

Chapter I

STATE OF THE ART: SOLVATION, MODELS AND METHODS

This chapter gives a brief review of all the basic concepts and previous work that this thesis is based on.

In section 2, we begin by introducing the frequent models of solvent in a simulation, from the simplest implicit continuum model to the most complex explicit one. The overview of these models then helps to understand the choice of description scale used in our study, as well as its limits.

Once the model is chosen, all the theories become mathematical problems. Section 3 reviews some basic concepts of statistical mechanics for liquids (i.e. theory of liquids), which present some brief formalisms deduced for an atomic solvent model. Two frontier approaches are introduced with a few deductions: the classical density functional theory (cDFT), and the integral equation theory (IET). A mathematical equivalence between these two theories is also presented.

The following section 4 gives the extension of the two theories to the molecular solvent case, i.e. the molecular density functional theory (MDFT) that this thesis works upon, and the molecular Ornstein-Zernike (MOZ) approach for IET. The equivalence between these two theories gave us the idea to use the expansion techniques in IET to serve MDFT.

Section 5 gives some supplementary presentation of the initial code MDFT, which the development of this thesis is based on.

2

MODEL OF SOLUTION SYSTEM

Computing models of solvents are broadly divided into two types: those treating the solvent as a continuous medium (implicit models) and those describing the individual solvent molecules (explicit models). In the continuum model, the solvent is characterized by the dielectric constant ε and contains an artificially shaped cavity. The explicit models can have more specific microscopic scales. Within the scope of classical mechanics, the most detailed (and thus the most expensive) methods involve flexible and polarizable explicit models, while in computational chemistry, less detailed models often have wider usage (for example, the proteins are treated in the unity of residues). As the theory of liquids was initially established for spherical atom-like solvent particles, the model adopted by such a theory is a rigid entity carrying distributed point charges, characterized by their position and orientation, i.e. there is no internal movement considered. This approximation has been proven reasonable [13]. There also exist models in which the scale lies somewhere between the implicit and explicit models; for example, so-called coarse-grained models [34], which gather groups of atoms into a single interaction site.

In this section, we will give a brief introduction of the implicit model in order to facilitate later discussion on solvation free energy corrections. We will then focus on the rigid solvent models and discuss the limits of such approximations. The flexible and polarizable models will also be briefly mentioned.

2.1 CONTINUUM SOLVATION MODELS

Continuum models [7–9], which are popular in QM calculations, consider the solvent as a polarizable medium with dielectric constant ε , with the solute M placed in the cavity within this medium (figure 2.1)

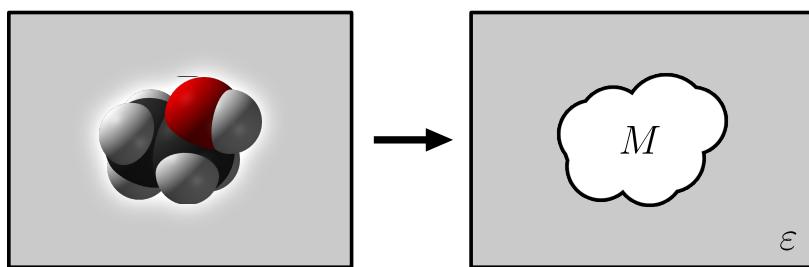


Figure 2.1: Continuum solvent model

The dielectric constant ε is the key parameter characterizing the solvent. It is normally a constant value, but that can depend on the distance from the solute M . (see §2.1.1)

The solvation Gibbs free energy according to this model is

$$\Delta G_{\text{solvation}} = \Delta G_{\text{cavity}} + \Delta G_{\text{dispersion}} + \Delta G_{\text{elec}} \quad (2.1)$$

where $\Delta G_{\text{cavity}} > 0$ is the energy needed to create a hole in the medium, and $\Delta G_{\text{dispersion}}$ is the dispersion interaction, which is roughly the van der Waals energy $\Delta G_{\text{vdW}} < 0$ between the solvent and solute. In principle, there may also be a repulsive component, and the dispersion term is sometimes denoted dispersion/repulsion. $\Delta G_{\text{elec}} < 0$ is the contribution of electrostatic interactions, introduced by electric charge distribution of M

which polarizes the medium, and the action back of the medium on the molecule (reaction field).

The initial two terms in eq. (2.1) are linked to the configuration of the first solvation shell (cavity). The definition of cavity varies from the simplest sphere or ellipsoid to the ensemble of atomic surfaces defined by the van der Waals radii in the solute. It is somewhat reasonable to consider the cavity area as proportional to the number of solvent molecules in the first solvation shell. This number can be calculated as the area passing through the middle region of first shell solvent. This area, named the solvent-accessible surface area (SASA) [35, 36], can be calculated by adding the radius of the probe solvent ball on the solvent excluded surface area (figure 2.2).

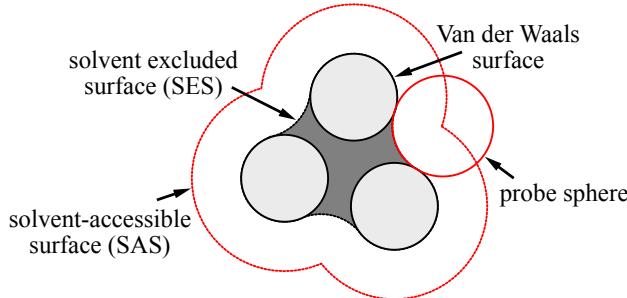


Figure 2.2: Definition of cavity surfaces. The solvent accessible surface (SAS) traced out by the center of the probe representing a solvent molecule. The solvent excluded surface (SES) is the topological boundary of the union of all possible probes that do not overlap with the molecule.

The energy required to create such a cavity and the stabilization due to van der Waals interactions between the solute and solvent, assumed to be proportional to the surface area of the cavity, is expressed as

$$\Delta G_{\text{cavity}} + \Delta G_{\text{dispersion}} = \gamma S_{\text{SASA}} + \beta \quad (2.2)$$

or parameterized by having a constant ξ specific for each atom type, with the ξ parameters being determined by fitting to experimental solvation data:

$$\Delta G_{\text{cavity}} + \Delta G_{\text{dispersion}} = \sum_i^{\text{atoms}} \xi_i S_i \quad (2.3)$$

The models and methods employed to calculate the electrostatic contribution ΔG_{elec} have varied greatly according to their usage. The sections below list the most common examples. On another topic, the integration of continuum models into QM calculations is also a very important field; these developments will not be detailed here as they do not connect yet to our work. Such kinds of methods are called the self-consistent reaction field (SCRF) models, which integrate the calculation of the solute-solvent interaction in addition to that of the solute wave function by an iterative procedure. Some examples are presented in the list of Gaussian keyword SCRF [37], and the field is well reviewed by, for example, Tomasi [9, 10] and Jensen [7].

2.1.1 Poisson-Boltzmann methods

The Poisson-Boltzmann equation (PBE) [38] makes it possible to calculate the position-dependent electrostatic potential $V_{\text{elec}}(\mathbf{r})$ in the continuum model, such that the electrostatic component of the free energy can be written as

$$\Delta G_{\text{elec}} = \frac{1}{2} \int d\mathbf{r} \rho_q(\mathbf{r}) V_{\text{elec}}(\mathbf{r}) \quad (2.4)$$

where ρ_q is the charge distribution of the solute.

The Maxwell-Gauss equation in SI units convention gives

$$\nabla \cdot D(\mathbf{r}) = \frac{\rho_q(\mathbf{r})}{\varepsilon_0} \quad (2.5)$$

where $D(\mathbf{r}) = \varepsilon_0 E(\mathbf{r}) + P(\mathbf{r})$ is the electric displacement field, $P(\mathbf{r})$ is the system polarization, $E(\mathbf{r})$ the electric field, and ε_0 the vacuum permittivity. $D(\mathbf{r})$ can also be expressed in terms of the position-dependent dielectric constant $\varepsilon(\mathbf{r})$, $D(\mathbf{r}) = \varepsilon(\mathbf{r})E(\mathbf{r})$, which thus gives

$$\nabla \cdot \varepsilon(\mathbf{r})E(\mathbf{r}) = \frac{\rho_q(\mathbf{r})}{\varepsilon_0} \quad (2.6)$$

or in terms of electrostatic potential:

$$\nabla \cdot [\varepsilon(\mathbf{r}) \nabla V_{\text{elec}}(\mathbf{r})] = -\frac{\rho_q(\mathbf{r})}{\varepsilon_0} \quad (2.7)$$

This second-order differential equation (2.7) is called the Poisson equation.

This equation cannot be solved analytically for complex geometries (such as a protein). Therefore it is done numerically using appropriate methods; for example, mentioned in the article of Roux and Simonson [39] or Holst [38]. A density functional approach based on the minimization of the polarization density can also be used to solve this equation [40, 41].

If the solvent is ionic, the Poisson equation can be modified by taking into account a (thermal) Boltzmann distribution of ions in the solvent, i. e., (equation to be verified with a factor 2 before the ionic contribution. I don't find the definition of n_{ion} .)

$$\rho_{\text{tot}}(\mathbf{r}) = \rho_q(\mathbf{r}) + q n_{\text{ion}} \sinh\left(\frac{q q_e}{k_B T} V_{\text{elec}}(\mathbf{r})\right) \quad (2.8)$$

for a salt composed of ions of charge $+q$ and $-q$ and of density n_{ion} . Replacing in eq. (2.7) leads to the Poisson-Boltzmann Equation: (here q is of which unity?)

$$\nabla \cdot (\varepsilon(\mathbf{r}) \nabla V_{\text{elec}}(\mathbf{r})) - \frac{q n_{\text{ion}}}{\varepsilon_0} \sinh\left(\frac{q V_{\text{elec}}(\mathbf{r})}{k_B T}\right) = -\frac{\rho(\mathbf{r})}{\varepsilon_0} \quad (2.9)$$

2.1.2 Born / Onsager / Generalized Born models

For simple geometries, the Poisson equation (2.7) can be solved analytically. The simplest model is a spherical cavity. For a net charge q in a cavity of radius a , the electrostatic free energy of a medium with a dielectric constant of ε is given by the Born formula:

$$\Delta G_{\text{elec}}(q) = -\frac{1}{8\pi\varepsilon_0} \left(1 - \frac{1}{\varepsilon}\right) \frac{q^2}{2a} \quad (2.10)$$

Other similar models include the Onsager model, in which a point dipole (characterized by the dipole moment μ) is put in a spherical cavity. The Kirkwood model refers to a general multipole expansion in a spherical cavity, while the Kirkwood-Westheimer model arises for an ellipsoidal cavity. Those simplified models are not fully able to predict the solvent behavior in many realistic cases [7].

The generalized Born (GB) model is an empirical model based on the superposition of several net charges in spherical cavities as the Born model describes, with a similar formula:

$$\Delta G_{\text{elec}} = -\frac{1}{8\pi\varepsilon_0} \left(1 - \frac{1}{\varepsilon}\right) \sum_i \sum_j \frac{q_i q_j}{f_{ij}} \quad (2.11)$$

where the function f_{ij} depends on the internuclear distance r_{ij} between the centers of atoms i and j and on the Born radii for each pair of atoms a_i and a_j :

$$f_{ij} = \sqrt{r_{ij}^2 - a_i a_j \exp\left(\frac{r_{ij}^2}{4a_i a_j}\right)} \quad (2.12)$$

The key (empirical) point is to be able to attribute an effective Born radius a_i to each atom inside the complex, non-spherical cavity formed by the solute. Once this is accomplished, the GB model provides a very fast method, with an overall accuracy comparable to that of Poisson-Boltzmann calculations. That makes it widely used in computational structural biology to perform structure optimization and molecular dynamics simulations.

2.2 MODEL POTENTIAL OF EXPLICIT MOLECULES

The model potential frequently used in the theory of liquids is a classical, rigid, pairwise additive model [12, 13]. It is based on three assumptions.

1. Firstly, the quantum effects should be ignored. It is assumed that the rotational and transitional motion of solvent particles are continuous and classical, which means the separation of both transitional and rotational states are largely inferior of $k_B T$. For light molecules, that is not always convincing. Some molecules containing hydrogen (e. g. H₂O, NH₃, and particularly H₂) exhibit obvious quantum effects at low temperature in the liquid state. Gaseous H₂O and NH₃ also need quantum effect corrections. However, for the liquid of most interest to us, H₂O at room temperature, the contribution of this effect is small enough to be neglected. And obviously, there should not be any chemical interaction of the solvent with the solute.
2. Secondly, the intramolecular movement (vibration and internal rotation) should be either independent of transitional and rotational movement or absent. This rigid molecule approximation assumes that the intermolecular potential $\mathcal{U}(\mathbf{r}^N, \Omega^N)$ for N particles only depends on the positions of the N molecular centers $\mathbf{r}^N \equiv (\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_N)$ and on their orientations $\Omega^N \equiv (\Omega_1, \Omega_2, \dots, \Omega_N)$, where $\Omega \equiv (\Theta, \Phi, \Psi)$ represents the Euler angles (figure 2.3). The natural choice for the molecular center is the center of mass. This is, however, arbitrary if only equilibrium properties are considered.
3. Finally, the intermolecular forces have to be assumed as pairwise additive:

$$\mathcal{U}(\mathbf{r}^N, \Omega^N) = \frac{1}{2} \sum_{i \neq j} u(\mathbf{r}_{ij}, \Omega_i, \Omega_j) = \sum_{i < j} u(\mathbf{r}_{ij}, \Omega_i, \Omega_j) \quad (2.13)$$

This means that the model potential only depends on the intermolecular separation \mathbf{r} and on the molecular orientations Ω_1 and Ω_2 . This approximation is quasi-exact for low density gases, where the contribution of the three and more body terms decreases rapidly. But for dense fluids, in most of the cases the multi-body potential cannot be ignored. The complete model potential with higher-order corrections can be written in the form of

$$\mathcal{U}(\mathbf{r}^N, \Omega^N) = \sum_{i < j} u(ij) + \sum_{i < j < k} u(ijk) + \sum_{i < j < k < l} u(ijkl) + \dots \quad (2.14)$$

Compared to atomic models that only depend on \mathbf{r}^N , the angular correlations can give influence on both structural and thermodynamic proprieties. That is why our theory is extended to linear case, $\Omega \equiv (\Theta, \Phi)$, then molecular case, $\Omega \equiv (\Theta, \Phi, \Psi)$.

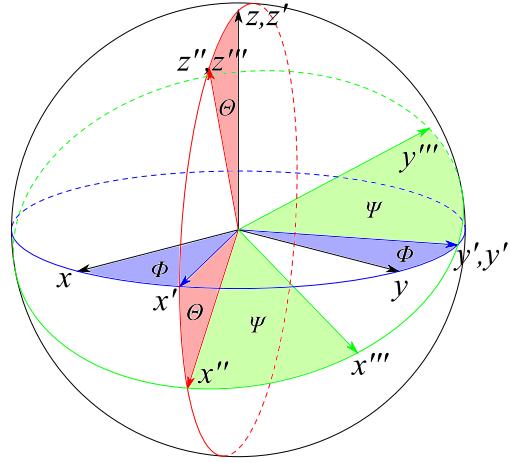


Figure 2.3: Euler angles. The basis vectors of the new orientation are obtained by 3 sequential operations: (1) A rotation ϕ ($0 < \phi < 2\pi$) about the z -axis, bringing the frame of axes from the initial position \mathbf{S} into the position \mathbf{S}' (2) A rotation θ ($0 < \theta < \pi$) about the y -axis of the frame \mathbf{S}' , which is transformed into \mathbf{S}'' (3) A rotation ψ ($0 < \psi < 2\pi$) about the z -axis of the frame \mathbf{S}'' .

where $u(ij) = u(\mathbf{r}_{ij}, \Omega_i, \Omega_j)$ and $u(ijk) = u(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{ki}, \Omega_i, \Omega_j, \Omega_k)$, etc. The omission of the three-body and higher-order terms can cause error, for example, in surface tension and surface energy calculation [42]. However the higher-order terms are often accounted for by an effective pair potential (measured by experiments or calculated by simulations), which reduces considerably the computational cost for simulations, or the degree of theory needed. Such models are presented below going from simple to molecular liquids. For the molecular solvent considered in this thesis, water, most publications have stayed at this two-body level of description.

2.2.1 Interaction of spherical particle

The simplest model of a fluid is the hard sphere model. With d the hard-sphere diameter, the pair potential is defined as:

$$u(r) = \begin{cases} \infty & r < d \\ 0 & r > d \end{cases} \quad (2.15)$$

This model is indeed a fundamental reference model in statistical mechanics, and it can represent some physical systems, such as neutral colloidal suspensions [43]. However, the absence of attractive force, which precludes the existence of a liquid-gas transition, makes it too simple for realistic fluids. More realistic neutral particle models, like the Lennard-Jones (LJ) model, exhibit a potential energy curve that has the same shape as the real interaction of rare gas, as shown in figure 2.4.

The Lennard-Jones (LJ) interaction gives

$$u_{LJ}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^6 \right] \quad (2.16)$$

where r is the distance from centre to centre, σ is the collision diameter or the particles separation where $u(r) = 0$, and ϵ is the well depth of the potential (of unity of energy). The well minimum occurs at $r_{\min} = 2^{1/6}\sigma$ and $u(r_{\min}) = -\epsilon$. The parameters σ and ϵ can be extracted from experiments.

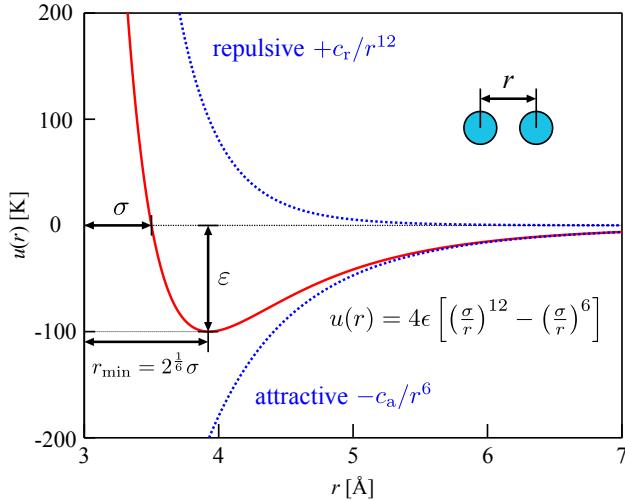


Figure 2.4: LJ pair potential. The plot gives the potential energy $u(r)$ versus internuclear distance r of two particles. At large distances, both attractive and repulsive interactions are small. As the distance between the atoms decreases, the attractive electron-proton interactions dominate, and the energy of the system decreases. At the observed bond distance, the repulsive electron-electron and proton-proton interactions just balance the attractive interactions, preventing a further decrease in the internuclear distance. At very short internuclear distances, the repulsive interactions dominate, making the system less stable than the isolated atoms.

Theoretically, all terms in the multipole series represent attractive contributions to the potential. The leading term, varying as r^{-6} , describes the quantum dipole-dipole interaction. Higher-order terms represent dipole-quadrupole (r^{-8}), quadrupole-quadrupole (r^{-10}) interactions, and so on, but these are negligible compared to r^{-6} . The short-range interaction is difficult to define properly, and for the sake of simplicity and numerical efficiency, it is defined as r^{-12} in the LJ model.

If the spherical particles are charged (as in molten salts), the electrostatic interaction between them is described by the Coulomb point charge interaction:

$$u_{\text{Coul}}(r) = \frac{q_1 q_2}{4\pi\epsilon_0 r} \quad (2.17)$$

For such charged simple fluids, the overall pair $u(r)$ is a sum of LJ and Coulomb interactions. Such decomposition can be extended to molecular fluids in terms of site-site interactions, which are discussed in the following section.

2.2.2 Site-site interactions

Indeed, a spherical description of interactions is not sufficient to fully describe molecular fluids. The site-site model is a further extension of atomic models, in which the solvent molecule is represented by a set of discrete interaction sites. The total potential energy is a sum of spherical interaction potentials:

$$u(1, 2) = \sum_{\alpha} \sum_{\beta} u_{\alpha\beta}(|\mathbf{r}_{2\beta} - \mathbf{r}_{1\alpha}|) \quad (2.18)$$

where \mathbf{r}_{is} is the coordinates of site s in molecule i , $u_{\alpha\beta}(r)$ the interatomic potential energy of pairs of sites α and β , as discussed above. More specifically, it is generally decomposed into a Lennard-Jones and a Coulombic contribution:

$$u(1, 2) = \sum_{\alpha} \sum_{\beta} \left\{ 4\epsilon_{\alpha\beta} \left[\left(\frac{\sigma_{\alpha\beta}}{r_{12}^{\alpha\beta}} \right)^{12} - \left(\frac{\sigma_{\alpha\beta}}{r_{12}^{\alpha\beta}} \right)^6 \right] + \frac{1}{4\pi\epsilon_0} \frac{q_{\alpha} q_{\beta}}{r_{12}^{\alpha\beta}} \right\} \quad (2.19)$$

where $r_{12}^{\alpha\beta} = |\mathbf{r}_{2\beta} - \mathbf{r}_{1\alpha}|$, $\epsilon_{\alpha\beta}$ and $\sigma_{\alpha\beta}$ are the site-site LJ parameters and q_α the partial charge on each site. This model is the most commonly adopted, and it will be so in this thesis for the calculation of external free energy functional.

2.2.3 Multipole and spherical harmonic expansion

To obtain fewer terms in the calculation, the model potential can be presented in convergent series via multipole or spherical harmonic expansion.

For polar liquids, the dipole-dipole interaction should be mainly taken into account. Thus the model considers dipole-dipole interactions in addition to a spherically symmetric Lennard-Jones-like potential:

$$u(1, 2) = u_0(r) - \boldsymbol{\mu}_1 \cdot \mathbf{T}(\mathbf{r}) \cdot \boldsymbol{\mu}_2 \quad (2.20)$$

where \mathbf{r} is the vector separation of the molecular centers, $u_0(r)$ is the spherically symmetric term discussed above, $\boldsymbol{\mu}_i$ is the dipole moment vector of particle i and $\mathbf{T}(\mathbf{r})$ is the dipole-dipole interaction tensor:

$$\mathbf{T}(\mathbf{r}) = \nabla^2 \left(\frac{1}{r} \right) = 3\mathbf{rr}/r^5 - \mathbf{I}/r^3 \quad (2.21)$$

and \mathbf{I} is the unit tensor. Note that this model can be made more realistic by including higher-order electrostatic interactions, such as dipole-quadrupole, quadrupole-quadrupole, etc. Such a systematic multipolar approach has been proposed for water [44].

Alternatively, the intermolecular potential can be expanded onto rotational invariants in the form:

$$u(\mathbf{r}_{12}, \Omega_1, \Omega_2) = \sum_{mnlm\mu\nu} u_{\mu\nu}^{mn}(\mathbf{r}_{12}) \Phi_{\mu\nu}^{mn}(\hat{\mathbf{r}}_{12}, \Omega_1, \Omega_2) \quad (2.22)$$

where the angular basis functions $\Phi_{\mu\nu}^{mn}(\hat{\mathbf{r}}_{12}, \Omega_1, \Omega_2)$ can be expressed in terms of generalized spherical harmonics (GSHs) [13]. A detailed description of rotational invariant transform is in appendix D.

2.2.4 SPC/E water model

As water cannot be perfectly described by a pair potential (due to multi-body effects, quantum effects, hydrogen bond, etc.), various models have been developed to fit a maximum number of properties. Those models contain several sites, which can be placed possibly elsewhere than at the center of atoms (figure 2.5). The more sites the model has, the more precise it can be. There is a great work done by Martin Chaplin [45] to summarize the most widely used water models.

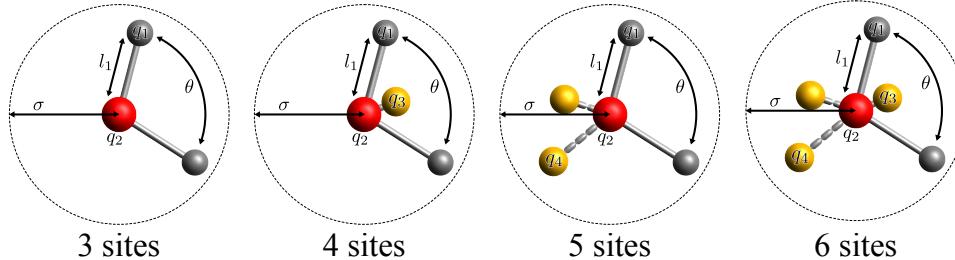


Figure 2.5: Water models

In this thesis, we use the extended simple point charge model (SPC/E) of water [46] as our solvent model through all this thesis. It is a 3-site model, the electrostatic interaction

being modeled using Coulomb's Law and the dispersion and repulsion forces using the Lennard-Jones potential, as described above.

With respect to the original SPC model, the SPC/E model takes into account the polarization in an implicit and phenomenological way, re-normalizing the dipole of the effective pair model, μ , and thus increasing the partial charge slightly compared to SPC (table 2.1; the center of water molecule has been placed at atom O, for convenience):

$$E_{\text{pol}} = \frac{1}{2} \sum_i \frac{(\mu - \mu^0)^2}{\alpha_i} \quad (2.23)$$

μ^0 being the dipole moment of an isolated water molecule, and α_i the isotropic scalar polarizability [46]. The SPC/E model gives a better radial distribution function and diffusion constant than the SPC model. It is the most commonly used model for applications.

It should be noted that any rigid solvent model is compatible with the theory that this thesis bases on, e.g. acetonitrile used in [31].

MODEL	σ [\AA^6]	ε [$\text{kJ} \cdot \text{mol}^{-1}$]	l_1 [\AA]	q_1 [e]	q_2 [e]	θ [$^\circ$]
SPC [47]	3.166	0.650	1.0000	+0.410	-0.8200	109.47
SPC/E [46]	3.166	0.650	1.0000	+0.4238	-0.8476	109.47
experiment [48]	-	-	0.991	-	-	105.5

Table 2.1: Structural parameters of SPC and SPC/E water

2.2.5 Flexible and polarizable models

Up to this point, molecules were considered as rigid bodies. Flexible models give extra degrees of freedom in vibration and internal rotation. In that case, the interaction potential contains several extra terms, yielding typically five kinds of forces: three for the direct interactions in addition to the two indirect interactions (LJ and Coulomb).

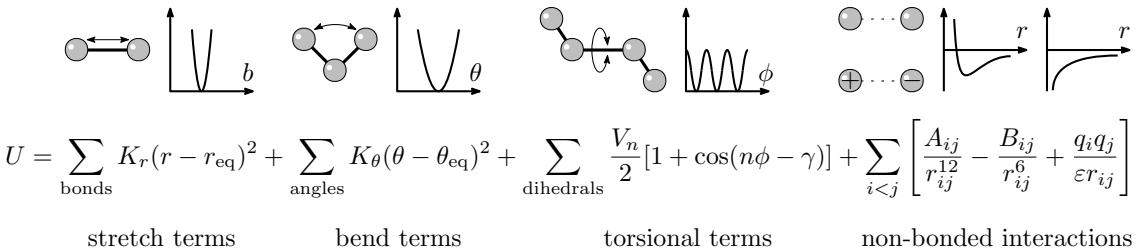


Figure 2.6: Interactions in a flexible model

The flexible yielding can deal with the non-rigidity of the solvent, which is partially polarized owing to the vibrational degrees of freedom (the so-called atomic polarizability). On the other hand, electronic polarizability (the deformation of the molecule electron cloud under the action of the external electric field) can be taken into account even in a rigid model. This polarizability can be described by introducing a modifiable charge distribution, for example by adding an induced dipole at the molecular center of the molecule, or even on each of its atomic sites, and by solving the set of induced dipoles self-consistently. Introducing variable atomic charges is possible too [ref? or idea?]. Optimizing the induced charges/dipoles has a large computational overhead compared to fixed charges.

Complex models require expensive computing cost, but still can have large fluctuations due to use of imposed small system size. There is a compromise between the choice of

model and the choice of system size. For this reason, the rigid models are still nowadays the most popular. On the other hand, computing technologies have greatly developed compared to the theories themselves, which makes it possible to use more and more precise models in computation.

2.3 MODEL OF SOLUTE

The model of solute also have a substantial influence on the predicted energy and structure of solvation. The solute can eventually be treated by QM calculations in terms of wave function and electron density. This is the case for the implicit SCRF method, which for apolar solvents (i.e. toluene) has been proven to work well. There is a clear mismatch, however, between the very refined description of the solute and the rather primitive continuous-medium treatment of the solvent. The compromise to have a better model of solvent or solute is debatable, and should vary according to the applications. On the other extreme, one never uses a quantum solvent model with an implicit solute; this would not be profitable even if the solute is of simple geometry (wall). In the case of molecular solutes, it is consistent to require the solute to have at least the same scale of description as the solvent. Within molecular force fields, this leads to the hierarchy of potential models described in figure (2.7).

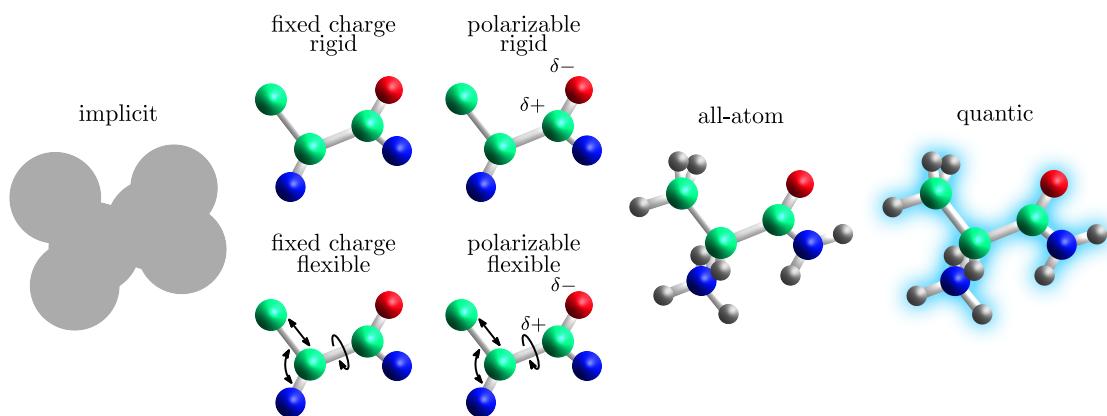


Figure 2.7: Hierarchy of solute models

In this thesis, our first step will be to use a rigid molecular model to describe the solute. This is coherent with IET, which cannot treat the solvent and solute at different scales of description. Polarizable and/or flexible models of solute, and the coupling of a QM solute to the molecular solvent, will be described in perspective.

In conclusion, the choice of model for the solute/solvent system is a compromise between the required precision according to the application, and the computing cost that the research can afford.

3

STATISTICAL MECHANICS OF ATOMIC FLUIDS

Statistical mechanics serves to deduce thermodynamic quantities from the Hamiltonian of any given system. In this section, we present some basic formalism for a classical atom-like spherical solvent model in grand canonical ensemble (μ, V, T) . Firstly, we introduce the relations between the statistical mechanics and thermodynamic quantities. Then we change the view to the structure of the solvent. The two theories we use in this thesis, here referred to as IET and cDFT, as well as their equivalency, are presented with brief derivations in the following sections. The majority of these sections is based on the book by Hansen & McDonald [12, 49], and the articles and notes of Evans [20, 50, 51]. A very detailed review is done by Wu *et al.* [52] to the same purpose, thus here we only introduce the concepts that will be useful to understand this thesis.

3.1 HAMILTONIAN AND ENSEMBLE PROPERTIES

Once we define a spherical solvent model, of which the movement only depends on its position and momentum (\mathbf{r}, \mathbf{p}) , the instantaneous state (phase point, micro-state) of an N -particle solvent system is specified by $3N$ coordinates $\mathbf{r}^N \equiv \mathbf{r}_1, \dots, \mathbf{r}_N$ and $3N$ momenta $\mathbf{p}^N \equiv \mathbf{p}_1, \dots, \mathbf{p}_N$. The internal energy of particles in a system is characterized by its Hamiltonian:

$$H_N(\mathbf{r}^N, \mathbf{p}^N) = K_N(\mathbf{p}^N) + V_N(\mathbf{r}^N) + V_N^{\text{ext}}(\mathbf{r}^N) \quad (3.1)$$

where

$$\begin{aligned} K_N(\mathbf{p}^N) &= \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m} \text{ is the kinetic energy;} \\ V_N(\mathbf{r}^N) &= \sum_{i < j}^N u(|\mathbf{r}_i - \mathbf{r}_j|) + 3 \text{ body} + \dots \text{ is the interatomic potential energy } \mathcal{U}(\mathbf{r}^N); \\ V_N^{\text{ext}}(\mathbf{r}^N) &= \sum_{i=1}^N V_{\text{ext}}(\mathbf{r}_i) \text{ is the potential energy arising from the interaction of the} \\ &\quad \text{particles with the external field (e.g. a solute).} \end{aligned}$$

The grand potential, characteristic thermodynamic state function for the grand canonical ensemble, which depends on the chemical potential μ , the volume V and the temperature T , is linked with the statistical mechanics quantities with the relation:

$$\Omega(\mu, V, T) = -k_B T \ln \Xi \quad (3.2)$$

where

$$\Xi = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{h^{3N} N!} \int d\mathbf{r}^N d\mathbf{p}^N e^{-\beta H_N(\mathbf{r}^N, \mathbf{p}^N)} \quad (3.3)$$

$$= \sum_{N=0}^{\infty} \frac{1}{N!} \int d\mathbf{r}^N e^{-\beta V_N(\mathbf{r}^N)} \left(\prod_{i=1}^N \frac{e^{\beta V_{\text{int}}(\mathbf{r}_i)}}{\Lambda^3} \right) \quad (3.4)$$

is the grand partition function, with $\Lambda = (2\pi\beta\hbar^2/m)^{-\frac{1}{2}}$ the de Broglie thermal wavelength, and

$$V_{\text{int}}(\mathbf{r}_i) = \mu - V_{\text{ext}}(\mathbf{r}_i) \quad (3.5)$$

the intrinsic chemical potential.

N = $\int d\mathbf{r}\rho(\mathbf{r})$ is the number of particles in canonical ensemble, but the formulae (3.6) and (3.8) are also available for grand canonical ensemble.

where

$$\rho(\mathbf{r}) = \langle \varrho(\mathbf{r}) \rangle = \left\langle \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i) \right\rangle \quad (3.6)$$

is the density profile of instantaneous density $\rho(\mathbf{r})$ distribution.

The differential form of \mathcal{F}_{int} is

$$\delta\mathcal{F}_{\text{int}} = -S\delta T + \int d\mathbf{r}\delta\rho(\mathbf{r})V_{\text{int}}(\mathbf{r}) \quad (3.8)$$

with S the entropy.

The internal energy of the solvent contains two contributions, one due to the kinetic energy of the particles, $K_N(\mathbf{p}^N)$, and the other linked to the interaction between particles, $V_N(\mathbf{r}^N)$. When the fluid is a perfect gas, which means $V_N = 0$, it can be easily derived from eq. (3.2-3.5) that \mathcal{F}_{int} has the following expression:

$$\mathcal{F}_{\text{id}} = k_B T \int d\mathbf{r}\rho(\mathbf{r}) \left(\ln [\Lambda^3 \rho(\mathbf{r})] - 1 \right) \quad (3.9)$$

When interactions between particles are accounted for, the total expression of \mathcal{F}_{int} can be

$$\mathcal{F}_{\text{int}} = \mathcal{F}_{\text{id}} + \mathcal{F}_{\text{exc}} \quad (3.10)$$

and the form of \mathcal{F}_{exc} will be detailed in later sections.

3.2 FUNCTIONAL DERIVATIVES AND DISTRIBUTION FUNCTIONS

The structure of the solvent in the grand canonical ensemble can be characterized by its n -particle density

$$\rho^{(n)}(\mathbf{r}^n) = \frac{1}{\Xi} \sum_{N=n}^{\infty} \frac{1}{(N-n)!} \int d\mathbf{r}^{(N-n)} e^{-\beta V_N(\mathbf{r}^N)} \left(\prod_{i=1}^N \frac{e^{\beta V_{\text{int}}(\mathbf{r}_i)}}{\Lambda^3} \right) \quad (3.11)$$

which means the probability to find n particles in a volume element $d\mathbf{r}^n$. In particular, the probability to find one particle in a volume element is the solvent density $\rho^{(1)}(\mathbf{r}) = \rho(\mathbf{r})$, that

$$\rho^{(1)}(\mathbf{r})d\mathbf{r} = \langle N \rangle \quad (3.12)$$

where $\langle N \rangle$ is the ensemble average of the number of particles, that is to say the average number of particles at equilibrium. $\rho^{(n)}(\mathbf{r}^n)$ becomes ρ^n if the system is homogeneous. It can be proven that

$$\frac{\delta\Omega}{\delta V_{\text{int}}(\mathbf{r})} = -\rho^{(1)}(\mathbf{r}) \quad (3.13)$$

The corresponding n -particle distribution function is defined as:

$$g^{(n)}(\mathbf{r}^n) = \frac{\rho^{(n)}(\mathbf{r}^n)}{\prod_{i=1}^n \rho^{(1)}(\mathbf{r}_i)} \quad (3.14)$$

such that $g^{(n)}(\mathbf{r}^n) \rightarrow 1$ when all pairs of particles becomes sufficiently large.

The two-particle pair distribution function (PDF), $g^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$, is one of the most important quantities in the theory of liquids. Its corresponding pair correlation function (PCF) is defined as:

$$h^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = g^{(2)}(\mathbf{r}_1, \mathbf{r}_2) - 1 \quad (3.15)$$

which vanishes when $|\mathbf{r}_1 - \mathbf{r}_2| \rightarrow \infty$.

If we define the density-density correlation function as:

$$H^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \rho^{(1)}(\mathbf{r}_1)\rho^{(1)}(\mathbf{r}_2)h^{(2)}(\mathbf{r}_1, \mathbf{r}_2) - \rho^{(1)}(\mathbf{r}_1)\delta(\mathbf{r}_1 - \mathbf{r}_2) \quad (3.16)$$

For any ensemble.

which means the correlation [53] between the instantaneous fluctuation of particle density from its ensemble average, it can be proven that

$$\frac{\delta\Omega^2}{\delta V_{\text{int}}(\mathbf{r}_1)\delta V_{\text{int}}(\mathbf{r}_2)} = -\beta H^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = -\frac{\delta\rho^{(1)}(\mathbf{r}_1)}{\delta V_{\text{int}}(\mathbf{r}_2)} \quad (3.17)$$

As an analogue, the direct correlation function (DCF) is defined as the derivative of the excess free energy functional $\mathcal{F}_{\text{exc}}[\rho]$:

$$c^{(1)}(\mathbf{r}) = -\frac{\delta(\beta\mathcal{F}_{\text{exc}}[\rho^{(1)}])}{\delta\rho^{(1)}(\mathbf{r})} \quad (3.18)$$

$$c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = \frac{\delta c^{(1)}(\mathbf{r}_1)}{\delta\rho^{(1)}(\mathbf{r}_2)} = -\frac{\delta^2(\beta\mathcal{F}_{\text{exc}}[\rho^{(1)}])}{\delta\rho^{(1)}(\mathbf{r}_1)\delta\rho^{(1)}(\mathbf{r}_2)} = c^{(2)}(\mathbf{r}_2, \mathbf{r}_1) \quad (3.19)$$

$$c^{(n)}(\mathbf{r}_1, \dots, \mathbf{r}_n) = \frac{\delta c^{(n-1)}(\mathbf{r}_1, \dots, \mathbf{r}_{n-1})}{\delta\rho^{(1)}(\mathbf{r}_n)} \quad (3.20)$$

According to the definition of F_{int} , as well as the expression of δF_{int} in eq. (3.8), we have

$$\begin{aligned} \beta V_{\text{int}}(\mathbf{r}) &= \beta \frac{\delta F_{\text{int}}[\rho^{(1)}]}{\delta\rho^{(1)}(\mathbf{r})} = \beta \frac{\delta F_{\text{id}}[\rho^{(1)}]}{\delta\rho^{(1)}(\mathbf{r})} + \beta \frac{\delta F_{\text{exc}}[\rho^{(1)}]}{\delta\rho^{(1)}(\mathbf{r})} \\ &= \ln [\Lambda^3 \rho^{(1)}(\mathbf{r})] + c^{(1)}(\mathbf{r}) \end{aligned} \quad (3.21)$$

The functional derivative chain rule leads to

$$\begin{aligned} \int d\mathbf{r}_3 \frac{\delta V_{\text{int}}(\mathbf{r}_1)}{\delta\rho^{(1)}(\mathbf{r}_3)} \cdot \frac{\delta\rho^{(1)}(\mathbf{r}_3)}{\delta V_{\text{int}}(\mathbf{r}_2)} &= \int d\mathbf{r}_3 \frac{\delta V_{\text{int}}[\rho^{(1)}(\mathbf{r}_1)]}{\delta\rho^{(1)}(\mathbf{r}_3)} \cdot \beta H^{(2)}(\mathbf{r}_3, \mathbf{r}_2) \\ &= \delta(\mathbf{r}_1 - \mathbf{r}_2) \end{aligned} \quad (3.22)$$

which in addition to the definition of H in eq. (3.16) gives

$$h^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) + \int d\mathbf{r}_3 \left(c^{(2)}(\mathbf{r}_1, \mathbf{r}_3) \rho^{(1)}(\mathbf{r}_3) h^{(2)}(\mathbf{r}_3, \mathbf{r}_2) \right) \quad (3.23)$$

which is called the Ornstein-Zernike (OZ) equation.

3.3 CLASSICAL DENSITY FUNCTIONAL THEORY

The density functional theory is based on two theorems :

1. For a given choice of V_N , T and μ , the intrinsic free energy functional \mathcal{F}_{int} is a unique functional of the equilibrium one-particle density $\rho^{(1)}(\mathbf{r})$, expressed by $\mathcal{F}_{\text{int}}[\rho^{(1)}]$.
2. Let $n(\mathbf{r})$ be some arbitrary one-particle microscopic density, and define the grand potential $\Omega[n]$ as

$$\Omega[n] = \mathcal{F}_{\text{int}}[n] - \int d\mathbf{r} n(\mathbf{r}) V_{\text{int}}(\mathbf{r}) \quad (3.24)$$

for a fixed external potential V_{ext} (or intrinsic chemical potential V_{int}), then the variational principle states that

$$\Omega[n] \geq \Omega[\rho^{(1)}] \quad (3.25)$$

with the equal sign takes at $n(\mathbf{r}) = \rho^{(1)}(\mathbf{r})$. The differentiation of eq. (3.24) with respect to $n(\mathbf{r})$ gives

$$\frac{\delta \Omega[n]}{\delta n(\mathbf{r})} \Big|_{n=\rho^{(1)}} = \frac{\delta \mathcal{F}_{\text{int}}[n]}{\delta n(\mathbf{r})} \Big|_{n=\rho^{(1)}} - V_{\text{int}}(\mathbf{r}) = 0 \quad (3.26)$$

The fact that the right hand vanishes at equilibrium is agreed with eq. (3.8).

Note that eq. (3.26) is then equivalent to

$$\frac{\delta \mathcal{F}[n]}{\delta n(\mathbf{r})} \Big|_{n=\rho^{(1)}} = 0 \quad (3.27)$$

where

$$\mathcal{F}[n(\mathbf{r})] = \mathcal{F}_{\text{int}} + \mathcal{F}_{\text{ext}} = \mathcal{F}_{\text{id}} + \mathcal{F}_{\text{exc}} + \mathcal{F}_{\text{ext}} \quad (3.28)$$

is the total Helmholtz free energy functional of the solvent; with

$$\mathcal{F}_{\text{ext}} = \int d\mathbf{r} n(\mathbf{r})(V_{\text{ext}}(\mathbf{r}) - \mu) = - \int d\mathbf{r} n(\mathbf{r}) V_{\text{int}}(\mathbf{r}) \quad (3.29)$$

being defined as the external functional. These theorems build an approach that, for a given choice V_N , T and μ , one can obtain the equilibrium density of solvent $\rho^{(1)}(\mathbf{r})$ by just minimizing the functional $\mathcal{F}[n]$.

The external term \mathcal{F}_{exc} of the functional can be given by Taylor expansion around a reference homogeneous bulk solvent density ρ_0 :

$$\begin{aligned} \mathcal{F}_{\text{exc}}[\rho^{(1)}] &\equiv \mathcal{F}_{\text{exc}}[\rho_0] + \int d\mathbf{r}_1 \frac{\delta \mathcal{F}_{\text{exc}}[\rho^{(1)}]}{\delta \rho^{(1)}(\mathbf{r}_1)} \Bigg|_{\rho^{(1)}=\rho_0} \Delta \rho(\mathbf{r}_1) \\ &\quad + \frac{1}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 \frac{\delta^2 \mathcal{F}_{\text{exc}}[\rho^{(1)}]}{\delta \rho^{(1)}(\mathbf{r}_1) \delta \rho^{(1)}(\mathbf{r}_2)} \Bigg|_{\rho^{(1)}=\rho_0} \Delta \rho(\mathbf{r}_1) \Delta \rho(\mathbf{r}_2) + \mathcal{O}(\Delta \rho^3) \\ &= \mathcal{F}_{\text{exc}}[\rho_0] - k_B T \int d\mathbf{r}_1 c^{(1)}(\mathbf{r}) \Delta \rho^{(1)}(\mathbf{r}_1) \\ &\quad - \frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \Delta \rho^{(1)}(\mathbf{r}_1) \Delta \rho^{(1)}(\mathbf{r}_2) + \mathcal{O}(\Delta \rho^3) \end{aligned} \quad (3.30)$$

where $c^{(n)}(\mathbf{r})$ is the corresponding bulk DCF defined in eq. (3.20), and $\Delta \rho = \rho^{(1)} - \rho_0$.

According to eq. (3.26), we can calculate $c^{(1)}(\mathbf{r})$, which is a constant term depending on $V_{\text{int}}(\mathbf{r})$. If we consider the boundary conditions that the solvation free energy for bulk solvent $\mathcal{F}[\rho_0] = 0$, as well as the internal chemical potential $V_{\text{int}}(\mathbf{r}) = 0$, the first two terms in eq. (3.30) vanish:

$$\mathcal{F}_{\text{exc}} \left[\rho^{(1)} \right] = -\frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) \Delta\rho^{(1)}(\mathbf{r}_1) \Delta\rho^{(1)}(\mathbf{r}_2) + \mathcal{O}(\Delta\rho^3) \quad (3.31)$$

If the reference fluid is homogeneous, the DCF only depends on the relative distance, i.e. $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) = c(r_{12})$, so that

$$\mathcal{F}_{\text{exc}} \left[\rho^{(1)} \right] \simeq -\frac{k_B T}{2} \int d\mathbf{r}_1 d\mathbf{r}_2 c(r_{12}) \Delta\rho(\mathbf{r}_1) \Delta\rho(\mathbf{r}_2) \quad (3.32)$$

This was called the homogenous reference fluid (HRF) approximation. The generalization to a molecular, non-spherical solvent for which orientations matter is described in §4.

3.4 INTEGRAL EQUATION THEORY

Similar to the DFT approach which aims to find the equilibrium solvent density ρ and the Helmholtz free energy \mathcal{F} , the integral equation theory (IET) aims to find the pair distribution function g and the gradient of energy γ . Both theories give complete information of solvation energy and structure.

IET is about to solve a pair of integral equations of $h^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$ and $c^{(2)}(\mathbf{r}_1, \mathbf{r}_2)$. One of these equations is the OZ equation shown as eq. (3.23). Other is a closure equation, which can be deduced from eq. (3.26), giving the minimum density

$$\rho^{(1)}(\mathbf{r}_1) = \rho_0 \exp \left(-\beta V_{\text{int}}(\mathbf{r}_1) + \int d\mathbf{r}_2 \Delta\rho^{(1)}(\mathbf{r}_2) c^{(2)}(\mathbf{r}_1, \mathbf{r}_2) + \mathcal{O}(\Delta\rho^2) \right) \quad (3.33)$$

(in Hansen eq. (4.3.16) it is V_{ext} , but what I deduced is V_{int}) which gives for example, when $\mathcal{O}(\Delta\rho^2) = 0$, one of the simplest closure equation, the hypernetted-chain HNC approximation:

$$g(1, 2) = \exp [-\beta u(1, 2) + h(1, 2) - c(1, 2)] \quad (3.34)$$

Here u corresponds to V_{int} when the particle 1 and 2 are respectively the solute and solvent.

The general form of OZ closure is:

$$g(1, 2) = \exp [-\beta u(1, 2) + h(1, 2) - c(1, 2) + b(1, 2)] \quad (3.35)$$

where the b is the bridge function. Other closures are also possible, such as Percus-Yevick (PY) approximation (a linear expansion of the second exponential term in HNC) specifically for systems with short-range forces, or mean-spherical approximation (MSA) in the limit of low density.

3.5 EQUIVALENCE BETWEEN CDFT AND IET FOR A DILUTE SOLUTION SYSTEM

The generalization of the OZ equation in eq. (3.23) to n components can be written as

$$h_{\nu\mu}(1, 2) = c_{\nu\mu}(1, 2) + \rho \sum_{\lambda} x_{\lambda} \int c_{\nu\lambda}(2, 3) h_{\lambda\mu}(1, 3) d3 \quad (3.36)$$

where $x_{\nu} = N_{\nu}/N$ is the number concentration of species $\nu \in [1, n]$.

For a two-component homogeneous solute-solvent mixture, where the solute (M) is infinitely diluted in the solvent (S) ($x_S \rightarrow 1$), the coupled OZ relations are written as

$$h_{SS}(1, 2) = c_{SS}(1, 2) + \rho \int h_{SS}(1, 3)c_{SS}(2, 3)d3 \quad (3.37)$$

$$h_{SM}(1, 2) = c_{SM}(1, 2) + \rho \int h_{SS}(1, 3)c_{SM}(2, 3)d3 \quad (3.38)$$

$$h_{MS}(1, 2) = c_{MS}(1, 2) + \rho \int h_{MS}(1, 3)c_{SS}(2, 3)d3 \quad (3.39)$$

$$h_{MM}(1, 2) = c_{MM}(1, 2) + \rho \int h_{MS}(1, 3)c_{SM}(2, 3)d3 \quad (3.40)$$

Eq. (3.37) is the OZ equation for bulk solvent. Eqs. (3.38) and (3.39) describe the correlations between the solute and solvents, which are equivalent. From eq. (3.39) we can deduce eq. (3.31) for the DFT approach, if we impose $\mathcal{O}(\Delta\rho^3) = 0$, i.e. the HNC approximation. And in IET, eq. (3.38) is normally used for two-component solution. In eq. (3.40), $r_{MM} \rightarrow \infty$, that means the correlation between the solutes are infinity weak. The difficulty to solve such equation is to find a proper closure equation. As the approximations like HNC is already quantitatively far from sufficient to describe solute-solvent correlation, it becomes very bad for solute-solute.

4

APPROACH TO MOLECULAR SOLVENTS

In the case of non-spherical solvent like water, the solvent particle carries a molecular structure described by a collection of distributed atomic interaction sites (LJ and Coulombic). The two theories mentioned in the previous section are now formulated in the molecular picture in which each solvent molecule is considered as a rigid body and characterized by its position \mathbf{r} (e.g. the position of center of mass), and its orientation $\boldsymbol{\Omega}$. In MDFT, the solvent is characterized by an angle-dependent inhomogeneous density, $\rho(\mathbf{r}, \boldsymbol{\Omega})$; in IET, an angle-dependent form of the pair distribution function $g(\mathbf{X}_1, \mathbf{X}_2)$ ($X \equiv (\mathbf{r}, \boldsymbol{\Omega})$) is proposed, while the molecular OZ equation is expanded on rotational invariants. The reference interaction site model (RISM) [15], which provides another way for IET to treat molecular solvent, will not be discussed in this thesis.

4.1 MOLECULAR DENSITY FUNCTIONAL THEORY

In molecular density functional theory (MDFT), the grand potential is rewritten as:

$$\Omega[\rho(\mathbf{r}, \boldsymbol{\Omega})] = \Omega[\rho_0] + \mathcal{F}[\rho(\mathbf{r}, \boldsymbol{\Omega})] \quad (4.1)$$

where $\Omega[\rho_0]$ is the correspondent reference bulk fluid grand potential (with $\rho = \rho_0$ and $V_{\text{int}} = 0$). $\rho(\mathbf{r}, \boldsymbol{\Omega})$ is the angle-dependent fluid density function, depending on 3 variables for spatial coordinates \mathbf{r} , and also 3 for orientation $\boldsymbol{\Omega} \equiv (\Theta, \Phi, \Psi)$. In case of linear solvent, this number can reduce to 2, i.e. $\boldsymbol{\Omega} \equiv (\Theta, \Phi)$. The homogeneous bulk density ρ_0 is normalized to $n_0 / \int d\boldsymbol{\Omega}$, to keep coherent with the relation

$$\int d\boldsymbol{\Omega} \rho(\mathbf{r}, \boldsymbol{\Omega}) = \int d\cos\Theta d\Phi d\Psi \rho(\mathbf{r}, \boldsymbol{\Omega}) = n(\mathbf{r}) \quad (4.2)$$

which reduces eq. (4.1) to eq. (3.24) in §3.3.

According to the variation principle described in §3.3, the equilibrium density can be found by minimizing the free energy functional

$$\mathcal{F}[\rho] = \mathcal{F}_{\text{id}}[\rho] + \mathcal{F}_{\text{ext}}[\rho] + \mathcal{F}_{\text{exc}}[\rho] \quad (4.3)$$

regarding to $\rho(\mathbf{r}, \boldsymbol{\Omega})$:

$$\left. \frac{\delta \mathcal{F}[\rho]}{\delta \rho(\mathbf{r}, \boldsymbol{\Omega})} \right|_{\rho=\rho_0} = 0 \quad (4.4)$$

which remains exactly the same formalism with those in §3.3.

4.1.1 The ideal term

The ideal term $\mathcal{F}_{\text{id}}[\rho]$ deduced from the particle interaction-free condition is:

$$\mathcal{F}_{\text{id}}[\rho] = k_B T \int d\mathbf{r} d\boldsymbol{\Omega} \left[\rho(\mathbf{r}, \boldsymbol{\Omega}) \ln \left(\frac{\rho(\mathbf{r}, \boldsymbol{\Omega})}{\rho_0} \right) - \rho(\mathbf{r}, \boldsymbol{\Omega}) + \rho_0 \right] \quad (4.5)$$

(incoherent with §3.3 with an additional term ρ_0 , and disappeared Λ^3 .)

The differentiation of $\mathcal{F}_{\text{id}}[\rho]$ used for the minimization, which will be discussed later, has form:

$$\frac{\delta \mathcal{F}_{\text{id}}[\rho]}{\delta \rho(\mathbf{r}, \boldsymbol{\Omega})} = \beta^{-1} \ln \left(\frac{\rho(\mathbf{r}, \boldsymbol{\Omega})}{\rho_0} \right) \quad (4.6)$$

4.1.2 The external term

The solute, like the solvent, is described in microscopic detail by a molecular non-polarizable force field involving atomic Lennard-Jones and partial charge parameters, creating at each point in space an external potential $V_{\text{ext}}(\mathbf{r}, \boldsymbol{\Omega})$, containing two components:

$$V_{\text{ext}}(\mathbf{r}, \boldsymbol{\Omega}) = V_{\text{LJ}}(\mathbf{r}) + V_{\text{coul}}(\mathbf{r}, \boldsymbol{\Omega}) \quad (4.7)$$

The external potential term calculates the contribution of V_{ext} :

$$\mathcal{F}_{\text{ext}}[\rho] = \int d\mathbf{r} d\boldsymbol{\Omega} \rho(\mathbf{r}, \boldsymbol{\Omega}) (V_{\text{ext}}(\mathbf{r}, \boldsymbol{\Omega}) - \mu) \quad (4.8)$$

with μ the chemical potential.

The Lennard-Jones potential is given by:

$$V_{\text{LJ}}(\mathbf{r}) = \sum_u \sum_v 4\epsilon_{uv} \left[\left(\frac{\sigma_{uv}}{r_{uv}} \right)^{12} - \left(\frac{\sigma_{uv}}{r_{uv}} \right)^6 \right] \quad (4.9)$$

where u stands for solute, v stands for solvent, $\epsilon_{uv} = \sqrt{\epsilon_u \epsilon_v}$ and $\sigma_{uv} = (\sigma_u + \sigma_v)$ are the geometric and arithmetic average Lennard-Jones parameters between solute and solvent, according to the Lorentz-Berthelot mixing rules. r_{ij} is the norm of relative site-site vector

$$\mathbf{r}_{uv} = \mathbf{r}_v + \mathbf{R}(\boldsymbol{\Omega}) \mathbf{s}_v - \mathbf{r}_u \quad (4.10)$$

where \mathbf{r}_u and \mathbf{s}_v are the coordinates of solute/solvent molecules in the molecular frame, and $\mathbf{R}(\boldsymbol{\Omega})$ is the rotation matrix of the Euler angles $\boldsymbol{\Omega}$. In cases where the solvent site wears only one LJ centre, eq. (4.10) reduces to

$$\mathbf{r}_{uv} = \mathbf{r}_v - \mathbf{r}_u \quad (4.11)$$

which is actually what we use in the code as the solvent is SPC/E water.

The Coulomb interaction is calculated by solving the Poisson equation. The charge density of the solute is projected onto a space grid \mathbf{r} ,

$$\rho_q(\mathbf{r}) = \sum_u q_{ijk}/\Delta v \quad (4.12)$$

where q_{ijk} is the charge on the space grid distributed by its nearby point charge as shown in figure 4.1, and Δv the volume of the unit cube that this point charge situates in.

The electrostatic potential created by the charge distribution $\rho_q(\mathbf{r})$, $V_q(\mathbf{r})$, can be thus computed using a periodic Poisson Solver. The Poisson equation (2.7)

$$\nabla^2 V_q(\mathbf{r}) = -\frac{\rho_q(\mathbf{r})}{\epsilon_0} \quad (4.13)$$

(where is the $\epsilon(\mathbf{r})$?) gives in Fourier space

$$\hat{V}_q(\mathbf{k}) = \frac{\hat{\rho}_q(\mathbf{k})}{\epsilon_0 k^2} \quad (4.14)$$

where $\hat{V}_q(\mathbf{k})$ and $\hat{\rho}_q(\mathbf{k})$ are the Fourier transform of $V_q(\mathbf{r})$ and $\rho_q(\mathbf{r})$ respectively. These two equations provide a fast way to calculate $V_q(\mathbf{r})$ from $\rho_q(\mathbf{r})$.

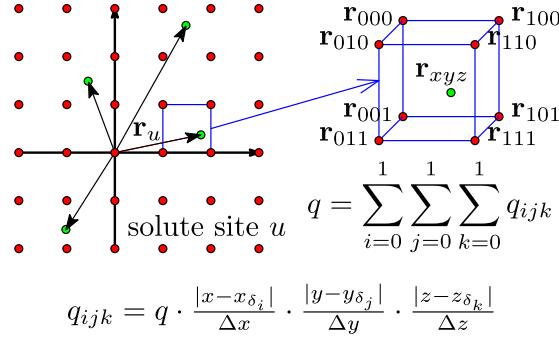


Figure 4.1: Solute charge density projected onto grids

The Coulomb potential is expressed as a sum of solvent partial charge contributions at each grid node:

$$V_{\text{coul}}(\mathbf{r}, \Omega) = \sum_v q_v V_q(\mathbf{r}_v) \quad (4.15)$$

where q_v is the point charge of solvent, and

$$\mathbf{r}_v = \mathbf{r} + \mathbf{R}(\Omega) \mathbf{s}_v \quad (4.16)$$

is the cartesian coordinate of a solvent site v ; $V_q(\mathbf{r}_u)$ is the electrostatic potential, given by a linear interpolation of the nearby point of $V_q(\mathbf{r})$ obtained in the last step from the Poisson solver.

4.1.3 The excess term

The two terms $\mathcal{F}_{\text{id}}[\rho]$ and $\mathcal{F}_{\text{ext}}[\rho]$ are physically exact, while the excess term $\mathcal{F}_{\text{exc}}[\rho]$ depends on the exact correlation function, which is a priori unknown. As shown in §3.3, we invoke here the HRF approximation which amounts to a second-order Taylor expansion around the homogeneous fluid at density ρ_0 :

$$\mathcal{F}_{\text{exc}}[\rho] = -\frac{k_B T}{2} \int d\mathbf{r}_1 d\Omega \gamma(\mathbf{r}_1, \Omega) \rho(\mathbf{r}_1, \Omega) \quad (4.17)$$

where γ is the normalized gradient of the excess functional:

$$\gamma(\mathbf{r}_1, \Omega_1) = -\frac{\delta \beta F_{\text{exc}}}{\delta \rho} = \int d\mathbf{r}_2 d\Omega_2 \Delta \rho(\mathbf{r}_2, \Omega_2) c(\mathbf{r}_{12}, \Omega_1, \Omega_2) \quad (4.18)$$

(attention the negative sign) which can be related to the solute-solvent 2-component integral equation theory (IET) with its definition:

$$\gamma_{\text{MS}}(1, 2) = h_{\text{MS}}(1, 2) - c_{\text{MS}}(1, 2) \quad (4.19)$$

To evaluate the integration $\int d\mathbf{r}_2 d\Omega_2$ for each gradient $\gamma(\mathbf{r}_1, \Omega_1)$ in eq. (4.18), a total number of $N^2 \equiv N_{\mathbf{r}}^2 N_{\Omega}^2 = O(N^2)$ function evaluations (FE) are required, which, with typically $N_{\mathbf{r}} = 64^3$ and $N_{\Omega} = 50 \sim 100$, is far too costly for current computing technology. For this reason, Fourier transform is used to treat the spatial convolution.

A convolution

$$h(x_1) \equiv f(x_2) \otimes g(x_2) \equiv \int_a^b f(x_2) g(x_1 - x_2) dx_2 \quad (4.20)$$

has the property that

$$\mathfrak{F}[h(x_1)] = \mathfrak{F}[f(x_2)] \mathfrak{F}[g(x_2)] \quad (4.21)$$

\mathfrak{F} being the Fourier transform operation. As $\mathbf{r}_{12} = \mathbf{r}_1 - \mathbf{r}_2$, eq. (4.18) is a 3D convolution, which leads to

$$\hat{\gamma}(\mathbf{k}, \Omega_1) = \int d\Omega_2 \Delta\hat{\rho}(\mathbf{k}, \Omega_2) \hat{c}(\mathbf{k}, \Omega_1, \Omega_2) \quad (4.22)$$

Here we put the hat symbol on the physical quantities to represent the Fourier transform of their original function.

In eq. (4.22), the integral $\int d\mathbf{r}_2$ of eq. (4.18) is transformed into a simple product, only $N_r N_\Omega^2$ FE are needed to obtain $\hat{\gamma}(\mathbf{k}, \Omega_1)$ with given $\Delta\hat{\rho}(\mathbf{k}, \Omega_2)$. To this computational cost should be added the transform from $\Delta\rho(\mathbf{r}, \Omega)$ to $\Delta\hat{\rho}(\mathbf{k}, \Omega)$ and the backward transform from $\hat{\gamma}(\mathbf{k}, \Omega)$ to $\gamma(\mathbf{r}, \Omega)$ which are both of order $N_\Omega \cdot O(N_r \log_2 N_r)$ due to the properties of Fast Fourier Transforms (FFT). The total number of FE is thus reduced from quadratic complexity $O(N_r^2 N_\Omega^2)$ to $N_r N_\Omega^2 + 2N_\Omega \cdot O(N_r \log_2 N_r) = O(N_r \log_2 N_r N_\Omega^2)$. As the total number of spatial grid N_r is of magnitude $10^5 \sim 10^6$, this procedure, which is mathematically equivalent to the direct evaluation (4.18), offers a great advantage in terms of computational efficiency (figure A.1 in appendix A).

The angular-dependent DCF of the homogeneous solvent, $\hat{c}(\mathbf{k}, \Omega_1, \Omega_2)$, is an input data which can be obtained from MD or MC simulations. A detailed presentation of the DCFs used in this thesis is available in appendix B.

4.2 MOLECULAR INTEGRAL EQUATION THEORY

To adapt the IET formalism to non-spherical solvent, Blum [16, 17] proposed to expand the angle-dependent correlation functions $F(\mathbf{X}_1, \mathbf{X}_2) \equiv F(\mathbf{r}_1, \mathbf{r}_2, \Omega_1, \Omega_2)$ onto rotational invariants, such that the OZ equation can be reduced to only a few FE. This theory is then adopted by Fries & Patey [18], who proposed a numerical solution for full HNC closure. The test below describes the theory of Blum, but based on the convention of Fries & Patey, where Messiah's definition of generalized spherical harmonics (GSHs) is used. A detailed explication of different conventions of GSH is given in appendix F. (I have a lot of difficulty to write this paragraph...)

4.2.1 Translational and rotational invariance

If F describes a homogeneous fluid, the translational invariance ($\mathbf{r}_{12} \equiv \mathbf{r}_1 - \mathbf{r}_2$) should be presented, then the number of independent variables is reduced from 12 to 9:

$$F(\mathbf{X}_1, \mathbf{X}_2) = F(\mathbf{r}_{12}, \Omega_1, \Omega_2) = F(r, \hat{\mathbf{r}}_{12}, \Omega_1, \Omega_2) \quad (4.23)$$

We can further expand F on Wigner GSHs of the three orientations, then F becomes a sum of infinite number of projections that depending on r and 8 indices:

$$F(\mathbf{X}_1, \mathbf{X}_2) = \sum_{m,n,l=0}^{\infty} \sum_{|\mu',\mu| \leq m, |\nu',\nu| \leq n, |\lambda'| \leq l} F_{\mu'\mu\nu'\nu\lambda'}^{mn} (r) R_{\mu'\mu}^m (\Omega_1) R_{\nu'\nu}^n (\Omega_2) R_{\lambda'0}^l (\hat{\mathbf{r}}_{12}) \quad (4.24)$$

Assuming that this expansion converges, which is normally the case for correlation functions, the expansion can be expressed in limit number of projections. If we also take into account the rotational invariance by recombine some terms, only r and 5 independent indices are necessary to describe all the projections:

$$F_{\mu'\nu'\lambda'}^{mn} (r) = \sum_{\mu' \nu' \lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} F_{\mu'\mu\nu'\nu\lambda'}^{mn} (r) \quad (4.25)$$

The projections $F_{\mu\nu}^{mnl}(r)$ with a finite order of expansion have much fewer of terms compared to the angular form in eq. (4.23) within the same precision of description. They provide a better way to describe the molecular structure by taking advantage of these invariance.

We can define a basis set of rotational invariant as:

$$\Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}_{12}) = f^{mnl} \sum_{\mu'\nu'\lambda'} \begin{pmatrix} m & n & l \\ \mu' & \nu' & \lambda' \end{pmatrix} R_{\mu'\mu}^m(\Omega_1) R_{\nu'\nu}^n(\Omega_2) R_{\lambda'0}^l(\hat{\mathbf{r}}_{12}) \quad (4.26)$$

where the normalization factor f^{mnl} can be any arbitrary nonzero constant, depending only on indices m, n, l . In Blum's convention, it is taken as $[(2m+1)(2n+1)]^{\frac{1}{2}}$.

With these definition, relation between the projections and the original function is:

$$F(\mathbf{X}_1, \mathbf{X}_2) = \sum_{mnl\mu\nu} \tilde{F}_{\mu\nu}^{mnl}(r) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}_{12}) \quad (4.27)$$

where $\tilde{F}_{\mu\nu}^{mnl}(r) = F_{\mu\nu}^{mnl}(r) / f^{mnl}$.

4.2.2 Blum's reduction of molecular OZ equation

The molecular Ornstein-Zernike (MOZ) equation is defined as:

$$\gamma(\mathbf{X}_1, \mathbf{X}_2) = h(\mathbf{X}_1, \mathbf{X}_2) - c(\mathbf{X}_1, \mathbf{X}_2) = \frac{\rho}{8\pi^2} \int d\mathbf{X}_3 h(\mathbf{X}_1, \mathbf{X}_3) c(\mathbf{X}_3, \mathbf{X}_2) \quad (4.28)$$

which takes advantage of the convolution properties to give

$$\begin{aligned} \hat{\gamma}(\mathbf{k}, \Omega_1, \Omega_2) &= \frac{\rho}{8\pi^2} \int d\Omega_3 \hat{h}(\mathbf{k}, \Omega_1, \Omega_3) \hat{c}(\mathbf{k}, \Omega_3, \Omega_2) \\ &= \frac{\rho}{8\pi^2} \int d\Omega_3 [\hat{\gamma}(\mathbf{k}, \Omega_1, \Omega_2) + \hat{c}(\mathbf{k}, \Omega_1, \Omega_3)] \hat{c}(\mathbf{k}, \Omega_3, \Omega_2) \end{aligned} \quad (4.29)$$

where $\hat{\gamma}$, \hat{h} and \hat{c} are respectively the Fourier transform of γ , h and c .

The rotational invariant expansion gives:

$$c(\mathbf{X}_1, \mathbf{X}_2) = \sum_{mnl\mu\nu} c_{\mu\nu}^{mnl}(r) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}_{12}) \quad (4.30)$$

$$\gamma(\mathbf{X}_1, \mathbf{X}_2) = \sum_{mnl\mu\nu} \gamma_{\mu\nu}^{mnl}(r) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{r}}_{12}) \quad (4.31)$$

and also in k -space:

$$\hat{c}(\mathbf{k}, \Omega_1, \Omega_2) = \sum_{mnl\mu\nu} \hat{c}_{\mu\nu}^{mnl}(k) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{k}}_{12}) \quad (4.32)$$

$$\hat{\gamma}(\mathbf{k}, \Omega_1, \Omega_2) = \sum_{mnl\mu\nu} \hat{\gamma}_{\mu\nu}^{mnl}(k) \Phi_{\mu\nu}^{mnl}(\Omega_1, \Omega_2, \hat{\mathbf{k}}_{12}) \quad (4.33)$$

The relation between these projections in r and k -space are built by the Hankel transform:

$$\hat{c}_{\mu\nu}^{mnl}(k) = 4\pi i^l \int dr r^2 j_l(kr) c_{\mu\nu}^{mnl}(r) \quad (4.34)$$

$$\hat{\gamma}_{\mu\nu}^{mnl}(k) = 4\pi i^l \int dr r^2 j_l(kr) \gamma_{\mu\nu}^{mnl}(r) \quad (4.35)$$

where $j_l(kr)$ are the spherical Bessel functions of order l . As an analogue to FFT, the fast Hankel transform is available for such a process.

The projections $f_{\mu\nu}^{mnl}(r)$ are real, owing to the property of rotational invariants in condition that $f(\mathbf{X}_1, \mathbf{X}_2)$ is real and processes a symmetry axis C_{2v} . Therefore $\hat{f}_{\mu\nu}^{mnl}(k)$ is real if l is even, and pure imaginary if l is odd. (Which is not the exactly case for $\rho_{\mu\nu}^{mnl}(r)$ calculated in MDFT!!! There are some projections that are totally imaginary.)

The MOZ equation based on the rotational invariants $\hat{f}_{\mu\nu}^{mnl}(k)$ can be found in the article of Blum, by the form is a bit complicate. To provide a simpler form, Blum defined the χ -transform:

$$\hat{c}_{\mu\nu,\chi}^{mn}(k) = \sum_{l=|m-n|}^{m+n} \begin{pmatrix} m & n & l \\ \chi & -\chi & 0 \end{pmatrix} \hat{e}_{\mu\nu}^{mnl}(k) \quad (4.36)$$

$$\hat{\gamma}_{\mu\nu,\chi}^{mn}(k) = \sum_{l=|m-n|}^{m+n} \begin{pmatrix} m & n & l \\ \chi & -\chi & 0 \end{pmatrix} \hat{\gamma}_{\mu\nu}^{mnl}(k) \quad (4.37)$$

where we try to use the apostrophe represents the functions in an intermolecular frame. (It is not convenient to use the same symbol for $\hat{e}_{\mu\nu}^{mnl}(k)$ and $\hat{c}_{\mu\nu,\chi}^{mn}(k)$, because they are not the same quantity. However, to be coincident with Blum, $\hat{C}_{\mu\nu,\chi}^{mn}(k)$ should be used instead. Therefore, γ' is Γ , ρ' is P .)

The result MOZ equation is:

$$\hat{\gamma}_{\mu\nu,\chi}^{mn}(k) = \rho \sum_{n_1} \sum_{\nu_1=-n_1}^{n_1} (-)^{\chi+\nu_1} [\hat{\gamma}_{\mu\nu_1,\chi}^{mn_1}(k) + \hat{c}_{\mu\nu_1,\chi}^{mn_1}(k)] \hat{c}_{\underline{\nu_1}\nu,\chi}^{n_1 n}(k) \quad (4.38)$$

This simple form of MOZ equation reduces the calculation of each $(\mathbf{X}_1, \mathbf{X}_3)$ for $(\mathbf{X}_3, \mathbf{X}_2)$ in eq. (4.28) to only a sum of terms of n_1, ν_1 .

5

CODE MDFT

The code MDFT upon which all the development during this thesis is based is a Fortran 95 sequential code developed by Maximilien Levesque, Daniel Borgis *et al.* [22–31], which implement the MDFT theory. It reads the force field (Lennard-Jones and Coulomb parameters) describing the solute and the solvent as input, as well as necessary parameters like the temperature T , number density of solvent n_0 , etc. It minimizes the functional and gives the equilibrium density $\rho(\mathbf{r}, \Omega)$, then computes output properties.

5.1 SUPERCELL DISCRETIZATION

$L_x \times L_y \times L_z [\text{\AA}^3]$ space is discretized on a regular grid of $\text{nfft}_1 \times \text{nfft}_2 \times \text{nfft}_3$ nodes. The solute center is at $\mathbf{r}_T = \left(\frac{L_x}{2}, \frac{L_y}{2}, \frac{L_z}{2} \right)$ of the box. If the internal coordinates of solute \mathbf{r}_M , the solute coordinates in the box $\mathbf{r} = \mathbf{r}_M + \mathbf{r}_T$.

Angular grid is discretized with Lebedev (L) quadrature for $\Omega \equiv (\Theta, \Phi)$, $\Theta \in [0, \pi]$, $\Phi \in [0, 2\pi]$, or Gauss-Legendre (GL) quadrature for Θ and trapezoidal quadrature for Φ . $\Psi \in [0, \pi]$, as we used the code mainly for water, is discretized with trapezoidal quadrature. The number of each angular dimension is linked to the order of quadrature, m_{\max} , which is discussed mainly in the chapter of theory.

5.2 MINIMIZER L-BFGS-B

The minimizer adopted by MDFT is the L-BFGS-B [54, 55] package version 3.0 written in Fortran 77, implementing the limited-memory Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm with constraints of the form $l \leq x \leq u$ to the variable x .

The functional $\mathcal{F}[x_i]$ and the gradient of functional $\nabla \mathcal{F}[x_i] = \frac{\delta \mathcal{F}}{\delta x}(x_i)$ are required by L-BFGS to minimize the functional. It saves the variables x_i and gradients of the past m iterations, which requires a lot of memory.

The functional in MDFT to be minimized is eq. (4.3):

$$\mathcal{F}[\rho] = \mathcal{F}_{\text{id}}[\rho] + \mathcal{F}_{\text{ext}}[\rho] + \mathcal{F}_{\text{exc}}[\rho] \quad (5.1)$$

and its gradient is

$$\frac{\delta \mathcal{F}[\rho]}{\delta \rho(\mathbf{r}, \Omega)} = k_B T \ln \left(\frac{\rho(\mathbf{r}, \Omega)}{\rho_0} \right) + V_{\text{exc}}(\mathbf{r}, \Omega) + (V_{\text{ext}}(\mathbf{r}, \Omega) - \mu) \quad (5.2)$$

where ρ_0 is the angular density of bulk solvent,

$$\rho_0 = \begin{cases} n_0 & \text{if atomic, } \Omega \equiv 1 \\ n_0/4\pi & \text{if linear, } \Omega \equiv (\Theta, \Phi) \\ n_0/8\pi^2 & \text{if non-linear, } \Omega \equiv (\Theta, \Phi, \Psi) \end{cases} \quad (5.3)$$

5.3 TREATMENT TO AVOID UNPHYSICAL DENSITY

During minimization, the density variable $\rho(\mathbf{r}, \Omega)$ can have unphysical negative numbers, which also cause the divergence of the minimization. To avoid this phenomenon, a normalized $\varphi(\mathbf{r}, \Omega)$ is used as the variable during the minimization in place of $\rho(\mathbf{r}, \Omega)$, so that:

$$\rho(\mathbf{r}, \Omega) = \rho_0 \varphi^2(\mathbf{r}, \Omega) \quad (5.4)$$

According to the definition (5.4), we see:

$$\frac{\delta \rho(\mathbf{r}, \Omega)}{\delta \varphi} = 2\rho_0 \varphi(\mathbf{r}, \Omega) \quad (5.5)$$

Therefore the gradient to feed the L-BFGS minimizer is:

$$\frac{\delta \mathcal{F}}{\delta \varphi} = \frac{\delta \mathcal{F}}{\delta \rho} \cdot \frac{\delta \rho}{\delta \varphi} = 2\rho_0 \varphi(\mathbf{r}, \Omega) \cdot [\beta^{-1} \ln \varphi^2 + V_{\text{exc}} + (V_{\text{ext}} - \mu)] \quad (5.6)$$

The main structure of the code is shown in figure 5.1.

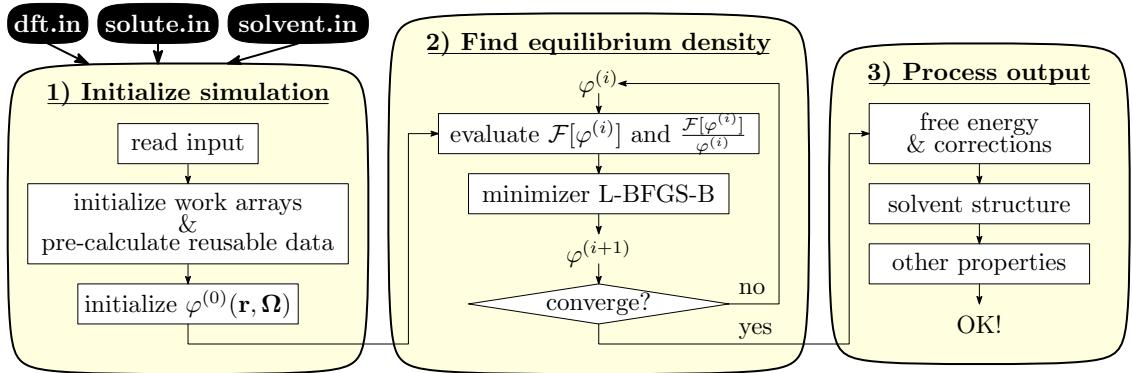


Figure 5.1: Main structure of code MDFT