

High-Performance Platform of Physical Simulations by Using Low Computational Cost Methods

March 25, 2019

Part I

The tight-binding method

Chapter 1

General Eigenvalue Equation.

The problem of generalized eigenvalue in which the tight-binding method is based on consists in solving the matricial equation

$$\mathbf{H} \cdot \mathbf{x} = \lambda \mathbf{S} \cdot \mathbf{x}, \quad (1.1)$$

where the eigenvalue results λ give the energy levels of each electron in the periodic crystalline structure. Based on the approximation of two bodies proposed by Sater-Koster[7], each element from the matrix \mathbf{H} consists on the electron's probability of jumping from one atom's orbit i to another j under the influence of an interaction described by the hamiltonian \hat{H} while the the elements of the matrix \mathbf{S} describes the overlap between the orbitals ψ_i e ψ_j ,

$$h_{j_i m_i, j_j m_j} = \left\langle R(r_i); j_i, m_i \left| \hat{H} \right| R(r_j); j_j, m_j \right\rangle, \quad (1.2)$$

$$s_{j_i m_i, j_j m_j} = \langle R(r_i); j_i, m_i | R(r_j); j_j, m_j \rangle. \quad (1.3)$$

If $s = 0$, we can consider $j \equiv l$, where $l = 0, 1, 2$ e $m = -l, \dots, 0, \dots, l$. However it is common to define the vectors $|l, m\rangle$ as the combination of the orbitals that describe the standard bonds σ , π e δ , to that end we use the expansion

$$|l, m\rangle = \sum_{m'} d_{m'm}^l |l, m'\rangle. \quad (1.4)$$

The projection of each vector $|l, m'\rangle$ is given by the rotation operator $d_{m'm}^l(\alpha, \beta, \gamma)$ [5], where

$$d_{m'm}^l(\alpha, \beta, \gamma) = e^{-im\alpha - im'\gamma} d_{m'm}^l(\beta). \quad (1.5)$$

In this case, α, β e γ are the Euler's rotation angles of each orbital in a absolute reference. Due to the cylindrical symmetry of the atomic orbitals, the rotation of the angle γ around the z axis makes the system unchanged in face of this rotation, which implies $\gamma = 0$. According to Wigner, if the system is unchanged in face of the γ rotation, the matrix $d_{m'm}^l(\alpha, \beta, \gamma = 0)$ is defined as

$$\begin{aligned} d_{m'm}^l(\alpha, \beta) &= e^{-im\alpha} \sum_k (-1)^{k-m+m'} \frac{\sqrt{(l+m)!(l-m)!(l+m')!(l-m')!}}{(l+m-k)!k!(l-k-m')!(k-m+m')!} \\ &\times \cos\left(\frac{\beta}{2}\right)^{2l-2k+m-m'} \sin\left(\frac{\beta}{2}\right)^{2k-m+m'}. \end{aligned} \quad (1.6)$$

Therefor, $h_{l_i m_i, l_j m_j}$ can be described as

$$\begin{aligned} h_{l_i m_i, l_j m_j} &= \sum_{m' m''} d_{m' m_i}^{l_i*}(\alpha, \beta) d_{m' m_j}^{l_j}(\alpha, \beta) \left\langle R(r_i); l_i, m'' \left| \hat{H} \right| R(r_j); l_j, m' \right\rangle, \\ &= \sum_{m'=-\min(l_i, l_j)}^{m'=\min(l_i, l_j)} d_{m' m_i}^{l_i*}(\alpha, \beta) d_{m' m_j}^{l_j}(\alpha, \beta) V_{l_i l_j m'}. \end{aligned} \quad (1.7)$$

In a simmilar manner the overlap $s_{l_i m_i, l_j m_j}$ is defined as

$$s_{l_i m_i, l_j m_j} = \sum_{m'=-\min(l_i, l_j)}^{m'=\min(l_i, l_j)} d_{m' m_i}^{l_i*}(\alpha, \beta) d_{m' m_j}^{l_j}(\alpha, \beta) S_{l_i l_j m'}. \quad (1.8)$$

$V_{l_i l_j m'}$ and $S_{l_i l_j m'}$ are named Slater-Koster (SK) parameters, where m' represents the bonds σ, π and δ ($V_{ss\sigma}, V_{sp\sigma}, V_{pp\sigma}, V_{pd\sigma}, V_{dd\sigma}, V_{pp\pi}, V_{pd\pi}, V_{dd\pi}$ e $V_{dd\delta}$ and $S_{ss\sigma}, S_{sp\sigma}, S_{pp\sigma}, S_{pd\sigma}, S_{dd\sigma}, S_{pp\pi}, S_{pd\pi}, S_{dd\pi}$ e $S_{dd\delta}$).

From a practical perspective, in order to obtain a better analysis of the physical system, the atomic orbital are commonly represented in Cartesian coordinates (s, p_x, p_y, p_z , etc), where the vectors $|l_j, m_j\rangle$ from Eq. 1.4 are replaced by the combinations

$$\begin{aligned}
|s\rangle &= |0, 0\rangle, \\
|p_x\rangle &= \frac{\sqrt{2}}{2} (-|1, 1\rangle + |1, -1\rangle), \\
|p_y\rangle &= \frac{\sqrt{2}}{2} i (|1, 1\rangle + |1, -1\rangle), \\
|p_z\rangle &= |1, 0\rangle, \\
|d_{xz}\rangle &= \frac{\sqrt{2}}{2} (-|2, 1\rangle + |2, -1\rangle), \\
|d_{yz}\rangle &= \frac{\sqrt{2}}{2} i (|2, 1\rangle + |2, -1\rangle), \\
|d_{xy}\rangle &= \frac{\sqrt{2}}{2} i (-|2, 2\rangle + |2, -2\rangle), \\
|d_{x^2-y^2}\rangle &= \frac{\sqrt{2}}{2} (|2, 2\rangle + |2, -2\rangle), \\
|d_{z^2}\rangle &= |2, 0\rangle, \\
\ldots &= \ldots
\end{aligned}$$

Represented in a summarized manner

$$|l\mu\rangle = \sum_{m=-l}^{m=l} c_{\mu m}^l |l, m\rangle, \quad (1.9)$$

where $\mu = 1, \dots, l+1$. Therefore, the Eqs. 1.7 and 1.8 are presented as

$$\begin{aligned}
h_{l_i\mu_i, l_j\mu_j} &= \sum_{m_i=-l_i}^{m_i=l_i} \sum_{m_j=-l_j}^{m_j=l_j} c_{\mu_i m_i}^{l_i*} c_{\mu_j m_j}^{l_j} \sum_{m'=-\min(l_i, l_j)}^{m'=\min(l_i, l_j)} d_{m' m_i}^{l_i*}(\alpha, \beta) d_{m' m_j}^{l_j}(\alpha, \beta) V_{l_i l_j m'}, \\
s_{l_i\mu_i, l_j\mu_j} &= \sum_{m'=-l_i}^{m'=l_i} \sum_{m''=-l_j}^{m''=l_j} c_{\mu_i m_i}^{l_i*} c_{\mu_j m_j}^{l_j} \sum_{m'=-\min(l_i, l_j)}^{m'=\min(l_i, l_j)} d_{m' m_i}^{l_i*}(\alpha, \beta) d_{m' m_j}^{l_j}(\alpha, \beta) S_{l_i l_j m'}. \quad (1.10)
\end{aligned}$$

In addition, after the rotation operation on the conjugated complex $h_{l_i\mu_i, l_j\mu_j}^* = h_{l_j\mu_j, l_i\mu_i}$ must be applied the condition

$$h_{l_j l_i} = (-)^{(l_i + l_j)} h_{l_i l_j}. \quad (1.11)$$

The SK parameters can be obtained comparing the tight-binding curves obtained through the

eigenvalue equation with the band structures of the DFT calculation.

1.1 Distance dependency in the Slater-Koster parameters

The terms of the main diagonal (onsite) are defined as the average energy of the atomic orbitals s and p of each individual atom. However, the onsite term can present a dependency from the neighboring atom's orbitals. In this paper, this dependency is described by the contribution of electronic density of neighboring orbitals played by Gaussian functions, Eq. 1.12,

$$\rho = \sum_j e^{-\lambda r_{ij}^2}. \quad (1.12)$$

Presuming that for the isolated atom, the electronic density of each orbital can be defined by the Gaussian function, after expansion in power series, that is,

$$e^{\alpha r^2} \approx \alpha^1 + \alpha^2 \rho + \alpha^3 \rho^2, \quad (1.13)$$

a representation of the onsite terms of each i atom that depend on the ρ density can be defined with good approximation using the above expression. The neighboring atom's count was done by multiplying the above expression by a Fermi-Dirac cut function that selects the closest atoms, in which the proximity has the most influence on the energy of each atom. Therefore, the onsite terms are defined according to the Eq. 1.14,

$$h_l = \frac{\alpha_l^1 + \alpha_l^2 \rho + \alpha_l^3 \rho^2}{1 + e^{-\frac{(r_{ij} - r_c)}{\kappa}}}. \quad (1.14)$$

r_c represents the cut radius, where the contribution of the farther atoms strongly drops following the γ parameter. In this paper we adopt the values $r_c = 2, 5 \text{ \AA}$ and $\kappa = 0, 15 \text{ \AA}$, in which remains four distinct parameters ($\rho, \alpha^1, \alpha^2, \alpha^3$) to h_l .

The terms outside of the main diagonal (hoppings) are defined as the electron's probability of jumping from one atom's orbit to a neighboring atom's orbit due to the action of an effective

potential. In this case, each hopping has a direct dependency to the SK parameters in which in turn depend on the interatomic distance of both atoms that the electron transitions and also the influence of other orbitals around it. In order to consider the contribution of the closest atoms in the hopping of each individual atom, the terms V_{lm} and S_{lm} were multiplied by the Fermi-Dirac function. In this sense, the Eqs. 1.2 and 1.3 define the SK parameters that have an explicit dependency to the interatomic distance,

$$V_{lm}(r) = \frac{(\beta_{lm}^1 + \beta_{lm}^3 r) e^{-\beta_{lm}^2}}{1 + e^{-\frac{(r-r_c)}{\kappa}}}, \quad (1.15)$$

$$S_{lm}(r) = \frac{(\gamma_{lm}^1 + \gamma_{lm}^3 r) e^{-\gamma_{lm}^2}}{1 + e^{-\frac{(r-r_c)}{\kappa}}}, \quad (1.16)$$

that represents another six parameters to each angular moment l and bond m . In graphene's case, constituted by carbon atoms that have the s and p orbitals, we have 24 parameters regarding the offsite terms and 7 regarding the onsite terms, that account for 31 parameters to be determined.

1.2 Secular matrix diagonalization

Assuming the generalized eigenvalue problem, Eq. 1.1, as long as the matrices \mathbf{H} and \mathbf{S} are symmetrical and \mathbf{S} is considered a positive-defined matrix, the Eq.1.1 can be transformed after the matrix's \mathbf{S} , $\mathbf{S}^{-1}\mathbf{S} = 1$ orthogonality property is applied, in the form

$$(\mathbf{S}^{-1} \cdot \mathbf{H}) \cdot \mathbf{x} = \lambda \mathbf{x}, \quad (1.17)$$

where λ are the equation's eigenvalues and can be attained after diagonalizing the matrix $\mathbf{S}^{-1} \cdot \mathbf{H}$. Nonetheless, although \mathbf{H} and \mathbf{S} are symmetrical, the matrix $\mathbf{S}^{-1} \cdot \mathbf{H}$ is not necessarily symmetrical, which makes difficult the diagonalization process. To convert it into the symmetric form the Cholesky decomposition is applied in the matrix \mathbf{S} ,

$$\mathbf{S} = \mathbf{L} \cdot \mathbf{L}^T, \quad (1.18)$$

being \mathbf{L} the inferior diagonal matrix and \mathbf{L}^T the superior diagonal matrix. Multiplying by \mathbf{L}^{-1} we have

$$\mathbf{L}^{-1} \cdot \mathbf{H} \cdot \mathbf{x} = \lambda (\mathbf{L}^{-1} \cdot \mathbf{L}) \cdot \mathbf{L}^T \cdot \mathbf{x}, \quad (1.19)$$

$$\mathbf{L}^{-1} \cdot \mathbf{H} \cdot \mathbf{x} = \lambda \cdot \mathbf{L}^T \cdot \mathbf{x}. \quad (1.20)$$

Defining the matrix \mathbf{H} as

$$\mathbf{H} = \mathbf{L} \cdot \mathbf{C} \cdot \mathbf{L}^T \quad (1.21)$$

We have

$$(\mathbf{L}^{-1} \cdot \mathbf{L}) \cdot \mathbf{C} \cdot \mathbf{L}^T \cdot \mathbf{x} = \lambda \cdot \mathbf{L}^T \cdot \mathbf{x}, \quad (1.22)$$

$$\mathbf{C} \cdot \mathbf{L}^T \cdot \mathbf{x} = \lambda \cdot \mathbf{L}^T \cdot \mathbf{x}, \quad (1.23)$$

that lead to the eigenvalue equation, being $\mathbf{L}^T \cdot \mathbf{x}$ it's eigenvector. Knowing that \mathbf{H} is a symmetrical matrix, to satisfy the Eq. 1.21 \mathbf{C} must also be symmetrical. In this equation, λ represents the same eigenvalues of Eq. 1.1. Moreover, the \mathbf{C} matrix can be attained after solving the matrical equation

$$\mathbf{L} \cdot \mathbf{C} = \mathbf{Y}, \quad (1.24)$$

being \mathbf{Y} given by the equation

$$\mathbf{Y} \cdot \mathbf{L}^T = \mathbf{H}. \quad (1.25)$$

Therefore, to solve the eigenvalue problem we have to just diagonalize the \mathbf{C} matrix. However, according to the tight-binding method the \mathbf{C} matrix has to be complex and hermitian, where the eigenvalues λ are real. To that end \mathbf{C} defined as $\mathbf{C} = \mathbf{C}_R + i\mathbf{C}_i$ where \mathbf{C}_R and \mathbf{C}_i are real matrices, the eigenvalue equation

$$[\mathbf{C}_R + i\mathbf{C}_i] \cdot [\mathbf{u} + i\mathbf{v}] = \lambda [\mathbf{u} + i\mathbf{v}] \quad (1.26)$$

can be represented in the matrix form

$$\begin{bmatrix} \mathbf{C}_{\mathbf{R}} & -\mathbf{C}_i \\ \mathbf{C}_i & \mathbf{C}_R \end{bmatrix} \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix} = \lambda \begin{bmatrix} \mathbf{u} \\ \mathbf{v} \end{bmatrix}, \quad (1.27)$$

where the λ values can be attained after diagonalizing the real matrix of Eq. 1.27.

Part II

Molecular Dynamics

Chapter 2

The Molecular Dynamics method to work with molecular systems and nanostructures

Taking into account a system in where there is a need to calculate energy levels to a substantial amount of electrons, the solution through the Eq. 1.1 becomes unfeasible from a computational stand point. In this case, methods that use Molecular Mechanics like Monte Carlo and Molecular Dynamics can be alternatives to this situation. The key principle that governs molecular dynamics involves concepts related to the dynamic of the participants that compose a physical system at a molecular level. Due to this characteristic where the mechanical variables have a direct dependency to the system's temporal evolution, several applications are based on this principle. Therefore, the solution of the particles' movement equations that compose the system are necessary in order to obtain these mechanical variables that represent the phase space. For a system constituted by enough particles, through the thermodynamic limit, the connection with thermodynamics can be established, covering applications that encompass studies that fit into balanced and unbalanced thermodynamic systems. However, due to the limitations of mechanical statistics, studies of unbalanced thermodynamic systems are difficult to be analyzed, which makes balanced systems much more viable from a theoretical stand point. Therefore, molecular systems are reduced to a limited class of cases, where certain thermodynamic variables remain constant along the simulation's time. Such variables determine the thermodynamic state of the system and carry on a series of

limitations. To this end these variables are controlled by specific control methods, so that they remain invariant along the time. The molecular dynamic calculation in thermodynamic balance is based upon three basic principles:

- Integrators of movement equations to obtain the canonical variables: The canonical variables of position and moment are determined solving the movement equations through algorithms called movement integrators.
- Force field to describe the interactions between the participants: The interactions between the participants that constitute the system are defined by empirical potentials, which together are called Force Field.
- Statistical Ensemble in order to handle the system from a thermodynamic stand point: The connection with thermodynamic is made through algorithms that control the temperature and pressure along the simulation.

The Fig. 2.1 shows the flowchart that contain the necessary steps for the development of the Molecular Dynamic. The first block refers to the initialization step. In this step the system's topology such as super cell periodicity, temperature, pressure, maximum simulation time, timestep and thermodynamic ensemble are determined by the user. The initial values of of coordinate positions, velocity of each particle and the crystal lattice vectors are attained from the input file HICOLM.md. At the initialization the user can also define how the interactions between the system's participants will proceed. After initialization, the relaxation of the system is proceeded and the calculation of the canonical variables through the MD cycle. Firstly, the forces for each particle of the system is calculated, and after this step, the canonical variables of position and moment are calculated using the Velocit-Verlet algorithm (if necessary, the particles' velocity and crystal lattice values are recalculated in order to reproduce a determined thermodynamic ensemble). Such procedure is repeated while the simulation time is less than the maximum time stipulated by the user when initializing. In each MD cycle, essential values such as position, velocity and thermodynamic quantities are printed in the output file HICOLM.md, where the user can use them to obtain auto correlation calculations (ACF), thermodynamic and structural properties like

radial distribution function.

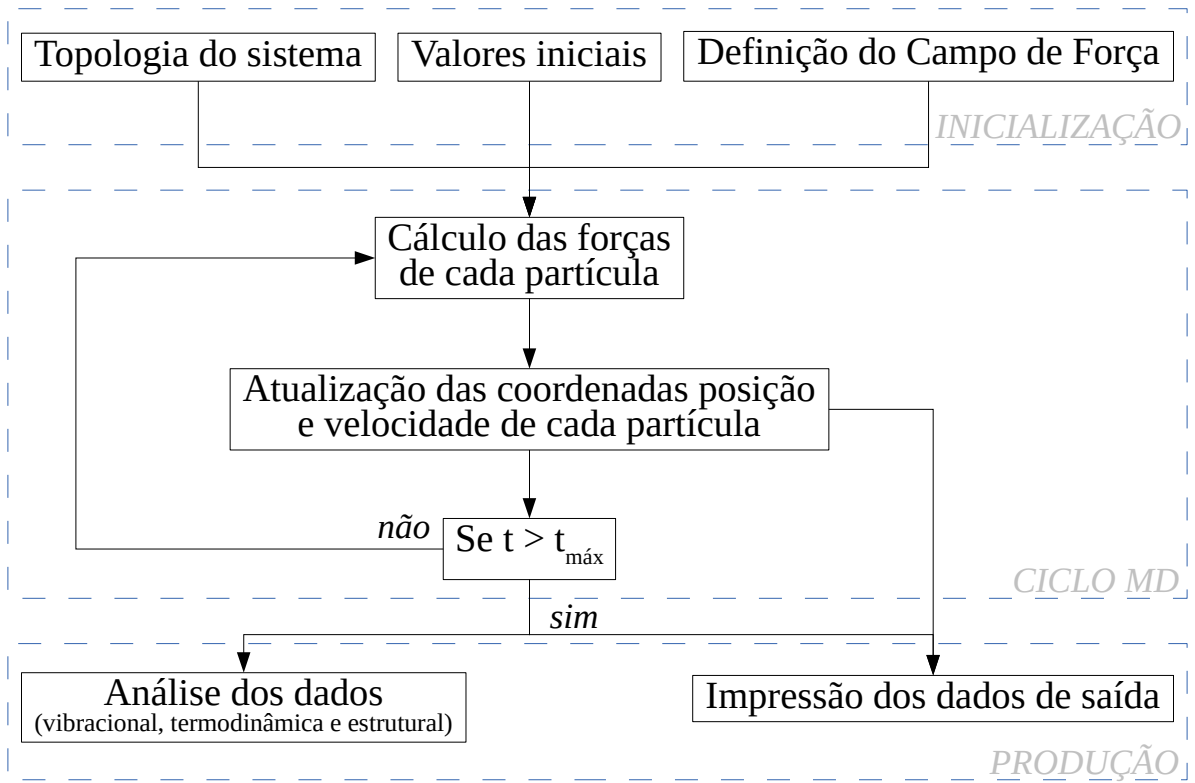


Figure 2.1: Flowchart representing the functioning of the molecular dynamic.

From a theoretical stand point, during each step of the MD cycle, not only the position and velocity of each particle are determined but also its force and potential energy. This requires an update of the list of intra and intermolecular interactions that consumes up to 40% of the time designated to each cycle. To reduce this consumption, in this paper is used a strategy called Verlet's neighbors list. Considering that the intermolecular interaction is of small range we can say with some approximation that only the atoms that are at a certain distance, $r_c + dr_c$, where r_c is called cut radius, will interact effectively with its neighbor. For those that are beyond the cut radius, its contribution is not considered. Therefore, for each atom in the system a neighbor's list is created and it is called Verlet's neighbors list. The update of the Verlet's list is done at each

interval $verlchk$, where

$$verlchk = \frac{dr_c}{\left(\mathbf{v}\delta t + \frac{\mathbf{F}}{2m}\delta t\right)}. \quad (2.1)$$

Chapter 3

Measurement Units

To make the calculations at an atomic scale, the measurement units of the fundamental values were standardized. Therefore, the program standardizes the measurement units so that all of the calculations are done using the atomic units system (a.u.) that are the internal system's units. In the atomic units system, the fundamental units and the conversion factors related to the international units system are defined as shown in Tab. 3.1,

Table 3.1: Fundamental Units applied in the atomic units system.

| Property | Symbol | Corresponding Value in IS |
|------------------------|---------|---|
| Mass | m_e | $9.10938291 \times 10^{-31}$ kg |
| Charge | e | $1.602176565 \times 10^{-19}$ C |
| Angular moment | \hbar | $1.054571726 \times 10^{-34}$ J·s |
| Electrostatic Constant | K | 8.9875517874×10^9 N·m ² C ⁻² |

To get the conversion factors of the physical values, just combine the fundamental units above, also considering the conversion values of the IS for the input units given by the user, as shown in Tab. 3.2. In this table, the first column shows the input units given by the user and the output units printed on the HICOLM.out file. The second column refers to the conversion to a.u., where k_B is Boltzmann's constant and N_0 the Avogadro number.

Table 3.2: Input and output units and their conversions to the a.u. system

| Property | Input Unit | Internal Unit | Conversion |
|-------------|---------------|----------------------------------|---|
| Length | Å | a_0 | $\left(\frac{\hbar^2}{m_e \hbar c^2}\right) [10^{-10}]$ |
| Energy | electron-volt | E_h | $\left(\frac{m_e K^2 e^4}{\hbar^2}\right) [6, 241506 \times 10^{18}]$ |
| Pressure | atmosphere | $\left(\frac{E_h}{a_0^3}\right)$ | $\left(\frac{E_h}{a_0^3}\right) [0.9872 \times 10^{-5}]$ |
| Mass | Atomic mass | m_e | $\left(\frac{m_e}{N_0}\right) [10^3]$ |
| Time | picoseconds | $\left(\frac{\hbar}{E_h}\right)$ | $\left(\frac{\hbar}{E_h}\right) [10^{12}]$ |
| Temperature | Kelvin | $\frac{E_h}{k_B}$ | $\frac{E_h}{k_B}$ |
| Charge | e | e | e |
| Angle | degrees | radians | $\frac{180}{\pi}$ |

Chapter 4

Minimum image convention

The molecular dynamic method generally is used in systems with a limited amount of atoms and molecules that suffer the surface effects caused by the existing interface between liquid and vacuum. With the purpose of representing real systems, where is desired to provide information about the system as a whole, being their “bulk”, such effects don’t take place. The bulk behavior is attained by replicating the primary cell in all directions, where those replicas are named image cells. These image cells have the same dimension, form and the same amount of atoms as the primary cell. Thus, the primary cell can be expanded in space annulling the surface effect and producing the “bulk” state that occurs in real systems when extending the positions and moment of each particle of the primary cell to the respective images in the image cells. The method considers the analyzed system as part of its “bulk” state, where the region in which it is confined is called primary cell. However, multiplying image cells in order to reproduce a bulk state implies in a significant increase in computational cost, making the calculation nonviable. As an alternative we can apply a method called minimum image convention. In this method, a super cell is considered as dimensions large enough so that just the neighboring atoms found on the borders of the single cells can be projected with good approximation. Such procedure is justified if considered short distance interactions, like intramolecular forces and Van der Waals. To that end, in case of an orthorhombic cell, we impose that the maximum distance between the atom and its neighbor can not be larger than the half of the lattice’s constant. In case that happens, the atom considered as a neighboring one is projected

to the image on the side, and the relative distance is changed according to the relation

$$\delta\mathbf{x} = \delta\mathbf{x} - a \cdot \text{int}\left(\frac{2\delta\mathbf{x}}{a}\right), \quad (4.1)$$

$$\delta\mathbf{y} = \delta\mathbf{y} - b \cdot \text{int}\left(\frac{2\delta\mathbf{y}}{b}\right), \quad (4.2)$$

$$\delta\mathbf{z} = \delta\mathbf{z} - c \cdot \text{int}\left(\frac{2\delta\mathbf{z}}{c}\right), \quad (4.3)$$

where a , b and c are the lattice's constants and δr is the relative distance between the atoms in question. However, it is crucial that the simulation box's dimension does not surpass the double of the value of the cut radius, to avoid the counting of each atom with its own image.

Chapter 5

Force Field

The force field of the physical system of interest is based on a group of empirical potentials, with the purpose of reproducing the intra and intermolecular interactions between the system's atoms due to quantum effects. In this case, the system's hamiltonian can be defined as

$$H = K + U, \quad (5.1)$$

where K is kinetic energy and U potential energy. For a system that has N particles and L molecules we have

$$U = \sum_l^L \sum_i^{M-1} \sum_{j<i}^M U_{ij}^{bond}(r_{ij}) + \sum_l^L \sum_i^{M-2} \sum_{j<i}^{M-1} \sum_{k<j}^M U_{jik}^{bend}(\theta_{jik}) + \sum_i^{N-1} \sum_{j<i}^N [U_{ij}^{vw}(r_{ij}) + U_{ij}^{coul}(r_{ij})]. \quad (5.2)$$

The first term refers to the potential energy of stretching between intramolecular bonds in each molecule l . U_{ij}^{bond} depends exclusively on the relative distance r_{ij} of the atoms i and j that form the molecule. The second term refers to the potential energy of the angular deformity. In this case, U_{jik}^{bend} depends exclusively on the θ_{jik} angle between the atoms i , j and k that form each molecule. The third term refers to the intermolecular interactions (between atoms of different molecules) coulombic and Van der Waals, considered as short range.

5.1 Electrostatic Energy

The coulombic interaction allows the interaction between two electrically charged particles continues to act even in long distances, because its expression follows Coulomb's law ($U \propto \frac{1}{r}$), where the potential decays with the inverse of the relative distance. This effect causes an inconvenient, where the interactions between the particles, even when in relatively larger distances, keeps acting considerably. Which in return increases the computational cost and simulation time. However, Ewald [3] has shown that the charged particle among a charge distribution, guides the charges around it so that a shield effect appears, that is a charge distribution of opposite sign neutralizing the coulombic effect due to the electric charge. Ewald suggested that this charge distribution could be mathematically represented by a Gaussian function ρ [referencia] dependent on the q charge and the α parameter, that defines the Gaussian width and is related to the dielectric behavior of the mean,

$$\rho(r) = \frac{q_i \alpha^3}{\pi^{3/2}} e^{-\alpha^2 r^2}. \quad (5.3)$$

Therefore, the coulombic potential energy between particles i and j , including the effect of the Gaussian charge distribution around particle i , is given by the following expression

$$U^{coul} = \sum_i^{N-1} \sum_{j < i}^N \frac{q_j}{4\pi\epsilon_0 r_{ij}} \int_0^\infty q_i \rho(r) d\mathbf{r}, \quad (5.4)$$

$$= \sum_i^{N-1} \sum_{j < i}^N \frac{q_i q_j}{4\pi\epsilon_0} \frac{\text{erfc}(r_{ij})}{r_{ij}}, \quad (5.5)$$

where erfc is the error function. Howeverm the proposed method by Ewald suggests that the long range interaction of the coulombic potential goes beyond the limits of the single cell. In this manner, the radius range of the potential can be expanded replicating the single cell with periodic contour conditions. Therefore, the Eq. 5.5 should be redefined in order to include the sum n that shows the replication number of the single cell,

$$U^{coul} = \sum_i^{N-1} \sum_{j < i}^N \sum_{\{n_1, n_2, n_3\}=0} \frac{q_i q_j}{4\pi\epsilon_0} \frac{\text{erfc}(R_{ij})}{R_{ij}}. \quad (5.6)$$

In this case, $\mathbf{R}_{ij} = \mathbf{r}_{ij} + n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3$, where \mathbf{a}_i , $i = \{1, 2, 3\}$ are the single cell vectors.

However, in certain systems, the periodicity can be problematic, with the increase of computational cost. But the shield effect, caused by the Gaussian distribution can in turn smooth the coulombic interaction as the distance increses, which allows to reduce the effective interaction potential to a a certain distance named cut radius, r_c . Considering that the potential can tend to zero on the limit $r \rightarrow r_c$, the coulombic potential can be redefined as [9]

$$V^{coul}(r) = \begin{cases} v(r), & r \leq r_c \\ 0, & r > r_c \end{cases}. \quad (5.7)$$

In order to avoid the potential's discontinuity and gradient when $r = r_c$, we should include an additional term in the expression. Thus having

$$U_{ij}^{coul}(r_{ij}) = \begin{cases} q_i q_j \left[\frac{1}{r_{ij}} - \frac{1}{r_c} + \frac{1}{r_c^2} (r_{ij} - r_c) \right], & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases}. \quad (5.8)$$

Knowing that $\mathbf{F}_j = -\left(\frac{1}{r} \frac{d}{dr} V_{ij}\right) \mathbf{r}_{ij}$ we have for the force due to Coulomb's law acting on the j particle

$$\mathbf{F}_j(r_{ij}) = \begin{cases} q_i q_j \left(\frac{1}{r_{ij}^2} - \frac{1}{r_c^2} \right), & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases}, \quad (5.9)$$

where the relation $\mathbf{F}_i = -\mathbf{F}_j$ e $\mathbf{r}_{ij} = \mathbf{r}_j - \mathbf{r}_i$ is also valid. Fennel and associates [4] included in the potential energy's term a buffer function in order to speed up the function's converging and decrease the effects due to energy fluctuations closer to the cut radius. The Eq. 5.10 shows the expression used to define the coulombic contribution to the potential energy, and Eq. 5.11 the

expression of the force acting on the j particle due to the coulombic potential,

$$U_{ij}^{coul}(r_{ij}) = \begin{cases} q_i q_j \left[\frac{\text{erfc}(\alpha r_{ij})}{r_{ij}} - \frac{\text{erfc}(\alpha r_c)}{r_c} + \left(\frac{\text{erfc}(\alpha r_c)}{r_c^2} - \frac{2\alpha}{\sqrt{\pi}} \frac{e^{(-\alpha^2 r_c^2)}}{r_c} \right) (r_{ij} - r_c) \right], & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases} \quad (5.10)$$

$$\mathbf{F}_j(r_{ij}) = \begin{cases} \frac{q_i q_j}{r_{ij}^2} \left[\left(\frac{\text{erfc}(\alpha r_{ij})}{r_{ij}^2} - \frac{2\alpha}{\sqrt{\pi}} \frac{e^{(-\alpha^2 r_{ij}^2)}}{r_{ij}} \right) \left(\frac{\text{erfc}(\alpha r_c)}{r_c^2} - \frac{2\alpha}{\sqrt{\pi}} \frac{e^{(-\alpha^2 r_c^2)}}{r_c} \right) \right] \mathbf{r}_{ij}, & r_{ij} \leq r_c \\ 0, & r_{ij} > r_c \end{cases} \quad (5.11)$$

erfc is the complementary error function, to reproduce the shield effect acting on the i particle.

5.2 Van der Waals energy

On the case of Van der Waal's interaction between different particles and molecules, we consider the Lennard Jones potential. Therefor, its contribution to the potential energy and to the force \mathbf{F}_j acting on the j particle can be defined by Eqs. 5.12 and 5.13,

$$U^{vw}(r_{ij}) = 4\varepsilon \left[\left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right], \quad (5.12)$$

$$\mathbf{F}_j = \frac{24\varepsilon}{r_{ij}^2} \left[2 \left(\frac{\sigma}{r_{ij}} \right)^{12} - \left(\frac{\sigma}{r_{ij}} \right)^6 \right] \mathbf{r}_{ij}, \quad (5.13)$$

In this case, ε represents the balance energy and σ the range of the repulsive core between the particles' pairs.

5.3 Intramolecular Terms

5.3.1 Stretching and deformation intramolecular energy.

In the case of the stretching intramolecular interaction, was suggested the harmonic potential

$$U^{bond}(r_{ij}) = \frac{k}{2}(r_{ij} - r_0)^2,$$

$$\mathbf{F}_j = -k(r_{ij} - r_0).$$

Where k represents the vibration mode and r_0 the balance distance between particles i and j. For the angular deformation intramolecular interaction, the harmonic potential was also used as shown in the following equation

$$U^{bend}(\theta_{jik}) = \frac{k}{2}(\theta_{jik} - \theta_0)^2, \quad (5.14)$$

where

$$\theta_{jik} = \cos^{-1} \left(\frac{\mathbf{r}_{ij} \cdot \mathbf{r}_{ik}}{r_{ij}r_{ik}} \right). \quad (5.15)$$

where r_{ij} and r_{ik} define the relative distance between atoms i and j and i and k of each molecule and θ_0 the balance angle. Through the gradient $\mathbf{F} = -\nabla U^{bend}$ we can obtain the contribution for the force of the j particle j,

$$\begin{aligned} \nabla_j U^{bend} &= \frac{\partial U^{bend}}{\partial \theta} \nabla_j \theta, \\ \frac{\partial U^{bend}}{\partial \theta} &= k(\theta_{jik} - \theta_0), \\ \nabla_j \theta &= -\frac{1}{\sin \theta_{jik}} \left[\frac{\mathbf{r}_{ik}}{r_{ij}r_{ik}} - \frac{\mathbf{r}_{ij}}{r_{ij}^2} \cos \theta_{jik} \right]. \end{aligned} \quad (5.16)$$

Calculating \mathbf{F}_k in a similar manner, and knowing that $\mathbf{F}_i = -(\mathbf{F}_j + \mathbf{F}_k)$ we find the contribution for the force of the i, j and k particles of each molecule,

$$\mathbf{F}_l = \frac{k(\theta_{jik} - \theta_0)}{\sin \theta_{jik}} \left[(\delta_{lj} - \delta_{li}) \frac{\mathbf{r}_{ik}}{r_{ij}r_{ik}} + (\delta_{lk} - \delta_{li}) \frac{\mathbf{r}_{ij}}{r_{ij}r_{ik}} - \cos \theta_{jik} \left[(\delta_{lj} - \delta_{li}) \frac{\mathbf{r}_{ij}}{r_{ij}^2} + (\delta_{lk} - \delta_{li}) \frac{\mathbf{r}_{ik}}{r_{ik}^2} \right] \right], \quad (5.17)$$

being δ represented by the Kronecker's delta and $l = \{i, j, k\}$.

5.3.2 Torsion intramolecular energy.

For the intramolecular angular deformation interaction, the harmonic potential was also used, as shown in the following equation

$$U^{tor\tilde{\phi}ao}(\phi) = \frac{k}{2}(\phi_{ijkn} - \phi_0)^2, \quad (5.18)$$

where

$$\phi_{ijkn} = \cos^{-1} \left[\frac{(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \right]. \quad (5.19)$$

Where $\mathbf{r}_{ij} \times \mathbf{r}_{jk}$ defines the vector normal to the plane ijk parallel to the vectors \mathbf{r}_{ij} and \mathbf{r}_{jk} and $\mathbf{r}_{jk} \times \mathbf{r}_{kn}$ and vector normal to the jkn plane parallel to vectors \mathbf{r}_{jk} e \mathbf{r}_{kn} . ϕ_0 the balance angle.

Through the gradient $\mathbf{F}_l = -\nabla_l U^{tor\tilde{\alpha}o}$ we have the contribution to the particle's force l , where

$$\nabla_l U^{tor\tilde{\alpha}o} = -\frac{1}{\sin(\phi_{ijkn})} \frac{\partial U(\phi_{ijkn})}{\partial \phi_{ijkn}} \nabla_l B(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{kn}), \quad (5.20)$$

onde $l = \{i, j, k, n\}$ and

$$B(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{kn}) = \frac{(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|}. \quad (5.21)$$

The B gradient gives us

$$\begin{aligned} \nabla_l B(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{kn}) &= \frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \nabla_l [(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})] + \\ &+ (\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn}) \nabla_l \left[\frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \right], \end{aligned} \quad (5.22)$$

$$\begin{aligned} \nabla_l B(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{kn}) &= \frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \nabla_l [(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})] + (\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn}) \times \\ &\times \left[\frac{1}{|\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \nabla_l [|\mathbf{r}_{ij} \times \mathbf{r}_{jk}|^{-1}] + \frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}|} \nabla_l [|\mathbf{r}_{jk} \times \mathbf{r}_{kn}|^{-1}] \right], \end{aligned} \quad (5.23)$$

$$\begin{aligned} \nabla_l B(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{kn}) &= \frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \nabla_l [(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})] + \frac{(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \times \\ &\times \left[-\frac{1}{2 |\mathbf{r}_{ij} \times \mathbf{r}_{jk}|^2} \nabla_l [|\mathbf{r}_{ij} \times \mathbf{r}_{jk}|^2] - \frac{1}{2 |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|^2} \nabla_l [|\mathbf{r}_{jk} \times \mathbf{r}_{kn}|^2] \right], \end{aligned} \quad (5.24)$$

$$\begin{aligned} \nabla_l B(\mathbf{r}_{ij}, \mathbf{r}_{jk}, \mathbf{r}_{kn}) &= \frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}| |\mathbf{r}_{jk} \times \mathbf{r}_{kn}|} \nabla_l [(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})] - \\ &- \frac{\cos(\phi_{ijkn})}{2} \left[\frac{1}{|\mathbf{r}_{ij} \times \mathbf{r}_{jk}|^2} \nabla_l [|\mathbf{r}_{ij} \times \mathbf{r}_{jk}|^2] + \frac{1}{|\mathbf{r}_{jk} \times \mathbf{r}_{kn}|^2} \nabla_l [|\mathbf{r}_{jk} \times \mathbf{r}_{kn}|^2] \right]. \end{aligned} \quad (5.25)$$

Therefore, the issue in solving the ∇B gradient is summarized by solving the gradients $\nabla [|\mathbf{r}_{ij} \times \mathbf{r}_{jk}|^2]$, $\nabla [|\mathbf{r}_{jk} \times \mathbf{r}_{kn}|^2]$ and $\nabla [(\mathbf{r}_{ij} \times \mathbf{r}_{jk}) \cdot (\mathbf{r}_{jk} \times \mathbf{r}_{kn})]$.

5.4 Tersoff

The bonds between carbon and hydrogen atoms are relatively present in certain structures like hydrocarbons and depend, so to speak, of the mean. Tersoff [8] proposed a model where the covalent bonds between two atoms are played by an attraction and a repulsion term. Tersoff sought to include the mean effects under the potential in which it is inserted. Thus, in this model the attractive character of the chemical bond can vary depending on the number of bonds that the atom shares and the number of neighboring atoms in each bond. The Eq. 5.26 shows the function that plays the potential, in which the Tersoff model is based on,

$$U^{Tersoff} = \sum_i U_i^{Tersoff}, \quad (5.26)$$

$$U_i^{Tersoff} = \sum_{j \neq i} f_c(r_{ji}) [V_R(r_{ji}) - \bar{b}_{ij} V_A(r_{ji})]. \quad (5.27)$$

V_R represents the potential's repulsion term and V_A the attraction term. \bar{b}_{ij} measures the intensity of the attraction term, which depends on the number of bonds shared by each atom in the structure. The HICOLM program adopts the Tersoff model adapted to hydrocarbons proposed by Brenner [6]. According to Brenner, the repulsion and attraction terms are played by functions similar to the Morse potential,

$$V_R(r) = \frac{D}{S-1} e^{-\sqrt{2S}\beta(r-R_e)}, \quad (5.28)$$

$$V_A(r) = \frac{DS}{S-1} e^{-\sqrt{\frac{2}{S}}\beta(r-R_e)}, \quad (5.29)$$

while

$$\bar{b}_{ij} = \frac{1}{2} (b_{ij} + b_{ji}). \quad (5.30)$$

f_c is the cut function, that restricts the potentials action for just the first neighbors,

$$f_c(r) = \begin{cases} 1, & r \leq R_1 \\ \left[1 + \frac{1}{2} \cos \left(\pi \frac{r-R_1}{R_2-R_1} \right) \right], & R_1 < r \leq R_2 \\ 0, & r > R_2 \end{cases} \quad (5.31)$$

b_{ij} is the term responsible to replicate the energy bond, length bond and formation energy between the carbon and hydrogen atoms. b_{ij} depends exclusively on the distribution between the neighboring atoms and the C-C and C-H bond, represented by the variable θ_{jik} , the angle between the atoms i, j and k, Eq. 5.15. A Eq. 5.32 represents the term b_{ij} and its dependency to the angle θ_{jik} inserted in $g(\theta_{jik})$,

$$b_{ij}(r_{ik}, \theta_{jik}) = \left(1 + \sum_{k \neq j} g(\theta_{jik}) f_c(r_{ik}) \right)^{-\delta}, \quad (5.32)$$

$$g(\theta_{jik}) = a_0 \left(1 + \frac{c_0^2}{d_0^2} - \frac{d_0^2}{d_0^2 + (1 + \cos(\theta_{jik}))^2} \right). \quad (5.33)$$

The parameters $R_1, R_2, D, \beta, S, R_e, a_0, c_0, d_0$ and δ are given by the user at the beginning of the simulation.

The force acting in each atom i can be calculated through the expression

$$\mathbf{F}_i = -\nabla_i U_i^{Tersoff}, \quad (5.34)$$

$$= \sum_{j \neq i} - \left[f_c(r) \nabla_i (V_R(r) - \bar{b}_{ij} V_A(r)) + (V_R(r) - \bar{b}_{ij} V_A(r)) \nabla_i f_c(r) \right], \quad (5.35)$$

$$= \sum_{j \neq i} - \left[f_c(r) \nabla_i V_R(r) - f_c(r) \nabla_i [\bar{b}_{ij} V_A(r)] + (V_R(r) - \bar{b}_{ij} V_A(r)) \nabla_i f_c(r) \right], \quad (5.36)$$

$$= \sum_{j \neq i} - \left[f_c(r) \nabla_i V_R(r) - f_c(r) \bar{b}_{ij} \nabla_i V_A(r) - f_c(r) V_A(r) \nabla_i \bar{b}_{ij} + (V_R(r) - \bar{b}_{ij} V_A(r)) \nabla_i f_c(r) \right], \quad (5.37)$$

$$= \sum_{j \neq i} - \left[f_c(r) (\nabla_i V_R(r) - \bar{b}_{ij} \nabla_i V_A(r)) + (V_R(r) - \bar{b}_{ij} V_A(r)) \nabla_i f_c(r) - f_c(r) V_A(r) \nabla_i \bar{b}_{ij} \right], \quad (5.38)$$

where

$$\nabla V_R(r) = \frac{1}{r} \left[-\frac{D\beta\sqrt{2S}}{S-1} e^{-\sqrt{2S}\beta(r-R_e)} \right] \mathbf{r}, \quad (5.39)$$

$$\nabla V_A(r) = \frac{1}{r} \left[-\frac{D\beta\sqrt{2S}}{S-1} e^{-\sqrt{\frac{2}{S}}\beta(r-R_e)} \right] \mathbf{r}, \quad (5.40)$$

$$\nabla f_c(r) = \begin{cases} 0, & r \leq R_1 \\ -\frac{\pi}{2r(R_2-R_1)} \text{sen} \left(\pi \frac{r-R_1}{R_2-R_1} \right) \mathbf{r}, & R_1 < r \leq R_2 \\ 0, & r > R_2 \end{cases} \quad (5.41)$$

Notice that $\mathbf{r}_{ji} = -\mathbf{r}_{ij}$. The derivation of \bar{b}_{ij} is relatively more complex, because we have to consider the variable θ_{jik} that implicitly depends of \mathbf{r}_{ij} . Calculating the gradient in \bar{b}_{ij} we have

$$\nabla_i \bar{b}_{ij}(r_{ik}, \theta_{jik}) = -\delta \bar{b}_{ij}^{-1} \nabla_i \left[1 + \sum_{k \neq i, j} f_c(r_{ik}) g(\theta_{jik}) \right], \quad (5.42)$$

$$= -\delta \bar{b}_{ij}^{-1} \sum_{k \neq i, j} [g(\theta_{jik}) \nabla_i f_c(r_{ik}) + f_c(r_{ik}) \nabla_i g(\theta_{jik})]. \quad (5.43)$$

The first term gives $g(\theta_{jik})$ e $\nabla_i f_c(r_{ik})$ that follows the Eqs. 5.33 e 5.41. Calculating the gradient

of $g(\theta_{jik})$ we have the second term,

$$\nabla_i g(\theta_{jik}) = -\frac{2\sin\theta_{jik}(1+\cos\theta_{jik})a_0d_0^2}{[d_0^2+(1+\cos\theta_{jik})^2]^2}\nabla_i\theta_{jik}, \quad (5.44)$$

Where $\nabla\theta$ is given by Eq. 5.16. Having thus

$$\nabla_i g(\theta_{jik}) = \frac{2(1+\cos\theta_{jik})a_0d_0^2}{[d_0^2+(1+\cos\theta_{jik})^2]^2} \left[\frac{\mathbf{r}_{ik}}{r_{ij}r_{ik}} + \frac{\mathbf{r}_{ij}}{r_{ij}r_{ik}} - \cos\theta_{jik} \left(\frac{\mathbf{r}_{ij}}{r_{ij}^2} + \frac{\mathbf{r}_{ik}}{r_{ik}^2} \right) \right], \quad (5.45)$$

being $l = \{i, j, k\}$ e $k \neq \{i, j\}$. Therefore, the force action on the group $\{i, j, k\}$ can be calculated from the sum of the terms of Eq. 5.38.

Chapter 6

Molecular dynamics and the thermodynamic ensemble

6.1 Pressure and temperature

The temperature of a molecular system containing N particles can be attained through the kinetic gas theory,

$$K = \frac{1}{2} \sum_{i=1}^N m_i v_i^2 = \frac{f}{2} k_B T, \quad (6.1)$$

$$T = \frac{2}{k_B f} K. \quad (6.2)$$

f represents the number of degrees of freedom, where $f = 3(N - N_{fx} - 1)$, being N the total of atoms and N_{fx} the number of fixed atoms. In the calculation of f was also subtracted the value 3, due to the fact that the system's mass center moment is zero in any case. For pressure, it can be calculated through the virial theorem,

$$\begin{aligned} P &= \frac{N k_B T}{V} + \frac{1}{3V} \left\langle \sum_i \sum_{i \leq j} \mathbf{r}_{ij} \cdot \mathbf{F}_j \right\rangle, \\ &= \frac{1}{3V} [2K + W], \end{aligned} \quad (6.3)$$

where

$$W = \sum_i \sum_{i \leq j} \mathbf{r}_{ij} \cdot \mathbf{F}_i + W_{corr} \quad (6.4)$$

represents the sum of the contributions of the virial terms regarding the interactions between the particles and the correction term W_{corr} due to the potential truncation effect, where

$$W_{corr} = \frac{1}{3V} \frac{2\pi}{V} \sum_i \sum_{i \leq j} N_i N_j \int_{r_c}^{\infty} g(r) r^3 \frac{dV_{ij}(r)}{dr} dr, \quad (6.5)$$

$$= \frac{2\pi}{3V^2} \sum_i \sum_{i \leq j} N_i N_j \int_{r_c}^{\infty} r^3 \frac{dV_{ij}(r)}{dr} dr. \quad (6.6)$$

$V(r)$ represents the potential defined by the Force Field, N the number of particles of each type and V the super cell volume. In this case, we can consider $g(r) \approx 1$ when $r \geq r_c$. The first term of Eq. 6.3 refers to the ideal gas equation, that represents the average of the collisions between the particles and the simulation's box wall.

6.2 The microcanonical ensemble

In isolated systems where the energy and volume are also unchanged we name it microcanonical ensemble (NVE), and no controls are used during the simulation process, once that the definition represents a isolated system where the energy is preserved during the process. The variables position and moment are given by solving the Hamilton equations,

$$\frac{d\mathbf{r}}{dt} = \mathbf{v}, \quad (6.7)$$

$$\frac{d\mathbf{H}}{dt} = \frac{\mathbf{F}}{m}, \quad (6.8)$$

where \mathbf{F} is the acting force in each particle, m is mass, \mathbf{r} the position of each particle and \mathbf{v} its velocity. H is the adiabatic system's hamiltonian, given by Eq. 6.9,

$$H = K + V, \quad (6.9)$$

where V represents the potential energy according to the Force Field. However, in the caso of a system containing N particles, is necessary to solve $3N$ second-order equations, making the solution practically unviable from a computational stand point. Yet, it is possible to obtain these values over time using the artifice proposed by Verlet[1]. Considering initially the expansion in power series fo velocity at an intermediate instant $t_n + \frac{1}{2}\delta t$,

$$\mathbf{v}(t_n + \frac{1}{2}\delta t) = \mathbf{v}(t_n) + \frac{1}{2} \left. \frac{d\mathbf{v}(t)}{dt} \right|_{t_n} \delta t + \frac{1}{2 \cdot 2!} \left. \frac{d^2\mathbf{v}(t)}{dt^2} \right|_{t_n} \delta t^2 + \dots, \quad (6.10)$$

$$= \mathbf{v}(t_n) + \frac{1}{2} \left[\left. \frac{d\mathbf{v}(t)}{dt} \right|_{t_n} \delta t + \frac{1}{2} \left. \frac{d^2\mathbf{v}(t)}{dt^2} \right|_{t_n} \delta t^2 + \dots \right]. \quad (6.11)$$

Identifying $\frac{\mathbf{F}}{m} = \frac{d\mathbf{v}(t)}{dt}$ we have

$$\mathbf{v}(t_n + \frac{1}{2}\delta t) = \mathbf{v}(t_n) + \frac{1}{2m} \left[\mathbf{F}(t_n) + \frac{1}{2} \left. \frac{d\mathbf{F}(t)}{dt} \right|_{t_n} \delta t + \dots \right] \delta t. \quad (6.12)$$

Se $\delta t \ll 1$ we can consider

$$\mathbf{v}(t_n + \frac{1}{2}\delta t) \approx \mathbf{v}(t_n) + \frac{1}{2m} \mathbf{F}(t_n) \delta t. \quad (6.13)$$

In the case of velocity at the following instant $t_n + \delta t$, we have

$$\mathbf{v}(t_n + \delta t) = \mathbf{v}(t_n) + \left. \frac{d\mathbf{v}(t)}{dt} \right|_{t_n} \delta t + \frac{1}{2} \left. \frac{d^2\mathbf{v}(t)}{dt^2} \right|_{t_n} \delta t^2 + \dots, \quad (6.14)$$

$$= \mathbf{v}(t_n) + \frac{1}{2m} \mathbf{F}(t_n) \delta t + \frac{1}{2m} \left[\mathbf{F}(t_n) + \left. \frac{d\mathbf{F}(t)}{dt} \right|_{t_n} \delta t + \dots \right] \delta t, \quad (6.15)$$

Expanding the \mathbf{F} force around instant t_n we have

$$\mathbf{F}(t_n + \delta t) = \mathbf{F}(t_n) + \left. \frac{d\mathbf{F}(t)}{dt} \right|_{t_n} \delta t + \frac{1}{2!} \left. \frac{d^2\mathbf{F}(t)}{dt^2} \right|_{t_n} \delta t^2 + \dots. \quad (6.16)$$

Therefore

$$\mathbf{v}(t_n + \delta t) = \mathbf{v}(t_n) + \frac{1}{2m} \mathbf{F}(t_n) \delta t + \frac{1}{2m} \mathbf{F}(t_n + \delta t) \delta t. \quad (6.17)$$

Comparing Eqs. 6.13 and 6.17 we can obtain the exact solution that defines the evolution of velocity in each δt time step, Eq. 6.18,

$$\mathbf{v}(t_n + \delta t) = \mathbf{v}(t_n) + \frac{1}{2}\delta t \mathbf{F}(t_n) + \frac{1}{2m}\mathbf{F}(t_n + \delta t)\delta t. \quad (6.18)$$

In the case of each particle's position, we can also expand in power series

$$\mathbf{r}(t_n + \delta t) = \mathbf{r}(t_n) + \left. \frac{d\mathbf{r}(t)}{dt} \right|_{t_n} \delta t + \frac{1}{2!} \left. \frac{d^2\mathbf{r}(t)}{dt^2} \right|_{t_n} \delta t^2 + \mathcal{O}(\delta t^4), \quad (6.19)$$

where the evolution of the position \mathbf{r} of each particle at t_n instant is considered to the following instant $t_n + \delta t$. Identifying the term $\mathbf{v} = \frac{d\mathbf{r}(t)}{dt}$ and applying the Eq. 6.13 we have the exact expression that defines each particle's position evolution in each δt time step, Eq. 6.20,

$$\mathbf{r}(t_n + \delta t) = \mathbf{r}(t_n) + \mathbf{v}(t_n) \delta t + \frac{1}{2}\delta t^2 \mathbf{F}(t_n). \quad (6.20)$$

The Eqs. 6.18 and 6.20 represent the scheme proposed by Verlet to obtain the canonical variables without the need to directly solve the movement equations of the system in question.

However, when recalculating the position of each particle is crucial to consider the boundaries of the single cell. To that end, the periodic countour conditions are applied, in this manner we ensure that all of the atoms are within the boundaries of the single cell in each step of the simulation. In case of the orthorhombic cell, the atomic positions are redefined according to the following equation,

$$\mathbf{x}(t) = \mathbf{x}(t) - a \cdot \text{nint} \left(\frac{\mathbf{x}(t)}{a} \right), \quad (6.21)$$

$$\mathbf{y}(t) = \mathbf{y}(t) - b \cdot \text{nint} \left(\frac{\mathbf{y}(t)}{b} \right), \quad (6.22)$$

$$\mathbf{z}(t) = \mathbf{z}(t) - c \cdot \text{nint} \left(\frac{\mathbf{z}(t)}{c} \right), \quad (6.23)$$

where a , b and c are the lattice constants. Fig. 1 shows the algorithm that defines the Verlet's strategy to obtain the position and velocity of each particle in each time step.

Algoritmo 1: Movement integrator Velocity Verlet

```
início
  se First step então
    | Force calculus  $\mathbf{F}(t)$ 
  fim
  para Atom i faça
    |  $\mathbf{v}_i(t + \frac{1}{2}\delta t) = \mathbf{v}_i(t) + \frac{1}{2m_i}\mathbf{F}_i(t)\delta t;$ 
    |  $\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{1}{2}\delta t)\delta t$ 
  fim
  para Atom i faça
    |  $\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t + \delta t) - \mathbf{a} \cdot \text{nint}\left(\frac{\mathbf{r}_i(t + \delta t)}{\mathbf{a}}\right)$ 
  fim
  Force calculus,  $\mathbf{F}(t + \delta t);$ 
  para Atom i faça
    |  $\mathbf{v}_i(t + \delta t) = \mathbf{v}_i(t + \frac{1}{2}\delta t) + \frac{1}{2}\frac{\mathbf{F}_i(t + \delta t)}{m_i}\delta t$ 
  fim
   $t \rightarrow t + \delta t$ 
fim
```

6.3 Canonical ensemble

In the case of the canonical ensemble, where we want to set not only the volume but also the temperature, the velocities are redefined including a re-scaling velocity parameter χ ,

$$\mathbf{v}(t) = \chi(t)\mathbf{v}(t). \quad (6.24)$$

One of the applied method to adjust the particles' velocities was adopted by Berendsen[2]. In that, the instant temperature is adjusted to the wanted temperature T_e using Eq. 6.25,

$$\chi(t) = \sqrt{1 + \frac{\delta t}{\tau_T} \left(\frac{0.5fk_B T_e}{K} - 1 \right)}, \quad (6.25)$$

where τ_T represents the user defined adjustment. The canonical variables can be obtained in a simmlar manner as the one used in the previous section. However, the Velocity Verlet algorithm is adapted including the parameter χ in the calculation of the velocity at the instant $t_n + \delta t$,

$$\mathbf{v}(t_n + \delta t) = \chi(t) \left[\mathbf{v}(t_n + \frac{1}{2}\delta t) + \frac{1}{2m}\mathbf{F}(t_n + \delta t)\delta t \right]. \quad (6.26)$$

Algoritmo 2: Velocity Verlet integrator adapted to the Canonical ensemble.

```
início
  se First step então
    | Force calculus  $\mathbf{F}(t)$ 
  fim
  para Atom i faça
    |  $\mathbf{v}_i(t + \frac{1}{2}\delta t) = \mathbf{v}_i(t) + \frac{1}{2m_i}\mathbf{F}\delta t;$ 
    |  $\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{1}{2}\delta t)\delta t$ 
  fim
  para Atom i faça
    |  $\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t + \delta t) - \mathbf{a} \cdot \text{nint}\left(\frac{\mathbf{r}_i(t+\delta t)}{\mathbf{a}}\right)$ 
  fim
  Force calculus,  $\mathbf{F}(t + \delta t);$ 
  para Atom i faça
    |  $\mathbf{v}_i(t + \delta t) = \mathbf{v}_i(t + \frac{1}{2}\delta t) + \frac{1}{2m_i}\mathbf{F}_i(t + \delta t)\delta t$ 
  fim
  Kinetic energy calculus,  $K = \sum_i \frac{1}{2}m_i v_i^2;$ 
  Scaling parameter calculus  $\chi(t + \delta t);$ 
  para Atom i faça
    |  $\mathbf{v}_i(t + \delta t) = \chi(t + \delta t)\mathbf{v}_i(t + \delta t)$ 
  fim
   $t \rightarrow t + \delta t$ 
fim
```

Thus, the velocity of each particle is redefined according to the χ parameter, and for the position, the Eq. 6.20 is maintained. Fig. 2 shows the Velocity Verlet scheme adapted to the canonical ensemble.

6.4 The Gibbs ensemble

In case of the Gibbs ensemble, where we want to set the temperature and pressure, the volume and the single cell shape are dynamically adjusted through algorithms being thermostat for the temperature control and barostat for pressure control. As well as in the Canonical ensemble, to redefine the system's temperature we include a scaling parameter in the velocity calculation, where Eqs. 6.25 e 6.26 are still valid. For the pressure adjustment, we can define the volume V and the

lattice parameters \mathbf{a} of the single cell through a η parameter,

$$V(t) = \eta(t)^3 V(t), \quad (6.27)$$

$$\mathbf{a} = \eta(t) \mathbf{a}, \quad (6.28)$$

where Berendsen[2] proposed a formula for the parameter η , in which depends of τ_p defined initially by the user,

$$\eta(t) = \left[1 + \frac{\beta \delta t}{\tau_p} (p(t) - p_e) \right]^{1/3}. \quad (6.29)$$

However, when adjunting the single cell shape in necessary to redefine the positions of each particle in order to follow its expansion or contraction. According to Eq. 6.20 we have

$$\mathbf{r}(t_n + \delta t) = \eta(t) \mathbf{r}(t_n) + \mathbf{v}(t_n + \frac{1}{2} \delta t) \delta t. \quad (6.30)$$

In this expression, p_e defines the desired pressure and $p(t)$ the instant pressure, calculated in each step of the simulation. β is a constant, that the value was determined as being $4,9 \times 10^{\{-5\}} \text{ atm}^{-1}$. Fig. 3 shows the Velocity Verlet scheme adapted for the Gibbs ensemble.

Algoritmo 3: Velocity Verlet integrator adapted for the Gibbs ensemble.

```
início
  se First step então
    | Force calculus  $\mathbf{F}(t)$ 
  fim
  Parameter's Calculus  $\eta(t)$ ;
  para Atom i faça
    |  $\mathbf{v}_i(t + \frac{1}{2}\delta t) = \mathbf{v}_i(t) + \frac{1}{2m_i}\mathbf{F}_i(t)\delta t$ ;
    |  $\mathbf{r}_i(t + \delta t) = \eta(t)^{1/3}\mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{1}{2}\delta t)\delta t$ 
  fim
  Redefining vectors of the single cell,  $\mathbf{a}(t + \delta t) = \eta(t)^{1/3}\mathbf{a}(t)$ ;
  Redefining volume of the single cell,  $V(t + \delta t) = \eta(t)V(t)$ ;
  para Atom i faça
    |  $\mathbf{r}(t + \delta t) = \mathbf{r}(t + \delta t) - \mathbf{a}(t + \delta t) \cdot \text{nint}\left(\frac{\mathbf{r}(t + \delta t)}{\mathbf{a}(t + \delta t)}\right)$ 
  fim
  Force calculus,  $\mathbf{F}(t + \delta t)$ ;
  para Atom i faça
    |  $\mathbf{v}_i(t + \delta t) = \mathbf{v}_i(t + \frac{1}{2}\delta t) + \frac{1}{2m_i}\mathbf{F}_i(t + \delta t)\delta t$ 
  fim
  Kinetic energy calculus,  $K = \sum_i \frac{1}{2}m_i v_i^2$ ;
  Parameter calculus  $\chi(t + \delta t)$ ;
  para Atom i faça
    |  $\mathbf{v}_i(t + \delta t) = \chi(t + \delta t)\mathbf{v}_i(t + \delta t)$ 
  fim
   $t \rightarrow t + \delta t$ 
fim
```

Bibliography

- [1] M P Allen and D J Tildesley. *Computer Simulation of Liquids*. Clarendon Press, New York, NY, USA, 1989.
- [2] 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, oct 1984.
- [3] P. P. Ewald. Die Berechnung optischer und elektrostatischer Gitterpotentiale. *Annalen der Physik*, 369(3):253–287, jan 1921.
- [4] Christopher J. Fennell and J. Daniel Gezelter. Is the Ewald summation still necessary? Pairwise alternatives to the accepted standard for long-range electrostatics. *The Journal of Chemical Physics*, 124(23):234104, jun 2006.
- [5] A. Grüneis, C. Attaccalite, L. Wirtz, H. Shiozawa, R. Saito, T. Pichler, and A. Rubio. Tight-binding description of the quasiparticle dispersion of graphite and few-layer graphene. *Physical Review B*, 78(20):205425, nov 2008.
- [6] Kaveh Samadikhah, Ragnar Larsson, Faranak Bazooyar, and Kim Bolton. Continuum-molecular modelling of graphene. *Computational Materials Science*, 53(1):37–43, feb 2012.
- [7] J. C. Slater and G. F. Koster. Simplified LCAO method for the periodic potential problem. *Physical Review*, 94(6):1498–1524, 1954.
- [8] J. Tersoff. Modeling solid-state chemistry: Interatomic potentials for multicomponent systems. *Physical Review B*, 39(8):5566–5568, mar 1989.

- [9] D. Wolf, P. Keblinski, S. R. Phillpot, and J. Eggebrecht. Exact method for the simulation of Coulombic systems by spherically truncated, pairwise r-1 summation. *The Journal of Chemical Physics*, 110(17):8254, apr 1999.