Computational biophysics

Giacomo Fantoni
telegram: @GiacomoFantoni
Github: https://github.com/giacThePhantom/mathematical-modelling-in-biology

November 20, 2022

Contents

Chapter 1

Introduction and proteins

1.1 Introduction

Biomolecular modelling has seen a recent increase in its use in the recent years, with a field still destined to expand. Most of this models take a top-down approach, starting from the macroscopic rather than to build simulation from the fundamental and quantistic concepts. Examples of systems studied through biomolecular modelling are:

- Channels.
- DNA/RNA interactions.
- Photosynthetic systems.
- Viruses.
- Inorganic systems.

Through biomolecular modelling it is possible to obtain:

- Molecular rationale for biological processes like proteins' function or its misfolding.
- A prediction of properties of macromolecular structures and architectures.
- A quantitative evaluation of molecular driving forces.
- A comparative assessment of molecular affinities through the binding free energy.

1.2 Proteins

Proteins have different functions within a cell:

- Give structure.
- Catalytic.
- Provide exchange of materials.
- Movement.
- Code for messages.
- Storage.
- Transport ions.
- Act as toxins.

Proteins are a polymer of amino-acids and occupy a space-scale of $10nm$. The amino-acids are in the range of $1nm$. They are built through a polymerization reaction as chain of amino-acids coded through a degenerate code of RNA nucleotides. Three bases of RNA code for an amino-acid.

1.2.1 Amino-acids

Amino-acids are the monomers of a protein. They have a general structure with an amino and a carboxyl terminal group for all of them. They are distinguished by a residue on the α-carbon which gives them different chemical and physical properties.

1.2.2 Structure

There are four level of a protein structure.

- Primary structure: the amino-acid sequence.
- Secondary structure: here α-helices and β-sheet can be distinguished.
- Tertiary structure: the spatial, 3D dynamic configuration of a protein which arise during protein folding.
- Quaternary structure: the interaction of multiple correctly-folded proteins.

Chapter 2

Proteins' geometry

2.1 Introduction

The study of the geometry of proteins involve what can be learned from protein coordinates.

2.2 The peptide bond

A protein is a collection of amino acids linked together by a peptide bond. A carboxylic end and an amino end of two amino acid react together losing a water molecule and forming a peptide bond. Beside the α-carbon there is another one bonded to the oxygen in the carboxylic group and a nitrogen bond in the amino group. The α carbon is linked to the nitrogen in the amino group of another amino acid. The carbon and nitrogen display $sp2$ hybridization, the central atom and the 3 that form a bond with it form a plane, so the peptide bond is planar. A plane of the peptide bond is formed and rotation of the plane is allowed only around one axis.

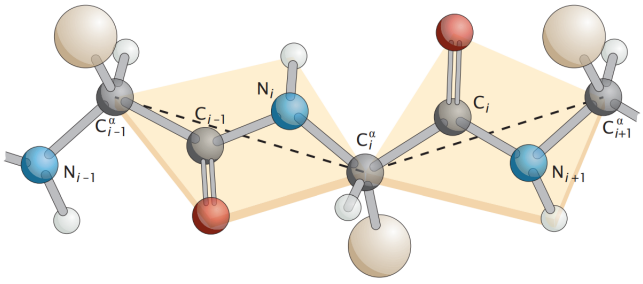

Figure 2.1: sp2 hybridization of backbone C and N (not C^{α}. This is an experimental finding which needs to be incorporated into the system: in this way (and by taking into consideration all the knowledge we have on protein), we are able to reduce the degree of freedom of the system.

2.2.1 Trans and cis

Looking at the peptide bond the carbon atom of the carboxylic group C' and the nitrogen N are each bonded to a different α-C and a trans or cis conformation can happen. Trying to visualize the atoms that belong to the molecules these repel through the Van der Waals interactions, that can be computed through the Lennard-Jones potential (represented in figure2.2:

$$U_{Lj}(r) = E_0\left[\left(\frac{r_0}{r}\right)^{12} - 2\left(\frac{r_0}{r}\right)^{6}\right]$$

Where:

- r_0 is the distance where the energy is minimum.
- r is the distance between two atoms.
- r_{min} is the distance at which the energy becomes high.

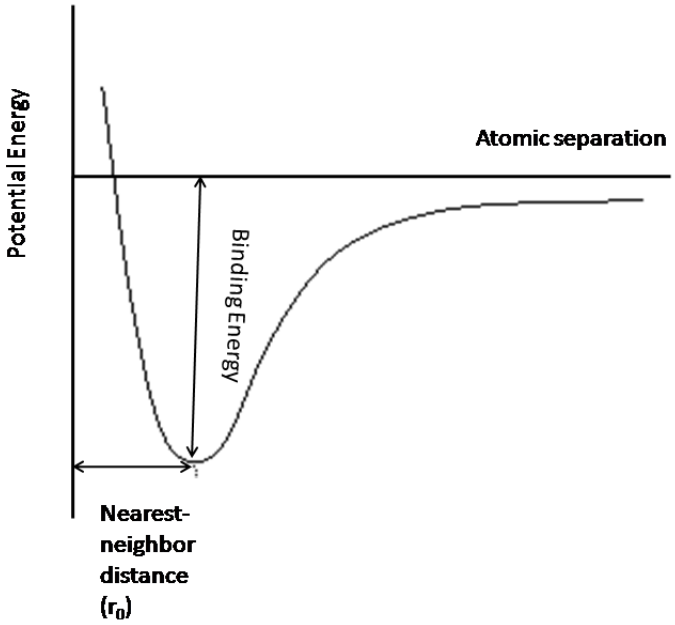

Figure 2.2: Graphical representation of the energy plotted along the distance between atoms. r_0 is the equilibrium distance and it explains the reason why the trans configuration is preferred over the cis configuration

With atoms most of the time the distance between them will be close to r_0. When decreasing the distance a lot of energy is needed and strain is introduced in the molecule. Plotting the values for the energy, r_0 and r_{min} the expected distance for each couple of atoms can be seen. Focusing on the C-C interaction:

- $r_0 = 3.4$Å.
- $r_{min} = 3.0$Å.

When two carbons atoms are below the minimum value the conformation is strained. Looking back at the conformation of the peptide bond it can be seen that the cis conformation creates a distance of 2.8Å between the two αC, so it is not favourable. So the trans conformation is the least energy-hungry and the most present.

Note that the Lennard-Jones potential is not the only one possible. Although it is preferred in most scenarios, other types of potential (e.g., buckingham potential or LJ with different powers) there exist.

2.3 The Ramachandran angles

The planes formed by the peptide bonds can rotate with respect to each other. So the Ramachandran angles ϕ and ψ can be defined between these planes. For each αC:

- ϕ describes the rotation around its bond with the nitrogen.
- ψ describes the rotation around its bond with the carboxylic group.

These are the angles between the subsequent planes. Some of the angles will require more energy.

2.3.1 Difficulty of rotation

It can be seen how a rotation of the ϕ angle could cause the two C' to come at a distance of 2.9Å (where $r_{min} = 3.0$Å. On the other hand a rotation of the ψ angle could cause the two N to come at the same distance, but in this case $r_{min}(N \dots N) = 2.7$Å. In the case of carbon atoms the distance is less than the minimum distance, while in the case of nitrogen it is greater than the minimum allowed value. Looking at this it can be seen how the ψ rotation is easier.

2.3.2 Ramachandran plot

A Ramachandran plot is a map with the ϕ angle on the x axis and the ψ angle on the y axis. Because a rotation along the ϕ angle is highly disfavoured the angle 0 is strongly disfavoured and is represented like a black stripe (disallowed region). If the amino acids where composed only by carbon and nitrogen atom the Ramachandran map would be the one represented in figure**??**, where:

- A forbidden region in the middle.
- Some strained region like for $\psi = 0$.

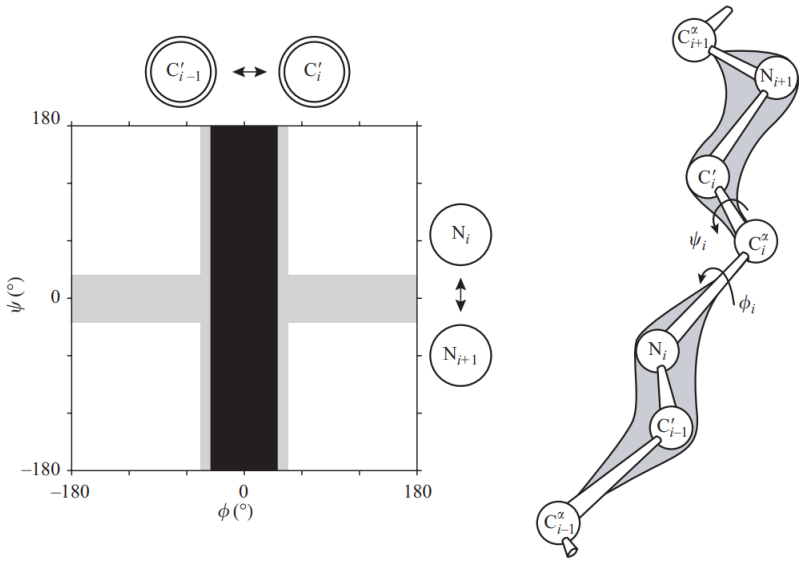

Figure 2.3: This is how Ramachandran plots of the disallowed (black stripe), strained (grey stripe), and fully allowed (white regions) (ϕ, ψ) conformations of the fragment $C^\alpha C'N$—C^α—$C'N - C^\alpha$ would look, provided all these atoms had no other atoms attached (right) and atoms of residues $i-1$ and $i+1$ had no interactions.

Looking at a real protein the complexity is increased and the other oxygen and nitrogen atoms are included **??** and other regions become disallowed due to steady clashes. It can be seen how the regions are quite complex. Looking at a glycine and alanine complex it can be seen in **??** the space becomes even more complex.

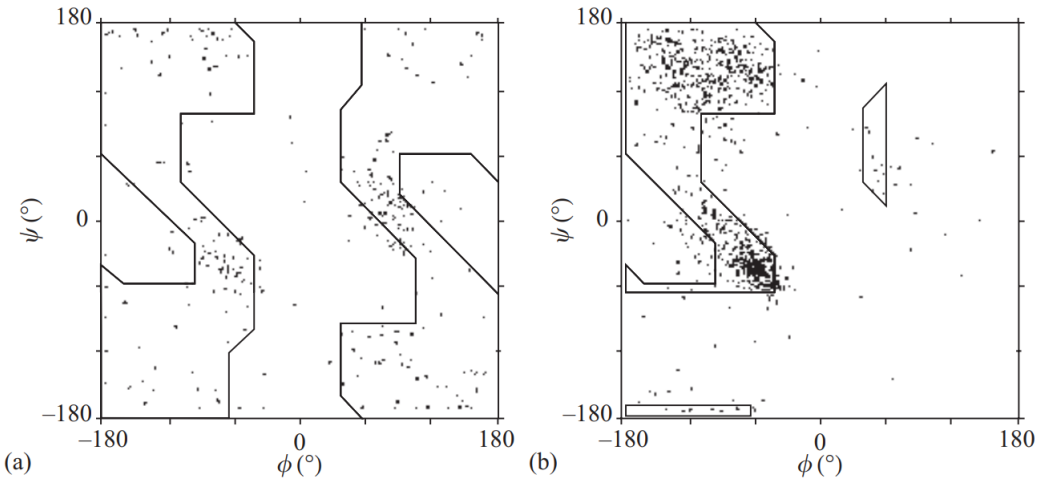

Figure 2.4: Observed conformations (dots) of glycine (a) and of other amino acid residues (b) in proteins. The sterically allowed regions are contoured.

In this case the white regions is very small and a strained region can be seen and the black one. Including other residues the allowed region reduces **??**. This is due to the presence of larger residues.

2.3.2.1 Observed Ramachandran plot

Trying to plot for each amino acid its angles an amino acid is represented as a dot. Most of the points fall inside of the allowed regions but there are some outliers. In some conformation the protein forces the amino acid to assume strange conformations. This is done to check if the structure places the amino acids in a proper way.

2.4 Contact map of proteins

Starting from the coordinates a contact can be built. It is a matrix that map all the contact between the amino acids. A primary structure can be represented as a collection of beads which will be in contact in the 3D structure. A square matrix can be built such that each entry in the matrix will determine whether there is a contact or not. This matrix will be symmetric with diagonal elements with value 1 and two parallel diagonals for the neighbouring amino acids (figure 2.5). Secondary structures will have specific signatures:

- α-helices: is usually represented by a line parallel to the diagonal. This is because the amino acids i is interacting with $i + 4$.
- β-strands: the situation is complicated. For parallel β sheets can be parallel to the diagonal. For anti-parallel it can be anti-parallel to the diagonal.

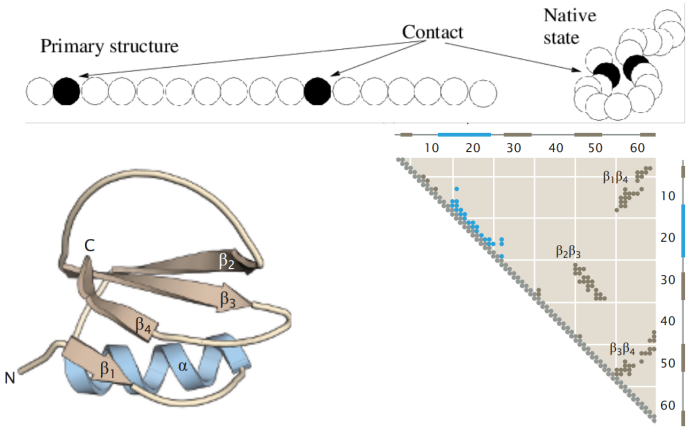

Figure 2.5: Example of a contact map for a protein.

2.4.1 Defining a contact

The contact between two amino acids needs to be defined. To do so the distance between α-C or the distance between the tail of the residue and an α-C. There is also the need to make a trade-off between computational speed and cost. Also the dimension of the protein need to be considered when choosing the distance.

2.5 Topology diagram

Having found the secondary structures with a contact map a topology diagram help to understand how those interact with each other (figure 2.6). In a topology diagram the start is the N terminus and the end the C terminus. β-strands are represented as arrows. If the strands always change direction they will form an anti-parallel β-sheet. α-helices are represented as small cylinder. Usually color codes represent the nature of the structure. This helps with numbering of the secondary structures.

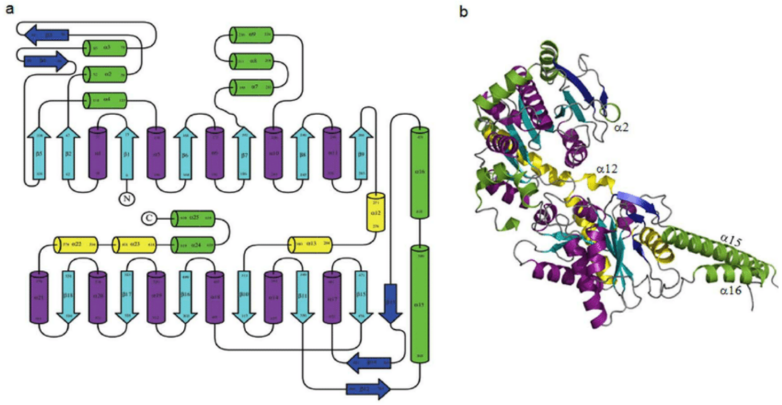

Figure 2.6: The topology of a protein structure is a highly simplified description of its fold including only the sequence of secondary structure elements, and their relative spatial positions and approximate orientations. This information can be embodied in a two-dimensional diagram of protein topology, called a TOPS cartoon. (from pubmed)

2.6 Coordinates

The coordinates of all atoms in a protein are described in a PDB file. This is a tabulated file containing different columns:

- Atom record: ATOM.
- Atom number: a unique identifier for the atom.
- Atom identifier: an identifier for the type of atom.
- Amino type: the amino acid from which the atom is from.
- Chain identifier: identifier for the chain.
- Residue sequence number: the number of the residue in the chain.
- x, y, z: the coordinates in angstrom.
- Occupancy: the probability of an atom to be in that space (confidence space).
- B-factor: how mobile that atom is in the crystal, it represent the noise in the x-ray diffraction map (it represents the fluctuations for the coordinate of the atoms. For example, loops in trans-membrane proteins are very mobile and a high B-factor is to be expected).
- Element symbol: the symbol of the element of the atom.

Once the coordinates of a protein is obtained, some geometrical properties can be directly computed.

2.6.1 Protein center of mass

The protein center of mass is the average position for the protein center. It is an average weighted by the mass of the atom.

$$\vec{R}_{cm} = \frac{\sum_{i=1}^{N} m_i \vec{r}_i}{\sum_{i=1}^{N} m_i}$$

2.6.2 Radius of gyration

Once the center of mass is known the radius of gyration can be computed. This measures the size of the protein as if it was a sphere. It is a good indication of the globular size of a protein. It also indicates an elongating/shrinking behavior, and useful to check whether a protein is going toward equilibrium in the simulation. The distance of each atom and the center of mass is computed and the square is taken, weighted with the mass of the atom.

$$R_g = \sqrt{\frac{\sum_{i=1}^{N} m_i (\vec{r}_i - \vec{R}_{cm})^2}{\sum_{i=1}^{N} m_i}}$$

2.6.3 Comparing protein structures

Proteins have structures that loop in a similar way, with similar regions within each other. To quantify the similarity between the protein structure a procedure needs to be followed:

- Select common regions: a 1-1 correspondence between amino acid need to be found: the parts present only in one protein are not considered. A correspondence is built between the common regions on the single amino-acids. These can be different, usually the coordinates are confronted between the α-C atom and the residue is not considered. One of the things that can be done is to look at the secondary structures and add loops only when they look similar.
- Align the two structures: compute the centre of mass of the two proteins and translate the proteins so the centre of masses coincide.
- Finding the optimal rotation: (add algorithm) the principal axes are computed and the proteins are rotated so that they superimpose. Once the optimal rotation is obtained the difference can be quantified.
- Compute RMSD: root mean square deviation: take the coordinates of the amino-acid i in protein A and A, their squared difference is computed and an average over all amino acid is computed and squared:

$$RMSD = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\vec{r}_{Ai} - \vec{r}_{Bi})^2}$$

the RMSD is a length, usually in angstrom scale. This can be done for two proteins or for the protein taken at two different time step in a molecular dynamics simulation:

$$RMSD(t) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(\vec{r}_{i}(t) - \vec{r}_{i}(0))^2}$$

Now the $RMSD$ can be plotted with respect to time. It can be seen how at $t = 0$ $RMSD = 0$ and after the value will increase. When the number reaches a plateau the protein should be in equilibrium. The plateau can jump to another value, meaning that the state is a meta-stable state of the protein, or the protein has more stable states or a loop is making something. This value is assigned to very complicated structures and different structures can have the same $RMSD$. Clearly, a reference structure needs to be picked. In MD of a protein evolving in time, the reference structure is the initial configuration of the protein. One can also decide to focus on a specific region. The $RMSD$ is an indicator of equilibrium: it is a necessary but not sufficient condition.

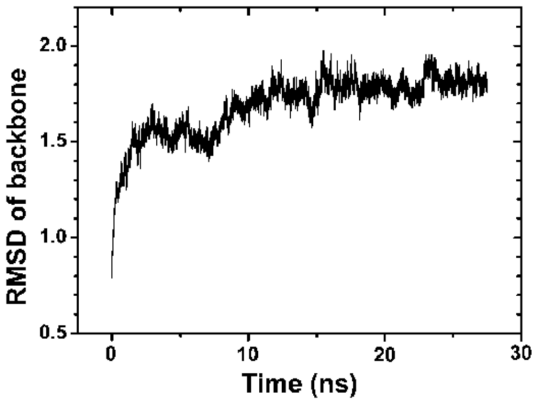

Figure 2.7: RMSD of the backbone (C^{α}, N, ...). We can see the typical steep jump a fluctuations towards the end of the plot. However, this is not enough to state that the protein reached equilibrium.

2.6.4 Native state

The native state is the functional state of a protein. It is not the state found for a crystallized protein, but only closely related to it, it is rather an ensemble in space and time of structures. This is due to proximity effect and the fact that the protein is not a static system. Proteins are extremely flexible and are moving a lot because the temperature corresponds to a constant movement of water molecule around causes movement in the protein. The native state is an ensemble of closely related states, all compatible with the conditions of the situation studied. Proteins need to be studied in

the isothermal-isobaric ensemble. All the calculation need to be done at constant temperature and pressure. The native state is so a collection of functional state.

In the case of the unfolded state the possibilities are too many to sample all of them.

2.6.5 RMSF

The flexibility of each amino acid can be computed. With flexibility is intended the movement of amino acid with respect to one another. In the α-helix, for example, less fluctuation is expected, while in loops more fluctuation is expected. This quantity is computed in the root mean squared fluctuation, which will be computed for each amino acid in the protein. With f referring to the frame, let:

- $\langle \vec{r}_i \rangle = \frac{1}{M} \sum_{f=1}^{M} \vec{r}_{i,f}$, the average position of atom i.
- $\Delta \vec{r}_{i,f} = \vec{r}_{i,f} - \langle \vec{r}_i \rangle$, the displacement of each atom in each frame with respect to its average.
- $\langle \Delta \vec{r}_o^{\,w} \rangle = \frac{1}{M} \sum_{f=1}^{M} (\vec{r}_{i,f} - \langle \vec{r}_i \rangle)^2$, the average squared distance over the frames.

So, the root mean squared fluctuation is:

$$RMSF_i = \sqrt{\langle \Delta \vec{r}_i \rangle^2}$$

Plotting the $RMSF$ with respect to residue number and the more mobile residue can be identified. This can be mapped onto the sequence so loops, helices and strands can be recognized. Usually the fluctuating part correspond to loops. The terminus have the highest $RMSF$.

2.6.5.1 B-factors

$RMSF$ can be translated into B-factors. They are the Debye-Waller factors and are a scaled version of the $RMSF$ squared. So the result of a simulation can be compared with the B-factor and a strong correspondence can be seen. The differences are due to the fact that the crystal is a different environment with respect to the normal one and packing effect can happen (some regions of the protein can interact with the image of the protein in the crystal).

$$B_i = \frac{8\pi^2}{3} \langle \Delta \vec{r}_i^{\,2} \rangle = \frac{8\pi^2}{3} RMSF_i^2$$

Chapter 3

Semi-empirical force fields

3.1 Introduction

A protein system can be modelled using a force field. Whenever a simulation is started the interactions between atoms need to be determined. To do so a topology file is built. This file will contain the connecting information. This is important because the kind of force field is specified here. Semi-empirical force field will be used because a very complicated process is modelled using formula that can be easily computed. These formulae are not rigorous and what it is done is to compare the result of the simulation with an experiment. Because of these the used force fields will be called semi-empirical.

3.2 Potential energy surface

Any system will correspond to a given potential energy surface. There are states which correspond to minima of this energy where the system will spend most of its time. In the global minimum the system is expected to be in equilibrium. There are also local minima where the system can spend some time and the system will go from the global to some other local minima and the time spent there depends on how deep that minimum is. There will be transitions from one minimum to another exploring the potential energy surface passing through a set of points. In principle the potential energy surface is known because it is assumed that there is a potential energy for all the interaction that there are in the system. This is not that simple because of temperature: at 0 temperature the potential energy surface will tell everything: the system goes through the global minimum and stay there. At finite and high temperatures the system will jump with a certain frequency, which depends on temperature, from one state to another. In this case statistical mechanics is necessary not the potential energy surface is necessary but also the free energy surface and the entropy have to be taken into account. The entropy part is the most difficult. In order to compute entropy the conformational space need to be explored: so the compatible conformation in certain condition need to be found. The phase space, the possibilities of a molecule need to be explored. These are so many in the case of a protein that estimating entropy will be the more costly process. In order to estimate entropy long simulation are needed as the conformational space will be explored as much as possible. In this chapter entropy will not be considered.

The potential energy surface or PES is the landscape of what values the potential energy of an atom can assume. Different points can be recognized like:

- Saddle point.
- Local maximums.
- Local minimum.

3.2.1 Bond stretching

In figure **??** the typical potential energy for a chemical bond can be seen. Let r_{eq} the distance for which the energy is minimal, where the equilibrium is. Moving from it the energy increases. There are transitions between different states and in that case quantum physics should be used. Everything will be assumed to be described using classical mechanics. This is valid when dealing with conformational transitions. SO all variables are continuous. Using this assumption light absorption and a chemical reactions cannot be described. The potential energy surface to describe an unbreakable bond can be seen in figure **??**. When close to the minimum the well can be assume symmetric, but some asymmetry later: the left part is steeper. An approximation for these potential energy is built into the force field. The potential energy surface can be reconstructed using vibrational spectroscopy and then it is reconstructed using a mathematical model. Using a Taylor expansion the potential energy is approximated at point r by taking the value at the equilibrium and then constructing all the corrections. The first one is the first derivative at the value at equilibrium multiplied by the distance. Then for the second correction the second derivative and so on. If the potential is symmetric the first derivative is 0, as it happens near the equilibrium. So the first correction can be neglected. Also the third order term is either 0 because the third derivative is 0 if the minimum is really shallow. This term is really small if r is really close to the equilibrium so it can neglected.

$$U(r) = U(r_{eq}) + \frac{dU}{dr}|_{r=r_{eq}}(r - r_{eq}) + \frac{1}{2!}\frac{d^2U}{dr^2}|_{r=r_{eq}}(r - r_{eq}^2) + \frac{1}{3!}\frac{d^3U}{dr^3}|_{r=r_{eq}}(r - r_{eq}^3) + \cdots$$

$$U(r) = U(r_{eq}) + \cancel{\frac{dU}{dr}|_{r=r_{eq}}(r - r_{eq})} + \frac{1}{2!}\frac{d^2U}{dr^2}|_{r=r_{eq}}(r - r_{eq}^2) + \cancel{\frac{1}{3!}\frac{d^3U}{dr^3}|_{r=r_{eq}}(r - r_{eq}^3)} + \cdots$$

So that in the end the typical harmonic potential is obtained.

$$U(r_{AB}) = \frac{1}{2}k_{AB}(r_{AB} - r_{AB,eq})^2$$

The constant k is related to the second derivative with respect to the distance. The distances and the k constants are labelled with A and B, which stand for the fact that this interaction has to be described for each couple of atom. Transferability of these parameter is an issue: using a particular force field then the parameters cannot be used for another. So each force field will come with its own set of parameters.

3.2.1.1 Anharmonic force constant

Considering that the shape is not completely symmetric the third order term can be inserted because it can be important. This introduces an asymmetry in the system: an anharmonic force constant.

$$U(r_{AB}) = \frac{1}{2}[k_{AB} + k_{AB}^{(3)}(r_{AB} - r_{AB,eq})](r_{AB} - r_{AB,eq})^2$$

3.2.1.2 Quartic correction

Also the fourth order term can be inserted.

$$U(r_{AB}) = \frac{1}{2}[k_{AB} + k_{AB}^{(3)}(r_{AB} - r_{AB,eq}) + k_{AB}^{(4)}(r_{AB} - r_{AB,eq})^2](r_{AB} - r_{AB,eq})^2$$

3.2.1.3 Morse potential

The Morse potential is used in implicit solvent simulation and uses the exponential because it can describe screen interaction that happens with implicit solvent. The exponential is quite expensive for a computer to compute. It is useful also for soft-system or coarse grained system.

$$U(r_{AB}) = D_{AB}[1 - e^{-\alpha_{AB}(r_AB - r_{AB,eq})^2}]$$

3.3 Valence angle bending

Chemical bonds are not described only by strings and beads but also bending of the bonds can happen. These are valence angle bending. These bonds cannot be broken during a classical molecular dynamic simulation. In order to describe them a Taylor expansion introducing and equilibrium angle is built. The first order term is not considered as it will be equal to 0. Then the second introduces the harmonic and the third for the anharmonic one, in principle also the quartic correction could be added. This formula is similar to the previous one but all the constant have to be described between each triplet of atoms.

$$U(\theta_{ABC}) = \frac{1}{2}[k_{ABC} + k_{ABC}^{(3)}(\theta_{ABC} - \theta_{ABC,eq}) + k_{ABC}^{(4)}(\theta_{ABC} - \theta_{ABC,eq})^2 + \cdots](\theta_{ABC} - \theta_{ABC,eq})^2$$

3.3.1 Multiple minima

There is another problem with valence angle bending: angles are not varying continuously: the Taylor expansion is good if the angles are not varying to much. If the angles invert the Taylor expansion cannot take track of it. In order to take this into account the potential energy is built using a Fourier expansion. A Fourier expansion introduces a periodic function like cos that contains oscillations in a period, allowing to model any possible periodic potential. Since the angles are being considered the potential is periodic. So in the parametrization of the angle bending interaction the Fourier term is introduced, labelling terms with the value j. So the amplitude, the value that multiplies each Fourier component is decreasing with j and a cut-off of the given value of j is given. In this case only low-frequency component are considered.

$$U(\theta_{ABC}) = \sum_{\{j\}_{ABC}} k_{j,ABC}^{fourier}[1 + \cos(j\theta_{ABC} + \psi_j)]$$

Where the amplitude:

$$k_{j,ABC}^{fourier} = \frac{2k_{ABC}^{harmonic}}{j^2}$$

Dividing by j^2 makes the contribution smaller for higher j.

3.4 Torsions

In order to describe torsion 4 atoms bonded by subsequent bonds are considered $A->B->C->D$. In this case two planes ABC and BCD are built and the angle ω between two plane is built. If they are on the same plane the angle is 0. So an angle ω is computed to describe the torsion that is around the bond $B->C$. In order to compute the angle the perpendicular vectors between the planes is computed and the angle between these two vector is ω. In this case the harmonic oscillation and the Taylor expansion are not used but the Fourier expansion and the periodic potential is used. This is because the value of the energy will be the same after a rotation. In principle torsions will be explored during a simulation. All the possible groups with permutations and repetitions of 4 atoms need to be considered to parametrize torsions.

$$U(\omega_{ABCD}) = \frac{1}{2} \sum_{\{j\}_{ABCD}} V_{j,ABCD}[1 + (-1)^{j+1} \cos(j\omega_{ABCD} + \psi_{j,ABCD})]$$

3.4.1 Improper torsions

Improper torsions happen when the bonds are $A->B->C$ and $B->D$. B can be in the same plane of A, C, D or it can pop out. In this case two planes are built, usually ABC and BCD and the angle between them will be built. The potential will describe the energy of the angle. The angle can be described as the out of plane angle OOP, obtaining the equation for the plane ACD and how much B is out of that plane. Here all possibilities of an atoms connected with other 3 atoms need to be considered. In the case of sp^2 the four atoms need to be in a plane.

$$U(\omega_{ABCD}) = \frac{1}{2} \sum_{\{j\}_{ABCD}} V_{j,ABCD}[1 + (-1)^{j+1} \cos(j\omega_{ABCD} + \psi_{j,ABCD})]$$

3.5 Van der Waals interactions

Van der Waals interactions happen whenever two atoms come close to one another withouth any chemical bond connecting them. This interaction is called dispersion interaction and depends on the interaction between the electron clouds that become correlated. Even if electrons are not shared the fluctuation of the electron clouds of an atom will interact with the one of an atom close to it. This can be computed and described by an attraction that goes with the 6th power of the distance. This interaction is the Van der Waals interaction. This attraction is an attractive force and will decrease with the 6th power of the distance. Also a repulsion happen. The repulsion is due to the occupancy of the possible orbitals. The problem is then Pauli's exclusion principle: whenever two atoms come close the energy level becomes occupied and the atoms cannot come to close to one another. This is difficult to compute but it can be modelled assuming that is a repulsion with an hard limit. The energy becomes very high when the distance becomes less than the equilibrium. If the attraction is modelled with the 6th power this interaction can be modelled by the 12th power. In this case if the distance is too small the energy becomes very high.

3.5.1 Lennard-Jones potential

The discussion done above lead to the introduction of the Lennard-Jones potential. This models Van der Waals interaction using the 6th power for the dispersion and the repulsion due to Pauli's

exclusion principle is modelled by the 12th power. This power is used to make the computation cheap and has no physical basis. In this way this formula is obtained:

$$U_(r_{AB}) = \frac{a_{AB}}{r_{AB}^{12}} - \frac{b_{AB}}{r_{AB}^{6}} = \\ = 4\epsilon_{AB}\left[\left(\frac{\sigma_{AB}}{r_{AB}}\right)^{12} - \left(\frac{\sigma_{AB}}{r_{AB}}\right)^{6}\right]$$

And the distance with minimum energy or equilibrium.

$$r_{AB}^{*} = 2^{\frac{1}{6}}\sigma_{AB}$$

And it can be seen how it is a multiple of σ_{AB} or the Van der Waals radius, the distance where the value of the energy if exactly equal to 0. $-\epsilon$ is the minimum value of the energy (equilibrium). In this way dispersion and repulsion of any couple of atoms can be described. A value of σ_{AB} have to be introduced for any couple of atom type.

3.5.2 Morse potential

The Morse potential and the Hill potential can be used with coarse grained system. Dealing only with all-atom simulation these will be rarely used. Working in soft-matter simulation these could be used. These have the same features of the Lennard-Jones potential.

$$U(r_{AB}) = D_{AB}[1 - e^{-a_{AB}(r_{AB}-r_{AB,eq})^2}]$$

3.5.3 Hill potential

$$U(r_{AB}) = \epsilon_{AB}\left[\frac{6}{\beta_{AB}-6}e^{\beta_{AB}\frac{1-r_{AB}}{r_{AB}^{*}}} - \frac{\beta_{AB}}{\beta_{AB}-6}\left(\frac{r_{AB}^{*}}{r_{AB}}\right)^{6}\right]$$

In some force fields 1-4 interactions (successive bonds atoms can be numbered) are reduced by a scaled factor. Atom 1 and 4 are cose to each other in a proper torsion but attraction and repulsion has been already described by the torsion, so in some force fields, the torsions takes into account the proximity of those atoms. So the Van der Waals interactions are excluded for 1-4 atoms as they are already been described by the torsion interaction.

3.6 Electrostatic interactions

Electrostatic interactions, together with Van der Waals one are two kind of non-bonded interactions. Moreover interactions between molecules that might be charged or that can be polar. The distribution of charges need to be described. In principle a multiple expansion should be performer: the cloud of positive and negative charges should be described by using given shapes of the clouds. And each molecule should be described and all the molecules should be described by these multiple expansion and they should be multiplied as matrix.

$$U_{AB} = \vec{M}^{(A)}V^{(B)}$$

Now, summing over all molecules:

$$U_{AB} = \sum_{A} \sum_{B>A} \vec{M}^{(A)} \vec{V}^{(B)}$$

This is a very costly procedure but accurate.

3.6.1 Point like charges

To make the computation easier molecules are represented as point-like partial charges. In a chemical molecule partial charges can be assumed that are point-like and are located on the atom. The electron cloud is displaced toward the negative cloud and away from the positive one. In this way the distribution of charges can be represented by number placed on each atom. Once this is done an electric charge is associated with each bead and the Coulomb interaction is used:

$$U_{AB} = \frac{q_A q_B}{\epsilon_{AB} r_{AB}}$$

The dielectric constant is 1 in the case of explicit solvent, but in the case of the implicit one it will be the one of the solvent.

3.6.2 Dipolar interactions

Bipolar interaction can be included when considering electrostatic forces. These formulae are mostly used in corse grained models. Because in that case the approximation of single charges on a bead cannot be made and dipoles are considered. In this case a functional form and the parameters need to be included. μ is the dipole of two molecules. Then there is the dielectric constant, the cube of the distance and the cosine term represent the orientation of the two dipoles.

$$U_{AB/CD} = \frac{\mu_{AB}\mu_{CD}}{\epsilon_{AB/CD} r^3_{AB/CD}} (\cos \chi_{AB/CD} - 3 \cos \alpha_{AB} \cos \alpha_{CD})$$

3.6.3 Dielectric constants

$$U_{AB} = \frac{q_A q_B}{\epsilon_{AB} r_{AB}}$$

The dielectric constant can assume different values. If they are connected by a chemical bond or if there is a valence angle between the atoms the dielectric constant if ∞ and the interaction is not computed as it is already taken into account by other terms. In several force field a factor of 2 is added for 1-4 interaction. In the last case the value depend on the force field.

$$\epsilon_{AB} = \begin{cases} \infty & \text{if } A \wedge B \text{are 1,2- or 1,3-related} \\ 3.0 & \text{if } A \wedge B \text{are 1,4-related} \\ 1.5 & \text{otherwise} \end{cases}$$

3.7 Cross terms

There are also cross terms: angle bending and bond stretching and other movement are not independent: all the possible degrees of freedom should be considered together. This is done only by some force fields.

$$
\begin{aligned}
U(\vec{q}) =& U(\vec{q}_{eq}) + \sum_{i=1}^{3N-6} (q_i - q_{i,eq}) \frac{\partial U}{\partial q_i}|_{\vec{q}=\vec{q}_{eq}} + \\
&+ \frac{1}{2!} \sum_{i=1}^{3N-6} \sum_{j=1}^{3N-6} (q_i - q_{i.eq})(q_j - q_{j,eq}) \frac{\partial^2 U}{\partial q_i \partial q_j}|_{\vec{q}=\vec{q}_{eq}} + \\
&= \frac{1}{3!} \sum_{i=1}^{3N-6} \sum_{j=1}^{3N-6} \sum_{k=1}^{3N-6} (q_i - q_{i,eq})(q_j - q_{j.eq})(q_k - q_{k,eq}) \frac{\partial^3 U}{\partial q_i \partial q_j \partial q_k}|_{\vec{q}=\vec{q}_{eq}} + \cdots
\end{aligned}
$$

This is done by starting from a Lagrangian describing all the interaction and compute all the terms. In practice everything is done without considering them.

$$U(r_{AB}, \theta_{ABC}) = \frac{1}{2} k_{AB,ACB}(r_{AB} - r_{AB,eq})(\theta_{ABC} - \theta_{ABC,eq})$$

3.8 Parametrization

All the interactions need a large set of parameters. The number of parameters increases with the number of atoms N:

$$p = N + (N-1) + (N-2) + \cdots = N\frac{N+1}{2}$$

In order to obtain these parameters data from experiments are used. This is the reason why the force-field are semi-empirical. To do so the results of the experiments are compared from the results of simulations. Then the penalty function:

$$Z = \left[\sum_{i}^{observables} \sum_{j}^{occurrences} \frac{(calc_{i,j} - expt_{i,j})^2}{w_i^2} \right]^{\frac{1}{2}}$$

Sums over all the data from all the observables and all possibilities. Then all the observable are measured and are compared with the simulation result. Then the square is taken to not compensate the error and it is divided by a weight. This penalty function is minimized changing the parameters until some reasonable compromise is obtained. The more experimental data is obtained the more accurate the description. This is the reason for the constant update of the force fields. A way to reduce the number of parameters is to employ a strategy so that a parameter to each atom is computed and then it is computed the interaction between two atoms:

$$\sigma_{AB} = \sigma_A + \sigma_B \epsilon_{AB} \qquad\qquad = (\epsilon_A \epsilon_B)^{\frac{1}{2}}$$

3.9 Force field energies

In molecules the ground state of one molecule is different from the ground state of another. Molecules can have different states and when they come in contact the ground state have to be taken into account. So parametrizing the force field is very complicated.

3.9.1 Geometry optimization

Once the force fields are obtained some atoms can be missing from the model and have to be included by hand. Sometimes it may happen that two atoms are too close to each other and by using the potential energy it can be high leading to high forces, causing the atom that experience them to be kicked out. So the simulation will explode. In order to avoid this a geometry optimization is done to obtain the parameters or to minimize so to avoid high forces values at the start of a simulation. When starting a simulation the system will be at an energy value which be greater than the equilibrium. The gradient of the energy is computed and the system is moved toward the minimum energy. The derivative need to be computed of U with respect to each component in the system.

$$\vec{g}(\vec{q}) = \begin{bmatrix} \frac{\partial U}{\partial q_1} \\ \frac{\partial U}{\partial q_2} \\ \vdots \\ \frac{\partial U}{\partial q_n} \end{bmatrix}$$

Such that the cost reaches a global minimum $J_{min}(\vec{w})$.

3.9.2 Derivative of the potential function

First the derivatives need to be computed. All the potentials are given in term of the mutual distance of two atoms. So, considering x the coordinates of a point:

$$\frac{\partial U}{\partial x_A} = \sum_{i \in A} \frac{\partial U}{\partial r_{Ai}} \frac{\partial r_{Ai}}{\partial x_A}$$

Taking the derivative of this:

$$U(r_{AB}) = \frac{1}{2}[k_{AB} + k_{AB}^{(3)}(r_{AB} - r_{AB,eq}) + k_{AB}^{(4)}(r_{AB} - r_{AB,eq})^2](r_{AB} - r_{AB,eq})^2$$

With the respect on the distance:

$$\frac{\partial U}{\partial r_{Ai}} = \frac{1}{2}[2k_{Ai} + 3k_{Ai}^{(3)} r_{Ai} - r_{Ai,eq}) + 4k_{Ai}^{(4)}(r_{Ai} - r_{Ai,eq})^2](r_{Ai} - r_{Ai,eq})$$

In order to take the derivative with respect to r with respect to x need to be known:

$$\frac{\partial r_{Ai}}{\partial x_A} = \frac{x_A - x_i}{\sqrt{(x_A - x_i)^2 + (y_A - y_i)^2 + (z_A - z_i)^2}}$$

Then this formula can plugged in according to the chain rule.

3.9.2.1 Newton-Raphson

To perform the minimization an iterative procedure, the Newton-Raphson to find a local minimum is employed. Here (n) refers to the iteration. This allow to obtain iteration $k+1$ starting from iteration k and is a Taylor expansion of coordinates at iteration k:

$$U(\vec{q}^{(k+1)}) = U(\vec{q}^{(k)}) + (\vec{q}^{(k+1)} - \vec{q}^{(k)})\vec{g}^{(k)} + \frac{1}{2}(\vec{q}^{(k+1)} - \vec{q}^{(k)})H^{(k)}(\vec{q}^{(k+1)} - \vec{q}^{(k)})$$

This is done to find eventually the coordinates at iteration $k+1$ which are unknown. Where H is the Hessian matrix built:

$$H_{ij}^{(k)} = \frac{\partial^2 U}{\partial q_i \partial q_j}|_{\vec{q}=\vec{q}^{(k)}}$$

Everything is computed from a Taylor expansion from coordinates at iteration k. And:

$$\frac{\partial U(\vec{q}^{(k+1)})}{\partial q_i^{(k+1)}} = \frac{\partial \vec{q}^{(k+1)}}{\partial q_i^{(k+1)}}\vec{g}^{(k)} + \frac{1}{2}\frac{\partial \vec{q}^{(k+1)}}{\partial q_i^{(k+1)}}H^{(k)}(\vec{q}^{(k+1)} - \vec{q}^{(k)}) + \frac{1}{2}(\vec{q}^{(k+1)} - \vec{q}^{(k)})H^{(k)}\frac{\partial \vec{q}^{(k+1)}}{\partial q_i^{(k+1)}}$$

This formula tells that:

$$\vec{g}_i^{(k+1)} = \vec{g}_i^{(k)} + [H^{(k)}(\vec{q}^{(k+1)} - \vec{q}^{(k)})]_i$$

Where at the end it should be obtained a stationary condition, where the system does not change after an iteration:

$$\vec{0} = \vec{g}^{(k)} + H^{(k)}(\vec{q}^{(k+1)} - \vec{q}^{(k)}) \Rightarrow \vec{q}^{(k+1)} = \vec{q}^{(k)} - [H^{(k)}]^{-1}\vec{g}^{(k)}$$

Because the starting point was a Taylor expansion the new coordinates will not be exact, and so other iterations will be needed. This process is repeated until convergence is reached, choosing a given tolerance on $\vec{0}$.

3.9.3 Types of force fields

Force fields can be categorized as:

- All atoms: one atom corresponds to one bead.
- More atoms: more atoms correspond to one bead.
- Corse grained: groups of atoms correspond to one bead.
- Polarizable force fields: point charges are variables.

As a golden rule parameters from different force fields should never be mixed.

Chapter 4

Classical mechanics

4.1 Newton's laws

So the mass times the acceleration is equal to the force. A dot on top of a vector is a derivative with respect to the time.

$$m_i \frac{d^2 \vec{r}_i}{dt^2} = m_i \ddot{\vec{r}}_i = \vec{F}_i$$

The index i in a simulation represent a particle, usually atoms.

4.1.1 Forces

In all the force fields the forces come always from pairs of atoms: all the forces depend on the distance between atom. This is not true for angles and torsions, but they can be re-conducted to forces that involve two atoms, respecting Newton's third law. The forces acting on each particle depend on the coordinates of all the atoms and they might depend on the velocity of atom i. Sometimes some frictional forces depend on the velocity. All the forces can be represented as forces between pairs of atom that depend on their mutual distance, in particular $\vec{f}_{ij} = -\vec{f}_{ji}$. All the external forces like friction or some fields depend on the position of atom i and its velocity. In the case of biological system viscosity and temperature are the major external forces.

$$\vec{F}_i(\vec{r}_1, \ldots, \vec{r}_N, \dot{\vec{r}}_i) = \sum_{j \neq i} \vec{f}_{ij}(\vec{r}_i - \vec{r}_j) + \vec{f}^{(ext)}(\vec{r}_i, \dot{\vec{r}}_i)$$

The forces are computed starting from the interactions:

- Bond stretching: $U = \frac{k_l}{2}(l - l^0)^2$.
- Bond bending: $U = \frac{k_\theta}{2}(\theta - \theta^0)^2$.
- Bond torsion: $U = k_\phi[1 + \cos(n\phi - \phi^0)]$.
- Van der Waals interactions: $U = \left[\frac{a_{ij}}{r_{ij}^{12}} - \frac{b_{ij}}{r_{ij}^{6}}\right]$.
- Electrostatic interactions: $U = \frac{332 q_i q_j}{\epsilon r_{ij}}$. Where factor 332 is necessary to compute in in $\frac{kcal}{mol}$.

4.1.2 Phase space

When dealing with a system with N atoms, each with 3 dimensions their position and momenta need to be considered:

$$\vec{p}_i = m_i \vec{v}_i = m_i \dot{\vec{r}}_i$$

Newton's law can be re-written in term of particle momenta.

$$\vec{F}_i = m_i \ddot{\vec{r}}_i = \dot{\vec{p}}_i$$

Then the full dynamics of a system in 3 dimension is specified by $6N$ functions, where N is the numbers of the body in the system:

$$\{\vec{r}_1(t), \ldots, \vec{r}_N(t), \vec{p}_1(t), \ldots, \vec{p}_N(t)\}$$

In this way at each instant of time each position and momenta of each particle is specified. Each microscopic state at time t is completely specified by $6N$ numbers, as each position and momenta need to be described. This creates the phase space vector which is $6N$ dimensional:

$$\vec{x} + \{\vec{r}_1, \ldots, \vec{r}_N, \vec{p}_1, \ldots, \vec{p}_N\}$$

A state of a system is a point in this $6N$ dimensional space. The trajectory represent a curve in the space:

$$\vec{x}_t = \{\vec{r}_1(t), \ldots, \vec{r}_N(t), \vec{p}_1(t), \ldots, \vec{p}_N(t)\}$$

4.1.3 One particle in one dimension

Considering one particle in one dimension the phase space is 2 dimensional. On the x axis there are the coordinates and on the y axis the momentum. In the case of motion at constant velocity the trajectory will be an horizontal line. This is the case of a free particle.

4.1.3.1 Harmonic oscillator

In the case of an harmonic oscillator, it will be described by equation:

$$m\ddot{x} = -kx$$

Where the acceleration is equal to minus the elastic force. And:

$$\omega = \sqrt{\frac{k}{m}}$$

Then the solution for the trajectory:

$$x(t) = x(0)\cos\omega t + \frac{p(0)}{m\omega}\sin\omega t$$

That depends on the initial velocity and initial position. The momentum as a function of time and the coordinate verify the equation:

$$\frac{p^2(t)}{2m} + \frac{1}{2}m\omega^2 x^2(t) = C$$

This equation represent an ellipse and is the conservation of energy. The constant C will have an effect on the length of the two axes on the ellipse, which will be: $(2mC)^{frac12}$ and $(2\frac{C}{m}\omega^2)^{\frac{1}{2}}$.

4.1.3.2 Hill potential

The Hill potential is important in chemistry. A bead may approach the hill of potential from the left with positive momentum or from the right with negative momentum. If it is slow the kinetic energy will be not sufficient to overcome the hill, so the particle will stop and go back. For a particle coming from the right the momentum is negative and if it is slow it will not overcome the hill and go back. If the velocity is enough to reach the top of the hill the particle will stay forever on top of it. If the velocity is more than that the bead will traverse the hill and fall on the other side. A trajectory in the phase space can be different with the one considering just the position.

4.2 Lagrangian formulation

Classical mechanics can be formulated using the Lagrangian formalism. In order to have the Lagrangian formulation only conservative forces are considered, from which potential forces are obtained:

$$\vec{F_i}(\vec{r}_1, \ldots, \vec{r}_N) = -\nabla_i U(\vec{r}_1, \ldots, r_N)$$

Where ∇_i is the gradient of the potential and is the force. For conservative forces the work done is computed as the difference of the potential energy and the path can be excluded:

$$W_{AB} = \int_A^B \vec{F_i} d\vec{l} = U_A - U_B = -\Delta U_{AB}$$

And on closed pathways, according to the previous formula:

$$\oint \vec{F_i} d\vec{l} = 0$$

To define the Lagrangian, one term is the kinetic energy, which in the case for N particles, the total kinetic energy of the system is:

$$K(\dot{\vec{r}}_1, \ldots, \dot{\vec{r}}_N) = \frac{1}{2} \sum_i m_i \dot{\vec{r}}_i^2$$

Which depends on the mass and their velocity. And the Lagrangian is defined as the difference of the potential and kinetic energy:

$$\mathcal{L}(\vec{r}_1, \ldots, \vec{r}_N, \dot{\vec{r}}_1, \ldots, \dot{\vec{r}}_N) = K(\dot{\vec{r}}_1, \ldots, \dot{\vec{r}}_N) - U(\vec{r}_1, \ldots, \vec{r}_N)$$

It can be seen how the Lagrangian depends on all the coordinates and all the velocities. Here momenta is not included, but only velocity and position. It can be seen how the kinetic energy depends only on the velocity and the potential one only from the position.

4.2.1 Euler-Lagrange equations

Once the Lagrangian is obtained the Euler-Lagrange equation can be computed. Where the difference of the derivative with respect to time of the derivative of the Lagrangian with respect to the velocity and the derivative of the Lagrangian with respect to the position is equal to 0:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_i}\right) - \frac{\partial \mathcal{L}}{\partial r_i} = 0$$

Now, as an example consider:

$$\mathcal{L} = \underbrace{\frac{1}{2}\sum_i m_i \dot{\vec{r}}_i^2}_{\text{Kinetic energy}} - U(\vec{r}_1, \ldots, \vec{r}_N)$$

And computing the equation of the Euler-Lagrange equation:

$$\frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_i} = m_i \dot{\vec{r}}_i$$

$$\frac{\partial \mathcal{L}}{\partial \vec{r}_i} = -\frac{\partial U}{\partial \vec{r}_i}$$

So that, putting everything together:

$$m_i \ddot{\vec{r}}_i + \underbrace{\frac{\partial U}{\partial \vec{r}_i}}_{=\vec{F}_i} = 0 \Rightarrow \overbrace{m_i \ddot{\vec{r}}_i = \vec{F}_i}^{\text{Newton's equation}}$$

4.2.1.1 Harmonic oscillator

Considering for example the harmonic oscillator the Lagrangian will be:

$$\mathcal{L}(x, \dot{x}) = \underbrace{\frac{1}{2} m \dot{x}^2}_{K} - \underbrace{\frac{1}{2} k x^2}_{U}$$

And, taking all the derivatives:

$$\frac{d}{dt}(m\dot{x}) + kx = 0 \Rightarrow m\ddot{x} - kx$$

The equation for the harmonic oscillator is obtained.

4.2.2 Conservation of energy

The Lagrangian equation can be applied even to very complex systems. Another issue involves the energy. Let the equation of the energy be:

$$E = \underbrace{\frac{1}{2}\sum_i m_i \dot{\vec{r}}_i^2}_{K} + U(\vec{r}_i, \ldots, \vec{r}_N)$$

The sum of the potential and kinetic energy.
Then, taking the derivative of energy with respect to time:

$$\begin{aligned}\frac{dE}{dt} &= \sum_i m_i \dot{\vec{r}}_i \ddot{\vec{r}}_i + \sum_i \frac{\partial U}{\partial \vec{r}_i} \dot{\vec{r}}_i = \\ &= \sum_i \dot{\vec{r}}_i \left[m_i \ddot{\vec{r}}_i + \frac{\partial U}{\partial \vec{r}_i} \right] = \\ &= \sum_i \dot{\vec{r}}_i [m_i v\ddot{e}cr_i - \vec{F}_i] = 0\end{aligned}$$

So that the energy is conserved in time. Following Lagrangian mechanics then the energy will be conserved. If energy is not conserved serious mistake are done in a molecular simulation.

4.2.3 Generalized coordinates

Introducing generalized coordinates, for example when studying small molecules in chemistry:

$$q_\alpha = f_\alpha(\vec{r_1}, \ldots, \vec{r}_N) \qquad \alpha = 1, \ldots, 3N$$

These are functions of the coordinates and are $3N$ generalized coordinates that take the place of the Cartesian coordinates. To go back from generalized to Cartesian coordinates:

$$\vec{r}_i = \vec{g}_i(q_1, d \ldots, q_{3N}) \qquad i = 1, \ldots, N$$

Trying to write the Lagrangian in generalized coordinates the kinetic part has to be recomputed. The velocity of particle i with respect to the new generalized coordinates. So the derivative of $\vec{r}_i$ with respect to q_α multiplied by the derivative in time of q_α. This is summed over all the generalized coordinates.

$$\dot{\vec{r}}_i = \sum_{\alpha=1}^{3N} \frac{\partial \vec{r}_i}{\partial q_\alpha} \dot{q}_\alpha$$

Writing the kinetic energy as a function of the generalized coordinates the velocity has to be included twice.

$$\tilde{K}(q, \dot{q}) = \frac{1}{2} \sum_{\alpha=1}^{3N} \sum \beta = 1^{3N} \left[\sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_\alpha} \frac{\partial \vec{r}_i}{\partial q_\beta} \right] \dot{q}_\alpha \dot{q}_\beta$$

And introducing the metric mass tensor $G_{\alpha\beta} = \left[\sum_{i=1}^{N} m_i \frac{\partial \vec{r}_i}{\partial q_\alpha} \frac{\partial \vec{r}_i}{\partial q_\beta} \right]$:

$$\tilde{K}(q, \dot{q}) = \frac{1}{2} \sum_{\alpha=1}^{3N} \sum \beta = 1^{3N} G_{\alpha\beta} \dot{q}_\alpha \dot{q}_\beta$$

Then the Lagrangian in generalized coordinates becomes:

$$\mathcal{L}(q, \dot{q}) = \frac{1}{2} \sum_{\alpha=1}^{3N} \sum \beta = 1^{3N} G_{\alpha\beta} \dot{q}_\alpha \dot{q}_\beta - U(q_1, \ldots, q_{3N})$$

And the Euler-Lagrange equations:

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha}\right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$$

In a molecular simulations the objective is to solve the Lagrangian's equation in the Hamiltonian form.

4.2.4 Legendre transforms

Legendre transforms will be used to go from a Lagrangian to an Hamiltonian description and are also useful in thermodynamics. Thermodynamics potentials are all Legendre transforms of one another. To go from one ensemble to another a Legendre transform is used. The aim of a Legendre transform is to express the function $f(x)$ in terms of its first derivative s:

$$s = f'(x) \equiv g(x)$$

The derivative of a function is the slope of the function in each point. In principle from the slope the function at each point it is not possible to reconstruct the original function. The reason for that is that translation information is lost when taking the derivative. To be able to perform this operation the intersection of the tangent at each point with the x axis in the curve the full function can be reconstructed. So at each point the slope and where the tangent intersect the x axis is needed to reconstruct the original function:

$$f(x_0) = f'(x_0)x_o + b(x_0)$$

Where $b(x_0)$ is the intersection of the tangent with the x axis. Considering the general form:

$$f(x) = f'(x)x + b(x)$$

b represent the functions in term of s. There is a need to isolate the b contribution and then write everything in term of s. So there is a need to invert g to obtain x as a function of s.

$$f'(x) = g(x) = s \Rightarrow x = g^{-1}(s)$$

Hence $b(x)$ contains the same information as $f(x)$:

$$b(g^{-1}(s)) = f(g^{-1}(s)) - sg^{-1}(s) \equiv \tilde{f}(s)$$

The Legendre transom is then:

$$\tilde{f}(s) = f(x(s)) - sx(s)$$

4.2.4.1 Legendre transform for multiple variables

Considering n variables:

$$s_1 = \frac{\partial f}{\partial x_1} = g_1(x_1, \ldots, x_n), \ldots, s_n = \frac{\partial f}{\partial x_n} = g_1(x_1, \ldots, x_n)$$

So the Legendre transform:

$$\tilde{f}(s_1,\dots,s_n) = f(x_1(s_1,\dots,s_n),\dots,x_n(s_1,\dots,s_n)) - \sum_i s_i x_i(s_1,\dots,s_n)$$

This holds for a subset of variables. This is important as most of the thermodynamics quantities can be expressed as derivatives of some other quantity, so everything is expressed in term of another variable.

4.3 Hamiltonian formulation

In the Hamiltonian formulation the conjugate momentum is obtained from the Lagrangian: the derivative of the Lagrangian with respect to velocity is the momentum.

$$\vec{p}_i \equiv \frac{\partial \mathcal{L}}{\partial \dot{\vec{r}}_i} = \frac{\partial}{\partial \dot{\vec{r}}_i}\left[\frac{1}{2}\sum_{j=1}^{N} m_j \dot{\vec{r}}_j^2 - U(\vec{r}_1,\dots,\vec{r}_N)\right] = m_i \dot{\vec{r}}_i$$

To express everything as a function of coordinate and momenta a Legendre transform of the Lagrangian needs to be performed, observing that the momentum is the derivative of the Lagrangian with respect to velocity.

$$\begin{aligned}\tilde{\mathcal{L}}(\vec{r}_1,\dots,\vec{r}_N,\vec{p}_1,\dots,\vec{p}_N) &= \mathcal{L}(\vec{r}_1,\dots,\vec{r}_N,\dot{\vec{r}}_1(\vec{p}_1),\dots,\dot{\vec{r}}_N(\vec{p}_N)) - \sum_i \vec{p}_i \dot{\vec{r}}_i(\vec{p}_i) = \\ &= \underbrace{\frac{1}{2}\sum_{i=1}^{N} m_i \left(\frac{\vec{p}_i}{m_i}\right)^2 - U(\vec{r}_1,\dots,\vec{r}_N)}_{\text{Lagrangian}} - \sum_{i=1}^{N} \vec{p}_i \frac{\vec{p}_i}{m_i} = \\ &= -\underbrace{\sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m_i}}_{\text{Kinetic energy}} - U(\vec{r}_1,\dots,\vec{r}_N)\end{aligned}$$

It can be seen how the Legendre transform of the Lagrangian is the opposite of the total energy of the system. The Hamiltonian is then defined as the opposite of the Legendre transform of the Lagrangian:

$$\mathcal{H}(\vec{r}_1,\dots,\vec{r}_N,\vec{p}_1,\dots,\vec{p}_N) = -\tilde{\mathcal{L}}(\vec{r}_1,\dots,\vec{r}_N,\vec{p}_1,\dots,\vec{p}_N)$$

So that it is the sum of the kinetic and potential energy:

$$\mathcal{H}(\vec{r}_1,\dots,\vec{r}_N,\vec{p}_1,\dots,\vec{p}_N) = \sum_i \vec{p}_i \dot{\vec{r}}_i(\vec{p}_i) - \mathcal{L}(\vec{r}_1,\dots,\vec{r}_N,\dot{\vec{r}}(\vec{p}_1),\dots,\dot{\vec{r}}_N(\vec{p}_N))$$

Or the total energy of the system:

$$\mathcal{H}(\vec{r}_1,\dots,\vec{r}_N,\vec{p}_1,\dots,\vec{p}_N) = \sum_{i=1}^{N} \frac{\vec{p}_i^2}{2m_i} + U(\vec{r}_1,\dots,\vec{r}_N)$$

The Hamiltonian is related to the Legendre transform of the Lagrangian. Using the Hamiltonian description, it is expected that the Hamiltonian formulation has the same results from the Lagrangian formulation. Another equation needs to be added.

4.3.1 Generalized coordinates

The same thing can be done to generalized coordinates:

$$p_\alpha = \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} = \sum_\beta G_{\alpha\beta}\dot{q}_\beta \Rightarrow \dot{q}_\alpha = \sum_\beta G^{-1}_{\alpha\beta} p_\beta$$

So that the Hamiltonian becomes:

$$\begin{aligned}\mathcal{H}(q_1,\dots,q_{3N},p_1,\dots,p_{3N}) &= \sum_\alpha p_\alpha \dot{q}_\alpha - \mathcal{L}(q_1,\dots,q_{3N},\dot{q}_1,\dots,\dot{q}_{3N}) = \\ &= \frac{1}{2}\sum_\alpha \sum_\beta p_\alpha G^{-1}_{\alpha\beta} p_\beta + U(q_1,\dots,q_{3N})\end{aligned}$$

4.3.2 Hamilton's equations

Now two sets of equations are obtained, but both are first order derivative in time, so that they are easier to solve both analytically and numerically with respect to the Euler-Lagrangian equation. Because of this the simulations will use Hamilton's equation.

$$\dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} \qquad \dot{p}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha}$$

Taking the derivative of the Hamiltonian with respect to time:

$$\begin{aligned}\frac{d\mathcal{H}}{dt} &= \sum_\alpha \left[\frac{\partial \mathcal{H}}{\partial q_\alpha}\dot{q}\alpha + \frac{\partial \mathcal{H}}{\partial p_\alpha}\dot{p}_\alpha\right] = \\ &= \sum_\alpha \left[\frac{\partial \mathcal{H}}{\partial q_\alpha}\frac{\partial \mathcal{H}}{\partial p_\alpha} - \frac{\partial \mathcal{H}}{\partial p_\alpha}\frac{\partial \mathcal{H}}{\partial q_\alpha}\right] = \\ &= \sum_\alpha \left[\cancel{\frac{\partial \mathcal{H}}{\partial q_\alpha}\frac{\partial \mathcal{H}}{\partial p_\alpha}} - \cancel{\frac{\partial \mathcal{H}}{\partial p_\alpha}\frac{\partial \mathcal{H}}{\partial q_\alpha}}\right] = 0\end{aligned}$$

Because the Hamiltonian is the total energy of the system it should be conserved as it describe the same system as the Lagrangian. In conclusion Hamiltonian mechanics keeps energy constant:

$$\mathcal{H}(q_1,\dots,q_{3N},p_1,\dots,p_{3N}) = const$$

So energy conservation should hold for any simulation written for Hamiltonian mechanics.

4.3.3 Conservation laws

This process can be generalized for any conserved quantity. Let $a(x_t)$ a property that depend on the trajectory of the system. In order to check if the property is conserved or not the derivative with respect to time have to be taken:

$$\frac{da}{dt} = \frac{\partial a}{\partial x_t}\dot{x}_t = \sum_\alpha \left[\frac{\partial a}{\partial q_\alpha}\dot{q}_\alpha + \frac{\partial a}{\partial p_\alpha}\dot{p}_\alpha\right] = \sum_\alpha \left[\frac{\partial a}{\partial q_\alpha}\frac{\partial \mathcal{H}}{\partial p_\alpha} - \frac{\partial a}{\partial p_\alpha}\frac{\partial \mathcal{H}}{\partial q_\alpha}\right] = \{a,\mathcal{H}\}$$

Where the Poisson brackets of two quantities depending on q and p:

$$\{a, b\} = \sum_{\alpha} \left[\frac{\partial a}{\partial q_\alpha} \frac{\partial b}{\partial p_\alpha} - \frac{\partial a}{\partial p_\alpha} \frac{\partial b}{\partial q_\alpha} \right]$$

To check if a quantity is conserved its Poisson bracket with the Hamiltonian need to be considered. If it is 0, then the quantity is conserved:

$$\{a, \mathcal{H}\} = 0 \Rightarrow \frac{da}{dt} = 0$$

So conservation laws are directly translated into this property of the Poisson brackets.

4.3.4 Compressibility

Compressibility of an equation is the divergence of the velocities. Define:

$$\dot{x} = \eta(x) = (\dot{q}_1, \dots, \dot{q}_{3N}, \dot{p}_1, \dots, \dot{p}_{3N})$$

The time derivative of all points in the phase space. This vector η because of Hamilton equation will be equal to:

$$\eta(x) = \left(\frac{\partial \mathcal{H}}{\partial p_1}, \dots, \frac{\partial \mathcal{H}}{\partial p_{3N}}, -\frac{\partial \mathcal{H}}{\partial q_1}, \dots, -\frac{\partial \mathcal{H}}{\partial q_{3N}} \right)$$

The divergence of this quantity is computed, taking the derivative of each component with respect to the coordinate itself.

$$\begin{aligned}
\nabla_x \dot{x} &= \sum_{\alpha} \left[\frac{\partial \dot{p}_\alpha}{\partial p_\alpha} + \frac{\partial \dot{q}_\alpha}{\partial q_\alpha} \right] = \\
&= \sum_{\alpha} \left[-\frac{\partial}{\partial p_\alpha} \frac{\partial \mathcal{H}}{\partial q_\alpha} + \frac{\partial}{\partial q_\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} \right] = \\
&= \sum_{\alpha} \left[-\frac{\partial^2 \mathcal{H}}{\partial p_\alpha \partial q_\alpha} + \frac{\partial^2 \mathcal{H}}{\partial q_\alpha \partial p_\alpha} \right] = \\
&= \sum_{\alpha} \left[\cancel{-\frac{\partial^2 \mathcal{H}}{\partial p_\alpha \partial q_\alpha}} + \cancel{\frac{\partial^2 \mathcal{H}}{\partial q_\alpha \partial p_\alpha}} \right] = 0
\end{aligned}$$

So, in conclusion $\nabla_x \dot{x} = 0$ and the divergence of the velocity is 0, then Hamilton's equations are incompressible: they lead to an incompressible flow of points in phase-space. This is a way to check whether a system is Hamiltonian.

4.3.5 Symplectic structure

Hamilton's equations have a symplectic structure and can be written as:

$$\dot{x} = M \frac{\partial \mathcal{H}}{\partial x} \qquad M = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix}$$

A trajectory in phase space $x_t = x_t(x_0)$ is a function of the initial positions and momenta or velocity. However it can be viewed also as a transformation of variables. This is done through the Jacobian:

$$J_{kl} = \frac{\partial x_t^k}{\partial x_o^l}$$

Applying the Jacobian transformation to the symplectic matrix M the original matrix is obtained: $M = J^T M J$. This is called the symplectic property and Hamilton's equations satisfy it.

Chapter 5

Theoretical foundations of statistical mechanics

5.1 Introduction

5.1.1 Loschmidt's paradox

When studying Newton's equation in any formulation, inverting the arrow of time, it is always possible to trace back any trajectory and to obtain the solutions. This is intrinsic in the Newton's equations. In reality this is not the case: some of the processes cannot be traced back and are irreversible. This is Loschmidts' paradox: the problem is that macroscopic systems are composed of a too large number of particles. And although in principle, the trajectory of each of them can be known, in the macroscopic process things are not obvious and a reasoning in term of probability have to be done. This is the objective of statistical mechanics: try to describe system with a very large number of degrees of freedom.

5.2 Thermodynamics

5.2.1 Thermodynamic system

A thermodynamic system is any macroscopic system. In general a macroscopic system has a number of particles similar to the Avogadro number. This is not the case for molecular simulations, but the system is close to the thermodynamic limits but it is not a macroscopic system. The system is not really a thermodynamic system.

5.2.1.1 Isoloted

A system is defined as isolated whenever no energy or mass transfer can be seen across its boundary. Considering the system's boundaries in molecular simulations periodic boundaries conditions are used and the reason that they are used is because boundaries effects have to be avoided. The boundaries will be more important the smaller the system is.

5.2.2 Thermodynamic equilibrium

A thermodynamic system is in thermodynamic equilibrium if its thermodynamic state does not change in time. When dealing with biological system the equilibrium is never reached, but only stationary or steady-states. This means that the system is in a local equilibrium and will remain there most of the time.

5.2.3 Thermodynamic state

A thermodynamic state is specified in terms of macroscopic parameters that are measurable quantities like:

- Pressure P.
- Volume V.
- Temperature T.
- Total mass M.
- Number of particles N.

This is important because some procedures are needed to measure starting from the coordinates and the velocity of the system these quantities in a simulation. These are the parameters that are compared with the experimental results.

5.2.4 Equation of state

Whenever a thermodynamic system is being described an equation can be written and that equations establishes a relations between the measurables:

$$g(N, P, V, T) = 0$$

As an example the equation of state for the ideal gas is:

$$PV - nRT = 0$$

The equation of state deals with thermodynamics state, which are assumed to be at equilibrium. SO the equation of state describes only equilibrium states.

5.2.5 Thermodynamic transformations

A thermodynamic transformation is a change of thermodynamic state. At equilibrium the only way to change it is to change the external conditions, like changing pressure. These transformations can be:

- Reversible: following equilibrium states and it is always possible to have the reverse transformation.
- Irreversible: non-equilibrium states are reached and when going back a different path is created. It is possible to reverse the process but the transformation is different from the forward one.

5.2.6 State function

A state function is defined as:

$$f(n, P, V, R)$$

It does not depend on the pathway of a transformation. And its change depends only on the initial and final states. There are many state functions like entropy or total energy. They are important because they can describe the effect of a transformation without knowing the nature of the transformation. The state functions itself depend on some parameters, which are the thermodynamics variables, which are not independent to each other.

5.2.7 Work

A reversible work performed on a system is defined as:

$$dW_{rev} = -PdV + \mu dN$$

So an infinitesimal amount of work come from a pressure contribution and the other is the contribution coming from the change of particles in the system. In this case the work is positive when the volume is decreasing, adding particles there is work performed on a system. μ is the chemical potential.

5.2.8 Heat

The heat added to the system is:

$$dQ_{rev} = CdT$$

Heat is assumed to be positive when it is added to the system and that it is added in a reversible manner. A change in heat correspond to a change in temperature if there is not a change transition. During a phase transition the temperature remains fixed but heat is still added. C is the specific heat.

5.2.9 First law of thermodynamics

In any thermodynamic transformation if a system absorbs an amount of heat ΔQ and has an amount of work ΔW performed on it, its internal energy will change by an amount ΔE given by:

$$\Delta e = \Delta Q + \Delta W$$

This law is the conservation of energy: adding heat or performing work on the system energy is being provided. Where E is a state function, although Q and W are not:

$$\Delta E = \Delta Q_{rev} + \Delta W_{rev} = \Delta Q_{irrev} + \Delta W_{irrev}$$

So the change in energy depend only on the initial and final state: it is always possible to compute the difference in energy for every transformation knowing only the initial and final state.

5.2.10 Second law of thermodynamics

There are two statements for the second law of thermodynamics which are completely equivalent.

5.2.10.1 Kelvin's formulation

There exists no thermodynamic transformation whose sole effect is to extract a quantity of heat from a high-temperature source and convert it entirely into work.

5.2.10.2 Clausius' formulation

There exists no thermodynamic transformation whose sole effect is to extract a quantity of heat from a cold source and deliver it to a hot source.

5.2.11 Entropy

The second law allow to define entropy. Entropy is a state function defined as:

$$\Delta S = S_2 - S_1 = \int_1^2 \frac{dQ_{rev}}{T}$$

The change in entropy depends only on the initial and the final state. It is a direct measure of the disorder of the system. There is no tool that can measure directly entropy.

5.2.12 Third law of thermodynamics

The third law of thermodynamics allow to find a reference state which can help to compute the absolute value of entropy. The reference state is the one at $0K$. The entropy of a system at the absolute zero of temperature is a universal constant, which can be taken to be zero. Finding a reversible transformation from $0K$ to any state an absolute value for any state and system can be found.

5.3 The ensemble

Defining thermodynamics from a macroscopic perspective the central concept is the ensemble. A collection of systems described by the same set of microscopic interactions and sharing a common set of macroscopic properties is said to be an ensemble. Only equilibrium ensembles will be considered. In equilibrium ensembles the systems in them evolve in time, but average quantities remain the same. So all particles move in time, but the average quantities of the macroscopic measurables remain the same. Now considering a macroscopic observable A:

$$A = \frac{1}{Z} \sum_{\lambda=1}^{N} \underbrace{a(x_\lambda)}_{\text{Phase space microscopic function}} \equiv \overbrace{\langle a \rangle}^{\text{Ensemble average}}$$

So any macroscopic observable is an average of the ensemble. So during a simulation there are several copies of a system, then a simulation of each copy must be taken and the average will be the result.

5.3.1 Phase space volume

Let a microstate of a system is a point in the phase space:

$$x_0 = (q_1(0), \ldots, q_{3N}(0), p_1(0), \ldots, p_{3N}(0))$$

Assuming that this is the initial state of the system and its evolution according to Hamilton's equation will be observed. The phase space volume element dx_0 contains a collection of microstates surrounding x_0. Thinking in term of ensemble the difference in macroscopic quantities will be small and will contain a number of microstate that can or not be part of the ensemble. This volume element evolves according to Hamilton's equation into dx_t, another volume element:

$$dx_t = J(x_t; x_0)dx_0$$

Where:

- $J(x_t; x_0) = detJ$.
- $J_{kl} = \frac{\partial x_t^k}{\partial x_0^l}$ is the Jacobian.

The time evolution of the phase space volume requires knowledge of the time evolution of the Jacobian.

5.3.1.1 Time evolution of the Jacobian

Consider the determinant of the Jacobian:

$$detJ = e^{Tr(\ln J)}$$

This is the exponential of the trace of the logarithm of J and this allow to take easier derivatives. The trace operator is the some of the diagonal element of the matrix and it has the property to be always the same of a matrix: it is an invariant. In the case of the Jacobian:

$$TrJ = \sum_k J_{kk} = \sum_k \lambda_k$$

Where λ_k are the eigenvalues. Then:

$$e^{Tr(\ln J)} = e^{\sum_k \ln \lambda_k} = \prod_k \lambda_k$$

Now the time evolution of the Jacobian:

$$\begin{aligned}\frac{d}{dt}J(x_t; x_0) = \frac{d}{dt}detJ = \frac{d}{dt}e^{Tr(\ln J)} = e^{Tr(\ln J)}Tr\left(\frac{dJ}{dt}J^{-1}\right) = \\ = J(x_t; x_0)\sum_{k,l}\left(\frac{dJ_{kl}}{dt}J^{-1}_{lk}\right) = J(x_t; x_0)\sum_{k,l}\left(\frac{\partial \dot{x}_t^k}{\partial x_0^l}\frac{\partial x_0^l}{\partial x_t^k}\right) = \\ = j(x_t; x_0)\sum_k \frac{\partial \dot{x}_t^k}{\partial x_t^k} = 0\end{aligned}$$

It is equal to zero because it is the compressibility of the system. So the time derivative of the quantity is equal to zero, so it means that the quantity itself is constant, so that the quantity, the determinant is equal to zero.

5.3.2 Liouville's theorem

In particular:

$$\frac{dJ(x_t;x_0)}{dt} = 0 \Rightarrow J(x_t;x_0) = const$$

However:

$$J(x_0;x_0) = 1 \Rightarrow J(x_t;x_0) = 1$$

This is because a transformation into itself is an identity matrix, so it is always equal to 1, so the determinant of the Jacobian is always equal to 1. The phase space element dx_0 does not change in time and $dx_t = dx_0$.

5.3.3 Ensemble distribution function

Focussing on a point x in phase space and on a small volume around it. The ensemble distribution function $f(x,t)dx$ is the fraction of the total ensemble members contained in the phase space volume element dx at time t. $f(x,t)$ verifies the typical properties of a probability density:

$$f(x,t) \geq 0$$

And integrating over the entirety of the phase space:

$$\int f(x,t)dx = 1$$

So there is a relation between the probability density and the ensemble distribution function. The objective of statistical mechanics is to find the shape of this function f.

5.3.4 Outward flux

Now the focus is how the ensemble distribution function evolves in time. The rate of decrease of ensemble members in a given volume Ω is:

$$-\frac{d}{dt}\int_\Omega f(x_t,t)dx_t = -\int_\Omega dx_t \frac{\partial f(x,t)}{\partial t}$$

So the integral over a volume element Ω the fraction of ensemble members in the volume is being counted. Minus the time derivative of that is the rate of decrease. Assuming that Ω is not changing in time, dx_t does not change because of Liouville's theorem. The only thing that might change is the ensemble distribution function, and it changes only if the ensemble distribution function has an explicit time dependence. The only way to change the number of the ensemble member is only if the ensemble distribution function has an explicit time dependence. There are no sink or sources of ensemble member and they change only if ensemble members pass through the surface S of the volume Ω. The outward flux of ensemble members leaving Ω is then

$$\int_S \dot{x}_t \cdot \hat{n} f(x_t,t)$$

Where $\dot{x}_t$ is the velocity in phase space with a dot product with a vector $\hat{n}$, a unit vector normal to the surface times the ensemble distribution function. With this quantity the divergence theorem can be applied:

$$\int_S \dot{x}_t \cdot \hat{n} f(x_t, t) = \int_\Omega \nabla_{x_t} \cdot [\dot{x}_t f(x_t, t)] dx_t$$

So the surface integral is translated in a volume integral taking the divergence of vector $\dot{x}_t f(x_t, f)$.

5.3.5 Liouville's equation

Writing down together the previous two results, the volume integral, the rate of decrease of ensemble element is equal to the flux:

$$\int_\Omega \nabla_{x_t} \cdot [\dot{x}_t f(x_t, t)] dx_t = -\int_\Omega \frac{\partial f(x_t, t)}{\partial t} dx_t$$

So the number inside the integral have to coincide:

$$\int_\Omega \left\{ \frac{\partial f(x_t, t)}{\partial t} + \nabla_{x_t} \cdot [\dot{x}_t f(x_t, t)] \right\} dx_t = 0 \Rightarrow \frac{\partial f(x_t, t)}{\partial t} + \nabla_{x_t} \cdot [\dot{x}_t f(x_t, t)] = 0$$

This implies that the argument of the integral has to be equal to 0 because it is valid for any value of Ω, obtaining a differential equation that will be satisfied for the ensemble distribution function. Writing down the equation more explicitly on the divergence:

$$\frac{\partial f(x_t, t)}{\partial t} + (\nabla_{x_t} \cdot \dot{x}_t) f(x_t, t) + \dot{x} \cdot \nabla_{x_t} f(x_t, t) = 0$$

Note that the divergence of the velocity is the compressibility, so that it is equal to zero, cancelling out the second element:

$$\frac{\partial f(x_t, t)}{\partial t} + \dot{x}_t \cdot \nabla_{x_t} f(x_t, t) = 0 \Rightarrow \frac{df(x_t, t)}{dt} = 0$$

So that the total derivative of f is equal to zero. This means that the ensemble distribution function, won't change in time.

5.3.5.1 Consequence on averages

Considering the previous result:

$$\frac{df(x_t, t)}{dt} = 0 \Rightarrow f(x_t, t) = f(x_0, 0)$$

And putting together Liouville's theorem and equation:

$$f(x_t, t) dx_t = f(x_0, 0) dx_0$$

So that there is a conserved fraction of ensemble members and averages can be performed at any time. So the result on averages is always the same because the ensemble members will be conserved during the trajectory.

5.3.5.2 A more elegant form

In a more elegant formalism the same formula can be written introducing η the velocity, written down in term of the Hamiltonian

$$\frac{\partial f(x,t)}{\partial t} + \dot{x} \cdot \nabla_x f(x,t) = \frac{\partial f(x,t)}{\partial f} + \eta(x,t) \cdot \nabla_x f(x,t) = 0$$

So that the function corresponds to the Poisson's brackets:

$$\frac{\partial f(x,t)}{\partial t} + \{f(x,t), \mathcal{H}(x,t)\} = 0$$

5.3.6 Equilibrium solutions

For a macroscopic observable A:

$$A = \langle a(x) \rangle = \int a(x) f(x,t) dx$$

This is an average over an ensemble of a microscopic coordinate. This can be written as the integral using the ensemble distribution function and that is the ensemble average. So to have A constant in time f cannot depend on time. An explicit time dependence of t implies an explicit time dependence of A hence at equilibrium it is required that:

$$\frac{\partial f(x,t)}{\partial t} = 0 \Rightarrow \{f(x,t), \mathcal{H}(x,t)\} = 0$$

So ensemble distribution functions must satisfy this equation. So there is a general solution:

$$f(x) \propto \mathcal{F}(\mathcal{H}(x))$$

Where the ensemble distribution function, a probability density must be proportional to some function of the Hamiltonian. So that the solution can be written exactly:

$$Z = \int dx \mathcal{F}(\mathcal{H}(x)) \Rightarrow f(x) = \frac{1}{Z} \mathcal{F}(\mathcal{H}(x))$$

So that the number Z a function of the thermodynamics parameters is a partition function and f can be written exactly. The integral is on the entire phase space. Most of statistical mechanics involves on computing the partition function, allowing to find the ensemble distribution function to find all the macroscopic observables allowing to compare with experimental results.

Chapter 6

Microcanonical ensemble

6.1 Introduction

All the other algorithm used in molecular simulation are evolution of the microcanonical ensemble. The microcanonical ensemble is a collection of systems which share the same macroscopic parameters and corresponds to different microstates. When comparing with an experimental observable the average of these quantities has to be taken. The microcanonical is the starting point of statistical mechanics. The number of particles, the total volume and the energy are kept fixed. This has some resemblance on Hamiltonian ensemble and so the microcanonical ensemble will be the natural ensemble for Hamiltonian mechanics.

6.2 State function depending on number of particle, volume and energy

Since volume, number of energy and number of particles are fixed a state function has to be defined. Starting from the first law of thermodynamics:

$$dE = dQ_{rev} + dW_{rev}$$

From the definition of entropy an infinitesimal variation of entropy will be:

$$dS = \frac{dQ_{rev}}{T} \Rightarrow dQ_{rev} = TdS$$

Looking now at dW:

$$dW_{rev} = -PdV + \mu dN$$

Where μ is the chemical potential. Putting everything together with the energy:

$$dE = TdS - PdV + \mu dN$$

So that the state function of N, V and E is entropy:

$$dS = \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN$$

6.2.1 Thermodynamic derivatives

Now, starting from the state function, the infinitesimal variation of S with respect to the variables can be written as:

$$\begin{aligned} dS &= \frac{1}{T}dE + \frac{P}{T}dV - \frac{\mu}{T}dN = \\ &= \left(\frac{\partial S}{\partial E}\right)_{V,N} dE + \left(\frac{\partial S}{\partial V}\right)_{N,E} dV + \left(\frac{\partial S}{\partial N}\right)_{V,E} dN \end{aligned}$$

Now a formula for the three elements of the state function can be found:

- $\left(\frac{\partial S}{\partial E}\right)_{V,N} = \frac{1}{T}$.
- $\left(\frac{\partial S}{\partial V}\right)_{N,E} = \frac{P}{T}$.
- $\left(\frac{\partial S}{\partial N}\right)_{V,E} = \frac{\mu}{T}$.

If entropy is known as a function of energy, volume and number of molecule, temperature, pressure and the chemical potential all the thermodynamics can be reconstructed.

6.2.2 Dirac's delta function

Dirac's delta function is a function in the form:

$$\delta(x) = \begin{cases} +\infty & x = 0 \\ 0 & otherwise \end{cases}$$

This function has some interesting properties:

- $\delta(x) = \delta(-x)$.
- $\delta(x) = \frac{d\theta(x)}{dx}$, where: $\theta(x) = \begin{cases} 1 & x \geq 0 \\ 0 & x < 0 \end{cases}$.
- $\int_{-\infty}^{+\infty} dx\delta(x) = 1$.
- $\int_{-\infty}^{+\infty} dx\delta(x)f(x) = f(0)$.

6.2.3 Computing entropy

The problem in the microcanonical ensemble is to find S. Considering Boltzmann's relation $S(N,V,E) = k\ln\Omega(N,V,E)$, where $\Omega(N,V,E)$ is the number of microscopic states of the system. The number of microscopic states of the system compatible with the condition of the system given by the values of N, V and E. To compute Ω the distribution function in the case of the microcanonical function: $f(x) = \mathcal{F}(\mathcal{H}(x)) = M\Delta(\mathcal{H}(x) - E)$. So that the energy is equal to the Hamiltonian. The constant M is a normalization constant to make f a distribution function. Now to count the number of the microscopic state all the microstate compatible with a given condition need to be summed over:

$$\Omega(N,V,E) = M_N \int d\vec{p}_1 \cdots \int d\vec{p}_N \int_{D(V)} d\vec{r}_1 \cdots \int_{D(V)} d\vec{r}_N \delta(\mathcal{H}(\vec{r},\vec{p}) - E)$$

So integrating over all the momenta and all the coordinates in the domain of V $D(V)$, so the coordinates vary in the coordinate of V which is fixed. Or, for simplicity the function can be written as:

$$\Omega(N,V,E) = M_N \int d\vec{p} \int_{D(V)} d\vec{r}\delta(\mathcal{H}(\vec{r},\vec{p}) - E) = M \int dx \delta(\mathcal{H}(x) - E)$$

Where x is a point in phase space and M_N is:

$$M_N = \frac{E_0}{N! h^{3N}}$$

Where h is Planck's constant and $N!$ is added in order to solve Gibbs paradox and is correct Boltzmann counting.

6.3 Average quantities

An average quantity can be obtained through:

$$A = \langle a \rangle = \frac{M_N}{\Omega(N,V,E)} \int dx a(x) \delta(\mathcal{H}(x) - E) = \frac{\int dx a(x) \delta(\mathcal{H}(x) - E)}{\int dx \delta(\mathcal{H} - E)}$$

There is no time dependence so equilibrium is assumed. The last term is a typical formula, with at the numerator an integral over the average of the microscopic property and at the denominator an integral of a partition function. Considering a given variable x_i, an object in phase space times the derivative on the Hamiltonian over another object in phase space. x_i and x_j are two distinct objects. To do so:

$$\begin{aligned}
\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle &= \frac{M_n}{\Omega(N,V,E)} \int dx x_i \frac{\partial \mathcal{H}}{\partial x_j} \delta(E - \mathcal{H}(x)) = \\
&= \frac{M_N}{\Omega(N,V,E)} \frac{\partial}{\partial E} \int dx x_i \frac{\partial \mathcal{H}}{\partial x_j} \theta(E - \mathcal{H}(x)) = \\
&= \frac{M_N}{\Omega(N,V,E)} \frac{\partial}{\partial E} \int_{\mathcal{H}(x)<E} dx x_i \frac{\partial \mathcal{H}}{\partial x_j} = \\
&= \frac{M_N}{\Omega(N,V,E)} \frac{\partial}{\partial E} \int_{\mathcal{H}(x)<E} dx x_i \frac{\partial (\mathcal{H} - E)}{\partial x_j}
\end{aligned}$$

6.3.1 Virial theorem

The Virial theorem allows to find the ensemble average of that quantity:

$$\begin{aligned}\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle &= \frac{M_N}{\Omega(N,V,E)} \frac{\partial}{\partial E} \int_{\mathcal{H}(x)<E} dx x_i \frac{\partial(\mathcal{H}-E)}{\partial x_j} = \\ &= \underbrace{\frac{M_N}{\Omega(N,V,E)} \frac{\partial}{\partial E} \int_{\mathcal{H}<E} dx \delta_{ij}(E-\mathcal{H})}_{\text{Integrating by part}} = \\ &= \frac{M_N}{\Omega(N,V,E)} \frac{\partial}{\partial E} \int dx \delta_{ij}(E-\mathcal{H})\theta(E-\mathcal{H}) = \\ &= \frac{E_0}{N!h^{3N}\Omega(N,V,E)} \delta_{ij} \int dx \theta(E-\mathcal{H}) = \\ &= \delta_{ij} \frac{\Sigma(E)}{\frac{\partial \Sigma(E)}{\partial E}}\end{aligned}$$

Where:

- $\Sigma(N,V,E) = \frac{1}{N!h^{3N}} \int dx \theta(E-\mathcal{H})$, the uniform ensemble and is all the points where the Hamiltonian is less than the energy.
- $\Omega(N,V,E) = E_0 \frac{\partial \Sigma(N,V,E)}{\partial E}$.

So:

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = \delta_{ij} \frac{\Sigma(E)}{\frac{\partial \Sigma(E)}{\partial E}} = \delta_{ij} \left(\frac{\partial \ln \Sigma(E)}{\partial E} \right)^{-1}$$

Considering Boltzmann's relation:

$$S(N,V,E) = k \ln \Omega(N,V,E) \simeq k \ln \Sigma(N,V,E) = \tilde{S}(N,V,E)$$

The difference between S and $\tilde{S}$ is close to the logarithm of N. If N is very big the difference is not very significative.

Then:

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = k\delta_{ij} \left(\frac{\partial \tilde{S}(E)}{\partial E} \right)^{-1} \simeq k\delta_{ij} \left(\frac{\partial S(E)}{\partial E} \right)^{-1} = kT\delta_{ij}$$

So that $\ln \Sigma(E)$ can be substituted with $k\tilde{S}(E)$

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = kT\delta_{ij}$$

So that the ensemble average of e quantity made in that way its average is $kT\delta_{ij}$.

6.3.2 Application of Virial theorem

This is useful because it allows to construct microscopic phase space functions whose ensemble averages yield macroscopic thermodynamics observables.

$$\left\langle x_i \frac{\partial \mathcal{H}}{\partial x_j} \right\rangle = kT\delta_{ij}$$

In this way calculation can be not performed and a direct exact result starting from macroscopic phase-space variables and compute ensemble averages exactly. This theorem need to be respected in simulations.

6.3.2.1 An example

Consider an Hamiltonian like the following and take the ensemble average of $\left\langle p_i \frac{\partial \mathcal{H}}{\partial p_i} \right\rangle$.

$$\mathcal{H} = \sum_i \frac{\vec{p_i}^2}{2m_i} + U(\vec{r}_1, \ldots, \vec{r}_N) \Rightarrow \left\langle p_i \frac{\partial \mathcal{H}}{\partial p_i} \right\rangle = kT$$

So that:

$$\left\langle p_i \frac{\partial \mathcal{H}}{\partial p_i} \right\rangle = \left\langle \frac{p_i^2}{m_i} \right\rangle = kT$$

So that the kinetic energy of a single particle in three dimensions:

$$\left\langle \frac{p_i}{2m_i} \right\rangle = \frac{3}{2}kT$$

And the total kinetic energy:

$$\sum_{i=1}^{N} \left\langle \frac{p_i^2}{2m_i} \right\rangle = \frac{3}{2}NkT$$

The temperature of the simulation can be computed starting from the kinetic energy of the particles, so that an instantaneous temperature of the system can be computed. In the microcanonical ensemble the energy is fixed, but the temperature is not.

6.4 Thermal contact

Consider an isolated system composed by two systems 1 and 2 be divided by a heat conducting divider: volumes are fixed and molecules do not travel across it, only heat. Considering them together:

$$N = N_1 + N_2 \qquad \wedge \qquad V = V_1 + V_2 \qquad \wedge \qquad \mathcal{H}(x) = \mathcal{H}_1(x_1) + \mathcal{H}_2(x_2)$$

And the state equations:

$$S_1(N_1, V_1, E_1) = k \ln \Omega_1(N_1, V_1, E_1) \qquad \wedge \qquad S_2(N_2, V_2, E_2) = k \ln \Omega_2(N_2, V_2, E_2)$$

Now considering the two ω functions:

$$\omega_1(N_1, V_1, E_1) = M_{N_1} \int dx_1 \delta(\mathcal{H}_1 - E_1) \qquad \wedge \qquad \omega_2(N_2, V_2, E_2) = M_{N_2} \int dx_2 \delta(\mathcal{H}_2 - E_2)$$

To compute the entropy of the total system Ω of the entire system needs to be computed:

$$\Omega(N, V, E) = M_N \int dx \delta(\mathcal{H}_1(x_1) + \mathcal{H}_2(x_2) - E) \neq \Omega_1(N_1, V_1, E_1)\Omega_2(N_2, V_2, E_2)$$

So Ω is not the product of the two Ω_1 and Ω_2. In particular, fixing the value of energy of system 1 in E_1 and doing so the energy for the second system is fixed as $E - E_1$. For each microstate in system 1 there are Ω_2 microstates for system 2. So the product of the combination for each single value of E_1 has to be taken into account for the total Ω:

$$\Omega(N, V, E) = C \int_0^E dE_1 \Omega_1(N_1, V_1, E_1)\Omega_2(N_2, V_2, E - E_1)$$

This integration should be solved exactly, however, assuming that $\bar{E}_1$ is the value that maximises the product $\Omega_1(N_1, V_1, E_1)\Omega_2(N_2, V_2, E - E_1)$. Now the sum will be greater than the maximum but it will be less than the maximum times the number of elements. In the case where the number of trial is great this difference becomes negligible:

$$S(N, V, E) = k \ln \Omega(N, V, E) \simeq k \ln[\Omega_1(N_1, V_1, \bar{E}_1)\Omega_2(N_2, V_2, E - \bar{E}_1)] + o(\ln N)$$

So, in the thermodynamics limit:

$$S(N, V, E) = k \ln \Omega_1(N_1, V_1, \bar{E}_1) + k \ln \Omega_2(N_2, V_2, E - \bar{E}_1) = S_1(N_1, V_1, \bar{E}_1) + S_2(N_2, V_2, E - \bar{E}_1)$$

So in the thermodynamic limit entropy is additive.

6.4.1 Temperature

Another consequence of the previous result is that considering the last equation $\bar{E}_1$ is the value of E_1 that maximizes the quantity:

$$k \ln \Omega_1(N_1, V_1, E_1)\Omega_2(N_2, V_2, E - E_1)$$

Since $\bar{E}_1 + \bar{E}_2 = E$ and E is fixed, $d\bar{E}_1 + d\bar{E}_2 = 0 \Rightarrow d\bar{E}_1 = -d\bar{E}_2$. Considering the formula:

$$S(N, V, E) = S_1(N_1, V_1, \bar{E}_1) + S_2(N_2, V_2, \bar{E}_2)$$

Taking the derivative with respect to $\bar{E}_1$ and $\bar{E}_2$L

$$0 = \frac{\partial S_1(N_1, V_1, \bar{E}_1)}{\partial \bar{E}_1} + \frac{\partial S_2(N_2, V_2, \bar{E}_2)}{\partial \bar{E}_1} = \frac{\partial S_1(N_1, V_1, \bar{E}_1)}{\partial \bar{E}_1} - \frac{\partial S_2(N_2, V_2, \bar{E}_2)}{\partial \bar{E}_2}$$

And so, remembering the derivative of the entropy with respect to energy:

$$\frac{1}{T_1} - \frac{1}{T_2} = 0 \Rightarrow T_1 = T_2$$

Obtaining that two system in thermal contact have the same temperature at equilibrium.

6.5 Some examples

6.5.1 Free particle in one dimension

In order to study the ideal gas the free particle in $1D$ have to be solved. This particle has Hamiltonian:

$$\mathcal{H} = \frac{p^2}{2m}$$

Considering a particle confined in a length L:

$$\Omega(1, L, E) = \frac{E_0}{h}\int_0^L dx \int_{-\infty}^{+\infty} dp\delta\left(\frac{p^2}{2m} - E\right) = \frac{E_0 L}{h}\int_{-\infty}^{+\infty} dp\delta\left(\frac{p^2}{2m} - E\right)$$

Considering $y^2 = \frac{p^2}{2m}$:

$$\int_{-\infty}^{+\infty} dp\delta\left(\frac{p^2}{2m} - E\right) = \sqrt{2m}\int_{-\infty}^{+\infty} dy\delta(y^2 - E)$$

And remembering another property of the δ function:

$$\delta(f(x)) = \sum_{i\in\ zeros\ of\ f(x)} \frac{1}{|f'(x_i)|}\delta(x - x_i) \Rightarrow \delta(y^2 - E) = \frac{1}{2\sqrt{E}}[\delta(y - \sqrt{E}) + \delta(y + \sqrt{E})]$$

Computing the integral:

$$\sqrt{2m}\int_{-\infty}^{+\infty} dy\delta(y^2 - E) = \sqrt{2m}E \Rightarrow \Omega(1, V, E) = \frac{E_0 L}{h}\sqrt{\frac{2m}{E}}$$

6.5.2 Classical ideal gas

Considering N free particles in $3D$, or the classical ideal gas, which it will be close to one particle in $1D$:

$$\Omega(N, V, E) = \frac{E_0}{N!h^{3N}}\int d^N\vec{p}\int_{D(V)} d^N\vec{r}\delta\left(\sum_{i=1}^{N}\frac{\vec{p}_i^2}{2m_i} - E\right) = \frac{E_0 V^N}{N!h^{3N}}\int d^N\vec{p}\left(\sum_{i=1}^{N}\frac{\vec{p}_i^2}{2m_i} - E\right)$$

Where $N!$ is correct Boltzmann counting and h is Plank's constant, the sum of all kinetic energy is because they are all free particles. Computing the integral and using Boltzmann relation the entropy can be computed:

$$\Omega(N, V, E) = \frac{1}{N!}\left[\frac{V}{h^3}\left(\frac{4\pi mE}{3N}\right)^{\frac{3}{2}}\right]^N e^{\frac{3N}{2}} \Rightarrow S(N, V, E) = Nk\ln\left[\frac{V}{h^3}\left(\frac{4\pi mE}{3N}\right)^{\frac{3}{2}}\right] + \frac{3Nk}{2} - k\ln N!$$

So the entropy for the ideal gas can be obtained starting from the Hamiltonian. Now the thermodynamics derivatives can be used to obtain all thermodynamic properties:

$$\frac{1}{T} = \left(\frac{\partial S}{\partial E}\right)_{N,V} = \frac{3Nk}{2E} \Rightarrow E = \frac{3}{2}NkT$$

$$\frac{p}{T} = \left(\frac{\partial S}{\partial V}\right)_{N,E} = \frac{Nk}{V} \Rightarrow pV = NkT$$

$$\frac{\mu}{T} = -\left(\frac{\partial S}{\partial N}\right)_{V,E} = k\ln N - k\ln\left[\frac{V}{h^3}\left(\frac{4\pi mE}{3N}\right)^{\frac{3}{2}}\right] \Rightarrow \mu = -kT\ln\left[\frac{V}{Nh^3}\left(\frac{4\pi mE}{3N}\right)^{\frac{3}{2}}\right]$$

6.6 Gibbs paradox

Considering Boltzmann equation the factorial factor is due to the fact that particles are undistinguishable.

$$S(N,V,E) = Nk\ln\left[\frac{V}{h^3}\left(\frac{4\pi mE}{3N}\right)^{\frac{3}{2}}\right] + \frac{3Nk}{2} - \cancel{k\ln N!}$$

If the factorial is not added Gibbs paradox comes into play, obtaining the formula for classical entropy.

$$S^{(cl)}(N,V,T) = Nk\ln\left[\frac{V}{h^3}(2\pi mkT)^{\frac{3}{2}}\right] + \frac{3Nk}{2}$$

Assume that there are N_1 particles in one box and N_2 particles in another box. Now, computing the entropy of the total system and of the mixed system:

$$S_1^{(cl)}(N_1,V_1,T) = N_1k\ln\left[\frac{V_1}{h^3}(2\pi mkT)^{\frac{3}{2}}\right] + \frac{3N_1k}{2} \quad S_2^{(cl)}(N_2,V_2,T) = N_2k\ln\left[\frac{V_2}{h^3}(2\pi mkT)^{\frac{3}{2}}\right] + \frac{3N_2k}{2}$$

$$S^{(cl)}(N_1+N_2,V_1+V_2,T) = (N_1+N_2)k\ln\left[\frac{V_1+V_2}{h^3}(23\pi mkT)^{\frac{3}{2}}\right] + \frac{3(N_1+N_2)k}{2}$$

Now the entropy of the mixed system:

$$\begin{aligned}\Delta S_{mix}^{(cl)} &= S^{(cl)}(N_1+N_2) - S_1^{(cl)}(N_1) - S_2^{(cl)}(N_2) = \\ &= (N_1+N_2)k\ln(V_1+V_2) - N_1k\ln V_1 - N_2k\ln V_2 = \\ &= N_1k\ln\frac{V}{V_1} + N_2k\ln\frac{V}{V_2} > 0\end{aligned}$$

This result is greater then 0, so mixing the two gases entropy will increase.

- $\Delta S_{mix}^{(cl)} > 0$: this results makes sense if the two gases are different.
- $\Delta S_{mix}^{(cl)} \neq 0$: if the same gas is present in the two boxes it should be equal to 0, but the formula predicts that the entropy increases. This is because when inserting back the wall the same situation at the start should be obtained and this is not possible because entropy is a state function.

6.6.1 Correct Boltzmann counting

To solve the paradox correct Boltzmann counting is introduced, the $N!$ factor. In fact inserting it the correct formula becomes:

$$S^{(ST)}(N,V,T) = Nk\ln\left[\frac{V}{Nh^3}(2\pi mkT)^{\frac{3}{2}}\right] + \frac{5Nk}{2}$$

Or the Sackur-Tetrode formula. Using Stirling's approximation for the factorial. Using this formula:

Now, in the following cases:

- $\Delta S_{mix} = N_1 k \ln \frac{V}{V_1} + N_2 k \ln \frac{V}{V_2} > 0$: for two different gases.
- $\Delta S_{mix} = N_1 k \ln \frac{V}{N} \frac{N_1}{V_1} + N_2 k \ln \frac{V}{N} \frac{N_2}{V_2} = 0$: for the same gas.

Chapter 7

Introduction to molecular dynamics

7.1 Introduction

7.1.1 Hamilton's equations

The starting point to a Molecular dynamics simulation is Hamilton's equation. Hamilton's equation will yield Newton's equation at the end, allowing to study the system as first order differential equations.

$$\dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} \qquad \dot{p}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha}$$

The objective is to obtain a numerical result from these equation. Remembering that solving Hamilton's equation means integrating them keeping the energy constant.

$$\frac{d\mathcal{H}}{dt} = \sum_\alpha \left[\frac{\partial \mathcal{H}}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \mathcal{H}}{\partial p_\alpha} \dot{p}_\alpha \right] = \sum_\alpha \left[\frac{\partial \mathcal{H}}{\partial q_\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} - \frac{\partial \mathcal{H}}{\partial p_\alpha} \frac{\partial \mathcal{H}}{\partial q_\alpha} \right] = 0$$

All the work is done in the microcanonical ensemble.

$$\mathcal{H}(q_1, \ldots, q_{3N}, p_1, \ldots, p_{3N}) = const$$

The quantities obtained through Hamilton's equations are representative of the microcanonical ensemble. The time-dependent solutions will be rigorous, conformations or state that are separated by an energy barrier, local minima cannot be escaped.

7.1.2 Ergodicity

When a property has to be measured in an ensemble, what is measured is the average over an ensemble:

$$A = \langle a \rangle = \frac{\int dx a(x) \delta(\mathcal{H}(x) - E)}{\int dx \delta(\mathcal{H}(x) - E)} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_0^\tau dt a(x_t) \equiv \bar{a}$$

Assuming that the system will visit during in dynamics all the possible state of its ensemble the average over the ensemble can be substituted over an average over time. Usually the time average

is indicated by a bar: $\bar{a}$. The numerical integrator will give the value of $a(x_t)$ at different time steps and from that a discretized time average will be obtained:

$$A = \langle a \rangle = \frac{1}{M} \sum_{n=1}^{M} a(x_{n\Delta t})$$

This discretized time average for sure is not the ensemble average. In order for this to be true the system has to have the property of ergodicity. Ergodicity means that in a simulation, while exploring different state and point in time, all possible states compatible with a macrostate are being explored. So all the state belonging to the ensemble are being explored in the phase-space. This is not easy to assume for complex states and depends on the energy profile of the system.

7.1.3 Basic components of a molecular dynamics simulation

The basic components of a MD simulation are:

- The model: the chosen force field, the model used to represent chemical bonds and reality.
- Calculation of energies and forces: accurate and efficient. So how to compute the energies and forces, numerical steps in the model characterized by some error.
- The algorithm: used to integrate the equations of motion.

7.2 Verlet algorithm

An algorithm used to integrate the equations of motion is the Verlet algorithm. Starting from the Taylor expansion of the coordinates at time $t + \Delta t$ and then they can be written using Newton's law.

$$\vec{r}_i(t + \Delta t) \approx \vec{r}_i(t) + \Delta t \dot{\vec{r}}_i(t) + frac12\Delta t^2 \ddot{\vec{r}}_i(t) \qquad \approx \vec{r}_i(t + \Delta t) + \Delta t \vec{v}_i(t) + \frac{\Delta t^2}{2m_i} \vec{F}_i(t)$$

Similarly, integrating backward in time:

$$\vec{r}_i(t - \Delta t) \approx \vec{r}_i(t) - \Delta t \vec{v}_i(t) + \frac{\Delta t^2}{2m_i} \vec{F}_i(t)$$

Summing up the two equations the first order terms will cancel:

$$\vec{r}_i(t + \Delta t) + \vec{r}_i(t - \Delta t) = 2\vec{r}_i(t) + \frac{\Delta t^2}{m_i} \vec{f}_i(t)$$

In these equations the third order terms will also cancel, so that this sum is correct until the forth order term. So this equation is correct up to the forth order term So the coordinates at time $t + \Delta t$ are:

$$\vec{r}_i(t + \Delta t) = 2\vec{r}_i(t) - \vec{r}_i(t - \Delta t) + \frac{\Delta t^2}{m_i} \vec{F}_i(t)$$

In order to get the coordinates at the previous instant in time have to be stored. The velocity can be computed using velocity's definition:

$$\vec{v}_i(t) = \frac{\vec{r}_i(t+\Delta t) - \vec{r}_i(t-\Delta t)}{2\Delta t}$$

The best thing to do is to compute the average velocity over 2 time step as to have a numerically more stable solution. This algorithm is time-reversible and will keep energy constant. Some problem about it is that the velocity are the kinetic energy and is related to the temperature of the system.

7.2.1 Velocity Verlet

A variation over Verlet providing the same trajectory, but computing velocities and coordinates at the same time. Again the starting point is the Taylor expansion:

$$\vec{r}_i(t+\Delta t) = \vec{r}_i(t) + \Delta t\vec{v}_i(t) + \frac{\Delta t^2}{2m_i}\vec{F}_i(t)$$

Now integrating backward in time considering the velocities:

$$\vec{r}_i(t) = \vec{r}_i * t + \Delta t) - \Delta t\vec{v}_i(t+\Delta t) + \frac{\Delta t^2}{2m_i}\vec{F}_i(t+\Delta t)$$

By substituting $\vec{r}_i * t + \Delta t)$ with the first equation:

$$\vec{v}_i(t+\Delta t) = \vec{v}_i(t) + \frac{\Delta t}{2m_i}[\vec{F}_i(t) + \vec{F}_i(t+\Delta t)]$$

With the average of the forces at the two time steps. The velocities and the coordinates are updated at the same time. Time reversibility is necessary because Hamilton's equations are being solved. Moreover these algorithms have a symplectic structure, a property related with their numerical stability: this means that the trajectories that are obtained using this algorithms although it is not exactly the same, the errors will not diverge from the classical trajectory.

7.2.2 Initial condition

When starting a simulation initial coordinates have to be provided. This can be done by taking data from experimental data or guessed. Also the initial velocities are needed for the Verlet algorithm. The best thing to guess the velocities are to take them randomly from a Maxwell-Boltzmann distribution:

$$f(v) = \left(\frac{m}{2\pi kT}\right)^{\frac{1}{2}} e^{-\frac{mv^2}{2kT}}$$

This is a Gaussian distribution whose variance depends on the temperature:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$$

The wider the spread the higher the temperature. So in a simulation an initial temperature is given and then the velocities are extracted from the Maxwell-Boltzmann distribution. In the microcanonical simulation the energy is fixed but the temperature is not, so most of the time the temperature will not be kept fixed and after some time it will reach a constant value, depending on the initial conditions.

7.2.3 Action integral

Because the molecules present bonds these will oscillates, but the characteristic time for their oscillation will be smaller of the Δt of the time step of the algorithm. If the oscillation needs to be represented the time step should be smaller than their oscillation, so smaller than a femtosecond. This time step should be larger so, the bonds can be treated as rigid bonds. Constraints will be posed on the chemical bonds so that their length will be kept fixed. This constraints allow to use a time step in the order of femtoseconds. In order to implement a constraint in the Verlet algorithm the Lagrangian formalism should be obtained from an action integral. Consider the coordinates such that Q is the coordinates and $\dot{Q}$ is the velocities, both generalized.

$$Q \equiv \{q_1, \ldots, q_{3N}\} \qquad \dot{Q} \equiv \{\dot{q}_1, \ldots, \dot{q}_{3N}\}$$

Now the action integral is:

$$A[Q] = \int_{t_1}^{t_2} \mathcal{L}(Q(t), \dot{Q}(t))dt$$

This is a number that integrates a path from $(Q(t_1), \dot{Q}(t_1))$ and $(Q(t_2), \dot{Q}(t_2))$ and summing over all the trajectory. The value of the action depends on the path taken when changing position in the phase-space. Each path will be characterized a different value of the action integral. So A is a function of a function, the path. The path Q that renders the action stationary is:

- $Q(t_1) = Q_1$.
- $Q(t_2) = Q_2$.
- $\dot{Q}(t_1) = \dot{Q}_1$.
- $\dot{Q}(t_2) = \dot{Q}_2$.

Stationary means that making a small variation to the path the corresponding variation of the action is zero in the first order. So the starting and ending coordinates will be exactly the same for all the paths. So that the variation at time t_1 and at time t_2 will be equal to zero:

- $\delta Q(t_1) = \delta Q(t_2) = 0$.
- $\delta\dot{Q}(t_1) = \delta\dot{Q}(t_2) = 0$.

So at the boundaries the variation will be equal to zero. Now computing the variation on the action when changing the path. Since the variation is small a Taylor expansion can be performed on the integral, considering all the partial derivatives, making the derivative equal to zero:

$$\begin{aligned}
\delta A &= \int_{t_1}^{t_2} \mathcal{L}(Q(t) + \delta Q(t), \dot{Q}(t) + \delta\dot{Q}(t))dt - \int_{t_1}^{t_2} \mathcal{L}(Q(t), \dot{(Q)}(t))dt = \\
&= \int_{t_1}^{t_2} \sum_{\alpha=1}^{3N} \left[\frac{\partial \mathcal{L}}{\partial q_\alpha} \delta q_\alpha(t) + \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \delta \dot{q}_\alpha(t) \right] dt = \\
&= \sum_{\alpha=1}^{3N} \frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \delta q_\alpha(t)|_{t_1}^{t_2} + \int_{t_1}^{t_2} \sum_{\alpha=1}^{3N} \left[\frac{\partial \mathcal{L}}{\partial q_\alpha} \delta q_\alpha(t) - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) \delta q_\alpha(t) \right] dt = 0
\end{aligned}$$

Thus, remembering the boundaries conditions and since the formula will be valid for any α, so the integral itself will be equal to zero, because all the δq_α are independent from each other if there are no constraints:

$$\frac{\partial \mathcal{L}}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) = 0 \Rightarrow \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = 0$$

Obtaining the Euler-Lagrangian equations from a variational principles. The Euler-Lagrangian equations correspond to the path that will render the action integral stationary. So the two formulation are completely equivalent to each other.

7.3 Constraints

Constraints refer to chemical bond that are kept fixed during the simulation time. Defining constraints:

- Holonomic constraints: $\sigma_k(q_1, \dots, q_{3N}, t) = 0 \qquad k = 1, \dots, N_C$. This means that the equation corresponding to the constraint can be written as above. There will be one equation for each constraint in the system, so that the contraint is an equation that depends on all coordinates and time. So fixing a bond is a holonomic contraints.
- Nonholonomic constraints: there is also a relation on the velocities $\zeta(q_1, \dots, q_{3N}, \dot{q}_1, \dots, \dot{q}_{3N}) = 0$. Coordinates and velocities are not independent.

Each constraint adds an equation to the system, reducing the degrees of freedom, from $3N$ to $3N - N_C$ because the coordinates will depend on each other. This means that the number of generalized coordinates will be $3N - N_C$. A particular condition from non-holonomic constraint is an example:

$$\frac{1}{2}\sum_i m_i \dot{\vec{r}}_i^2 - C = 0$$

Constraining the total kinetic energy to be constant. This is a nonholonomic constraint.

7.3.1 Differential forms

These constraints can be written in a differential form, defined as:

$$\sum_{\alpha=1}^{3N} a_{k\alpha} dq_\alpha + a_{kt} dt = 0 \qquad k = 1, \dots, N_C$$

Depending on the variation of coordinates and time. α refers to the coordinates ant k to the constraints.

7.3.1.1 Holomonic constraints

Going from their equation to their differential form, taking the derivative of the holonomic constraint equation:

$$\sum_{\alpha=1}^{3N} \frac{\partial \sigma_k}{\partial q_\alpha} dq_\alpha + \frac{\partial \sigma_k}{\partial t} dt = 0 \qquad k = 1, \dots, N_N \qquad a_{k\alpha} = \frac{\partial \sigma_k}{\partial q_\alpha} \qquad a_{kt} = \frac{\partial \sigma_k}{\partial t}$$

7.3.1.2 Nonholonomic constraints

For nonholonomic constraints is easy for some exception, for example on the one on kinetic energy:

$$\frac{1}{2}\sum_i m_i \dot{\vec{r}}_i^2 - C = 0 \Rightarrow \frac{1}{2}\sum_i m_i \dot{\vec{r}}_i \frac{d\vec{r}_i}{dt} - C = 0 \Rightarrow \frac{1}{2}\sum_i m_i \dot{\vec{r}}_i d\vec{r}_i - Cdt = 0$$

So that:

$$a_{1i} = \frac{1}{2} m_i \dot{\vec{r}}_i \qquad a_{1t} = -C$$

In general only constraint that can be written as:

$$\sum_{\alpha=1}^{3N} a_{k\alpha} dq_\alpha = 0$$

Will be considered. So Holonomic constraints without an explicit time dependence.

7.3.2 Lagrange multipliers

If the constraints can be written in an integrable form then the Euler-Lagrange equations can be written directly following a variational principle. This can be done writing Lagrange multipliers. The problem is that the derivation of the Euler-Lagrange equation was based on the fact that the coordinates were independent from each other. Dealing with constraints this cannot be done: there are $3N - N_C$ independent coordinates and the δq_α. To make them independent other variables need to be introduced: the Lagrange multipliers, so that the sum will be equal to zero because of the integrable form. Now looking inside the square brackets a sum introduced that is equal to zero. N_C unknowns (λ_k) are introduced so that there are $3N - N_C + N_C = 3N$ independent coordinates. The last thing to do so is to obtain the values for λ_k which will have specific values.

$$\int_{t_1}^{t_2} \sum_{\alpha=1}^{3N} \left[\frac{\partial \mathcal{L}}{\partial q_\alpha} - \frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) + \sum_{k=1}^{N_N} \lambda_k a_{k\alpha} \right] \delta q_\alpha(t) dt = 0$$

A modified version of the Euler-Lagrange equation is obtained.

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{q}_\alpha} \right) - \frac{\partial \mathcal{L}}{\partial q_\alpha} = \sum_{k=1}^{N_C} \lambda_k a_{k\alpha}$$

To obtain all the λ_k N_C additional equations are needed. These are the equations of the constraints:

$$\sum_{\alpha=1}^{3N} a_{k\alpha} \dot{q}_\alpha + a_{kt} = 0 \qquad k = 1, \ldots, N_C$$

So that all the velocities, coordinates and λ_k can be computed.
So there are $3N + N_C$ equations for $3N + N_c$ unknowns.

7.3.3 Hamiltonian formulation

The Hamiltonian formulation can be obtained for these constraints. For time-independent holonomic constraints:

$$\begin{cases} \dot{q}_\alpha = \frac{\partial \mathcal{H}}{\partial p_\alpha} \\ \dot{q}_\alpha = -\frac{\partial \mathcal{H}}{\partial q_\alpha} - \sum\limits_{k=1}^{N_C} \lambda_k a_{k\alpha} \\ \sum\limits_{\alpha=1}^{3N} a_{k\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} = 0 \end{cases}$$

The constraints equations appear in the same place of the forces: whenever constraints are introduced in the system constraints forces are added to the system. So that integrating the system of equation, its derivative with respect to time should be equal to 0:

$$\begin{aligned} \frac{d\mathcal{H}}{dt} &= \sum_\alpha \left[\frac{\partial \mathcal{H}}{\partial q_\alpha} \dot{q}_\alpha + \frac{\partial \mathcal{H}}{\partial p_\alpha} \right] = \\ &= \sum_\alpha \left[\frac{\partial \mathcal{H}}{\partial q_\alpha} \frac{\partial \mathcal{H}}{\partial p_\alpha} + \frac{\partial \mathcal{H}}{\partial p_\alpha} \left(\frac{\partial \mathcal{H}}{\partial q_\alpha} + \sum_k \lambda_k a_{k\alpha} \right) \right] = \\ &= \sum_k \lambda_k \sum_\alpha \frac{\partial \mathcal{H}}{\partial p_\alpha} a_{k\alpha} = 0 \end{aligned}$$

So no work is done on a system by the imposition of holonomic constraints.

7.3.4 Constraints in a simulation

To implement the constraints in a simulation, forces coming from the constraints are called the constraints forces, so the forces are:

$$m_i \ddot{\vec{r}}_i = \vec{F}_i + \underbrace{\sum_{k=1}^{N_C} \lambda_k \nabla_i \sigma_k}_{\text{Constraint forces}}$$

So these equations need to keep the constraints, adding a condition on the gradient of σ_k:

$$\frac{d}{dt}\sigma_k(\vec{r}_1, \ldots, \vec{r}_N) = 0 \Rightarrow \dot{\sigma}_k = \sum_{i=1}^{N} \nabla_i \sigma_k \cdot \dot{\vec{r}}_i = 0$$

Now a condition on the velocities is added: they need to be perpendicular to the gradient of the constraint equation. To compute λ_k while the simulation is running the constraints are implemented directly in the algorithm. Considering the velocity Verlet:

$$\vec{r}_i(\Delta t) = \vec{r}_i(0) + \Delta t \vec{v}_i(0) + \frac{\Delta t^2}{2m_i} \vec{F}_i(0) + \frac{\Delta t^2}{2m_i} \sum_k \lambda_k \nabla_i \sigma_k(0)$$

To obtain λ_k:

$$\vec{r}_i' = \vec{r}_i(0) + \Delta t \vec{v}_i(0) + \frac{\Delta t^2}{2m_i} \vec{F}_i(0)$$

$$\vec{r}_i(\Delta t) = \vec{r}_i' + \frac{1}{m_i}\sum_k \tilde{\lambda}_k \nabla_i \sigma_k(0)$$

Considering the notation:

$$\tilde{\lambda}_k = \frac{\Delta t^2}{2}\lambda_k$$

7.3.4.1 Constraint condition

Considering that the new coordinates need to satisfy the constraints equations. This is where the constraint equations come into place:

$$\sigma_l(\vec{r}_1(\Delta t), \ldots, \vec{r}_N(\Delta t)) = 0 \qquad l = 1, \ldots, N_C$$

Now there are N_C equations for the new coordinates, with N_C unknowns. These equations will allow to compute the λ_k and to solve the constraint equations:

$$\sigma_l\left(\vec{r}_1' + \frac{1}{m_1}\sum_k \tilde{\lambda}_k \nabla_1 \sigma_k(0), \ldots, \vec{r}_N' + \frac{1}{m_N}\sum_k \tilde{\lambda}_k \nabla_N \sigma_k(0)\right) = 0 \qquad l = 1, \ldots, N_C$$

There are different strategies to solve this system. The first thing to do is to assume that all the sum of the sum over the λ_k are slight corrections to the new coordinates. The new values of the coordinates are those obtained without constraints with a slight modification given by the constraints. Because of this a Taylor expansion can be done and considering the method SHAKE, an iterative solution is given from an initial guess $\tilde{\lambda}_k^{(1)}$:

$$\vec{r}_i^{(1)} = \vec{r}_i' + \frac{1}{m_i}\sum_k \tilde{\lambda}_k^{(1)} \nabla_1 \sigma_k(0)$$

In SHAKE an initial guess on $\tilde{\lambda}_k^{(1)}$ is done and then a Taylor expansion is done. The exact solution: $\tilde{\lambda}_k = \tilde{\lambda}_k^{(1)} + \delta\tilde{\lambda}_k^{(1)}$, so related with the initial guess with an assumed small variation, then the coordinates at δt will be:

$$\vec{r}_i(\Delta t) = \vec{r}_i^{(1)} + \frac{1}{m_i}\sum_k \delta\tilde{\lambda}_k^{(1)} \nabla_1 \sigma_k(0)$$

Now adding the guess to the initial formula:

$$\sigma_l\left(\vec{r}_1^{(1)} = \frac{1}{m_1}\sum_k \delta\tilde{\lambda}_k \nabla_1 \sigma_k(0), \ldots, \vec{r}_N^{(1)} + \frac{1}{m_N}\sum_k \delta\tilde{\lambda}_k \nabla_N \sigma_k(0)\right) = 0$$

And the Taylor expansion is performed assuming a small $\delta\tilde{\lambda}$ to be small and stopping at the first term (linearising the equation):

$$\sigma_l(\vec{r}_1^{(1)}, \ldots, \vec{r}_N^{(1)}) + \sum_{i=1}^{N}\sum_{k=1}^{N_C} \frac{1}{m_i} \nabla_i \sigma_l(\vec{r}_1^{(1)}, \ldots, \vec{r}_N^{(1)}) \nabla_i \sigma_k(\vec{r}_1(0), \ldots, \vec{r}_N(0)) \delta\tilde{\lambda}_k \approx 0$$

Now assuming that to be 0 the $\delta\tilde{\lambda}_k$ can be obtained inverting the resulting matrix. This will be a linear system that can be solved efficiently. So the matrix can be inverted or it can be simplified. This

guess is done iteratively to continuously obtain corrections on the guesses, making the constraints equations to become more and more accurate. The process will stop when the equation will be less than an arbitrary tolerance. This is important because if it is too strict this will require a lot of iterations, making the simulation too slow. If the tolerance is too high the constraints will not be satisfied any more and systems will explode, making the particles far apart, because the forces will be to high.

7.3.4.2 Possible algorithms

$$\sigma_l(\vec{r}_1^{(1)}, \ldots, \vec{r}_N^{(1)}) + \sum_{i=1}^{N} \sum_{k=1}^{N_C} \frac{1}{m_i} \nabla_i \sigma_l(\vec{r}_1^{(1)}, \ldots, \vec{r}_N^{(1)}) \nabla_i \sigma_k(\vec{r}_1(0), \ldots, \vec{r}_N(0)) \delta\tilde{\lambda}_k \approx 0$$

There are different techniques to solve this matrix.

7.3.4.2.1 Direct inversion In direct inversion method like Matrix-shake or M-SHAKE the matrix is inverted and the $\tilde{\lambda}$ are computed and this procedure must be repeated because the equation above is a linear approximation.

7.3.4.2.2 Quick trick In Quick trick only diagonal elements are considered without any matrix inversion.

7.3.4.2.3 RATTLE The same procedure can be employed for velocities like in RATTLE. The algorithm is the same thing as in Shake and the same thing is done for velocities, where the Lagrange multipliers for the velocities will be called $\tilde{\mu}$.

7.3.4.2.4 LINCS LINCS or linear constraint solver is based on the same principle and it is implemented in GROMACS. It is a slight variation on the Matrix-Shake.

Chapter 8

Direct translation

8.1 Introduction

The direct translation method is a formal representation of the Verlet and the Velocity Verlet algorithms. It has direct effect on those.

8.2 Liouville operator

The Liouville operator can be computed for any quantity a that depends on the coordinates in phase space. And its derivative is:

$$\frac{da}{dt} = \frac{\partial a}{\partial x_t}\dot{x}_t = \sum_a \left[\frac{\partial a}{\partial q_a}\dot{q}_a + \frac{\partial a}{\partial p_a}\dot{p}_a\right] = \sum_a \left[\frac{\partial a}{\partial q_a}\frac{\partial \mathcal{H}}{\partial p_a} - \frac{\partial a}{\partial p_a}\frac{\partial \mathcal{H}}{\partial q_a}\right] = \{a, \mathcal{H}\}$$

The iouville operator is defined as:

$$iLa = \{a, \mathcal{H}\} \Rightarrow \frac{da}{dt} = iLa$$

The operation of the Liouville operator is to compute the Poisson brackets of a with $\mathcal{H}$. So that the dependence on time of a can be written. In abstract terms the operator acts on something and computes the Poisson brackets with the Hamiltonian: $iL\cdots = \{\cdots, \mathcal{H}\}$. It can also be written as:

$$iL = \sum_a \left[\frac{\partial \mathcal{H}}{\partial q_a}\frac{\partial}{\partial q_a} - \frac{\partial \mathcal{H}}{\partial q_a}\frac{\partial}{\partial p_a}\right]$$

A formal solution to the equation can be written:

$$\frac{da}{dt} = iLa \Rightarrow a(x_t) = e^{\overbrace{iLt}^{\text{Classical propagator}}} a(x_0)$$

Obtaining the value of $a(x_t)$. It is called a classical propagator because it propagates a from time 0 to t.

8.2.1 Liouville operator split

This classical propagator is difficult to solve analytically so it needs to be approximated. The Liouville operator can be split in two operators:

$$iL = \sum_a \left[\frac{\partial \mathcal{H}}{\partial p_a} \frac{\partial}{\partial q_a} - \frac{\partial \mathcal{H}}{\partial q_a} \frac{\partial}{\partial p_a} \right] = iL_1 + iL_2$$

$$iL_1 = \sum_a \frac{\partial \mathcal{H}}{\partial p_a} \frac{\partial}{\partial q_a} \qquad iL_2 = -\sum_a \frac{\partial \mathcal{H}}{\partial q_a} \frac{\partial}{\partial p_a}$$

These two operators are non-commuting:

$$iL_1 iL_2 \phi(x) \neq iL_2 iL_1 \phi(x)$$

Formally, the commutator:

$$iL_1 iL_2 - iL_2 iL_1 \equiv [iL_1, iL_2]$$

If it is equal to zero the operators are commuting. In the case for the Liouville's operator split they are not commuting.

8.2.2 One dimensional example

To demonstrate that the two Liouville's operator are not commuting consider the Hamiltonian:

$$\mathcal{H} = \frac{p^2}{2m} + U(x)$$

So that the two operators are:

$$iL_1 = \frac{\partial \mathcal{H}}{\partial p} \partial \partial x = \frac{p}{m} \frac{\partial}{\partial x} \qquad iL_2 = -\frac{\partial \mathcal{H}}{\partial x} \frac{\partial}{\partial p} = -\frac{\partial U}{\partial x} \frac{\partial}{\partial p} = F(x) \frac{\partial}{\partial p}$$

Applying the two operators in sequence to a function $\phi(x, p)$. Here the operator iL_2 is applied first:

$$iL_1 iL_2 \phi(x,p) = \frac{p}{m} \frac{\partial}{\partial x} \left[F(x) \frac{\partial}{\partial p} \right] \phi(x,p) = \frac{p}{m} F(x) \frac{\partial^2 \phi}{\partial x \partial p} + \frac{p}{m} F'(x) \frac{\partial \phi}{\partial p}$$

Here the operator iL_1 is applied first:

$$iL_2 iL_1 \phi(x,p) = F(x) \frac{\partial}{\partial p} \left[\frac{p}{m} \frac{\partial}{\partial x} \right] \phi(x,p) = \frac{p}{m} F(x) \frac{\partial^2 \phi}{\partial x \partial p} + \frac{F(x)}{m} \frac{\partial \phi}{\partial x}$$

Computing the commutator:

$$[iL_1, iL_2] \phi(x,p) = \frac{p}{m} F'(x) \frac{\partial \phi}{\partial p} - \frac{F(x)}{m} \frac{\partial \phi}{\partial x} \neq 0$$

8.3 Trotter theorem

Since the two operators do not commute this equality is false:

$$iL_1 iL_2 \phi(x) \neq iL_2 iL_1 \phi(x) \Rightarrow e^{iLt} \neq e^{iL_1 t} e^{iL_2 t}$$

This is important because although the classical propagator cannot be written, in many situation the propagator corresponding to iL_1 and iL_2, so that the classical propagator could be obtained. However, according to the Trotter theorem:

$$e^{A+B} = \lim_{P\to\infty} [e^{\frac{B}{2P}} e^{\frac{A}{P}} e^{\frac{B}{2P}}]^P$$

Therefore, considering the Liouville's split:

$$e^{iLt} = e^{iL_1 t + iL_2 t} = \lim_{P\to\infty} [e^{\frac{iL_2 t}{2P}} e^{\frac{iL_1 t}{P}} e^{\frac{iL_2 t}{2P}}]^P$$

8.3.1 $\Delta t = \frac{t}{P}$

$$e^{iLt} = e^{iL_1 t + iL_2 t} = \lim_{P\to\infty} [e^{\frac{iL_2 t}{2P}} e^{\frac{iL_1 t}{P}} e^{\frac{iL_2 t}{2P}}]^P$$

This formula can be written as:

$$e^{iLt} = e^{iL_1 t + iL_2 t} = \lim_{P\to\infty, \Delta t\to 0} [e^{\frac{iL_2 \Delta t}{2}} e^{iL_1 \Delta t} e^{\frac{iL_2 \Delta t}{2}}]^P$$

Assuming that $\Delta t = \frac{t}{P}$. When a simulation there will be a finite value of P, which will be a natural value:

$$e^{iLt} \approx [e^{\frac{iL_2 \Delta t}{2}} e^{iL_1 \Delta t} e^{\frac{iL_2 \Delta t}{2}}]^P + \underbrace{o(P\Delta t^3)}_{\text{Global error}}$$

So that the classical propagator is equal to the term in the limit with an error due to the truncation to a finite value of P. Taking the $\frac{1}{P}$ power:

$$e^{iLt} \approx e^{\frac{iL_2 \Delta t}{2}} e^{iL_1 \Delta t} e^{\frac{iL_2 \Delta t}{2}} + \underbrace{o(\Delta t^3)}_{\text{Local error}}$$

For a small time step Δt. So following a trajectory for a huge number of step the error is $o(\Delta t^2)$. In this way an algorithm that comes from Trotter factorization can be built. This algorithm will have a limited global error of the order of $o(\Delta t^2)$, so it will have a strong numerical stability.

8.3.2 One-dimensional example

In one-dimensional example everything can be written as:

$$iL_1 = \frac{\partial \mathcal{H}}{\partial p} \partial\partial x = \frac{p}{m} \frac{\partial}{\partial x} \qquad iL_2 = F(X) \frac{\partial}{\partial p}$$

So the Trotter algorithm can be written as:

$$e^{iLt} \approx e^{\frac{iL_2 \Delta t}{2}} e^{iL_1 \Delta t} e^{\frac{iL_2 \Delta t}{2}} + o(\Delta t^3)$$

And making everything explicit:

$$e^{iL\Delta t} \approx e^{\frac{\Delta t}{2}F(x)\frac{\partial}{\partial p}} e^{\Delta t \frac{p}{m}\frac{\partial}{\partial x}} e^{\frac{\Delta t}{2}F(x)\frac{\partial}{\partial p}} + o(\Delta t^3)$$

Applying the operator to the coordinates in phase space:

$$\begin{pmatrix} x(\Delta t) \\ p(\Delta t) \end{pmatrix} = e^{iL\Delta t} \begin{pmatrix} x(\Delta t) \\ p(\Delta t) \end{pmatrix} \approx e^{\frac{\Delta t}{2}F(x)\frac{\partial}{\partial p}} e^{\Delta t \frac{p}{m}\frac{\partial}{\partial x}} e^{\frac{\Delta t}{2}F(x)\frac{\partial}{\partial p}} \begin{pmatrix} x(\Delta t) \\ p(\Delta t) \end{pmatrix}$$

8.3.3 Exponential operators

An exponential operator can be computed as:

$$e^{c\frac{\partial}{\partial x}} g(x) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(c\frac{\partial}{\partial x} \right)^k g(x) = \sum_{k=0}^{\infty} \frac{1}{k!} c^k \frac{d^k g(x)}{dx^k} = g(x+c)$$

This is another way to compute the Taylor expansion of g in $x + c$. So the operator translates the argument of g by a quantity c. Now applying the exponential operators to the points:

$$\begin{pmatrix} x(\Delta t) \\ p(\Delta t) \end{pmatrix} = e^{\frac{\Delta t}{2}F(x(0))\frac{\partial}{\partial p(0)}} e^{\Delta t \frac{p(0)}{m}\frac{\partial}{\partial x(0)}} e^{\frac{\Delta t}{2}F(x(0))\frac{\partial}{\partial p(0)}} \begin{pmatrix} x(0) \\ p(0) \end{pmatrix}$$

Computing the first exponential operator:

$$e^{\frac{\Delta t}{2}F(x(0))\frac{\partial}{\partial p(0)}} \begin{pmatrix} x(0) \\ p(0) \end{pmatrix} = \begin{pmatrix} x(0) \\ p(0) + \frac{\Delta t}{2}F(x(0)) \end{pmatrix}$$

Now doing the same for all the other operators:

$$\begin{aligned} \begin{pmatrix} x(\Delta t) \\ p(\Delta t) \end{pmatrix} &= e^{\frac{\Delta t}{2}F(x(0))\frac{\partial}{\partial p(0)}} e^{\Delta t \frac{p(0)}{m}\frac{\partial}{\partial x(0)}} \begin{pmatrix} x(0) \\ p(0) \\ p(0) + \frac{\Delta t}{2}F(x(0)) \end{pmatrix} = \\ &= e^{\frac{\Delta t}{2}F(x(0))\frac{\partial}{\partial p(0)}} \begin{pmatrix} x(0) + \Delta t \frac{p(0)}{m} \\ p(0) + \frac{\Delta t}{2}F(x(0) + \Delta t \frac{p(0)}{m}) \end{pmatrix} = \\ &= \begin{pmatrix} x(0) + \frac{\Delta t}{m}\left(p(0) + \frac{\Delta t}{2}F(x(0)) \right) \\ p(0) + \frac{\Delta t}{2}F(x(0)) + \frac{\Delta t}{m}\left(x(0) + \frac{\Delta t}{m}\left(p(0) + \frac{\Delta t}{2}F(x(0)) \right) \right) \end{pmatrix} \end{aligned}$$

8.4 Trotter algorithm

This is the result of Trotter algorithm:

$$\begin{pmatrix} x(\Delta t) \\ p(\Delta t) \end{pmatrix} = \begin{pmatrix} x(0) + \frac{\Delta t}{m}\left(p(0) + \frac{\Delta t}{2}F(x(0)) \right) \\ p(0) + \frac{\Delta t}{2}F(x(0)) + \frac{\Delta t}{m}\left(x(0) + \frac{\Delta t}{m}\left(p(0) + \frac{\Delta t}{2}F(x(0)) \right) \right) \end{pmatrix}$$

Looking at the first line:

$$x(\Delta t) = x(0) + \frac{\Delta t}{m}\left(p(0) + \frac{\Delta t}{2}F(x(0))\right) \Rightarrow x(\Delta t) = x(0) + v(0)\Delta t + \frac{\Delta t^2}{2m}F(0)$$

And considering the second line:

$$p(\Delta t) = p(0) + \frac{\Delta t}{2}F(x(0)) + \frac{\Delta t}{2}F\left(x(0) + \frac{\Delta t}{m}\left(p(o) + \frac{\Delta t}{2}F(x(0))\right)\right)$$

So that:

$$p(\Delta t) = p(0) + \frac{\Delta t}{2}[F(x(0)) + F(x(\Delta t))] \Rightarrow v(\Delta t) = v(0) + \frac{\Delta t}{2m}[F(0) + F(\Delta t)]$$

So that inside the square brackets it is an average of the forces at time 0 and at time Δt. The two final formulae are the one needed to update the positions and velocities of the system. These two formulas are the Velocity Verlet algorithm. This is an important result telling that the Velocity Verlet is the best algorithm for molecular dynamics.

8.4.1 Direct translation

There is a quick way to compute the results of the Trotter algorithm, which leads to direct translation. What there is to do is to work in a sloppy matter on the operators. In this way the formulae can be written directly into lines of code.

8.4.1.1 Momentum translation

Considering the first operator, it is translating the momentum by the quantity $\frac{\Delta t}{2}F(x)$. So this will translate p from time zero to time $\frac{\Delta t}{2}$. So when applying this operator p will change and x will remain the same.

$$e^{\frac{\Delta t}{2}F(x)\frac{\partial}{\partial p}}\begin{pmatrix}x(0)\\p(0)\end{pmatrix} = \begin{pmatrix}x(0)\\p\left(\frac{\Delta t}{2}\right) = p(0) + \frac{\Delta t}{2}F(x(0))\end{pmatrix}$$

Corresponding code:

```
p = p + 0.5 * Δt * F
```

8.4.1.2 Position translation

Applying the second operator only x will be translated:

$$e^{\frac{\Delta}{m}p\left(\frac{\Delta t}{2}\right)\frac{\partial}{\partial x}}\begin{pmatrix}x(0)\\p\left(\frac{\Delta t}{2}\right) = p(0) + \frac{\Delta t}{2}F(x(0))\end{pmatrix}$$

Corresponding code:

```
x = x + Δt * F * p/m
```

8.4.1.3 Momentum translation with updated forces

Now applying the third operator p will be translated by another $\frac{\Delta t}{2}$. Looking at this formula the same result of the Trotter theorem is obtained. In this third step also the forces are updated.

Corresponding code:

$$e^{\frac{\Delta t}{2}F(x)\frac{\partial}{\partial p}}\left(x(\Delta t)\left(\frac{\Delta t}{2}\right)\right) = \begin{pmatrix} x(\Delta t) \\ p(\Delta t) = p\left(\frac{\Delta t}{2}\right) + \frac{\Delta t}{2}F(x(\Delta t)) \end{pmatrix}$$

```
p = p + 0.5 * Δt * F
```

8.4.2 Multiple time step integration

The problem when studying a force field, there are several term that correspond to the bonded interactions and the non-bonded interactions:

$$U(\vec{r}_1,\ldots,\vec{r}_N) = \frac{1}{2}\sum_{bonds} K_{bond}(r-r_0)^2 + \frac{1}{2}\sum_{bends} K_{bend}(\theta-\theta_0)^2 + \sum_{torsions}\sum_{n=0}^{6} A_n[1+\cos(C_n\phi+\delta_n)] + \\ + \sum_{i,j\in nb}\left\{\left[4\epsilon_{ij}\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}}\right\}$$

It can be assumed that the atoms will be subject to forces with different time scales: the bonded interactions will change more rapidly than the non-bonded interactions. This is because the positions on the atoms do not change so much. It can be seen how the non-bonded interaction, which are the slow forces are very expensive to compute. So there is no reason to recompute the non-bonded interactions every time the fast forces are changing. So the Hamiltonian can be split into a reference Hamiltonian corresponding to the fast forces and another corresponding to the slow ones:

$$\mathcal{H}_{ref} = \frac{p^2}{2m} + U_{fast}(x) \qquad \dot{x} = \frac{p}{m} \qquad \dot{p} = F_{fast}(x) + F_{slow}(x)$$

In this case the Liouville;s operator will be:

$$iL = \frac{p}{m}\frac{\partial}{\partial x} + [F_{fast}(x) + F_{slow}(x)]\frac{\partial}{\partial p}$$

So that there will be a fast Liouville's operator and a slow one.

8.4.3 Reference system propagator

The reference system propagator is referred to the reference Hamiltonian. Considering:

$$iL = \frac{p}{m}\frac{\partial}{\partial x} + [F_{fast}(x) + F_{slow}(x)]\frac{\partial}{\partial p} = iL_{fast} + iL_{slow}$$

And:

$$\mathcal{H}_{ref} = \frac{p^2}{2m} + U_{fast}(x)$$

The reference system propagation for the fast forces and the slow part of the slow propagator:

$$iL_{fast} = \frac{p}{m}\frac{\partial}{\partial x} + F_{fast}(x)\frac{\partial}{\partial p} \qquad iL_{slow} = F_{slow}(x)\frac{\partial}{\partial p}$$

Applying Trotter factorisation:

$$e^{iL\Delta t} = e^{iL_{slow}\frac{\Delta t}{2}} e^{iL_{fast}\Delta t} e^{iL_{slow}\frac{\Delta t}{2}}$$

8.4.4 RESPA

RESPA is the reference system propagator algorithm, where Δt is chosen according to the time scale of the slow forces, allowing to use a much bigger Δt with respect to the one considering all the forces together. It comes from the split of the Hamiltonian into a slow and fast part. It allow to compute the electrostatic and Van der Waals forces less times than the calculation of the bonded interactions, saving a lot of computational power. Focussing on the fast part of the RESPA algorithm:

$$e^{iL_{fast}\Delta t} = \left[e^{\frac{\delta t}{2}F_{fast}\frac{\partial}{\partial p}} e^{\delta t \frac{p}{m}\frac{\partial}{\partial x}} e^{\frac{\delta t}{2}F_{fast}\frac{\partial}{\partial p}} \right]^n \qquad \delta t = \frac{\Delta t}{n}$$

This part can be factorized again using the fact that Δt is sub-divided in n intervals, where δt is the time-step where the fast forces will act. This will be applied n times so it corresponds to one time step. Writing down the entire classical propagator the fast operator is in-between the slow one.

$$e^{iL\Delta t} = e^{\frac{\Delta t}{2}F_{slow}\frac{\partial}{\partial p}} \left[e^{\frac{\delta t}{2}F_{fast}\frac{\partial}{\partial p}} e^{\delta t \frac{p}{m}\frac{\partial}{\partial x}} e^{\frac{\delta t}{2}F_{fast}\frac{\partial}{\partial p}} \right]^n e^{\frac{\Delta t}{2}F_{slow}\frac{\partial}{\partial p}}$$

And for a direct translation in pseudocode:

```
p = p + 0.5 * Δt * F_slow
for i = 1 to n
    p = p + 0.5 * *($\delta t$*) * F_fast
    x = x + δt * p/m
    Recompute only fast forces
    p = p + 0.5 * δt * F_fast
endfor
Recompute slow forces
p = p + 0.5 * δt * F_slow
```

8.5 Symplectic solver

By demonstrating the applicability of the Trotter algorithm, the symplectic property of the Verlet algorithm has been demonstrated. The Hamiltonian is not strictly conserved, however it conserves a shadow Hamiltonian $\tilde{\mathcal{H}}(x, \Delta t)$ that is close to the true Hamiltonian $\mathcal{H}(x)$ such that:

$$\lim_{\Delta t \to 0} \tilde{\mathcal{H}}(x, \Delta t) = \mathcal{H}(x)$$

This property guarantees that this algorithm is numerically stable for any amount of time.

Chapter 9

Evaluation of energies and forces

9.1 Periodic boundary conditions

$$U_{nb}(\vec{r}_1,\ldots,\vec{r}_N) = \sum_{i>j\in nb}\left\{4e_{ij}\left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^{6}\right] + \frac{q_i q_j}{r_{ij}}\right\}$$

Periodic boundary conditions are a Fourier representation. The idea is that any function that is long ranged in space will be short ranged in the reciprocal or Fourier space. The strategy is to divide en impera the long and short ranged contributions.

9.1.1 The error function

An error function is defined such that:

$$erf(x) = \frac{2}{\sqrt{\pi}}\int_o^x dte^{-t^2}$$

This has some properties:

- $\lim_{x\to\infty} erf(x) = 1.$
- $erf(0) = 0$

The complement error function is:

$$erfc(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}}\int_x^\infty dte^{-t^2}$$

Such that:

- $\lim_{x\to\infty} erfc = 0.$
- $erf(x) + erfc(x) = 1.$

$$\frac{1}{r} = \overbrace{\frac{erfc\alpha r}{r}}^{\text{short-ranged}} + \overbrace{\frac{erf(\alpha r)}{r}}^{\text{long-ranged}}$$

9.1.2 Divide et impera

$$U_{nb}(\vec{r}_1,\ldots,\vec{r}_N) = U_{short}(\vec{r}_1,\ldots,\vec{r}_N) + U_{long}(\vec{r}_1,\ldots,\vec{r}_n)$$

$$U_{short}(\vec{r}_1,\ldots,\vec{r}_N) = \sum_{\vec{S}} \sum_{i>j\in nb} \left\{ 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij,\vec{S}}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij,\vec{S}}} \right)^{6} \right] + \frac{q_i q_j erfc(\alpha r_{ij,\vec{S}})}{r_{ij,\vec{S}}} \right\}$$

$$U_{long}(\vec{r}_1,\ldots,\vec{r}_N) = \sum_{S} \sum_{i>j\in nb} \frac{q_i q_j erf(\alpha r_{ij,\vec{S}})}{r_{ij,\vec{S}}}$$

$$r_{ij,\vec{S}} = |\vec{r}_i - \vec{r}_j + \vec{S}| \qquad \vec{S} = \vec{m}L$$

So the original distribution of charges is equal to the sum in direct space of the screened electrostatic interactions and the sum in reciprocal space os the Gaussian clouds of charges.

Chapter 10

Canonical ensemble

Chapter 11

Thermostats

Chapter 12

The isobaric ensembles

Chapter 13

The grand canonical ensemble

Chapter 14

Quantifying uncertainties and sampling quality