinguished from each other by



(a) (b)

Figure 1.13: Schematic representation of the (a) normal and the (b) reverse hexagonal phases.

their locations in the phase diagram. Micellar cubic phases are found between micellar solutions and hexagonal phases, while the bicontinuous cubic phases are found between hexagonal and lamellar phases. The factors determining which cubic structure occurs are not understood [3].

Nematic phases

Surfactant nematic phase liquid crystals are less common than the other types of surfactant liquid crystal. If they do form for a particular system, they tend to form in the region between the micellar and hexagonal phases or the micellar and lamellar phases. There are two types of nematic liquid crystal for surfactants: one type is thought to be composed of small cylindrical micelles related to the hexagonal phase and the other is composed of planar disc type micelles related to the lamellar phase [32]. These phases are often found for short-chain surfactants and tend to have a low viscosity [3].

Gel Phases

The gel phase is similar to the lamellar phase in that it is also made up of surfactant layers. In this phase the surfactant layers are fixed in a rigid mostly all-*trans* configuration and the water phase is in a "liquid-like" state. There are three common types of gel phase; the first is the normal type where the surfactant layer is approximately twice the alkyl chain length [33], the second is the tilted type where the larger head groups force the alkyl chains to angle to pack efficiently [34], and the third is the inter-digitated type where the surfactant layer is closer to the alkyl chain length [35] (Figure 1.14).

The gel phase is normally formed upon cooling of a lamellar phase through the transition temperature. The rigid alkyl chains give rise to a high apparent viscosity.

Intermediate Phases

Intermediate phases cover any other type of structure that has been reported for any surfactant system. These phases, as the name suggests, seem to be structures that appear on the boundaries between two of the other phases and contain properties of both structures. The observed structures can be broadly divided into three types: rectangular ribbon structures, layered mesh structures, and bicontinuous structures that do not have cubic symmetry [36]. Ribbon structures can be considered as a distorted hexagonal phase, mesh phases are distorted lamellar phases where the surfactant layers have water filled defects, and the bicontinuous phases are distorted cubic structures.

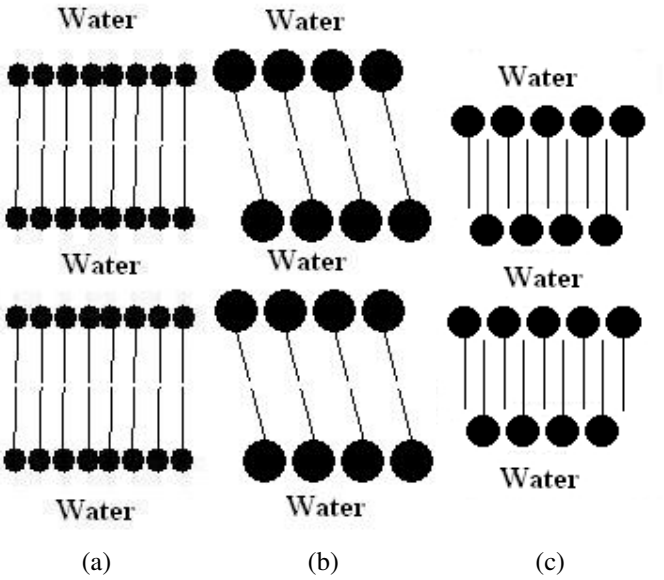


Figure 1.14: Schematic representation of the (a) normal, the (b) tilted and (c) interdigitated gel phases.

1.6.3 Phase Behaviour

Throughout the literature there is not a consistent system used to refer to the different liquid crystal phases. However some notations are more common than others, Table 1.4.

As mentioned in Section 1.4.2 the micelle shape is determined by the molecular structure of the surfactants. However, there is a critical volume fraction (i.e. concentration) of surfactant above which these individual micelles cannot be supported. This leads to the disordered solutions (micellar solutions) becoming ordered (liquid crystals). In addition, there is a maximum volume fraction of ordered micelles before they cannot fit together; e.g. spheres and cylinders cannot be packed together, no matter how tight, to totally fill a volume. (The existence of the micellar cubic phase also depends on the size and stability of the micelles. Micelles with a large energy barrier to size change will tend to increase in number rather than in size, which encourages packing of the micelles. However, if the stable micelle size is small there will be a large entropic barrier to overcome to form an ordered micellar structure, which will encourage the disordered solution to prevail until a higher concentration favours a larger aggregate phase, e.g. hexagonal or lamellar.) When all the available volume is occupied for a given shape, then the only way more surfactant can be accommodated is with a change of structure towards less curvature [37]. For the normal surfactant structure types these phase changes can be summarised by Figure 1.15 (modified from Mitchell *et al.* (1983) [37]).

In the phase diagram not all the phases may take part: e.g. the nematic and bicontinuous cubic phases may not occur, and there also may be added intermediate phases that are in a similar region to the bicontinuous cubic phase. The gel phase may exist in the lamellar region at lower temperatures.

The sequence for the reversed phase surfactant structures is much more complicated than for the normal phases due to less radius limitations on micellar shapes; water on the inside could swell almost indefinitely until packing limits are reached. This means that these phase changes are not yet understood in terms of the surfactant structure [3].

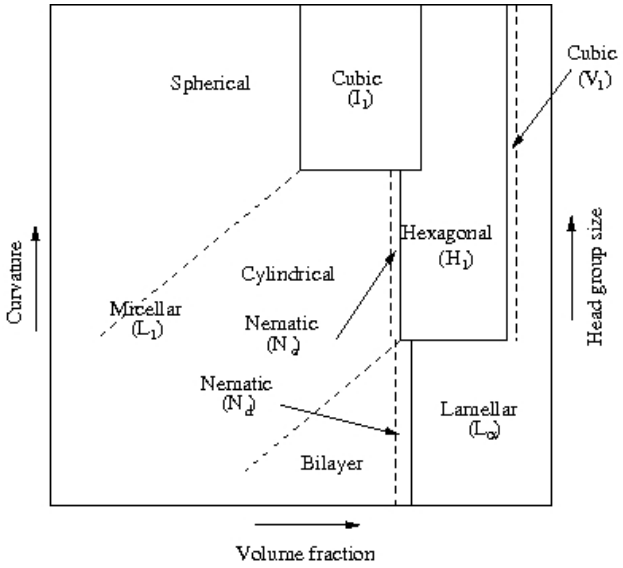


Figure 1.15: Schematic representation of the phase changes for the normal surfactant structure types.

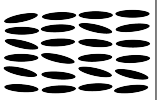
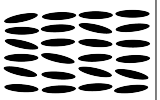
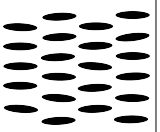
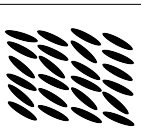
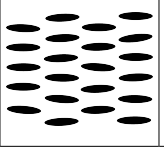
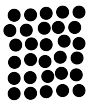
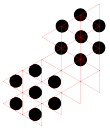
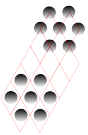
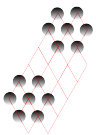
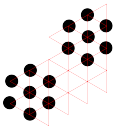
Group	A	B	C	D	E	F	G	H	I	J	K	L
Side				Molecules arrange in a micellar type shape and then pack into a body centred cubic phase								
Plan												

Table 1.3: Smectic phases.

Table 1.4: Notations used for the common phases.

Phase structure	Notation
Lamellar	L_α
Hexagonal	H_1
Reversed hexagonal	H_2
Micellar cubic	I_1
Reversed micellar cubic	I_2
Bicontinuous cubic	V_1
Reversed bicontinuous cubic	V_2
Nematic	N
Cylindrical Nematic	N_c
Discotic Nematic	N_D
Chiral Nematic	N*
Smectic	S
Chiral smectic C	C*
Gel	L_β
Micellar	L_1
Reversed micellar	L_2
Sponge phase	L_3
Reversed sponge phase	L_4
Vesicular	V
Microemulsion	L
Water	W

1.7 References

[1] American Institute of Physics. *http://www.aip.org/dbis/stories/2007/17141.html*, last access 10/02/2008.

[2] K. Holmberg, B. Jönsson, B. Kronberg, and B. Lindman. *Surfactants and Polymers in Aqueous Solution*. John Wiley & Sons, Chichester, England, 2nd edition, 2007.

[3] S. Hussan, W. Rowe, and G. J. T. Tiddy. *Handbook of Applied Surface and Colloid Chemistry*, chapter 21, pages 465–508. John Wiley & Sons, Chichester, England, 2001.

[4] D. Fennell Evans and H. Wennerström. *The Colloidal Domain - Where Physics, Chemistry, Biology and Technology meet*. Wiley-VCH, New York, 1999.

[5] J. N. Israelachvili, D. J. Mitchell, and B. W. Ninham. Theory of Self-Assembly of Hydrocarbon Amphiphiles into Micelles and Bilayers. *Journal of the Chemical Society - Faraday Transactions II*, 72:1525–1568, 1976.

[6] G. C. Shearman, O. Ces, R. H. Templer, and J. M. Seddon. Inverse Lyotropic Phases of Lipids and Membrane Curvature. *Journal of Physics: Condensed Matter*, 18:S1105–S1124, 2006.

[7] J. Israelachvili. *Intermolecular and Surface forces*. Academic Press, London, 2nd edition, 1992.

[8] C. Tanford. Micelle Shape and Size. *Journal of Physical Chemistry*, 73:3020–3024, 1972.

[9] F. Reiss-Husson and V. Luzzati. The Structure of the Micellar Solutions of Some Amphiphilic Compounds in Pure Water as Determined by Absolute Small-Angle X-Ray Scattering Techniques. *Journal of Physical Chemistry*, 68:3504–3511, 1964.

[10] K. S. Birdi. *Self-assembly Monolayer Structures of Lipids and Macromolecules at Interfaces*. Kluwer Academic, 1999.

[11] A. Bondi. van der Waals Volumes and Radii. *Journal of Physical Chemistry*, 68:441–451, 1964.

[12] B. Lindman and H. Wennerström. Amphiphile aggregation in aqueous solution. *Topics in Current Chemistry*, 87:1–83, 1980.

[13] A. El-Hamouz. Effect of Surfactant Concentration and Operating Temperature on the Drop Size Distribution of Silicon Oil Water Dispersion. *Journal of Dispersion Science and Technology*, 28:797–804, 2007.

[14] R. Zana, W. Binana-Limbelé, N. Kamenka, and B. Lindman. Ethyl(hydroxyethyl)cellulose-Cationic Surfactant Interactions: Electrical Conductivity, Self-Diffusion, and Time-Resolved Fluorescence Quenching Investigations. *Journal of Physical Chemistry*, 96:5461–5465, 1992.

[15] S. Miyagishi, H. Kurimoto, Y. Ishihara, and T. Asakawa. Determination of the Critical Micelle Concentration and Microviscosity with a Fluorescence Probe, Auramine. *Bulletin of the Chemical Society of Japan*, 67:2398–2402, 1994.

[16] J. A. Rego, J. A. A. Harvey, A. L. MacKinnon, and E. Gatdula. Asymmetric Synthesis of a Highly Soluble 'Trimeric' Analogue of the Chiral Nematic Liquid Crystal Twist Agent Merck S1011. *Liquid crystals*, 37:37–43, 2010.

[17] L. A. Madsen, T. J. Dingemans, M. Nakata, and E. T. Samulski. Thermotropic Biaxial Liquid Crystals. *Physical Review Letters*, 92:145505, 2004.

[18] J. A. Castellano, editor. *Liquid Gold: The Story of Liquid Crystal Displays and the Creation of an Industry*. World Scientific Publishing.

[19] P. G. de Gennes and J. Prost, editors. *The Physics of Liquid Crystals*. Oxford Claredon Press, 1993.

[20] S. K. Ghosh. A Model for the Orientational Order in Liquid Crystals. *Il Nuovo Cimento D*, 4:229, 1984.

[21] P. J. Collings and M. Hird, editors. *Introduction to Liquid Crystals*. Taylor & Francis, 1997.

[22] V. I. Kopp, B. Fan, H. K. M. Vithana, and A. Z. Genack. Low Threshold Lasing at the Edge of a Photonic Stop Band in Cholesteric Liquid Crystals. *Opt. Letters*, 23:1707–1709, 1998.

[23] I. Dierking, editor. *Textures of Liquid Crystals*. Wiley-VCH, 2003.

[24] H. J. Coles and M. N. Pivnenko. Liquid Crystal 'Blue Phases' with a Wide Temperature Range. *Nature*, 436:997–1000, 2005.

[25] D. Demus, J. W. Goodby, G. W. Gray, H. W. Spiess, and V. Vill, editors. *Handbook of Liquid Crystals*. Wiley-VCH, 1998.

[26] G. Pelzl, S. Diele, and W. Weissflog. *Advanced Materials*, 11:707, 1999.

[27] K. Lorenz, D. Holter, B. Stuhn, R. Mulhaupt, and H. Frey. *Advanced Materials*, 8:414, 1996.

[28] F. Vogtle, S. Gestermann, R. Hesse, H. Schwierz, and B. Windisch. *Progress in Colloid and Polymer Science*, 25:987, 2000.

[29] M. Makai, E. Csányi, I. Dékány, Z. Németh, and I. Erös. Structural Properties of Nonionic Surfactant/Glycerol/Paraffin Lyotropic Liquid Crystals. *Colloid and Polymer Science*, 281:839–844, 2003.

[30] K. Fontell. Cubic Phases in Surfactant and Surfactant-like Lipid Systems. *Colloid and Polymer Science*, 268:264–285, 1990.

[31] P. Sakya, J. M. Seddon, R. H. Templer, R. J. Mirkin, and G. J. T. Tiddy. Micellar Cubic Phases and Their Structural Relationships: The Nonionic Surfactant System C12EO12/Water. *Langmuir*, 13:3706 – 3714, 1997.

[32] B. Lühmann, H. Finkelmann, and G. Rehage. Phase Behaviour and Structure of Polymer Surfactants in Aqueous Solution. The Occurrence of Lyotropic Phases. *Makromolekulare Chemie*, 186:1059–1073, 1985.

[33] D. Chapman, R. M. Williams, and B. D. Ladbrooke. Physical Studies of Phospholipids. VI. Thermotropic and Lyotropic Mesomorphism of some 1,2-Diacyl-Phosphatidylcholines (Lecithins). *Chemistry and Physics of Lipids*, 1:445–475, 1967.

[34] K. Larsson. Structure of Mesomorphic Phases and Micelles in Aqueous Glyceride Systems. *Zeitschrift fur Physikalische Chemie*, 56:173–198, 1967.

[35] J. M. Vincent and A. Skoulios. "Gel" and "Coagel." I. Detection. Localization in a Phase Diagram and Determination of the "Gel" Structure of Potassium Stearate. *Acta Crystallographica*, 20:432–440, 1966.

[36] M. C. Holmes. Intermediate Phases of Surfactant-Water Mixtures. *Current Opinion in Colloid and Interface Science*, 3:485–492, 1998.

[37] D. J. Mitchell, G. J. T. Tiddy, L. Waring, T. Bostock, and M. P. McDonald. Phase Behaviour of Polyoxyethylene Surfactants: Mesophase Structures and Partial Miscibility (Cloud Points). *Journal of the Chemical Society: Faraday Transactions*, 79:975–1000, 1983.

Chapter 2

Review of Simulation Techniques

Contents

2.1 Introduction

Simulation is the imitation of the operation of a real-world process or system over time. It is one of the most widely used tool for decision making, usually performed on a computer with appropriate software. Simulation can by an analysis/descriptive tool (can answer what if questions) or a synthesis/prescriptive tool (can develop new systems). Simulations are generally applied to complex systems that are impossible to solve mathematically.

2.2 Simulation Time and Length Scales

When considering the simulation of a physical system using computers it is important to know what the intended output of the simulation is i.e. What is the purpose of this process. Once the purpose is known, usually to study a particular phenomena, the length scale on which this phenomena is likely to occur must be identified in order to recognise which type of simulation is most likely to produce relevant results. For instance a simulation technique where the finest detail is on the micrometre scale would be unlikely to yield useful results if studying the kinetics of a chemical reaction. Similarly a nanoscale simulation would be extremely inefficient for studying the dynamics of bulk fluids. Figure 2.1 shows a representation of simulation detail. Figure 2.2 shows examples of soft matter simulation techniques and their relevant time- and size-scales.

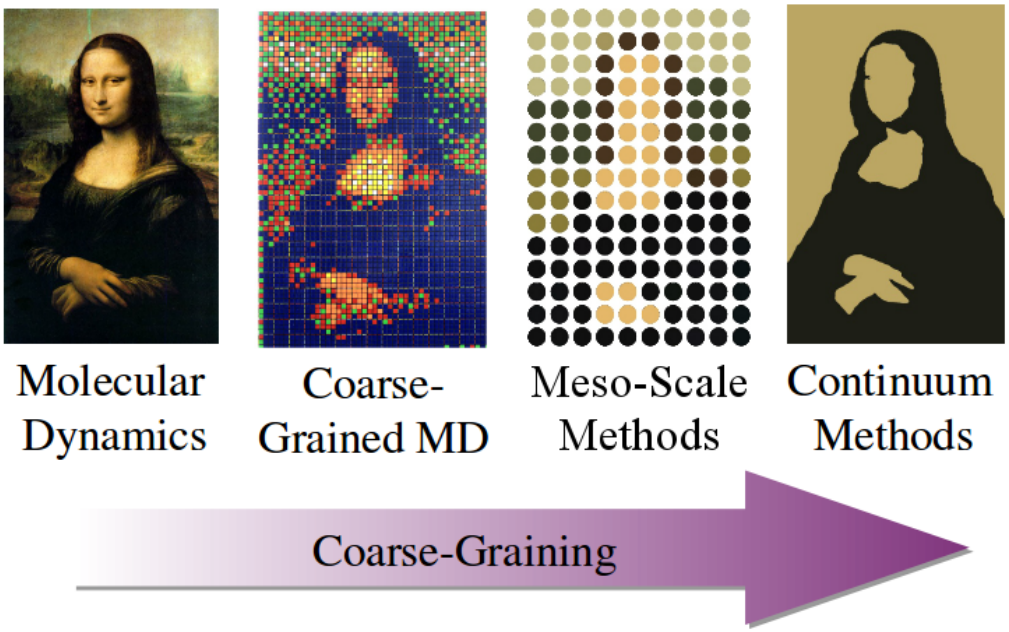


Figure 2.1: Schematic of simulation scales. The representation on the left has the most detail however would take a long time to paint, whereas the representation on the right still has the general features but would be a lot faster to produce.

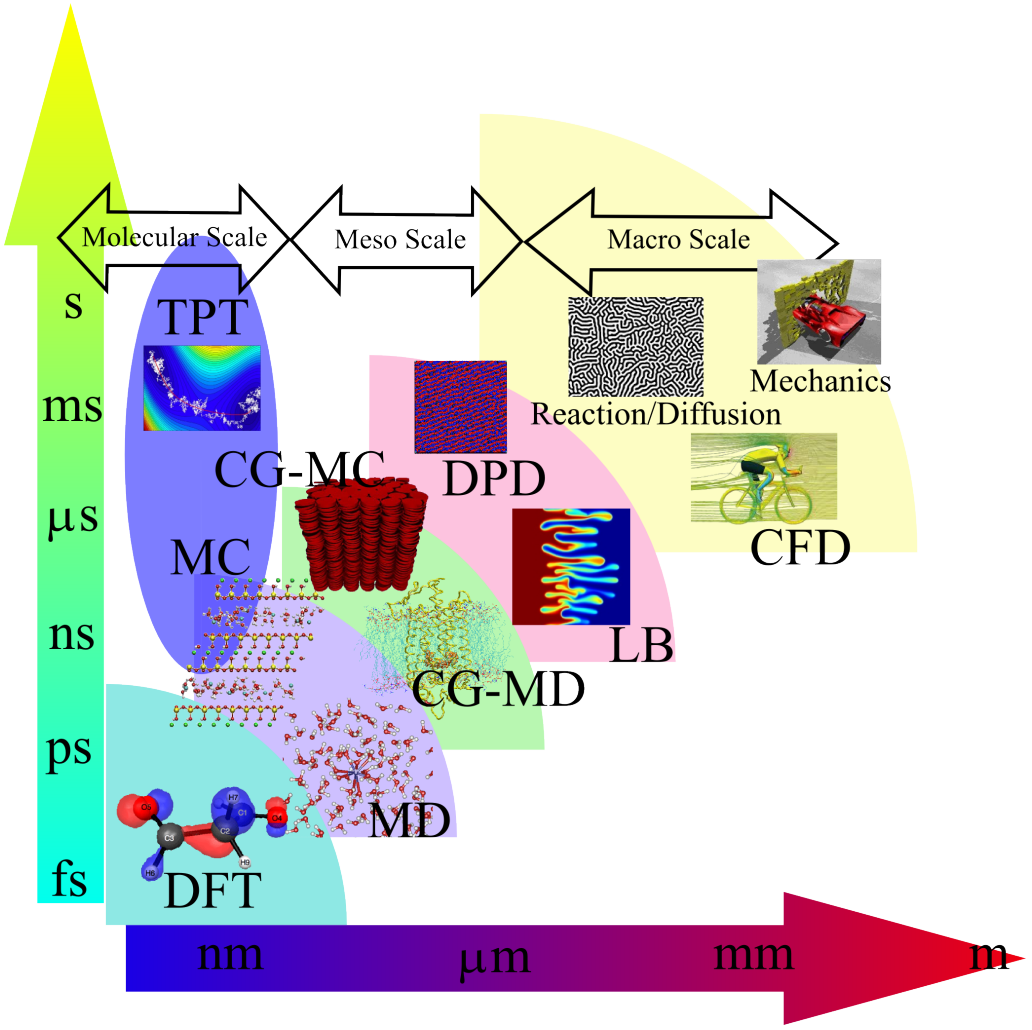


Figure 2.2: Examples of soft matter simulation techniques and their time- and size-scales.

2.3 Statistical Mechanics

When dealing with computer simulations only a small number of molecules are simulated. Chemical thermodynamics on the other hand deals with bulk properties of matter, typically of the order of 10^{23} particles. There should be some link between these two sets of particles. If 10^{23} particles are modelled then the number of interactions, positions, and velocities would be astronomical. Even if this were possible then a method to relate all of the molecular information to the bulk properties.

The pressure exerted by a gas on the container wall depends on the rate at which the particles collide with the wall. It is not necessary to know which particle underwent a particular collision. What is needed is the root mean square speed of the particles and their standard deviation about this mean. This is were statistical thermodynamics is useful. If the pressure exerted on the walls is measured many times at regular time intervals then the mean, $\langle P \rangle$, and standard deviation, σ_P, can be given by equation 2.3.1.

$$\begin{aligned} \langle P \rangle &= \frac{1}{n} \sum_{i=1}^{n} P\left(t_i\right) \\ \sigma_P &= \sqrt{\frac{1}{n} \sum_{i=1}^{n} \left(P\left(t_i\right) - \langle P \rangle\right)^2} \end{aligned} \tag{2.3.1}$$

The larger the number of samples the closer the sample mean will be to the true mean and the smaller the sample deviation will become.

Consider an ensemble of N^* cells, comprising the original cell and N^*-1 replications. Energy my flow between the cells but the total ensemble energy is constant. Suppose that the possible total energies of the N particles contained in each cell are E_1^*, E_2^*, and so on. If a snapshot is taken there is a distribution of energies amongst the cells as follows:

N_1^* cells have energy E_1^*,

N_2^* cells have energy E_2^*, etc.

According to Boltzmann, the E^* and the N^* are related by equation 2.3.2.

$$\frac{N_i^*}{N^*} = \frac{\exp\left(-\frac{E_i^*}{kT}\right)}{\sum_i \exp\left(-\frac{E_i^*}{kT}\right)} \tag{2.3.2}$$

The denominator is called the partition function and is generally given by Q as in equation 2.3.3, this is a simplification and strict ensemble partition function is dependent on the thermodynamic ensemble as shown in Section 2.4.

$$Q = \sum_i \exp\left(-\frac{E_i^*}{kT}\right) \tag{2.3.3}$$

Some useful thermodynamic properties are:

Helmholtz energy:	$A = U - TS$
Gibbs free energy:	$G = H - TS = A + PV$
Enthalpy:	$H = U + PV$
Entropy:	$S = -\left(\frac{\partial A}{\partial T}\right)_{N,V}$
Pressure:	$P = -\left(\frac{\partial A}{\partial V}\right)_{N,T}$
Constant volume heat capacity:	$C_v = \left(\frac{\partial U}{\partial T}\right)_{N,V}$
Constant pressure heat capacity:	$C_p = \left(\frac{\partial H}{\partial T}\right)_{N,P}$

2.4 Thermodynamic Ensembles

Statistically speaking, the method by which macroscopic properties can be obtained from a simulation is by the averaging of (in theory) an infinite number of microscopic configurations, each of which can exist with certain probability. In statistical mechanics, this collection of micro-states is usually referred to as an ensemble. The ensemble also describes the probability of seeing each region in phase space with an accompanying partition function. Considered together, this distribution of micro-states can be used to model a system, averaging the properties of a large number of micro-states to find properties which could be correlated to the thermodynamic state of a real system. In practice however, it is impossible to simulate an infinite number of states, and so averages must be calculated from a finite sample of micro-states. Each ensemble can be defined from a real life experimental situation with particular rules and limits. This allows each ensemble of states to be connected with a thermodynamic property, which can be calculated as a function of the partition function.

2.4.1 Canonical (NVT)

The canonical ensemble, describes a system which cannot exchange mass (i.e. number of molecules N) or volume V with the outside world, but exchanges energy such that the temperature T of the system is uniform and constant. The canonical partition function, which describes the distribution of energies of the microstates is given by an expression taking into account both the kinetic and potential energies, equation 2.4.1, where Λ is the thermal de Broglie wavelength.

$$\begin{aligned} Q_{\mathrm{NVT}} &= \frac{1}{N!\Lambda^{3N}} \int \mathrm{d}\,\mathbf{p}^N \mathrm{d}\,\mathbf{r}^N \exp\left[-\beta\mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)\right] \\ &= \sum_i \exp\left(-\frac{E_i}{kT}\right) \end{aligned} \tag{2.4.1}$$

The relationship between the partition function and the thermodynamic potential of the ensemble is the Helmholtz free energy, A, equation 2.4.2.

$$A = U - TS = -kT \ln Q \tag{2.4.2}$$

2.4.2 Grand-canonical (μVT)

The grand-canonical ensemble is similar to the canonical ensemble in the assumption of a system of constant temperature and volume which can pass energy into and out of the system. Additionally however, a grand canonical system may pass mass into and out of the system to keep the chemical potential, μ, constant. As a result the partition function is now also dependent on the number of particles, N, in the system, equation 2.4.3.

$$\begin{aligned} Q_{\mu\mathrm{VT}} &= \sum_{N=1}^{\infty} \frac{1}{N!\Lambda^{3N}} \exp\left(\beta N\mu\right) \int \mathrm{d}\,\mathbf{p}^N \mathrm{d}\,\mathbf{r}^N \exp\left[-\beta\mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)\right] \\ &= \sum_i \exp\left(-\frac{E_i - \mu N_i}{kT}\right) \end{aligned} \tag{2.4.3}$$

The thermodynamic potential of the grand-canonical ensemble can be described by equation 2.4.4.

$$\Omega_{\mu\mathrm{VT}} = A - N\mu = -PV = -kT \ln Q \tag{2.4.4}$$

2.4.3 Microcanonical (NVE)

The micro-canonical ensemble keeps system mass and volume constant. Where it differs from all other ensembles is that the system energy is controlled by changing the temperature of the system. The NVE ensemble's main advantage was that it does not require random numbers. Since modern everyday computers are capable of running effective random number generators in negligible time, this ensemble is hardly ever used for molecular simulation. It can however be useful during the preparation of an initial configuration for a simulation run in a different ensemble.

2.4.4 Isothermal-Isobaric (NPT)

The isothermal-isobaric ensemble is used to describe a system in which the temperature and number of molecules are constant as in the canonical ensemble. However, the important feature of the NPT ensemble is that the pressure of the system is maintained to

a constant value by changing the volume of the system. This ensemble often provides a more realistic simulation environment than the canonical and grand-canonical ensembles. The partition function, which includes a pressure-volume relationship PV is described by equation 2.4.5.

$$Q_{\text{NPT}} = \frac{\beta P}{N!\Lambda^{3N}} \int_0^\infty \mathrm{d}\,\mathbf{V}\mathbf{V}^N \exp\left[-\beta P\mathbf{V}\right] \int \mathrm{d}\,\mathbf{p}^N \mathrm{d}\,\mathbf{r}^N \exp\left[-\beta \mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)\right] \quad (2.4.5)$$

2.4.5 Other Ensembles

Isobaric-Isenthalpic (NPH)

This ensemble is characterised by a constant number of atoms, N, the pressure, P, and the enthalpy, H.

Gibbs Ensemble

The Gibbs ensemble is used to simulate phase coexistence, usually two liquid phases or a liquid and a vapour phase. This is accomplished by the use of two simulation boxes, each held at a constant temperature. The boxes may exchange both mass and volume between each other, but in such a way that there is no change in mass or volume for the overall system.

Semi-grand Canonical Ensemble

The semi-grand canonical ensemble keeps N_1, μ_2, V, and T constant. Its main use is for the study of solvent-solute systems, where the number of solvent molecules N_1 is kept constant, but the amount of solute N_2 is allowed to fluctuate, keeping μ_2 constant.

2.5 Monte Carlo Simulations

The Monte Carlo (MC) method is a technique to try to sample the infinite number of available configurations available for a system. Pure MC involves selecting a configuration at random, and then assigning a weight to a configuration $\mathbf{r}_N$ based on the probability p of the system existing in that state. MC is a generic term applied to calculations that involve the use of random numbers for sampling.

Consider a circle inscribed in a unit square. Given that the circle and the square have a ratio of areas that is $\pi/4$, the value of π can be approximated using the Monte Carlo method:

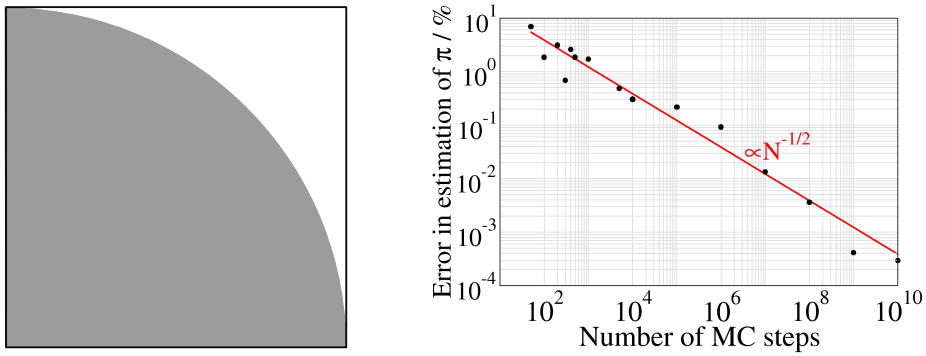
1. Draw a square on the ground, then inscribe a circle within it
2. Uniformly scatter some objects of uniform size (grains of rice or sand) over the square
3. Count the number of objects inside the circle and the total number of objects
4. The ratio of the number in the circle to the total number is an estimate of the ratio of the two areas, which is $\pi/4$

In this procedure the domain of inputs is the square that circumscribes the circle. Random inputs are generated by scattering grains over the square then a computation is performed on each input (test whether it falls within the circle). If grains are purposefully

dropped into only the centre of the circle, they are not uniformly distributed, so the approximation is poor. Second, there should be a large number of inputs, the approximation is generally poor if only a few grains are randomly dropped into the whole square. On average, the approximation improves as more grains are dropped.

This same example can be performed using a computational code, where random numbers are generated representing the coordinates of the dropped objects. This a more efficient way to perform multiple trials. The program used can be seen in Code example 2.1.

Figure 2.3 shows the setup of the circle within the square and the error to the value of π calculated with different numbers of samples.



(a) (b)

Figure 2.3: Monte Carlo example calculation for π (a) the quarter circle in the square and (b) the error in calculation of π with number of trials.

From the classical expression for the partition function, Q, equation **??** where $\mathbf{r}^N$ is the coordinates of all N particles and $\mathbf{p}^N$ is the corresponding momenta. The function $\mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)$ is the Hamiltonian of the system, which expresses the total energy of an isolated system as a function of the coordinates and the momenta: $\mathcal{H} = \mathcal{K} + \mathcal{U}$, where $\mathcal{K}$ is the kinetic energy and $\mathcal{U}$ is the potential energy. Therefore the classical ensemble for an observable $\mathcal{A}$ can be given by equation 2.5.1, where $\beta = 1/kT$.

$$\langle \mathcal{A} \rangle = \frac{\int \mathrm{d}\,\mathbf{p}^N \mathrm{d}\,\mathbf{r}^N \mathcal{A}\left(\mathbf{p}^N, \mathbf{r}^N\right) \exp\left[-\beta \mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)\right]}{\int \mathrm{d}\,\mathbf{p}^N \mathrm{d}\,\mathbf{r}^N \exp\left[-\beta \mathcal{H}\left(\mathbf{p}^N, \mathbf{r}^N\right)\right]} \tag{2.5.1}$$

In most cases equation 2.5.1 cannot be calculated analytically and must be calculated numerically. In fact for even very small systems $N\left(\mathcal{O}\left(100\right)\right)$ it is almost impossible to perform simple numerical integration, e.g. Simpson's rule. This is where use of the Monte Carlo method helps. The simplest Monte Carlo technique is that of random sampling (the method used for the calculation of π example) which samples with an even distribution through the configurational space. This is fine for simple situations, but for complex situations many of the computed values produce a Boltzmann factor that is negligible. For example, for a fluid of 100 hard spheres at the freezing point, the Boltzmann factor would be non-zero for 1 out of every 10^{260} configurations. Therefore, it makes sense to sample points distributed non-uniformly over the configurational space. This is where importance sampling is useful.

The simple one dimensional integral $I = \int_a^b \mathrm{d}\,x\, f\left(x\right)$ can be written as equation 2.5.2, where $\langle f\left(x\right)\rangle$ is the unweighted average of $f\left(x\right)$ over the interval $[a, b]$.

$$I = (b - a)\left\langle f\left(x\right)\right\rangle \tag{2.5.2}$$

```
PROGRAM MC
  IMPLICIT NONE

  INTEGER :: i   ! counter
  INTEGER :: n   ! number of MC steps
  INTEGER :: in  ! number inside circle
  REAL :: pos(2) ! random positions
  REAL :: pi

  WRITE(*,'(A)') "How many MC steps would you like?"
  READ(*,*) n

  in=0  ! set inital number inside circle to 0

  DO i=1,n
     CALL RANDOM_NUMBER(pos)  ! generate the random position
     ! Check to see if random position is inside the circle
     IF (pos(1)**2+pos(2)**2.le.1) THEN
        in=in+1
     END IF
  END DO

  pi=4.d0*real(in)/real(n) ! calculate pi

  WRITE(*,'(A,F8.6)') "The value of pi is ", pi

END PROGRAM MC
```

Code Example 2.1: Example code for the calculation of πusing a simple Monte Carlo program

In brute force Monte Carlo the average value is determined at a large number of x values randomly distributed over the range $[a, b]$. As this number of samples tends to infinity the correct value is produced for I. Now if the samples are distributed nonuniformly over the interval according to a non-negative probability density $w(x)$ then the integral solution can be written as equation 2.5.3.

$$I = \int_a^b \mathrm{d}\, x\, w(x) \frac{f(x)}{w(x)} \tag{2.5.3}$$

If $w(x)$ is the derivative of another function $u(x)$ with $u(0) = a$ and $u(0) = b$ then equation 2.5.3 can be normalised as equation 2.5.4.

$$I = \int_0^1 \mathrm{d}\, u \frac{f[x(u)]}{w[x(u)]} \tag{2.5.4}$$

If L random values of u are selected uniformly distributed in the interval $[0, 1]$ then I can be approximated as equation 2.5.5.

$$I = \frac{1}{L} \sum_{i=1}^{L} \frac{f[x(u_i)]}{w[x(u_i)]} \tag{2.5.5}$$

2.5.1 The Metropolis Method

The Metropolis implementation of the MC method [1] is the method that is commonly used in computational chemistry. It requires an algorithm for generating a new configuration by changing a previous configuration. Any change in the configuration or atomic

ensemble will give a change in energy. If the energy change is large and positive this may mean that the configuration is unrepresentative of the state the system is likely to be in; for instance the configuration overlaps, resulting in a very small Boltzmann weight. In order to sample more efficiently it is necessary to bias the sampling of microstates to significantly reduce the sampling of unphysical configurations. Metropolis [2] developed a sampling method where, instead of sampling states with an equal probability and designating the configuration a Boltzmann weight, the states would be generated from a probability distribution designated by the Boltzmann factor $\exp[-\beta E]$ and all assigned an equal weight [3]. This is achieved by; If the energy change ΔE is negative, leading to a lowering of the energy, the change is automatically accepted, on the other hand, if the energy change is positive, the configuration is only accepted with probability $\exp[-\beta \Delta E]$.

This procedure is repeated for a large number of steps, leading to an evolution of the ensemble through the multi-dimensional phase space. This approach ensures that the sampling procedure is consistent with thermodynamics; for example, states with any energy, E, occur with the relative probability $\exp[E/kT]$. The complete set of configurations, including duplicated configurations, can be analysed to give averaged quantities that have the correct thermodynamic weighting automatically ensured.

2.6 Molecular Dynamics

Molecular dynamics is a deterministic method of simulating the dynamic evolution of a molecular system. Since it is time dependant it has the distinct advantage over Monte Carlo of being useful to measure time-dependant quantities. The trajectory of a system is calculated by the solution of classical equations of motion [4]. Newton's equations of motion for a simple system are given by equation 2.6.1, $\mathbf{r}_i$ is the atomic coordinates and m_i is the mass of particle i, t is time, and $\mathbf{V}$ is the potential energy.

$$m_i \frac{\partial^2 \mathbf{r}_i}{\partial t^2} = -\nabla \mathbf{V}_i \tag{2.6.1}$$

The velocity of each particle is therefore given by equation 2.6.2, the acceleration by equation 2.6.3, and the kinetic energy by equation 2.6.4.

$$\mathbf{v}_i = \frac{\mathrm{d}\,\mathbf{r}_i}{\mathrm{d}\,t} \tag{2.6.2}$$

$$\mathbf{a}_i = \frac{\mathrm{d}\,\mathbf{v}_i}{\mathrm{d}\,t} = \frac{\mathbf{F}_i}{m_i} \tag{2.6.3}$$

$$KE_i = \frac{m_i}{2} \mathbf{v}_i \cdot \mathbf{v}_i \tag{2.6.4}$$

2.6.1 Integration Algorithms

In a system with a large number of particles the coupled differential equations 2.6.1 and 2.6.2 cannot be solved analytically, therefore must be solved numerically. Several different algorithms have been developed to integrate the equation of motion:

Verlet algorithm This algorithm was developed by Loup Verlet in 1967 [5] and became the base for most algorithms used in modern simulations. The algorithm calculates

positions forward and backward in time allowing a better calculation of the new positions than a simple single step numerical integration. This is performed via a third order Taylor expansion, equation 2.6.5.

$$\begin{aligned}\mathbf{r}\left(t+\delta t\right) &= \mathbf{r}\left(t\right)+\mathbf{v}\left(t\right)\delta t+\frac{1}{2}\mathbf{a}\left(t\right)\delta t^{2}+\frac{1}{6}\dddot{\mathbf{r}}\left(t\right)\delta t^{3}+\mathcal{O}\left(\delta t^{4}\right)\\ \mathbf{r}\left(t-\delta t\right) &= \mathbf{r}\left(t\right)-\mathbf{v}\left(t\right)\delta t+\frac{1}{2}\mathbf{a}\left(t\right)\delta t^{2}-\frac{1}{6}\dddot{\mathbf{r}}\left(t\right)\delta t^{3}+\mathcal{O}\left(\delta t^{4}\right)\end{aligned} \tag{2.6.5}$$

Adding two above expressions together yields, equation 2.6.6.

$$\mathbf{r}\left(t+\delta t\right)=2\mathbf{r}\left(t\right)-\mathbf{r}\left(t-\delta t\right)+\mathbf{a}\left(t\right)\delta t^{2}+\mathcal{O}\left(\delta t^{4}\right) \tag{2.6.6}$$

The main problem with this algorithm is that the velocities are not directly generated, though it makes it computationally cheaper as lower data storage is required.

The leap-frog algorithm The leap frog algorithm is a modification of the classical Verlet algorithm. It calculates velocities, $\mathbf{v}$, at a half time $\delta t/2$ step in advance of the reference time, t, equation 2.6.7.

$$\mathbf{v}\left(t+\frac{\delta t}{2}\right)=\mathbf{v}\left(t-\frac{\delta t}{2}\right)+\mathbf{a}\left(t\right)\delta t \tag{2.6.7}$$

These are then used to calculate the position at the full time step, δt, equation 2.6.8.

$$\mathbf{r}\left(t+\delta t\right)=\mathbf{r}\left(t\right)-\mathbf{v}\left(t+\frac{\delta t}{2}\right)\delta t \tag{2.6.8}$$

In this method the velocities leap over positions, which in turn leap over velocities. The advantage of this method is that velocities are calculated accurately. However, they are not calculated at the same time as positions. So if velocities at a given time, t, are required they must be averaged as equation 2.6.9.

$$\mathbf{v}\left(t\right)=\frac{1}{2}\left[\mathbf{v}\left(t-\frac{\delta t}{2}\right)+\mathbf{v}\left(t+\frac{\delta t}{2}\right)\right] \tag{2.6.9}$$

This algorithm is easily applied to complex molecules and may also be combined with constraint algorithms. The speed and simplicity of this algorithm means it often makes it the first choice for molecular simulation.

The velocity Verlet algorithm This is another modification of the Verlet algorithm, allowing for the calculation of positions, velocities and accelerations at a given time [6]. This results in more accurate trajectories and less drift of the total energy. The increase in precision also comes with increase in computational cost.

Beeman's algorithm Beeman's algorithm is also related to Verlet algorithm. Beeman's algorithm [7] uses more accurate expressions, that allows for more accurate velocities and provides better energy conservation. This method is computationally more expensive and is only used when high accuracy is required.

2.6.2 Constraints

For a realistic molecular model, many degrees of freedom need to be accounted for[1]. Internal degrees of freedom in the molecule vary in their frequencies, from about 10^8 to

[1] A molecule with N atoms has $3N-6$ normal modes of vibration.

10^{14} Hz (time periods of 1 ns to 1 fs), with the slowest being breathing modes of large molecules (e.g. proteins) and fastest being individual bond stretching. The latter limits the size of the time step as for accurate calculations the bond stretching should be samples between 10-100 times per stretch, i.e. a maximum time step of 100-10 as. As the properties that are being simulated are typically on a time scale larger that the bond stretching it is common practise to restrain the bond length (and occasionally bond angles within the molecule). This allows longer times steps of around 2 fs to be used. The typically used constrain algorithms are:

SHAKE was the first constraint algorithm developed and can be implemented as a modification of the Verlet and leap-frog algorithms where the calculated velocities, $\mathbf{v}(r)$, are modified to agree with the constraints and preserve bond (and rarely angle) lengths [8]. A parallel version of SHAKE called RD-SHAKE is implemented in the DL_POLY package.

LINKS is a newer algorithm [9] (1997), based on the EEM(Edberg, Evans and Morriss) algorithm [10], developed in 1986. LINCS has now been shown to be faster and more stable then SHAKE and, in addition, a parallel version has recently become available [11].

RATTLE is a modification of SHAKE algorithm for use with velocity Verlet algorithm [12]

SETTLE solves the system of non-linear equations analytically for 3 constraints in constant time. Although it does not scale to larger numbers of constraints, it is very often used to constrain rigid water molecules, which are present in almost all biological simulations and are usually modelled using three constraints [13].

2.6.3 Simulation Box

System sizes accessible to atomistic simulations are of the order of thousands of atoms. These can be typically enclosed in a cubic box with 15 nm sides. The molecules on the surface of the box will experience different forces to the ones in the centre. This can be solver by imposing periodic boundary condition, i.e. surrounding the simulation box with 'pseudo' simulation boxes, being a clone of the original one. This results in a simulation with bulk properties across the whole original box. When the molecule leaves the box on one side, as shown on the figure 2.4, it will enter the neighbouring 'pseudo' box, and by such re-enter the initial box from the opposite side.

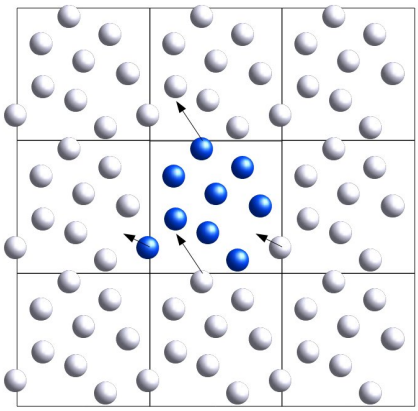


Figure 2.4: Periodic boundary condition, two-dimensional illustration: the box containing blue atoms is the simulation box, while the boxes containing white atoms are 'pseudo' boxes, that mimic the behaviour of the main simulation box.

To be able to perform calculations at the reasonable speed, the interaction of the molecule is only calculated with it's direct neighbours. The standard way is to use a Verlet list [14] that, for every particle, defines a surrounding shell at a distance slightly larger

than the cut-off radius for the interactions. The cut-off distance should be less or equal to the half of the box size.

2.6.4 Temperature Coupling

Temperature coupling is used to control the temperature of the system. The temperature of the system is related to the average kinetic energy, so the temperature can be controlled by scaling velocities. There are a number of methods developed for this purpose:

Berendsen Thermostat This algorithm mimics a weak coupling of the system to an external heat bath at a given temperature, T_0 [15]. Deviation of the system temperature from the T_0 is corrected by, equation 2.6.10, where τ is a time constant.

$$\frac{\mathrm{d}\,T}{\mathrm{d}\,t} = \frac{T_0 - T}{\tau} \tag{2.6.10}$$

This method does not generate a proper canonical ensemble, as it suppresses the fluctuation of the kinetic energy. This will effect the results only for very small systems or when calculating the kinetic energy itself. The velocities are rescaled by a scaling factor, λ, equation 2.6.11, where τ_T is close to but not equal to τ.

$$\lambda = \left[1 + \frac{\Delta t}{\tau_t}\left(\frac{T_0}{T\left(t - \frac{\Delta t}{2}\right)} - 1\right)\right]^{\frac{1}{2}} \tag{2.6.11}$$

In practice the scaling factor is limited to avoid too large scaling at an individual time step, which can lead to the calculation crashing.

Velocity Rescaling Thermostat This algorithm is similar to Berendsen thermostat, but it allows one to produce the correct ensemble by addition of a random force [16].

Nose-Hoover Thermostat This algorithm has the advantage that it can reproduce a correct canonical ensemble [17, 18]. The other thermostats above are weak coupling and are used for relaxing systems to a desired temperature, whereas this thermostat can be used for the final simulations. In this method, a thermal reservoir and friction coefficient are introduced into the equation of motion, and thereby allow kinetic energy changes in the system.

2.6.5 Pressure coupling

Similarly to temperature coupling, the system can be coupled to a pressure bath. There are a number of methods developed for this purpose:

Berendsen Barastat This algorithm rescales the coordinates and box vectors at every step towards a given reference pressure, P_0, equation 2.6.12.

$$\frac{\mathrm{d}\,P}{\mathrm{d}\,t} = \frac{P_0 - P}{\tau_P} \tag{2.6.12}$$

Rescaling is performed by applying the scaling matrix, μ, equation 2.6.13, where β is the isothermal compressibility of the system.

$$\mu = \delta_{ij} - \frac{\Delta t}{3\tau_P}\beta_{ij}\left[P_{0ij} - P_{ij}\left(t\right)\right] \tag{2.6.13}$$

Like the Berendsen thermostat, the barastat is a weak coupling so does not generate a proper thermodynamic ensemble.

Parrinello-Rahman Barastat This approach is similar to the Nose-Hoover for temperature coupling, where equation of motion for particles is updated at every step [19], and a correct thermodynamic ensemble is produced.

2.7 References

[1] N. Metropolis and S. Ulam. The Monte Carlo Method. *Journal of the American Statistical Association*, 44:335–341, 1949.

[2] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, and A. H. Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21:1087, 1953.

[3] R. J. Sadus, editor. *Molecular Simulation of Fluids: Theory, Algorithms and Object-Orientation*. Elsevier, 2002.

[4] B. J. Alder and T. E. Wainwright. Studies in Molecular Dynamics. I. General Method. *The Journal of Chemical Physics*, 31(2):459–466, 1959.

[5] Loup Verlet. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, 159:98–103, Jul 1967.

[6] W. C. Swope, H. C. Anderson, P. H. Berens, and K. R. Wilson. A Computer Simulation Method for the Calculation of Equlibrium Constants for the Formation of Physical Clusters of Molecules: Applications to Small Water Clusters. *Journal of Chemical Physics*, 76:637–649, 1982.

[7] D. and Beeman. Some multistep methods for use in molecular dynamics calculations. *Journal of Computational Physics*, 20(2):130 – 139, 1976.

[8] Jean-Paul Ryckaert, Giovanni Ciccotti, and Herman J.C Berendsen. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of Computational Physics*, 23(3):327 – 341, 1977.

[9] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, 1997.

[10] Roger Edberg, Denis J. Evans, and G. P. Morriss. Constrained molecular dynamics: Simulations of liquid alkanes with a new algorithm. *The Journal of Chemical Physics*, 84(12):6933–6939, 1986.

[11] Berk Hess. P-LINCS: A Parallel Linear Constraint Solver for Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(1):116–122, 2008.

[12] D. and Beeman. Some multistep methods for use in molecular dynamics calculations. *Journal of Computational Physics*, 20(2):130 – 139, 1976.

[13] Shuichi Miyamoto and Peter A. Kollman. Settle: An analytical version of the SHAKE and RATTLE algorithm for rigid water models. *Journal of Computational Chemistry*, 13(8):952–962, 1992.

[14] Loup Verlet. Computer "Experiments" on Classical Fluids. I. Thermodynamical Properties of Lennard-Jones Molecules. *Phys. Rev.*, 159:98–103, Jul 1967.

[15] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684–3690, 1984.

[16] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of Chemical Physics*, 126(1):014101, 2007.

[17] Shuichi NosÃl' and M.L. Klein. Constant pressure molecular dynamics for molecular systems. *Molecular Physics*, 50(5):1055–1076, 1983.

[18] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *The Journal of Chemical Physics*, 81(1):511–519, 1984.

[19] M. Parrinello and A. Rahman. Polymorphic transitions in single crystals: A new molecular dynamics method. *Journal of Applied Physics*, 52(12):7182–7190, 1981.

Chapter 3

Molecular Interactions

Contents

3.1 Introduction

When molecules are near each other, they can influence one another. Therefore, the balance between the forces of attraction and repulsion needs to be understood. Such forces must exist in reality otherwise there would be nothing to bring molecules together to form liquid and solid states. People have speculated about these intermolecular forces since the concept of atoms and molecules first existed. The present theory is that molecules attract at long range but repel strongly at short range.

3.2 Molecular Force Fields

Actual interactions between atoms is in the realm of physical chemistry. For computational simulations, these interactions are too complex to simulate systems larger than a couple of molecules on the current computers available. Therefore, it is common practise to represent these real interactions as a computational force field. The choice of computational method is dependent on the size of the system and the time scales of calculations. For systems with up to several hundred to several thousand atoms, the molecules can be represented as individual atoms with each atom represented by a sphere. There are many types of force field, but they all follow a similar pattern:

- Nuclei and electrons are combined in an atom, represented by a ball.
- A ball has a radius and a constant charge. The ball can vary in softness.
- Bonds are represented by springs.
- Springs have an equilibrium length and can vary in stiffness.

The interactions between these spheres form the molecular force field. The interactions are described by pre-assigned parameters, that are system dependent. These parameters can be obtained from experiment and/or higher level computational calculations, such as *ab-initio*. The force field is the set of parameters used in classical mechanical calculations. The potential energy of the system is described by the sum of all the interactions within the system, equation 3.2.1, where E_{str} is the energy function for stretching a bond between two atoms, E_{bend} is the energy required for bending an angle between three atoms, E_{tors} is the torsional energy for rotation around a bond, E_{VdW} is the Van der Waals energy, E_{el} is the electrostatic energy, and E_{Xterms} represents any coupling between the above terms.

$$E_{FF} = E_{str} + E_{bend} + E_{tors} + E_{VdW} + E_{el} + E_{Xterms} \tag{3.2.1}$$

The first three terms represent the bonded interactions. The electrostatic and Van der Waals energies are non bonded interactions between atoms. To be able to calculate the force field energy it is necessary to know atomic coordinates, geometries, and the relative energies of the atoms in the system.

The calculation of the total potential energy is dominated by the number of non-bonded interactions. The time for calculation of the bonded interactions increase on the order of $\sim N$, while the time for calculation of the non-bonded interactions increase on the order of $\sim N^2$.

3.2.1 Stretching energy

The stretching energy of a bond is the energy of the extension of the spring, bonding two atoms, Figure 3.1. The shape of the energy curve is well described by a Taylor expansion, equation 3.2.2, where r_0 is an equilibrium bond length for the given system and so corresponds to the minimal energy.

$$E_{str}\left(r^{AB}-r_0^{AB}\right)=E\left(r_0\right)+\frac{\mathrm{d}\,E}{\mathrm{d}\,r}\left(r^{AB}-r_0^{AB}\right) + \frac{1}{2}\frac{\mathrm{d}^2\,E}{\mathrm{d}\,r^2}\left(r^{AB}-r_0^{AB}\right)^2+\frac{1}{6}\frac{\mathrm{d}^3\,E}{\mathrm{d}\,r^3}\left(r^{AB}-r_0^{AB}\right)^3+\mathcal{O}\left(r^4\right) \quad (3.2.2)$$

At $r = r_0$, $E(r_0)$ is zero and is used as the zero point on the energy scale, the second term is a first order derivative that is also close to zero. The cubic term will drive the energy of the bond to negative value at infinity, so the formulation is often simplified to an harmonic interaction, equation 3.2.3, where k^{AB} is the force constant for stretching the bond between atoms A and B and will change depending on the system.

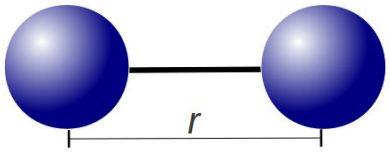


Figure 3.1: Model of diatomic molecule with bond length r_0.

$$E_{str}=\frac{1}{2}k^{AB}\left(r^{AB}-r_0^{AB}\right)^2 \quad (3.2.3)$$

Approximation to harmonic form is sufficient for most molecular dynamics calculations. For calculations of bond length dependent parameters, such as vibrational frequencies, or when the starting configuration is much different from equilibrium, terms up to the forth order may be included. The other common choice is the Morse potential [1], equation 3.2.4, where $\alpha=\sqrt{k/2D^{AB}}$ and D^{AB} is the bond dissociation energy, i.e. the depth of the well.

$$E_{str}=D^{AB}\left(1-\exp\left(-\alpha\left(r^{AB}-r_0^{AB}\right)^2\right)\right) \quad (3.2.4)$$

The Morse potential reproduces stretching well over a wide range of distances, at long distances tends to zero. Hence the Morse potential is slow in bringing atoms to equilibrium bond lengths, when the initial geometry has a large bond length. The comparison between harmonic and Morse potentials is shown in Figure 3.2.

The stretching energy has the highest contribution to overall energy of the system, followed by the bending energy.

3.2.2 Bending Energy

Similarly to the stretching energy, bending energy can also be expressed by a Taylor expansion, that can then be approximated to a harmonic potential, equation 3.2.5, where θ^{ABC} is an angle between three atoms *A, B* and *C*, as shown in Figure 3.3, and k^{ABC} is the force constant for the bending between three given atoms.

$$E_{bend}=\frac{1}{2}k^{ABC}\left(\theta^{ABC}-\theta_0^{ABC}\right)^2 \quad (3.2.5)$$

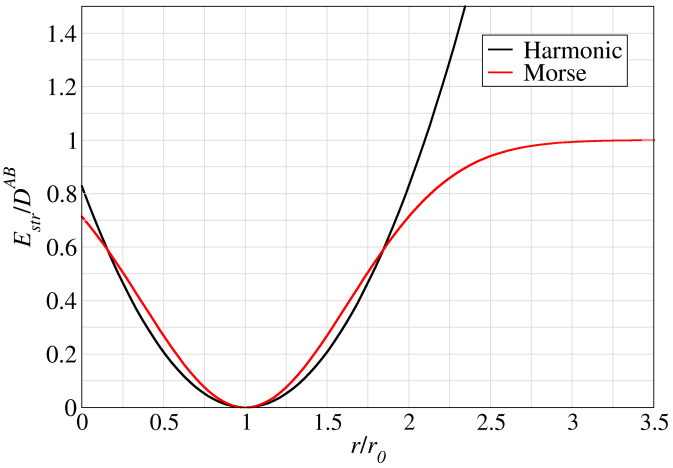


Figure 3.2: Bond stretching energy. Harmonic potential compared to the Morse potential.

When a higher accuracy is required for the calculation, the next highest order term in the expansion can be included, or a series of harmonic potentials can be used.

Changes in the bending energy are usually smaller than the changes in stretching energy, so less energy is required to distort bond angles in comparison to bond length. Hence, force constants for bending are proportionally smaller than those for stretching.

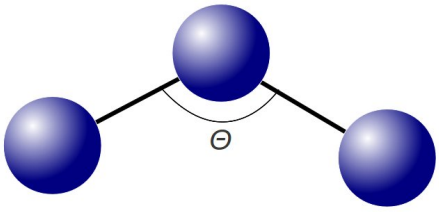


Figure 3.3: An angle between three bonded atoms.

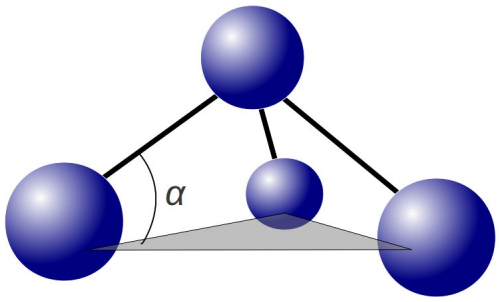


Figure 3.4: An illustration of out-of-plane bending, created by three atoms bonded to $sp2$ hybridised central atom.

When the central atom B is bonded to more than two other atoms, as shown on the Figure 3.4, a pyramidal structure is formed. This angle can no longer be calculated in the manner discussed above. A high force constant should be used to penalise out-of-plane bending, this will make planar structures very inflexible. By introducing a new term, the flexibility of the planar structure is maintained while out-of-plane movement is also taken into account. This bending is often referred to as the improper torsional energy. This also often takes the form of a harmonic potential, equation 3.2.6.

$$E_{improper} = \frac{1}{2}k^B(\alpha)^2 \tag{3.2.6}$$

3.2.3 Torsional Energy

Four aligned atoms frequently will rotate along the central bond, as shown in Figure 3.5.

Rotation along the bond is continuous, so when the bond rotates by $360°$ the energy should return to the initial value, making torsional energy periodic. The potential is well

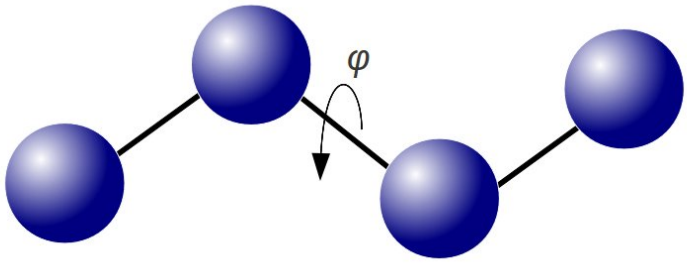


Figure 3.5: Torsional / dihedral angle, created by four linearly aligned atoms.

represented by the Fourier series, equation 3.2.7, where n is a periodicity term, depending on the allowed rotation it can be $n = 1$ for full $360°$ rotation, $n = 2$ for $180°$ periodicity, $n = 3$ for $120°$ periodicity, etc. V_n is a constant determining the barrier for the rotation, V_n is not equal to zero for allowed periodic rotations, φ is the torsional angle (also called the dihedral angle) and is shown on the Figure 3.5.

$$E_{tors} = \sum_n V_n \cos\left(n\varphi - \varphi_0\right) \tag{3.2.7}$$

Figure 3.6 shows a plot of the two torsional potentials with $n = 3$, $V_n = 4$ (dashed line), $n = 2$, $V_n = 3$ (solid line) and the corresponding total torsional potential in red if the two are combined.

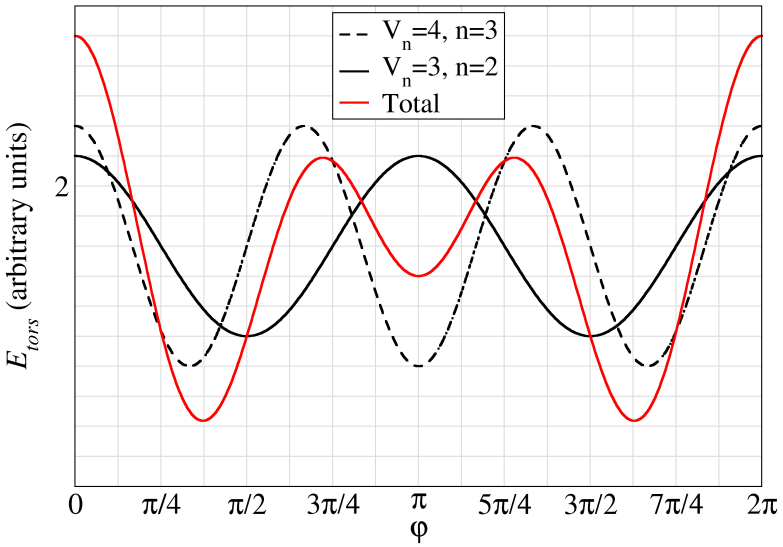


Figure 3.6: Variation of torsional potential for different values of n and V_n and the total potential (red).

3.2.4 Van der Waals

In addition to the bonded interactions described above, there are non bonded interactions in the system. These interactions appear between atoms in the same or neighbouring molecule and are normally described by two potentials, representing the Van der Waals energy and the electrostatic energy.

The Van der Waals energy describes the repulsion and attraction between non bonded atoms. At large interatomic distances the Van der Waals energy goes to zero, whereas at short distances it is positive and very repulsive. This mimics the overlap of the negatively charged electronic clouds of atoms. At intermediate distances there is a mild attraction between electronic clouds, due to electron-electron correlation. The attraction between two atoms arises because of an induced dipole–dipole moment created by the motion of electrons through the molecule, which effects the neighbouring molecule. The attraction between two fragments does not only depend on dipole–dipole interactions, but also on dipole–quadrupole, quadrupole–quadrupole etc. interactions. These interactions do not have a high contribution to the overall energy, so for simplicity they are neglected in calculations.

In 1903, Gustav Mie proposed an intermolecular pair potential [2], comprising two parts to represent attractive and repulsive forces, equation 3.2.8, where r is interatomic distance, ϵ is the depth of the well at σ, the interatomic separation at which repulsive and attractive terms balance out and m, n can be adjusted.

$$E_{pair}(r) = \left(\frac{n}{n-m}\right)\left(\frac{n}{m}\right)^{m/(n-m)} \epsilon \left[\left(\frac{\sigma}{r}\right)^{n} - \left(\frac{\sigma}{r}\right)^{m}\right] \tag{3.2.8}$$

There have been a number of models suggested, where m and n take different values and show an accurate representation of experimentally known interactions [3, 4]. Nevertheless, the most popular one is the 12–6 Lennard-Jones potential [5], equation 3.2.9, where σ is the interatomic separation at which repulsive and attractive terms balance out and ϵ is the depth of the well. The usage of 12 term reduces the amount of calculations needed, making it the lightest method available.

$$E_{VdW}^{LJ} = 4\epsilon^{AB}\left[\left(\frac{\sigma^{AB}}{r^{AB}}\right)^{12} - \left(\frac{\sigma^{AB}}{r^{AB}}\right)^{6}\right] \tag{3.2.9}$$

The 12–6 Lennard-Jones potential is shown on the Figure 3.7. Red shows the repulsive term, blue the attractive term and the overall potential is shown in black. Position of σ and ϵ is also shown on the graph.

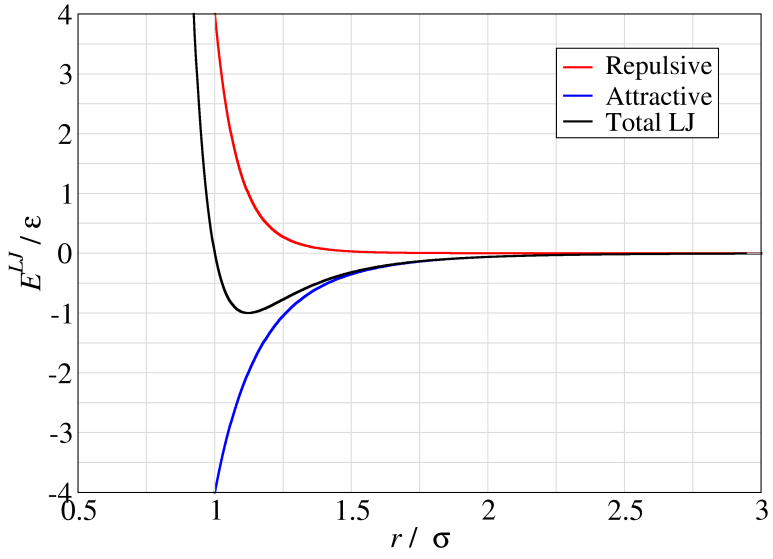


Figure 3.7: The Lennard-Jones Potential (black) is a combination of attractive (blue) and repulsive (red) terms.

It is possible to mix two Lennard-Jones potentials for single type atoms to obtain a potential describing interaction between two different atoms. There are three main approaches:

- Arithmetic or Lorentz-Berthelot rules, equations 3.2.10 and 3.2.11 [6].

$$\sigma_{ij} = \frac{\sigma_{ii} + \sigma_{jj}}{2} \tag{3.2.10}$$

$$\epsilon_{ij} = \sqrt{\epsilon_{ii}\epsilon_{jj}} \tag{3.2.11}$$

- Geometric rules equation 3.2.12, with ϵ_{ij} calculated as given in equation 3.2.11 [7].

$$\sigma_{ij} = \sqrt{\sigma_{ii}\sigma_{jj}} \tag{3.2.12}$$

- Fender-Halsey rule, equation 3.2.13 [8].

$$\epsilon_{ij} = \frac{2\epsilon_i\epsilon_j}{\epsilon_i + \epsilon_j} \tag{3.2.13}$$

A slightly more accurate potential, but up to 4 times more computationally demanding, is the Buckingham or Hill potential [4]. Repulsive forces take the form of an exponential, as the repulsion arises due to electron correlation, and the electron density reduces exponentially with the distance, equation 3.2.14, where A, B, and C are suitable constants.

$$E_{VdW}^{Hill} = A\exp\left(-Br_{AB}\right) - Cr_{AB}^{6} \tag{3.2.14}$$

3.2.5 Electrostatic Energy

The electrostatic energy calculates non bonded interactions that appear due to an uneven internal distribution of electrons. This leads to positively and negatively charged parts in molecule. The simplest way to model this behaviour is to place charges on atoms. The law of the interaction between two point charges was investigated by Charles Augustin Coulomb in 1780s and can be expressed by the Coulomb potential, equation 3.2.15, where ϵ_0 is the dielectric constant, the Q values are partial electronic charges on atoms A and B, and r is a distance between them.

$$E_{el} = \frac{1}{4\pi\epsilon_0}\frac{Q^A Q^B}{r^{AB}} \tag{3.2.15}$$

The atomic charges are mainly taken from electrostatic potential calculations carried out using higher precision methods. Another approach arises from assigning bond dipole moments. This gives similar results to the partial charge method, but these two methods will only give identical results for interactions at larger distances.

3.2.6 Cross Terms

There are also cross terms available for some force fields. These can cover the coupling between the fundamental stretching, bending and torsional interactions.

3.3 Choice of Force Field

The correct choice of the force field is essential for performing an accurate simulation. A large number of force fields have been developed through last decades. The force field consists of a set of parameters that depend on a particular the atom and the interaction.

Ideally force fields are designed to be transferable between a number of molecular systems. Nevertheless, some are more suitable for particular systems and states, then others. Comparison of the performance of the force filed should be done with a specific study in mind. Typical tests include the comparison to the experimental physical properties, such as density, boiling and melting temperatures, secondary protein structure, as well as the agreement with vapour-liquid coexistence, and the reproduction of thermodynamic properties.

Most force fields represent all atom systems, with a few being united atom where hydrogens are united to the neighbouring heavy atoms. United atom representation has less interactions, by such speeding up the calculations. Not every system can be accurately represented as united atom, for example systems having hydrogen bonding, benzene rings, polar and charged systems.

All atom force fields:

- OPLS-AA(Optimised Potentials for Liquid Simulations) [9]
- CHARMM (Chemistry at HARvard Macromolecular Mechanics) [10]
- AMBER (Assisted Model Building with Energy Refinement) [11]
- UFF (Universal Force Field) [12]

United atom force fields:

- OPLS (Optimised Potentials for Liquid Simulations)[13]
- CHARMM (Chemistry at HARvard Macromolecular Mechanics) [10]
- GROMOS (GROningen MOlecular Simulation) [14]
- TraPPE (Transferable Potentials for Phase Equilibria) [15]

3.3.1 Universal Force Field (UFF)

UFF [12] was developed by Rappe *et. al.* in 1992. It is an unusual force field, unlike others it covers the full periodic table, from hydrogen to lawrencium. The parameters are obtained by estimating from given general rules, that convert individual parameters for the element into the intermolecular and intramolecular parameters for calculations.

The performance of this force field with respect to the experimental results [16], is not very consistency, so it is not really suitable for high accuracy condensed state calculations. This force field is however useful for calculations of exotic molecules, where other force field have not been created.

3.3.2 Assisted Model Building with Energy Refinement (AMBER)

AMBER is both a molecular dynamics package and a set of force fields [11, 17]. The force field uses 12-6 Lennard-Jones, equation 3.2.9, with parameters computed by Lorentz-Berthelot mixing rules, equations 3.2.10 and 3.2.11, 1-4 interactions are scaled by $\frac{1}{2}$ and

Columbic interactions are represented by point charges and are scaled by $\frac{5}{6}$. Bond and angle interactions are expressed as harmonic potentials and torsional interactions are represented as a cosine series.

AMBER force fields are all atom only (though the newest force field, ff12sb, does include some united atom parameter), parametrised for biomolecular simulations and contains the 20 common amino acids as functional groups. More recently there have been some work put in to extend the AMBER force field to a wider range of systems through the development of the general AMBER force field (GAFF)[18].

3.3.3 Chemistry at HARvard Molecular Mechanics (CHARMM)

CHARMM is also molecular dynamics program [10], as well as a set of force fields. There are all atom CHARMM22 [19, 20], CHARMM27 [21], CgenFF [22], and united atom CHARMM19 [23] force fields. The potential function is of the same form as AMBER; however, non-bonded parameters are not scaled.

All of the force fields, apart from more recent CgenFF, were parametrised for biomolecular simulations, similarly to AMBER, containing the 20 common amino acid groups. CgenFF was created as a general force field for drug-type molecules, to accompany the rest of the force fields.

3.3.4 Optimised Potentials for Liquid Simulations (OPLS)

OPLS force field was developed by group of William L. Jorgensen and contains several sets of parameters, for both all atom and united atom calculations that can be mixed [9, 13]. The potential functions are the same as in AMBER, Lennard-Jones is computed by geometrical mean mixing rules, equation 3.2.12. Both, Lennard-Jones and Coulombic, parameters are scaled by a $1/2$.

OPLS at the beginning was parametrised not only for biomolecular systems, but also organic liquids. The parameters were optimized to fit experimental densities and heats of vaporisation of liquids, as well as gas-phase torsional profiles. The force field has being updated and many independent modifications to the force field have been developed [24, 25]. OPLS is, arguably, the most widely parametrised force field available.

3.3.5 GROningen MOlecular Simulation (GROMOS)

GROMOS is also both a molecular dynamics package [14] and the complementary force field. The first was also the base for GROMACS package [26]. GROMOS is a united atom force field, with ongoing improvements being done, currently most widely used version of the force field is GROMOS-96 [27] and its latest update [28]. This force field is also aimed at biomolecular calculations and is parametrised for proteins.

Unlike AMBER, the Lennard-Jones parameters for heteroatomic interactions are readily provided by the force field and should not be calculated via mixing rules. The force field uses a quartic expression for the bond length, a harmonic cosine potential for angle bending and a single cosine series for torsional angle potential.

3.3.6 Transferable Potentials for Phase Equilibria (TraPPE)

TraPPE force field evolved from the previous SKS (Smit-Karaborni-Siepmann) force field [**?**]. SKS was developed in Shell laboratories in the 1990s and is the first force field, aimed at reproducing properties of n-alkanes, the main components in fuels. The TraPPE

force field contains parameters for alkanes, aromatics, some oxygen, sulphur and nitrogen containing compounds.

TraPPE force field is mainly united atom [15, 29–34] with two explicit hydrogen models [35, 36]. Unlike most of united atom force fields, in TraPPE hydrogen atoms are also united in aromatic interactions.

The functional form is the same as in AMBER, including mixing rules, Coulombic terms are scaled by a $\frac{1}{2}$. Since the force field was developed with Monte Carlo calculations, there is no improper torsional angle term what can lead to incorrect conformations when performing molecular dynamics. The force field is parametrised to fit the vapour-liquid coexistence curve for small organic molecules.

3.4 Water Models

Water molecular models have been developed in order to help discover the structure of water. They are useful given the basis that if the (known but hypothetical) model can successfully predict the physical properties of liquid water then the (unknown) structure of liquid water is determined. There are two general forms for these models of water:

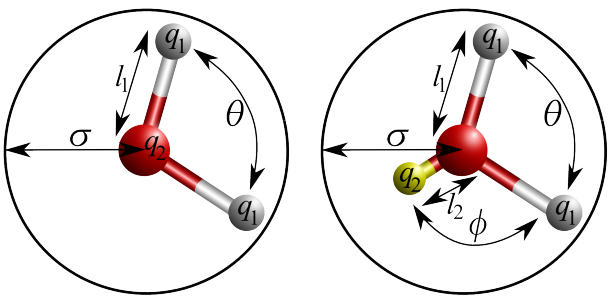
Explicit Water, where each water molecule is simulated directly.

Implicit Water, where the general effects of water are taken into account but the molecules are not directly simulated.

3.4.1 Explicit Water

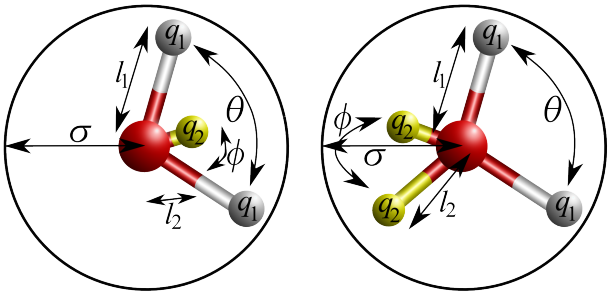
When simulating water explicitly it is important to use an accurate but efficient model. This is due to the computational expensiveness of simulating water in a solvated system, e.g. in simulating a protein in periodic box full of water at aqueous density about 90% of the computational time is used in simulating water-water interactions.

The original atomic-scale computational model for liquid water was proposed by Bernal and Fowler in 1933 [37]. This model took into account the position and the charge on the hydrogens and the position of the negative charge in the molecule. It was a highly insightful model and similar to the general view in Figure 3.8(c). At the time of the earliest developments of protein force fields the ST2 model of Stillinger and Rahman [38] was in wide use, which lead to the ST2 model being for some of the first protein simulations.



(a) 3-site

(b) 4-site



In the early 1980s two new force fields were developed, the SPC [39] and TIP3P [40] models. They involve orienting electrostatic effects and Lennard-Jones sites. The Lennard-Jones interaction accounts for the size of the molecules. It is repulsive at short distances, a ensuring that the structure does not completely collapse due to the electrostatic interactions. At intermediate distances it is significantly attractive but non-directional and competes with the directional attractive electrostatic interactions.

This competition ensures a tension between an expanded tetrahedral network and a collapsed non-directional network (for example, similar to that found in liquid noble gases).

Since then many models for water have been created around these general principles, although the charge and Lennard-Jones sites don't always coincide. The four main general model types are shown in Figure 3.8. Generally each model is developed to fit well with one particular physical structure or parameter (for example, the density anomaly, radial distribution function or the critical parameters) and it comes as no surprise when a model developed to fit certain parameters it gives good compliance with these same parameters [41]. In spite of the heavy computational investment in the calculations, the final agreement with experimental data is often 'by eye' and not statistically tested or checked for parametric sensitivity. Also, tests for 'quality' often use the radial distribution fit with diffraction data in spite of the major fitted peaks being derived from the tetrahedral nature of water that is built into every model and overpowers any disagreement in the fine detail. In particular, the O—O radial distribution function seems to be a poor discriminator between widely differently performing models [42]. Table 3.1 shows the values predicted for some key water properties by some of the most popular water force fields.

3.4.2 Implicit Water

There are many circumstances in molecular modeling studies in which a simplified description of solvent effects has advantages over the explicit modelling of each solvent molecule. Most of these are where the saving of computer time are needed over using explicit water. Most of the simulation techniques have models in which the solute degrees of freedom are treated explicitly but the solvent degrees of freedom are not. This requires that the energy surface used for the protein degrees of freedom be a potential of mean force (PMF) in which the solvent degrees of freedom are implicitly averaged over [66]. Assuming that the full potential energy function consists of a term, U_{vac} for the interactions within the protein, depending only on the protein degrees of freedom, $\mathbf{r}$, and an additional term for the protein-solvent and solvent-solvent interactions, the PMF is ideally equation 3.4.1, where $DeltaG_{\text{sol}}(\mathbf{r})$ is the free energy of transferring the protein from vacuum to the solvent with its internal degrees of freedom fixed at $\mathbf{r}$.

$$U_{\text{PMF}}(\mathbf{r}) = U_{\text{vac}}(\mathbf{r}) + \Delta G_{\text{sol}}(\mathbf{r}) \tag{3.4.1}$$

Some of the main models are:

COSMO Model [67] treats the solvent as a high dielectric continuum, interacting with charges that are embedded in solute molecules of lower dielectric. In spite of the severity of the approximation, this model often gives a good account of equilibrium solvation energetics.

Poisson-Boltzmann Model [68] describes electrostatic interactions in a multiple-dielectric environment and are typically solved by finite-difference or boundary element numerical methods. These can be efficiently solved for small molecules but may become expensive for proteins or nucleic acids. Although progress continues to be

made in numerical solutions, there is a clear interest in exploring more efficient, if approximate, approaches to this problem.

Generalized Born Model [69, 70] computes the electrostatic work required to move a charged sphere from a vacuum environment into a continuous dielectric region. The result is proportional to the square of the charge and is inversely proportional to the size of the ion. The basis of generalized Born theory is to extend these ideas to non-spherical molecules by casting the electrostatic contribution to solvation. The Born model have not traditionally considered salt effects, but the model has be extended to low-salt concentrations [71]. The key to making Generalized Born calculations more accurate (in the sense of agreement with Poisson-Boltzmann calculations) is improved estimation of the effective Born radii [72].

Table 3.1: Calculated physical properties of the water models at 25 °C and 1 atm.

Model	Type	Dipole Moment	Dielectric constant	Self Diffusion / 10^{-5}cm^2 s^{-1}	Average Configurational Energy / kJ mol^{-1}	Density Maximum Temperature /°C	Expansion Coefficient, 10^{-4} °C^{-1}
SSD [43]	-[1]	2.35 [43]	72 [43]	2.13 [43]	-40.2 [43]	-13 [43]	-
SPC [39]	a	2.27 [44]	65 [45]	3.85 [46]	-41.0 [45]	-45 [47]	7.3[2][48]
SPC/E [49]	a	2.35 [49]	71 [49]	2.49 [46]	-41.5 [49]	-38 [50]	5.14 [51]
SPC/Fw[4][51]	a	2.39 [51]	79.63 [51]	2.32 [51]	-	-	4.98 [51]
PPC[4][49]	b	2.52 [49]	77 [49]	2.6 [49]	-43.2 [49]	4 [52]	-
TIP3P [40]	a	2.35 [53]	82 [49]	5.19 [46]	-41.1 [53]	-91 [47]	9.2 [53]
TIP3P/Fw[4][51]	a	2.57 [51]	193 [51]	3.53 [51]	-	-	7.81 [51]
TIP4P [53]	c	2.18 [49, 53]	53 [49]	3.29 [46]	-41.8 [53]	-25 [53]	4.4 [53]
TIP4P-Ew [54]	c	2.32 [54]	62.9 [54]	2.4 [54]	-46.5 [54]	1 [54]	3.1 [54]
TIP4P-FQ [55]	c	2.64 [55]	79 [55]	1.93 [55]	-41.4 [56]	7 [55]	-
TIP4P/2005 [41]	c	2.305 [41]	60 [41]	2.08 [41]	-	5 [41]	2.8 [41]
TIP4P/2005f [57]	c	2.319 [57]	55.3 [57]	1.93 [57]	-	7 [57]	-
SWFLEX-AI[4][56]	c	2.69 [56]	116 [56]	3.66 [56]	-41.7 [56]	-	-
COS/G3 [48]	c	2.57[2][48]	88[2][48]	2.6[2][48]	-41.1[2][48]	-	7.0[2][48]
COS/D [58]	c	2.43 [58]	69.8 [58]	2.5 [58]	-41.8 [58]	-	-
GCPM[4][59]	c	2.723 [59]	84.3 [59]	2.26 [59]	-44.8 [59]	-13 [59]	-
SWM4-NDP[4][60]	c	2.461 [60]	79 [60]	2.33 [60]	-41.5 [60]	-	-
TIP5P [53]	d	2.29 [53]	81.5 [53]	2.62 [46]	-41.3 [53]	4 [53]	6.3 [53]
TIP5P-Ew [61]	d	2.29 [61]	92 [61]	2.8 [61]	-	8 [61]	4.9 [61]
TTM2-F [62]	c	2.67 [62]	67.2 [62]	1.4 [62]	-45.1 [62]	-	-
POL5/TZ[4][63]	d	2.712 [63]	98 [63]	1.81 [63]	-41.5 [63]	25 [63]	-
Six-site [64]	c+d	1.89[3][64]	33[3][64]	-	-	14[3][64]	2.4[3][64]
QCT [65]	a	1.85[2][65]	-	1.5[2][65]	-42.7[2][65]	10[2][65]	3.5[2][65]
Experimental	-	2.95	78.4	2.30	-41.5 [53]	3.984	2.53

[1]Has only a single, center of mass, interaction site with a tetrahedrally coordinated sticky potential that regulates the tetrahedral coordination of neighboring molecules

[2]Calculated at 27 °C

[3]Calculated at 20 °C

[4]Polarizable model

3.5 Coarse-Graining

Even though computational power is rapidly increasing, with hardware performance doubling every two years [73] (technically the number of transistors doubling), coupled to the development of ever more efficient algorithms to deliver faster and more accurate calculations; there is a pressing need to simulate chemical systems over timescales that go far beyond those currently obtainable with atomistic systems.

The basic concept of coarse graining involves a reduction of the number of sites in a system. This not only allows for smaller computational load, but also leads to a loss of fine detail in the system. There has to be a delicate balance between these two things, ideally leading to a considerable speed up of the calculation while still preserving the necessary chemical and physical information. Decreasing the number of the interaction sites in a model allows a significant speed up in the calculation. Computer time approximately scales as N^2 (N is a number of interaction sites) for pairwise potentials. Additionally, some intermolecular movements are lost, allowing the system to explore phase space faster. A third speed up is due to the increase of time step for the calculation, as the vibration frequency of larger particles is slower.

In chemistry, a coarse grained system will typically still be represented at the molecular level, as prediction of local properties of the system is dependent on the presence of molecules. A common type of coarse grained model represents of a group of atoms with a single interaction centre, a so called "super atom". These super atoms are then assigned interaction parameters. The parameters are derived from a smaller scale, higher resolution model of the system. Additionally, they may be fitted to the experimental values. Results of the calculation are dependant on them, so care should be taken when parametrizing the system.

Currently there is only one readily available coarse grained force field, MARTINI [74]. The force field is designed to be used in biomolecular simulations of lipids, proteins and carbohydrates [75–77]. The model uses a shifted Lennard-Jones 12-6 potential with parameters fitted to free energies of vaporisation, hydration and separation between water and organic solvents at room (300 K) temperature only. There are criticisms of water in this model due to its inability to remain in the liquid state in membrane pores [78, 79]. Recently a more sophisticated, polarizable water model was developed by the same group [80]. Water is proving to be one of the most difficult systems to coarse grain, as finding potentials capable of exhibiting the correct behaviour is difficult; though several efforts have been made recently [81–83].

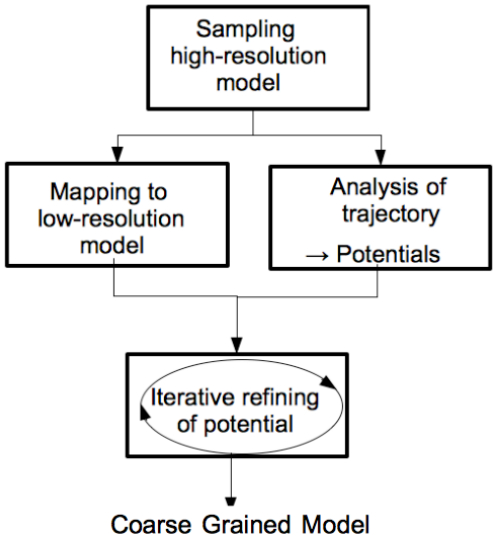


Figure 3.9: Illustration of the coarse graining procedure: sampling of a high resolution model, analysis of the trajectory, mapping to low resolution, iterative refining of the potential.

The main constraint of coarse grained potentials is the limited transferability between different systems and thermodynamic conditions. Transferability is highly dependent on the system parametrised, and is generally better for finer grained models; for polymer systems transferability is not greatly affected by a change in the length of the polymer chain [84].

To derive a suitable coarse grain potential, one would normally follow the general procedure, shown in the Figure 3.9. A smaller scale, higher resolution atomistic (sometimes united-atom) model is sampled using molecular dynamics or Monte Carlo simulations. When setting up the initial calculation and then mapping the system, it is important consider which properties are needed and at which time and length scales. The trajectories from high resolution calculation are then analysed and the starting potentials for the coarse grain model are obtained. For most of the cases it is necessary to refine the latter to obtain consistent results with the higher resolution method.

3.5.1 Mapping

In the mapping procedure, one represents a group of atoms with a single interaction centre. This centre is called a super atom and commonly positioned at the centre of mass of the group of initial atoms. Super atoms, as in atomistic calculations, are interlinked with springs and follow the same rules of Newtonian mechanics. The choice of atoms to group into a bead is typically done by geometrical examination of the molecule and leads to a coarse grained molecule still resembling the general shape of the initial molecule. This will allow a better agreement of non bonded properties. Additionally, one should consider which interactions form the simplest, if possible single well or simple shape potentials, that will be beneficial in refining the system in further iterative steps [85]. The parts of the molecule that are geometrically mobile (eg. undergoing gauche–trans conformational changes) will typically lead to potentials with more then one well.

3.5.2 Potentials

When the mapping is chosen, the next step is to evaluate the optimum parameters to describe the interactions between the coarse grained sites. Over the last few years a number of approaches have been developed:

Force matching , where a multibody potential of mean force of a high resolution system is matched by a low resolution one [86]. This approach looks at individual configurations, rather than the average properties of the system as encompassed by radial distribution functions [87, 88].

Inverse Monte Carlo , where the new parameters are fitted by the use of Monte Carlo calculations [89, 90]. The potentials are derived from distributions, e.g.% bond length, angle, dihedral angle, and radial distribution functions.

Iterative Boltzmann Inversion , where the new parameters are fitted with the aid of molecular dynamics calculations [91]. The potentials are derived from distributions, e.g. bond length, angle, dihedral angle, and radial distribution functions.

3.5.3 Iterative Boltzmann Inversion

Iterative Boltzmann Inversion is a structure based method and is aimed at matching distribution functions of a new coarse grained model to a higher resolution calculation or experiment. Bonded interactions (bond, angle and dihedral) incorporate neglected degrees of freedom and temperature effects. Non-bonded are calculated separately, creating physically sensible interactions between molecules. Generally, bonded potentials are stronger than non bonded ones. The fitting of the potentials should start with the strongest one

(deepest potential well), as it will change the least during fitting of the rest, hence the order:

$$P_{bond} \rightarrow P_{angle} \rightarrow P_{dihedral} \rightarrow P_{nonbonded}$$

Figure 3.10 illustrates the procedure for deriving an effective potential $V(x)$ from a known probability distribution, $P_{ref}(x)$. $P(x)$ is a general term for the distribution, that could be a radial distribution function, $g(r)$; bond length, $b(r)$; angle, $a(\theta)$; or dihedral angle, $d(\phi)$; distribution.

For the first step of the iteration, a reasonable guess of an effective potential should be obtained. This can be derived from a reference probability, $P_{ref}(x)$, of a high resolution simulation of a pure system, equation 3.5.1.

$$V_0(\mathbf{x}) = -k_B T \ln P_{ref}(\mathbf{x}) \qquad (3.5.1)$$

It should be mentioned that V_0 is not a potential energy, but rather a free energy, dependant on temperature and pressure. V_0 is a reasonable guess for the first coarse grained potential. A coarse grained simulation is performed and new probability distributions $P_i(\mathbf{x})$ evaluated. The new distribution will differ from $P_{ref}(\mathbf{x})$ and so the new potential should include a correction, in order to provide a better representation of the system equation 3.5.2 is used, where λ is a numerical scaling factor and should be $\lambda \in (0, 1]$.

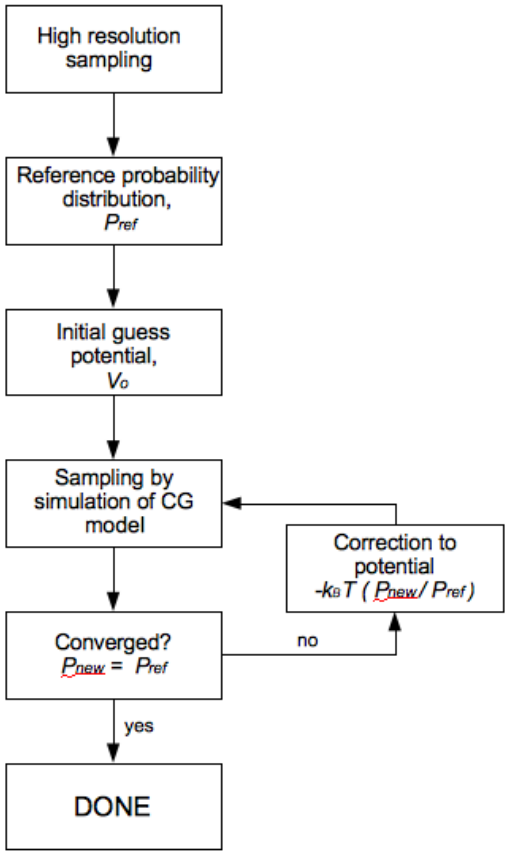


Figure 3.10: Illustrated procedure for deriving effective potentials by the iterative Boltzmann inversion (IBI) method.

$$V_{i+1}(\mathbf{x}) = V_i(\mathbf{x}) - \lambda k_B T \ln \frac{P_i(\mathbf{x})}{P_{ref}(\mathbf{x})} \qquad (3.5.2)$$

The new calculation will again result in a probability distribution, P_{i+1}. The iterative procedure should be repeated until P_{i+n} is sufficiently close to P_{ref}. It has been tested by numerous authors and shown that it can take up to 200 iterations to obtain a potential capable of representing aggregate distributions [92] and over 1000 iterations to recover correct potential functions [93].

3.6 References

[1] Philip M. Morse. Diatomic Molecules According to the Wave Mechanics. II. Vibrational Levels. *Phys. Rev.*, 34:57–64, Jul 1929.

[2] Gustav Mie. Zur kinetischen Theorie der einatomigen KÃűrper. *Annalen der Physik*, 316(8):657–697, 1903.

[3] A. Warshel and S. Lifson. Consistent force field calculations. II. Crystal structures, sublimation energies, molecular and lattice vibrations, molecular conformations, and enthalpies of alkanes. *The Journal of Chemical Physics*, 53:582, 1970.

[4] David N.J. White. A computationally efficient alternative to the Buckingham potential for molecular mechanics calculations. *Journal of Computer-Aided Molecular Design*, 11:517–521, 1997. 10.1023/A:1007911511862.

[5] J. E. Jones. On the Determination of Molecular Fields. I. From the Variation of the Viscosity of a Gas with Temperature. *Proceedings of the Royal Society of London. Series A*, 106(738):441–462, 1924.

[6] H. A. Lorentz. Ueber die Anwendung des Satzes vom Virial in der kinetischen Theorie der Gase. *Annalen der Physik*, 248(1):127–136, 1881.

[7] Robert J. Good and Christopher J. Hope. New Combining Rule for Intermolecular Distances in Intermolecular Potential Functions. *The Journal of Chemical Physics*, 53(2):540–543, 1970.

[8] B. E. F. Fender and Jr. G. D. Halsey. Second Virial Coefficients of Argon, Krypton, and Argon-Krypton Mixtures at Low Temperatures. *The Journal of Chemical Physics*, 36(7):1881–1888, 1962.

[9] William L. Jorgensen, David S. Maxwell, and Julian Tirado-Rives. Development and Testing of the OPLS All-Atom Force Field on Conformational Energetics and Properties of Organic Liquids. *Journal of the American Chemical Society*, 118(45):11225–11236, 1996.

[10] B. R. Brooks, C. L. Brooks, A. D. Mackerell, L. Nilsson, R. J. Petrella, B. Roux, Y. Won, G. Archontis, C. Bartels, S. Boresch, A. Caflisch, L. Caves, Q. Cui, A. R. Dinner, M. Feig, S. Fischer, J. Gao, M. Hodoscek, W. Im, K. Kuczera, T. Lazaridis, J. Ma, V. Ovchinnikov, E. Paci, R. W. Pastor, C. B. Post, J. Z. Pu, M. Schaefer, B. Tidor, R. M. Venable, H. L. Woodcock, X. Wu, W. Yang, D. M. York, and M. Karplus. CHARMM: The biomolecular simulation program. *Journal of Computational Chemistry*, 30(10):1545–1614, 2009.

[11] Wendy D. Cornell, Piotr Cieplak, Christopher I. Bayly, Ian R. Gould, Kenneth M. Merz, David M. Ferguson, David C. Spellmeyer, Thomas Fox, James W. Caldwell, and Peter A. Kollman. A Second Generation Force Field for the Simulation of Proteins, Nucleic Acids, and Organic Molecules. *Journal of the American Chemical Society*, 117(19):5179–5197, 1995.

[12] A. K. Rappe, C. J. Casewit, K. S. Colwell, W. A. Goddard, and W. M. Skiff. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *Journal of the American Chemical Society*, 114(25):10024–10035, 1992.

[13] William L. Jorgensen and Julian. Tirado-Rives. The OPLS [optimized potentials for liquid simulations] potential functions for proteins, energy minimizations for crystals of cyclic peptides and crambin. *Journal of the American Chemical Society*, 110(6):1657–1666, 1988.

[14] Markus Christen, Philippe H. Hunenberger, Dirk Bakowies, Riccardo Baron, Roland Burgi, Daan P. Geerke, Tim N. Heinz, Mika A. Kastenholz, Vincent Krautler, Chris Oostenbrink, Christine Peter, Daniel Trzesniak, and Wilfred F. van Gunsteren. The

GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*, 26(16):1719–1751, 2005.

[15] M.G. Martin and J.I. Siepmann. Transferable potentials for phase equilibria. 1. United-atom description of n-alkanes. *The Journal of Physical Chemistry B*, 102(14):2569–2577, 1998.

[16] Marcus G. and Martin. Comparison of the AMBER, CHARMM, COMPASS, GROMOS, OPLS, TraPPE and UFF force fields for prediction of vaporâĂŞliquid coexistence curves and liquid densities. *Fluid Phase Equilibria*, 248(1):50 – 55, 2006.

[17] Yong Duan, Chun Wu, Shibasish Chowdhury, Mathew C. Lee, Guoming Xiong, Wei Zhang, Rong Yang, Piotr Cieplak, Ray Luo, Taisung Lee, James Caldwell, Junmei Wang, and Peter Kollman. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. *Journal of Computational Chemistry*, 24(16):1999–2012, 2003.

[18] Junmei Wang, Romain M. Wolf, James W. Caldwell, Peter A. Kollman, and David A. Case. Development and testing of a general amber force field. *Journal of Computational Chemistry*, 25(9):1157–1174, 2004.

[19] A. D. MacKerell, D. Bashford, Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus. All-Atom Empirical Potential for Molecular Modeling and Dynamics Studies of ProteinsâĂă. *The Journal of Physical Chemistry B*, 102(18):3586–3616, 1998.

[20] A.D. Mackerell Jr, M. Feig, and C.L. Brooks III. Extending the treatment of backbone energetics in protein force fields: Limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *Journal of computational chemistry*, 25(11):1400–1415, 2004.

[21] A.D. MacKerell Jr, N. Banavali, and N. Foloppe. Development and current status of the CHARMM force field for nucleic acids. *Biopolymers*, 56(4):257–265, 2000.

[22] K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov, et al. CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields. *Journal of computational chemistry*, 31(4):671–690, 2010.

[23] III WH Reiher. *Theoretical studies of hydrogen bonding*. PhD thesis, Harvard University, 1985.

[24] S.V. Sambasivarao and O. Acevedo. Development of OPLS-AA force field parameters for 68 unique ionic liquids. *Journal of Chemical Theory and Computation*, 5(4):1038–1050, 2009.

[25] Z. Xu, H.H. Luo, and D.P. Tieleman. Modifying the OPLS-AA force field to improve hydration free energies for several amino acid side chains using new atomic charges and an off-plane charge model for aromatic residues. *Journal of computational chemistry*, 28(3):689–697, 2007.

[26] Berk Hess, Carsten Kutzner, David van der Spoel, and Erik Lindahl. GROMACS 4: Algorithms for Highly Efficient, Load-Balanced, and Scalable Molecular Simulation. *Journal of Chemical Theory and Computation*, 4(3):435–447, 2008.

[27] Wilfred F. van Gunsteren, S. R. Billeter, A. A. Eising, Philippe H. Hünenberger, P. Krüger, Alan E. Mark, W. R. P. Scott, and Ilario G. Tironi. *Biomolecular Simulation: The GROMOS96 manual and user guide*. 1996.

[28] Markus Christen, Philippe H. HÃijnenberger, Dirk Bakowies, Riccardo Baron, Roland BÃijrgi, Daan P. Geerke, Tim N. Heinz, Mika A. Kastenholz, Vincent Krautler, Chris Oostenbrink, Christine Peter, Daniel Trzesniak, and Wilfred F. van Gunsteren. The GROMOS software for biomolecular simulation: GROMOS05. *Journal of Computational Chemistry*, 26(16):1719–1751, 2005.

[29] M.G. Martin and J.I. Siepmann. Novel configurational-bias Monte Carlo method for branched molecules. Transferable potentials for phase equilibria. 2. United-atom description of branched alkanes. *The Journal of Physical Chemistry B*, 103(21):4508–4517, 1999.

[30] C.D. Wick, M.G. Martin, and J.I. Siepmann. Transferable potentials for phase equilibria. 4. United-atom description of linear and branched alkenes and alkylbenzenes. *The Journal of Physical Chemistry B*, 104(33):8008–8016, 2000.

[31] B. Chen, J.J. Potoff, and J.I. Siepmann. Monte Carlo calculations for alcohols and their mixtures with alkanes. Transferable potentials for phase equilibria. 5. United-atom description of primary, secondary, and tertiary alcohols. *The Journal of Physical Chemistry B*, 105(15):3093–3104, 2001.

[32] J.M. Stubbs, J.J. Potoff, and J.I. Siepmann. Transferable potentials for phase equilibria. 6. United-atom description for ethers, glycols, ketones, and aldehydes. *The Journal of Physical Chemistry B*, 108(45):17596–17605, 2004.

[33] C.D. Wick, J.M. Stubbs, N. Rai, and J.I. Siepmann. Transferable potentials for phase equilibria. 7. Primary, secondary, and tertiary amines, nitroalkanes and nitrobenzene, nitriles, amides, pyridine, and pyrimidine. *The Journal of Physical Chemistry B*, 109(40):18974–18982, 2005.

[34] N. Lubna, G. Kamath, J.J. Potoff, N. Rai, and J.I. Siepmann. Transferable potentials for phase equilibria. 8. United-atom description for thiols, sulfides, disulfides, and thiophene. *The Journal of Physical Chemistry B*, 109(50):24100–24107, 2005.

[35] B. Chen and J.I. Siepmann. Transferable potentials for phase equilibria. 3. Explicit-hydrogen description of normal alkanes. *The Journal of Physical Chemistry B*, 103(25):5370–5379, 1999.

[36] N. Rai and J.I. Siepmann. Transferable potentials for phase equilibria. 9. Explicit hydrogen description of benzene and five-membered and six-membered heterocyclic aromatic compounds. *The Journal of Physical Chemistry B*, 111(36):10790–10799, 2007.

[37] J. D. Bernal and R. H. Fowler. A Theory of Water and Ionic Solution, with Particular Reference to Hydrogen and Hydroxyl Ions. *Journal of Chemical Physics*, 1:515–548, 1933.

[38] F. H. Stillinger and A. Rahman. Improved simulation of liquid water by molecular dynamics. *Journal of Chemical Physics*, 60:1545–1557, 1974.

[39] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, and J. Hermans. *Intermolecular Forces*. Reidel, Dordrecht, 1981.

[40] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein. Comparison of simple potential functions for simulating liquid water. *Journal of Chemical Physics*, 79:926–935, 1983.

[41] J. L. F. Abascal and C. Vega. A general purpose model for the condensed phases of water: TIP4P/2005. *Journal of Chemical Physics*, 123:234505, 2005.

[42] P. E. Mason and J. W. Brady. "Tetrahedrality" and the relationship between collective structure and radial distribution functions in liquid water. *Journal of Physical Chemistry B*, 111:5669–5679, 2007.

[43] M-L. Tan, J. T. Fischer, A. Chandra, B. R. Brooks, and T.ÂăIchiye. A temperature of maximum density in soft sticky dipole water. *Chemical Physics Letters*, 376:646–652, 2003.

[44] K. Kiyohara, K. E. Gubbins, and A. Z. Panagiotopoulos. Phase coexistence properties of polarizable water models. *Molecular Physics*, 94:803–808, 1998.

[45] D. van der Spoel, P. J. van Maaren, and H. J. C. Berendsen. A systematic study of water models for molecular simulation: Derivation of water models optimized for use with a reaction field. *Journal of Chemical Physics*, 108:10220–10230, 1998.

[46] M. W. Mahoney and W. L. Jorgensen. Diffusion constant of the TIP5P model of liquid water. *Journal of Chemical Physics*, 114:363–366, 2001.

[47] C. Vega and J. L. F. Abascal. Relation between the melting temperature and the temperature of maximum density for the most common models of water. *Journal of Chemical Physics*, 123:144504, 2005.

[48] H. Yu and W. F. van Gunsteren. Charge-on-spring polarizable water models revisited: from water clusters to liquid water to ice. *Journal of Chemical Physics*, 121:9549–9564, 2004.

[49] H. J. C. Berendsen, J. R. Grigera, and T. P. Straatsma. The missing term in effective pair potentials. *Journal of Physical Chemistry*, 91:6269–6271, 1987.

[50] L. A. Baez and P. Clancy. Existence of a density maximum in extended simple point-charge water. *Journal of Chemical Physics*, 101:9837–9840, 1994.

[51] Y. Wu, H. L. Tepper, and G. A. Voth. Flexible simple point-charge water model with improved liquid state properties. *Journal of Chemical Physics*, 124Âă:024503, 2006.

[52] I. M. Svishchev, P. G. Kusalik, J. Wang, and R. J. Boyd. Polarizable point-charge model for water. Results under normal and extreme conditions. *Journal of Chemical Physics*, 105:4742–4750, 1996.

[53] M. W. Mahoney and W. L. Jorgensen. A five-site model for liquid water and the reproduction of the density anomaly by rigid, nonpolarizable potential functions. *Journal of Chemical Physics*, 112:8910–8922, 2000.

[54] H. W. Horn, W. C. Swope, J. W. Pitera, J. D. Madura, T. J. Dick, G. L. Hura, and T. Head-Gordon. Development of an improved four-site water model for biomolecular simulations. *Journal of Chemical Physics*, 120:9665–9678, 2004.

[55] S. W. Rick. Simulation of ice and liquid water over a range of temperatures using the fluctuating charge model. *Journal of Chemical Physics*, 114:2276–2283, 2001.

[56] P. J. van Maaren and D. van der Spoel. Molecular dynamics of water with novel shell-model potentials. *Journal of Physical Chemistry B*, 105:2618–2626, 2001.

[57] M. A. González and J. L. F. Abascal. A flexible model for water based on TIP4P/2005. *Journal of Chemical Physics*, 135:224516, 2011.

[58] A.-P. E. Kunz and W. F. van Gunsteren. Development of a nonlinear classical pPolarization model for liquid water and aqueous solutions: COS/D. *Journal of Physical Chemistry A*, 113:11570–11579, 2009.

[59] P. Paricaud, M. Predota, A. A. Chialvo, and P. T. Cummings. From dimer and condensed phases at extreme conditions: Accurate predictions of the properties of water by a Gaussian charge polarizable model. *Journal of Chemical Physics*, 122:244511, 2005.

[60] G. Lamoureux, E. Harder, I. V. Vorobyov, B. Roux, and A. D. MacKerell Jr. A polarizable model of water for molecular dynamics simulations of biomolecules. *Chemical Physics Letters*, 418:241–245, 2005.

[61] S. W. Rick. A reoptimization of the five-site water potential (TIP5P) for use with Ewald sums. *Journal of Chemcal Physics*, 120:6085–6093, 2004.

[62] G. S. Fanourgakis and S. S. Xantheas. The flexible, polarizable, Thole-type interaction potential for water (TTM2-F) revisited. *Journal of Physical Chemistry A*, 110:4100–4106, 2006.

[63] H. A. Stern, F. Rittner, B. J. Berne, and R. A. Friesner. Combined fluctuating charge and polarizable dipole models: Application to a five-site water potential function. *Journal of Chemical Physics*, 115:2237–2251, 2001.

[64] H. Nada and J. P. J. M. van der Eerden. An intermolecular potential model for the simulation of ice and water near the melting point: A six-site model of H_2O. *Journal of Chemical Physics*, 118:7401–7413, 2003.

[65] S. Y. Liem, P. L. A. Popelier, and M. Leslie. Simulation of liquid water using a high-rank quantum topological electrostatic potential. *International Journal of Quantum Chemistry*, 99:685–694, 2004.

[66] B Roux and T. Simonson. *Biophysical Chemistry*, 78:1–20, 1999.

[67] A. Klamt and G. Schüürmann. *J Chem. Soc. Perkin Trans*, 2:799–805, 1993.

[68] C. J. Cramer and D. G. Truhlar. *Chemical Reviews*, 99:2161–2200, 1999.

[69] M. Born. *Z. Phys.*, 1:45–48, 1920.

[70] D. Bashford and D. Case. *Annu. Rev. Phys. Chem.*, 51:129–152, 2000.

[71] J. Srinivasan, M. W. Trevathan, P. Beroza, and D. A. Case. *Theor. Chem. Ace.*, 101:126–434, 1999.

[72] M. S. Lee, F. R. Salsbury Jr, and C. L. Brooks III. *Journal of Physical Chemistry*, 116,:10606–10614, 2002.

[73] G. E. Moore. Cramming more components onto integrated circuits. *Electronics*, 38:114–117, 1965.

[74] S.J. Marrink, A.H. de Vries, and A.E. Mark. Coarse grained model for semiquantitative lipid simulations. *The Journal of Physical Chemistry B*, 108(2):750–760, 2004.

[75] S.J. Marrink, H.J. Risselada, S. Yefimov, D.P. Tieleman, and A.H. De Vries. The MARTINI force field: coarse grained model for biomolecular simulations. *The Journal of Physical Chemistry B*, 111(27):7812–7824, 2007.

[76] L. Monticelli, S.K. Kandasamy, X. Periole, R.G. Larson, D.P. Tieleman, and S.J. Marrink. The MARTINI coarse-grained force field: extension to proteins. *Journal of Chemical Theory and Computation*, 4(5):819–834, 2008.

[77] C.A. López, A.J. Rzepiela, A.H. de Vries, L. Dijkhuizen, P.H. Hünenberger, and S.J. Marrink. Martini coarse-grained force field: extension to carbohydrates. *Journal of Chemical Theory and Computation*, 5(12):3195–3210, 2009.

[78] M. Winger, D. Trzesniak, R. Baron, and W.F. Van Gunsteren. On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models. *Phys. Chem. Chem. Phys.*, pages 1934–1941, 2009.

[79] Wilfred F. van Gunsteren and Moritz Winger. Reply to the 'Comment on "On using a too large integration time step in molecular dynamics simulations of coarse-grained molecular models"' by S. J. Marrink, X. Periole, D. Peter Tieleman and Alex H. de Vries, Phys. Chem. Chem. Phys., 2010, 12, DOI: 10.1039/b915293h. *Phys. Chem. Chem. Phys.*, 12:2257–2258, 2010.

[80] S.O. Yesylevskyy, L.V. Schäfer, D. Sengupta, and S.J. Marrink. Polarizable water model for the coarse-grained MARTINI force field. *PLoS computational biology*, 6(6):e1000810, 2010.

[81] Leonardo Darre, Matias R. Machado, Pablo D. Dans, Fernando E. Herrera, and Sergio Pantano. Another Coarse Grain Model for Aqueous Solvation: WAT FOUR? *Journal of Chemical Theory and Computation*, 6(12):3793–3807, 2010.

[82] AJ and Rader. Coarse-grained models: getting more with less. *Current Opinion in Pharmacology*, 10(6):753 – 759, 2010. <ce:title>Endocrine and metabolic diseases/New technologies - the importance of protein dynamics</ce:title>.

[83] S. Riniker and W.F. van Gunsteren. A simple, efficient polarizable coarse-grained water model for molecular dynamics simulations. *The Journal of chemical physics*, 134:084110, 2011.

[84] P. Carbone, H.A.K. Varzaneh, X. Chen, and F. Müller-Plathe. Transferability of coarse-grained force fields: The polymer case. *The Journal of chemical physics*, 128:064904, 2008.

[85] V.A. Harmandaris, D. Reith, N.F.A. van der Vegt, and K. Kremer. Comparison between coarse-graining models for polymer systems: Two mapping schemes for polystyrene. *Macromolecular chemistry and physics*, 208(19/20):2109, 2007.

[86] F. Ercolessi and J.B. Adams. Interatomic potentials from first-principles calculations: the force-matching method. *EPL (Europhysics Letters)*, 26:583, 1994.

[87] S. Izvekov and G.A. Voth. A multiscale coarse-graining method for biomolecular systems. *The Journal of Physical Chemistry B*, 109(7):2469–2473, 2005.

[88] S. Izvekov and G.A. Voth. Multiscale coarse graining of liquid-state systems. *The Journal of chemical physics*, 123:134105, 2005.

[89] A.P. Lyubartsev and A. Laaksonen. Calculation of effective interaction potentials from radial distribution functions: A reverse Monte Carlo approach. *Physical Review E*, 52(4):3730, 1995.

[90] C.D. Berweger, W.F. van Gunsteren, and F. Müller-Plathe. Force field parametrization by weak coupling. Re-engineering SPC water. *Chemical physics letters*, 232(5-6):429–436, 1995.

[91] D. Reith, M. Pütz, and F. Müller-Plathe. Deriving effective mesoscale potentials from atomistic simulations. *Journal of computational chemistry*, 24(13):1624–1636, 2003.

[92] A.J. Rzepiela, M. Louhivuori, C. Peter, and S.J. Marrink. Hybrid simulations: combining atomistic and coarse-grained force fields using virtual sites. *Phys. Chem. Chem. Phys.*, 2011.

[93] S. Jain, S. Garde, and S.K. Kumar. Do inverse Monte Carlo algorithms yield thermodynamically consistent interaction potentials? *Industrial & engineering chemistry research*, 45(16):5614–5618, 2006.

Chapter 4

Mesophase Models and Beyond

Contents

4.1 Introduction

Colloidal suspensions are dispersions of mesoscopic systems, these systems can consist of millions or even billions of atoms. If infinite computational power was available it would be possible to simulate these systems; however, as this is not an option. Therefore, techniques have been developed that allow more coarse-grained representations of systems.

4.2 Lattice Boltzmann

When simulating traditional fluid dynamics the continuity and Navier-Stokes equations are typically used, equations 4.2.1 and 4.2.2, where $\mathbf{u}$ is the fluid velocity, ρ is the fluid density, p is the pressure, ν is the kinematic viscosity, and $\mathbf{f}$ is an acceleration due to external forces.

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \tag{4.2.1}$$

$$\rho \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \nabla \cdot \mathbf{u} \right) = -\nabla p + \rho \nu \nabla^2 \mathbf{u} + \rho \mathbf{f} \tag{4.2.2}$$

Often the best methods available for solving this system involve numerical integration of the equations after the system has been discretized into small elements. The Lattice Boltzmann method is an alternative approach to these computational fluid dynamics simulations and is an extension of the lattice gas methods invented in the late 1980s [1]. These methods allowed particles to move on a discrete lattice and local collisions conserved mass and momentum. The continuity and Navier-Stokes equations are only continuous forms of the mass and momentum conservation statements and methods that locally conserve mass and momentum will (approximately) obey the continuity and Navier-Stokes equations.

If a density distribution is used instead of discrete particles then the solution to the continuity and Navier-Stokes contains less noise. More than this the method allows a very general collision operator to be used. This is the Lattice Boltzmann method which has been extraordinarily successful for many applications including turbulence, multicomponent and multi-phase flows, as well as additional applications. The Lattice Boltzmann algorithm can be summarized by the following:

- Fluid properties are mapped onto a discrete lattice.
- The physical state at each lattice point is described by a set of particle distribution functions.
- Macroscopic fluid variables are defined by moments of the distribution functions.
- The system evolution towards equilibrium is through the relaxation of the distribution function to itsequilibrium form.

4.2.1 Basic Definitions

Two of the most popular grid forms are triangular and rectangular lattices. There are generally referred to using the $D\alpha Q\beta$-notation, where α is an integer number denoting the space dimensionality (i.e. 1, 2, or 3), and β is the number of discrete velocities (including the possibility of a particle at rest).

The macroscopic variables are defined as functions of the particle functions according to equations 4.2.3 and 4.2.4, where ρ is the density, f_i is the i-th particle distribution function, $\mathbf{e}_i$ is the i-th speed vector, and $\mathbf{u}$ is the macroscopic velocity.

$$\rho = \sum_{i=0}^{\beta-1} f_i \tag{4.2.3}$$

$$\mathbf{u} = \frac{1}{\rho} \sum_{i=0}^{\beta-1} f_i \mathbf{e}_i \tag{4.2.4}$$

4.2.2 Derivation of Equilibrium

The Maxwell-Boltzmann single particle equlibrium distribution function is given by equation 4.2.5, where $\theta = kT/m$, T is the temperature, m is the molar mass, D is the space dimension, $\boldsymbol{\xi}$ is the thermal velocity, and $\mathbf{u}$ is the macroscopic velocity.

$$f^{eq} = \frac{\rho}{(2\pi\theta)^{D/2}} \exp\left[-\frac{(\boldsymbol{\xi} - \mathbf{u})^2}{2\theta}\right] \tag{4.2.5}$$

When $\|\boldsymbol{\xi} - \mathbf{u}\| \ll \sqrt{\theta}$ equation 4.2.5 can be expanded to equation 4.2.6.

$$f^{eq} = \frac{\rho}{(2\pi\theta)^{D/2}} \exp\left(-\frac{\xi^2}{2\theta}\right)\left[1 + \frac{\boldsymbol{\xi}\cdot\mathbf{u}}{\theta} + \frac{(\boldsymbol{\xi}\cdot\mathbf{u})^2}{2\theta^2} - \frac{u^2}{2\theta}\right] \tag{4.2.6}$$

For a microscopic quantity, $\psi(\xi)$, the associated macroscopic quantity, Ψ is calculated from equation 4.2.7.

$$\Psi = \int \psi(\xi) f^{eq} \mathrm{d}\,\xi \tag{4.2.7}$$

Letting $\boldsymbol{\xi} = \sqrt{2\theta}\mathbf{c}$, where $\mathbf{c}$ is a rescaled thermal velocity (similar rescaling $\mathbf{u}$), equations 4.2.6 and 4.2.7 can be combined to produce equation 4.2.8.

$$\Psi = \int \exp\left(-\mathbf{c}^2\right) \psi(\mathbf{c}) \frac{\sqrt{2\theta}\rho}{(2\pi\theta)^{D/2}} \left[1 + 2(\mathbf{c}\cdot\mathbf{u}) + 2(\mathbf{c}\cdot\mathbf{u})^2 - \mathbf{u}^2\right] \mathrm{d}\,\mathbf{c} \tag{4.2.8}$$

Using Gaussian quadrature on equation 4.2.8, produces a sum to find the integral, equation 4.2.9.

$$\Psi = \sum_i \psi(\mathbf{c}_i) \frac{\sqrt{2\theta}\rho}{(2\pi\theta)^{D/2}} w(\mathbf{c}_i) \left[1 + 2(\mathbf{c}\cdot\mathbf{u}) + 2(\mathbf{c}\cdot\mathbf{u})^2 - \mathbf{u}^2\right] \tag{4.2.9}$$

The value of $w(\mathbf{c}_i)$ can be obtained from Gauss-Hermite integration. Equation 4.2.10 is the equilibrium particle distribution function in the discrete regime, where w_i is called the weight factor for the speed vector $\mathbf{c}_i$.

$$f_i^{eq} = w_i \rho \left[1 + 2(\mathbf{c}\cdot\mathbf{u}) + 2(\mathbf{c}\cdot\mathbf{u})^2 - \mathbf{u}^2\right] \tag{4.2.10}$$

4.3 Dissipative Particle Dynamics

Dissipative particle dynamics (DPD) is a mesoscopic simulation technique for complex fluids that can study systems over larger length and time-scales than classical molecular dynamics and Monte Carlo simulations. Dissipative particle dynamics is a relatively

recent simulation technique where particles move according to Newtonian dynamics, interacting with each other via soft repulsive interactions and with the implicit solvent via random and dissipative forces, representing Brownian and frictional forces respectively [2]. Each particle represents several other atoms, e.g. a $-CH_2-CH_2-CH_2-$ group (Warren has in fact represented surfactants as a DPD dimer [3]).

4.3.1 General Equations

The total force, $\mathbf{F}_{ij}$, acting on a particular bead, i, due to another bead, j, can be represented by the sum of a conservative force, $\mathbf{F}_{ij}^{C}$, a particle-particle interaction; a dissipative force, $\mathbf{F}_{ij}^{D}$, representing the friction of the particles between each other (the fluid viscosity); and a random force, $\mathbf{F}_{ij}^{R}$, representing the Brownian motion of the molecules, as in equation 4.3.1.

$$\mathbf{F}_{ij} = \mathbf{F}_{ij}^{C}\left(\mathbf{r}_{ij}\right) + \mathbf{F}_{ij}^{D}\left(\mathbf{r}_{ij}, \mathbf{v}_{ij}\right) + \mathbf{F}_{ij}^{R}\left(\mathbf{r}_{ij}\right) \tag{4.3.1}$$

The conservative force contains two contributions, for both bonded and non-bonded interactions. The force due to non-bonded interactions, $F_{ij}^{C,NB}$, is represented by a soft repulsive force, equation 4.3.2, where a_{ij} is the maximum repulsion between beads i and j, and $w^{R}\left(\mathbf{r}_{ij}\right)$ is a function to describe the interaction. The conservative energy is taken from a pairwise-additive with a weight function and variable amplitude that sets the temperature scale in the system [4]. As long as the conservative force is sufficiently weak, it does not induce appreciable inhomogeneities in the density around a particle. One consequence on this approximation is that phase separation between disordered phases cannot occur in a pure system; at the least, a binary mixture of different kinds of particles is needed [5].

$$\mathbf{F}_{ij}^{C,NB} = \begin{cases} a_{ij} w^{R}\left(\mathbf{r}_{ij}\right) \hat{\mathbf{r}}_{ij} & \|r_{ij}\| < r_c \\ 0 & \|r_{ij}\| \geq r_c \end{cases} \tag{4.3.2}$$

Particles that are in a chain also have a force due to bonded interactions, $F_{ij}^{C,B}$, given by equation 4.3.3, which is based around a simple harmonic spring force, which is proportional to the particle separation [6].

$$\mathbf{F}_{ij}^{C,B} = \begin{cases} -C_{ij}\left(\|\mathbf{r}_{ij}\| - r_0\right) \hat{\mathbf{r}_{ij}} & \text{if } i \text{ and } j \text{ are bonded} \\ 0 & \text{otherwise} \end{cases} \tag{4.3.3}$$

As well as the simple Hookean spring potential, there are examples of more complex potentials being used, including the finitely extensible nonlinear elastic potential [7, 8].

The dissipative force and the random force are given by equations 4.3.4 and 4.3.5 respectively, where γ is a friction coefficient, σ is the noise amplitude, and ξ_{ij} is a Gaussian random number with zero mean and unit variance, equation 4.3.6. The difference in velocity between the beads is $\mathbf{v}_{ij} = \mathbf{v}_j - \mathbf{v}_i$, and $w^{D}\left(\mathbf{r}_{ij}\right)$ is a function to describe the interaction.

$$\mathbf{F}_{ij}^{D} = \begin{cases} -\gamma_{ij} w^{D}\left(\mathbf{r}_{ij}\right)\left(\hat{\mathbf{r}}_{ij} \cdot \mathbf{v}_{ij}\right) \hat{\mathbf{r}}_{ij} & \|r_{ij}\| < r_c \\ 0 & \|r_{ij}\| \geq r_c \end{cases} \tag{4.3.4}$$

$$\mathbf{F}_{ij}^{R} = \begin{cases} \sigma_{ij} w^{R}\left(\mathbf{r}_{ij}\right) \xi_{ij} \Delta t^{-1/2} \hat{\mathbf{r}}_{ij} & \|r_{ij}\| < r_c \\ 0 & \|r_{ij}\| \geq r_c \end{cases} \tag{4.3.5}$$

$$\begin{aligned} \langle \xi_{ij}(t) \rangle &= 0 \\ \langle \xi_{ij}(t)\, \xi_{i'j'}(t') \rangle &= (\delta_{ii'}\delta_{jj'} + \delta_{ij'}\delta_{ji'})\, \delta(t-t') \end{aligned} \tag{4.3.6}$$

To produce the correct canonical ensemble the friction coefficient, γ, and the noise amplitude, $sigma$, need to be related to each other as in equation 4.3.7. The functions w^R and w^D also need to be related to each other as in equation 4.3.8. This is further discussed in Section 4.3.3. The function w^R can actually be any function desired, the most common used for accuracy, simulation speed, and ease of calculation is equation 4.3.9.

$$\sigma^2 = 2\gamma kT \tag{4.3.7}$$

$$w^D(\mathbf{r}_{ij}) = w^R(\mathbf{r}_{ij})^2 \tag{4.3.8}$$

$$w^R(\mathbf{r}_{ij}) = \left(1 - \frac{\|\mathbf{r}_{ij}\|}{r_c}\right) \tag{4.3.9}$$

4.3.2 Repulsive Parameters

For the thermodynamic state of the arbitrary DPD liquid to be described correctly by the soft sphere model, the fluctuations in the liquid need to be correctly described, i.e. the compressibility of the system needs to be correct. Therefore a dimensionless compressibility, κ^{-1}, can be defined analogously to the Weeks-Chandler-Anderson perturbation theory of liquids as equation 4.3.10, where n is the number density and κ_T is the isothermal compressibility.

$$\kappa^{-1} = \frac{1}{nk_BT\kappa_T} = \frac{1}{k_BT}\left(\frac{\partial p}{\partial n}\right)_T \tag{4.3.10}$$

The pressure of the system can be obtained from the virial theorem as equation 4.3.11.

$$p = \rho k_BT + \frac{1}{3V}\left\langle \sum_{j>i} \mathbf{r}_i - \mathbf{r}_j \cdot \mathbf{F}_i \right\rangle \tag{4.3.11}$$

As $w^D(r) = \left[w^R(r)\right]^2$ and $\sigma^2 = 2\gamma k_BT$, equation 4.3.11 can be simplified to produce equation 4.3.12, where $g(r)$ is the radial distribution function.

$$\begin{aligned} p &= \rho k_BT + \frac{1}{3V}\left\langle \sum_{j>i} \mathbf{r}_{ij} \cdot \mathbf{F}_{ij}^C \right\rangle \\ &= \rho k_BT + \frac{2\pi}{3}\rho^2 \int_0^1 r f(r) g(r) r^2 \mathrm{d}r \end{aligned} \tag{4.3.12}$$

The value of $2\pi/3\int_0^1 r f(r) g(r) r^2 \mathrm{d}r$ has been found to be equal to $0.1a_{ii}$ for $\rho > 2.5$, where a_{ii} is the self-repulsion constant, Figure 4.1 [9].

The dimensionless compressibility is thus given by equation 4.3.13.

$$5\kappa^{-1} = 1 + \frac{0.2a_{ii}\rho}{k_BT} \tag{4.3.13}$$

The known dimensionless compressibility of water is 15.98, so the self repulsion constant can be given by equation 4.3.14.

$$a_{ii} = \frac{75k_BT}{\rho} \tag{4.3.14}$$

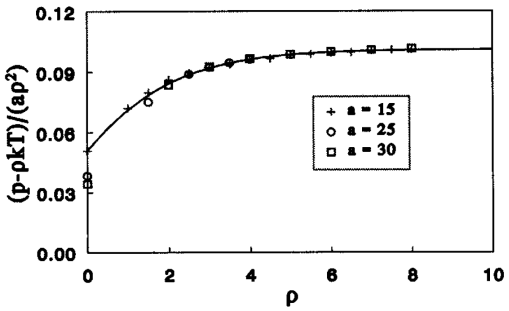


Figure 4.1: Variation of calculated pressure with system density [9].

The repulsion constant for the other types of particle can be given by equation 4.3.15, where Δa is the difference between the self and other type repulsion parameters.

$$a_{ij} = a_{ii} + \Delta a \tag{4.3.15}$$

Groot and Warren (1997) [9] comment that the method of DPD is similar to the Flory-Huggins theory of polymers, and can be viewed as a continuous version of this lattice model. This means that the effective Flory-Huggins parameter, $(\chi N)_{eff}$, can be used to predict the value of Δa. The general Flory-Huggins free energy for a two-component mixture is given by equation 4.3.16, where ϕ_A is the volume fraction of component A, and N_A and N_B are the number of beads of A and B respectively.

$$\frac{G}{k_B T} = \frac{\phi_A}{N_A} \ln \phi_A + \frac{1-\phi_A}{N_B} \ln (1-\phi_A) + \chi \phi_A (1-\phi_A) \tag{4.3.16}$$

The first term represents the free energy of the A type particles, the second term represents the free energy of the B type particles, and the third term represents the free energy interaction between the two particle types. Figure 4.2 shows the plot of the free energy of a mixture of polymers of type A and B with different values of χ. The larger the value of χ the more the different polymers dislike each other and want to phase separate. This is represented by a maxima in the free energy curve, producing a two phase system. If $N_A = N_B = N/2$ then the minimum free energy is found at $\mu = \partial G / \partial \phi = 0$, which means that the free energy is minimised when equation 4.3.17 is true.

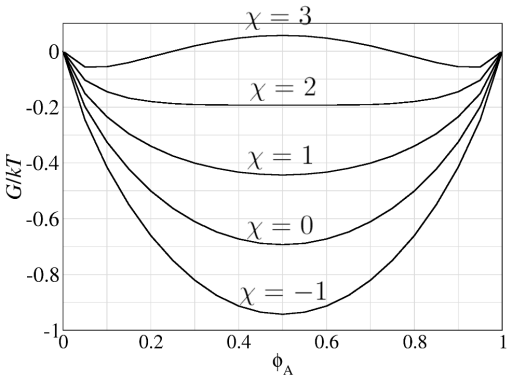


Figure 4.2: Variation of the mixture free energy for varying values of χ for polymers of $N_A = N_B$. Phase separation occurs as the χ parameter exceeds approximately 2.

$$\chi N_A = \frac{\ln \left[\frac{(1-\phi_A)}{\phi_A} \right]}{1 - 2\phi_A} \tag{4.3.17}$$

Equation 4.3.17 produces two value if the value of χ is larger than about 2. These values are the concentration of the two immiscible phases. Equation 4.3.17 is accurate for very long polymer systems. For DPD simulations the chains used are often of a small finite length. Horsch *et al.* (2004) [7] carried out simulations using hard potential Brownian Dynamics and found that they produced equation 4.3.18, when using short chains.

$$\frac{\ln \left[\frac{(1-\phi_A)}{\phi_A} \right]}{1 - 2\phi_A} \approx 1 + 3.9 N^{-0.51} (\chi N)_{eff} \tag{4.3.18}$$

Groot and Warren (1997) [9] carried out simulations with different values of Δa and have found that for values of the chain length, N, between two and ten and calculated the (χN) value, Figure 4.3. This produces the correlation given by equation 4.3.19.

$$\chi N_A = 0.306 N \Delta a \tag{4.3.19}$$

Combining equations 4.3.15, 4.3.17, 4.3.18, and 4.3.19 produces a value for the repulsion constant in terms of the effective Flory-Huggins parameter, equation 4.3.20.

$$a_{ij} = a_{ii} + \frac{1 + 3.9N^{-0.51}}{0.306N} (\chi N)_{eff} \tag{4.3.20}$$

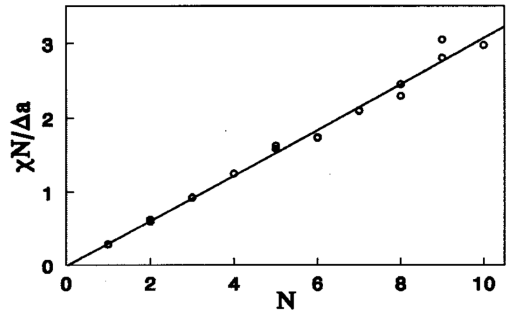


Figure 4.3: Variation of the (χN) for varying chain length and Δa [9].

The key process is now the setting of the effective Flory-Huggins parameter. This must be above the value of equation 4.3.21, which corresponds to the order-disorder transition (ODT) [10], if any kind of meso-phase ordering is to form.

$$(\chi N)_{\text{ODT}} = 10.5 + N^{-\frac{1}{3}} \tag{4.3.21}$$

Horsch *et al.* (2004) [7] look at a range of different effective Flory-Huggins parameters but do not comment on which they think is best as they achieve very similar results for any in the range of 23.7-57.9. Lísal and Brennan (2007) [6], however, suggest the use of $(\chi N)_{eff} = 35$ producing values of $a_{AA} = a_{BB} = 25$ and $a_{AB} = 50$ for their chains of 5A-5B. Even with the partial theoretical and partial simulation analysis of the calculation of a_{ij} there is still a lot of scope for the selection of almost any value desired, as there is no prediction of $(\chi N)_{eff}$ available. This means that the values should be tested and if possible checked before carrying out simulations in undiscovered areas. The other main problem with this approach is the equations were generated assuming there are only two types of particle in the simulation, and that these have equal numbers, i.e chains of 5A-5B and not 8A-2B etc.

The spring constant, K_{ij}, has to be chosen to represent the stiffness of the chain: the larger the value of the spring constant the stiffer the chain. The value of the spring constant also affects the length of the apparent bonds between particles in the chain. The larger the value of the spring constant the shorter the apparent bond length. Lísal and Brennan (2007) [6] use a value of $K_{ij} = 4$ which with their interaction parameters produces an apparent bond length of 0.75.

4.3.3 Dissipative and Random Force Parameters

The relationship between the dissipative and random force parameters can be determined by examining the probability of particle positions and velocities. The Chapman-Kolmogorov equation can be used to look at the probability of the particles being in the range of $\left(\mathbf{r}^N, \mathbf{v}^N\right)$ and $\left(\mathbf{r}^N + \Delta\mathbf{r}^N, \mathbf{v}^N + \Delta\mathbf{v}^N\right)$, equation 4.3.22, where W is the probability density function, Ψ is the transition probability of change during a time interval Δt, and N represents all the particles.

$$\begin{aligned} W\left(\mathbf{r}^N, \mathbf{v}^N, t\right) = \iint & W\left(\mathbf{r}^N - \Delta\mathbf{r}^N, \mathbf{v}^N - \Delta\mathbf{v}^N, t - \Delta t\right) \\ & \times \Psi\left(\mathbf{r}^N - \Delta\mathbf{r}^N, \mathbf{v}^N - \Delta\mathbf{v}^N; \Delta\mathbf{r}^N, \Delta\mathbf{v}^N\right) \mathrm{d}\left(\Delta\mathbf{r}^N\right) \mathrm{d}\left(\Delta\mathbf{v}^N\right) \end{aligned} \tag{4.3.22}$$

If the velocity does not change appreciably during the small time interval, the Ψ can be simplified by equation 4.3.23, where ψ is the transition probability which is independent

of $\Delta\mathbf{r}^N$.

$$\Psi\left(\mathbf{r}^N - \Delta\mathbf{r}^N, \mathbf{v}^N - \Delta\mathbf{v}^N; \Delta\mathbf{r}^N, \Delta\mathbf{v}^N\right) = \psi\left(\mathbf{r}^N - \Delta\mathbf{r}^N, \mathbf{v}^N - \Delta\mathbf{v}^N; \Delta\mathbf{v}^N\right)\delta\left(\Delta\mathbf{r}^N - \mathbf{v}^N\Delta t\right) \quad (4.3.23)$$

Substituting equation 4.3.23 into equation 4.3.22 produces equation 4.3.24.

$$W\left(\mathbf{r}^N + \mathbf{v}^N\Delta t, \mathbf{v}^N, t + \Delta t\right) = \iint W\left(\mathbf{r}^N, \mathbf{v}^N - \Delta\mathbf{v}^N, t\right) \times \psi\left(\mathbf{r}^N, \mathbf{v}^N - \Delta\mathbf{v}^N; \Delta\mathbf{v}^N\right)\mathrm{d}\left(\Delta\mathbf{v}^N\right) \quad (4.3.24)$$

Carrying out a Taylor series expansion on equation 4.3.24 and using equations 4.3.25 and 4.3.26, the Fokker-Planck equation in phase space (or the Chandrasekhar equation) can be defined, equation 4.3.27 [11].

$$\Delta\mathbf{r}_i = \mathbf{v}_i\Delta t \quad (4.3.25)$$

$$\Delta\mathbf{v}_i = \frac{1}{m}\left(\sum_{j(\neq i)}\mathbf{F}_{ij}^C - \sum_{j(\neq i)}\gamma w_D\left(\hat{\mathbf{r}}_{ij}\cdot\mathbf{v}_{ij}\right)\hat{\mathbf{r}}_{ij}\right)\Delta t + \frac{1}{m}\sum_{j(\neq i)}\sigma w_R\hat{\mathbf{r}}_{ij}\Delta t^{-\frac{1}{2}}\int_t^{t+\Delta t}\xi_{ij}\mathrm{d}\tau \quad (4.3.26)$$

$$\frac{\partial W}{\partial t} + \sum_i \mathbf{v}_i\cdot\frac{\partial W}{\partial \mathbf{r}_i} + \sum_{\substack{i\\(i\neq j)}}\sum_j \frac{\mathbf{F}_{ij}^C}{m}\cdot\frac{\partial W}{\partial \mathbf{v}_i} = \sum_{\substack{i\\(i\neq j)}}\sum_j \hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i}\left[\frac{1}{m}\gamma w_D\left(\hat{\mathbf{r}}_{ij}\cdot\mathbf{v}_{ij}\right)W\right] + \frac{1}{2}\sum_{\substack{i\\(i\neq j)}}\sum_j \frac{1}{m^2}\sigma^2 w_R^2\hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i}\left[\hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i} - \hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i}\right]W \quad (4.3.27)$$

If the system of dissipative particles is in equilibrium, the probability density function, W, can be represented as the canonical distribution (probability distribution of states) for an ensemble with a specified number of molecules, N, volume, V and temperature, T, which is given by equation 4.3.28, where Z is the partition function.

$$W = \frac{1}{Z}\exp\left[-\frac{1}{k_BT}\left(\sum_i m\frac{v_i^2}{2} + U\right)\right] \quad (4.3.28)$$

Using equation 4.3.28 it can be seen that the left hand side of equation 4.3.27 becomes equal to zero. This means that for the system to be able to reach an equilibrium then equation 4.3.29 must hold.

$$\sum_{\substack{i\\(i\neq j)}}\sum_j \hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i}\left[\frac{1}{m}\gamma w_D\left(\hat{\mathbf{r}}_{ij}\cdot\mathbf{v}_{ij}\right)W\right] = -\frac{1}{2}\sum_{\substack{i\\(i\neq j)}}\sum_j \frac{1}{m^2}\sigma^2 w_R^2\hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i}\left[\hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i} - \hat{\mathbf{r}}_{ij}\cdot\frac{\partial}{\partial \mathbf{v}_i}\right]W \quad (4.3.29)$$

Solving equation 4.3.29 shows that equations 4.3.8 and 4.3.7 are a simple but non-trivial solution.

$$w_D = w_R^2 \tag{4.3.8}$$

$$\sigma^2 = 2\gamma w_B T \tag{4.3.7}$$

The value of σ itself cannot be predicted but has to selected considering the properties of the system: a large value of σ indicates there is a large amount of random motion of the particles, and also due to equation 4.3.7 indicates the friction resistance to is high. This in turn produces a system where the particles cannot move easily and struggle to reach an equilibrium structure due to all the random forces. Conversely, a small value of σ indicates an unrealistically small amount of random motion and also a small friction resistance, which could mean the particles get stuck in a metastable structure. Groot and Warren (1997) [9] look at a range of values for σ and state that at values of σ too high the simulation may become unstable. They recommend the use of $\sigma = 3$ and say at this value there is good control of the temperature at fairly long time steps (0.04). Most other authors seem to use a value of $\sigma = 3$ [6, 8].

4.3.4 Time Step

The real time scale of a simulation can be calculated from equation 4.3.30, where σ is the Lennard-Jones parameter (in the case of DPD r_{cut}).

$$t = \sigma \left(\frac{m}{kT}\right)^{1/2} \tag{4.3.30}$$

Groot and Warren (1997) [9] recommend the dimensionless time step, δt^*, to be 0.04 as this gives little error, however they say that values of up to 0.05 are still fine. A dimensionless time step of 0.04-0.05 gives small errors of up to about 3% [14].

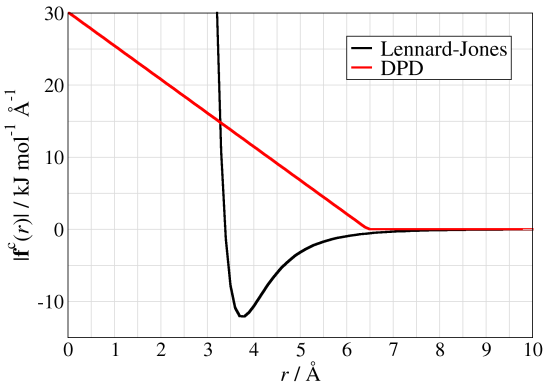


Figure 4.4: Variation of the conservative force for water: SSD model Lennard-Jones parameters [12] and DPD parameters [13].

Table 4.1 summarises the real simulated time for each time step for an typical MD simulation and a typical DPD simulation. The large time step for DPD in comparison to MD shows how much larger timescale and system sizes can be simulated.

Table 4.1: Simulation timescales: comparison of DPD with MD.

	MD	DPD
σ / Å	3.016[a]	6.46[b]
t / ps	~ 0.81	~ 3.02
δt^*	< 0.002	~ 0.04
δt / fs	< 2	~ 120

[a] From SSD model [12]
[b] From [13]

4.4 References

[1] S. Succi, editor. *The Lattice Boltzmann Equation, for Fluid Dynamics and Beyond*. Oxford Science Publications, 2001.

[2] D. Frenkel and B. Smit, editors. *Understanding Molecular Simulation From Algorithms to Applications*. Academic Press, 2002.

[3] P. Prinsen, P. B. Warren, and M. A. J. Michels. Mesoscale Simulations of Surfactant Dissolution and Mesophase Formation. *Physical Review Letters*, 89:148302–1–4, 2002.

[4] I. Pagonabarraga and D. Frenkel. Dissipative Particle Dynamics for Interacting Systems. *Journal of Chemical Physics*, 115:5015–5026, 2001.

[5] P. V. Coveney and K. E. Novik. Computer Simulations of Domain Growth and Phase Separation in Two-Dimensional Binary Immiscible Fluids using Dissipative Particle Dynamics. *Physical Reviews E*, 54:5134–5141, 1996.

[6] M Lísal and J. K. Brennan. Alignment of Lamellar Diblock Copolymer under Shear: Insight for Dissipative Particle Dynamics Simulations. *Langmuir*, 23:4809–4818, 2007.

[7] M. A. Horsch, Z. Zhang, C. R. Iacovella, and S. C. Glotzer. Hydrodynamics and Microphase Ordering in Block Copolymers: Are Hydrodynamics Required for Ordered Phases with Periodicity in more than One Dimension. *Journal of Chemical Physics*, 121:11455–11462, 2004.

[8] S. Chen, N. Phan-Thien, X.-J. Fan, and B. C. Khoo. Dissipative Particle Dynamics Simulation of Polymer Drops in a Periodic Shear Flow. *Jounal of Non-Newtonian Fluid Mechanics*, 118:65–81, 2004.

[9] R. D. Groot and P. B. Warren. Dissipative Particle Dynamics: Bridging the Gap Between Atomistic and Mesoscopic Simulation. *Journal of Chemical Physics*, 107:4423–4435, 1997.

[10] G. H. Fredrickson and E. Helfand. Fluctuation Effects in the Theory of Microphase Seperation in Block Copolymers. *Journal of Chemical Physics*, 87:697–705, 1987.

[11] A. Satoh. *Studies in Interface Science: Introduction to Molecular Microsimulation of Colloid Dispersion*. Elsevier Science, 2003.

[12] M-L. Tan, J. T. Fischer, A. Chandra, B. R. Brooks, and T.ÂăIchiye. A temperature of maximum density in soft sticky dipole water. *Chemical Physics Letters*, 376:646–652, 2003.

[13] R. D. Groot and K. L. Rabone. Mesoscopic Simulation of Cell Membrane Damage, Morphology Change and Rupture by Nonionic Surfactants. *Biophysical Journal*, 81:725–736, 2001.

[14] B. Hafskjold, C. C. Liew, and W. Shinoda. Can such Long Time Steps Really be used in Dissipative Particle Dynamics Simulations? *Molecular Simulation*, 30:879–885, 2007.