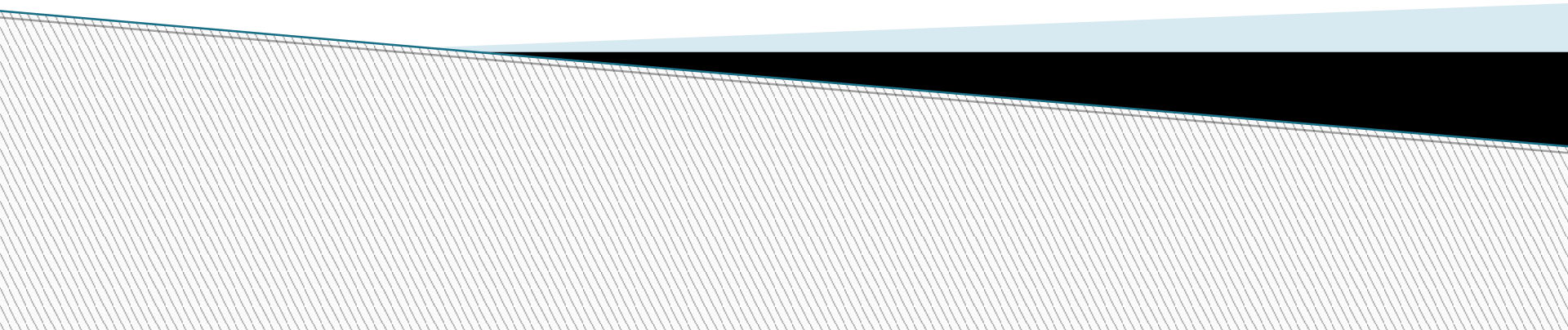


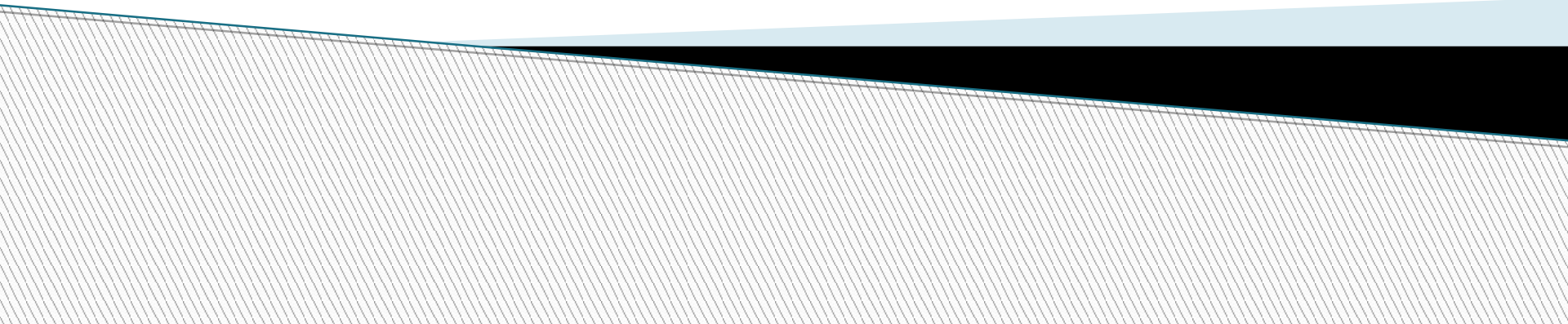
Molecular dynamics

Simulation methods

Matthys Mathieu & Vincent VonTammer



PART II



6. Time-dependant Properties

6.1 *Correlation functions*

- Quantifies the strenght of the correlation
- Commonly used :

$$C_{xy} = \frac{1}{M} \sum_{i=1}^M x_i y_i \equiv \langle x_i y_i \rangle$$

- If Normalised to a value between -1 and +1 :

$$c_{xy} = \frac{\frac{1}{M} \sum_{i=1}^M x_i y_i}{\sqrt{\left(\frac{1}{M} \sum_{i=1}^M x_i^2\right) \left(\frac{1}{M} \sum_{i=1}^M y_i^2\right)}} = \frac{\langle x_i y_i \rangle}{\sqrt{\langle x_i^2 \rangle \langle y_i^2 \rangle}}$$

- If x and y fluctuate about non-zero mean values, it is usual to consider just the fluctuating part :

$$c_{xy} = \frac{\frac{1}{M} \sum_{i=1}^M (x_i - \langle x \rangle)(y_i - \langle y \rangle)}{\sqrt{\left(\frac{1}{M} \sum_{i=1}^M (x_i - \langle x \rangle)^2 \right) \left(\frac{1}{M} \sum_{i=1}^M (y_i - \langle y \rangle)^2 \right)}} = \frac{\langle (x_i - \langle x \rangle)(y_i - \langle y \rangle) \rangle}{\sqrt{\langle (x_i - \langle x \rangle)^2 \rangle \langle (y_i - \langle y \rangle)^2 \rangle}}$$

- A useful formula that not require the mean values before the correlation coefficient :

$$c_{xy} = \frac{\sum_{i=1}^M x_i y_i - \frac{1}{M} \left(\sum_{i=1}^M x_i \right) \left(\sum_{i=1}^M y_i \right)}{\sqrt{\left[\sum_{i=1}^M x_i^2 - \frac{1}{M} \left(\sum_{i=1}^M x_i \right)^2 \right] \left[\sum_{i=1}^M y_i^2 - \frac{1}{M} \left(\sum_{i=1}^M y_i \right)^2 \right]}}$$

6.1.1. *Time correlation coefficients*

$$C_{xy}(t) = \langle x(t)y(0) \rangle$$

$$\lim_{t \rightarrow 0} C_{xy}(0) = \langle xy \rangle$$

$$\lim_{t \rightarrow \infty} C_{xy}(t) = \langle x \rangle \langle y \rangle$$

- *Cross correlation function* : x and y are different
 - *Autocorrelation function* : x and y are the same
- ⇒ indicates how the system retains a 'memory' of its previous values.

6.1.2. An example : *velocity autocorrelation coefficient*

$$C_{vv}(t) = \frac{1}{N} \sum_{i=1}^N \mathbf{v}_i(t) \cdot \mathbf{v}_i(0)$$

- And normalized :

$$c_{vv}(t) = \frac{1}{N} \sum_{i=1}^N \frac{\langle \mathbf{v}_i(t) \cdot \mathbf{v}_i(0) \rangle}{\langle \mathbf{v}_i(0) \cdot \mathbf{v}_i(0) \rangle}$$

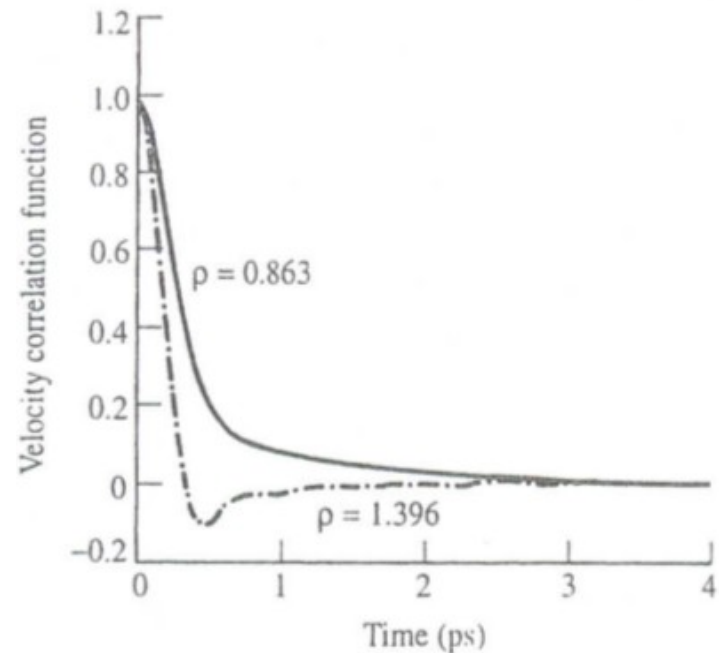
- Initial value of 1 and at long time, value of 0.
- The time to lose the correlation is the ***correlation time*** or the ***relaxation time***

6.1.3. *Relaxation time*

- P steps required for complete relaxation ;
- Simulation run for Q steps ;
- Then (Q - P) different sets of values could be used to calculate a value for the correlation function.
- If one uses M time origins (t_j) then :

$$C_{vv}(t) = \frac{1}{MN} \sum_{j=1}^M \sum_{i=1}^N \mathbf{v}_i(t_j) \cdot \mathbf{v}_i(t_j + t)$$

- An example : argon



- NB : the velocity autocorrelation function is a single-particle correlation function : the average is calculated over time origins and also over all atoms
- An example of a propertie calculated for the entire system is the net dipole moment of the system. This is the *total dipolar correlation function* :

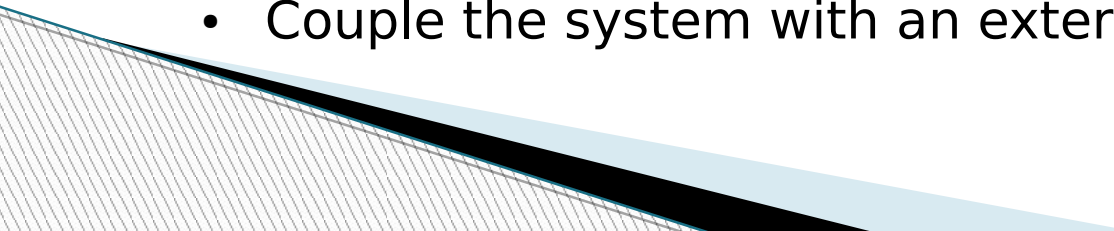
$$c_{\text{dipole}}(t) = \frac{\langle \mu_{\text{tot}}(t) \cdot \mu_{\text{tot}}(0) \rangle}{\langle \mu_{\text{tot}}(0) \cdot \mu_{\text{tot}}(0) \rangle}$$

$$\mu_{\text{tot}}(t) = \sum_{i=1}^N \mu_i(t)$$

7. Molecular Dynamics at Constant Temperature and Pressure

- Molecular dynamics traditionally performed in the constant NVE (or NVEP) ensemble ;
- But many experimental measurements made under isothermal-isobaric conditions
⇒ simulations under isothermal-isobaric conditions are often more relevant.
- Alternatives :
 - constant NVT ensemble ;
 - constant NPT ensembles.

7.1. Constant Temperature Dynamics

- Useful to determine how the system changes with temperature, such as the unfolding of a protein
 - Use of simulated annealing algorithms :
 - For searching conformational space and ;
 - Elucidation of macromolecular structure from NMR and X-ray data.
 - Two ways to control the temperature :
 - Multiply the velocities at each time step by a scaling factor or ;
 - Couple the system with an external heat bath.
- 

- Two ways → two different scaling factors :

$$\lambda = \sqrt{T_{\text{new}}/T(t)} \quad \lambda^2 = 1 + \frac{\delta t}{\tau} \left(\frac{T_{\text{bath}}}{T(t)} - 1 \right)$$

- Advantage of the 2nd approach : it allows the system to fluctuate about the desired temperature
- Both methods can lead to phenomenom of 'hot solvent, cold solute'

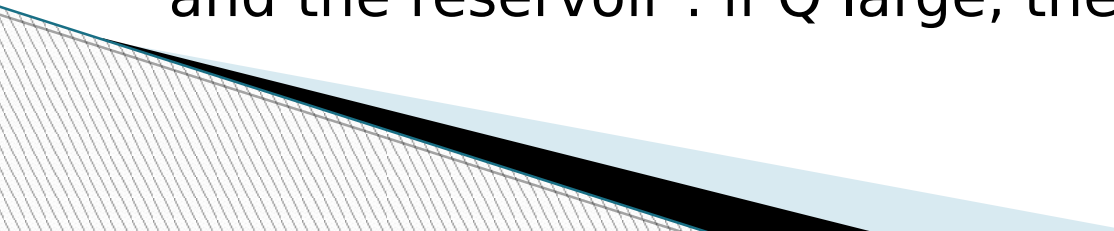
⇒ apply temperature coupling separatly to the solute and to the solvent BUT problem of unequal distribution of energy still remains

⇒ Two methods : *stochastic collisions* method and *extended system* method

7.1.1. Stochastic collisions method

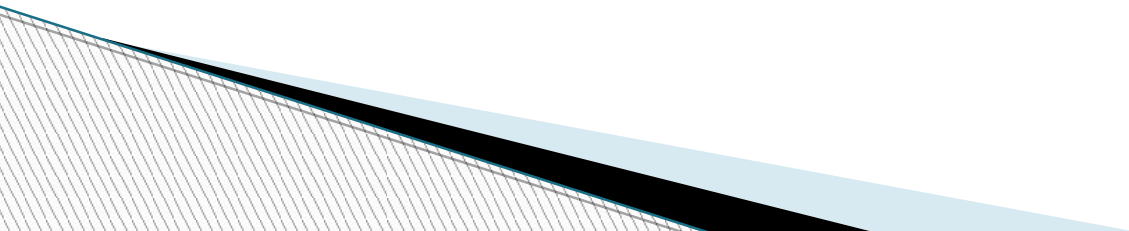
- Particles are randomly chosen and their velocities are reassigned from the Maxwell-Boltzmann distribution.
- Equivalent to the system being in contact with a heat bath that randomly emits 'thermal particles'.

7.1.2. Extended system methods

- Consider the thermal reservoir to be an integral part of the system.
 - The reservoir is represented by an additional degree of freedom with potential and kinetic energies.
 - The parameter Q (E/t^2 dimension) involved in the kinetic energy controls the energy flow between the system and the reservoir : if Q large, the energy flow is slow.
- 

7.2. Constant Pressure Dynamics

- Some structural rearrangements achieved more easily in isobaric simulations than in simulations at constant volume.
- Also useful if the numbers of particles in the system change.
- One maintains constant pressure by :
 - Scaling the volume or ;
 - Coupling the system with a pressure bath.

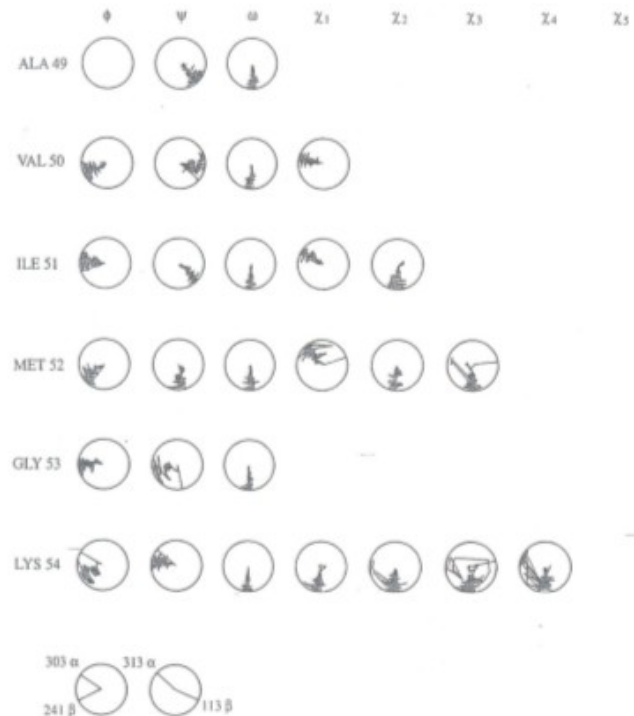


7.2.1. Extended pressure-coupling system methods

- Extra degree of freedom : volume of the box.
- The kinetic energy associated can be considered as piston of mass Q acting on the system :
 - If Q small, the system oscillate ;
 - If Q infinite, normal molecular dynamics behavior.

8. Conformational Changes from Molecular Dynamics Simulations

- The most direct way to show a conformational behavior is saving coordinates a regular interval and displaying it in sequence (graphs for exemple)
- Representation of bond rotations is difficult using x/y plots due to the 2π periodicity of a torsion angle \Rightarrow polar plot

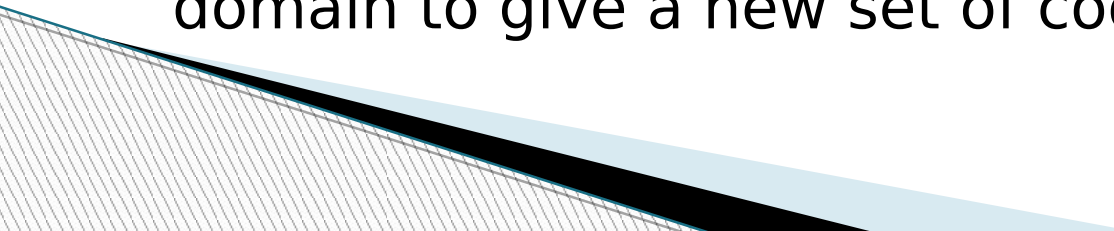


- The most direct way to show a conformational behavior is by saving coordinates at regular interval and displaying it in sequence (graphs for exemple)
- Representation of bond rotations is difficult using x/y plots due to the 2π periodicity of a torsion angle \Rightarrow polar plot

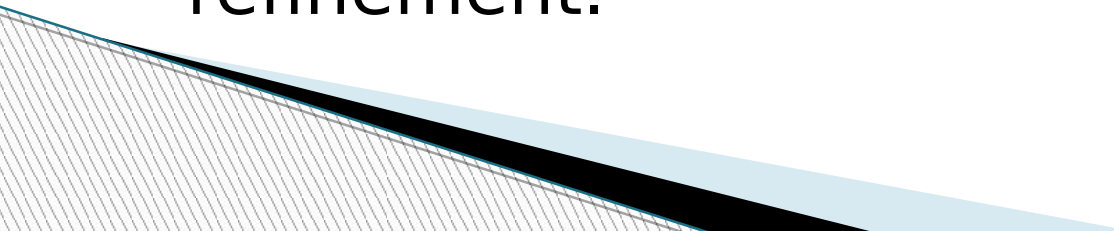
8.1. Fourier Analysis Techniques

- Motions of complex molecules are chaotic BUT underlying low-frequency motions which may correspond to conformational changes.
⇒ Fourier analysis techniques to filter out unwanted high-frequency motions.
- Fourier analysis introduced by *Fourier series* : considering functions that vary in a periodic manner with time as a superposition of sine and cosine functions
- BUT rarely relevant because the movements of the atoms are not periodic but chaotic
⇒ *Fourier transform*

8.1.1. Fourier Transform

- A *Fourier transform* enables to represents $x(t)$ as a function $X(v)$.
 - Can be developed from the Fourier series by period to infinity.
 - The frequency function is continuous.
 - At each step, for each Cartesian coordinates of each atoms, the variation with time is converted into a frequency function.
 - Then, the frequency spectrum is filtered.
 - Finally the spectrum is converted back to the time domain to give a new set of coordinate values.
- 

9. Solving Protein Structures Using Restrained Molecular Dynamics and Simulated Annealing

- Molecular dynamics used in conjunction with simulated annealing method to refine experimental data.
 - Use of *penalty functions* added to the potential energy function \Rightarrow Restrained MD.
 - Simulated annealing useful to explore the conformational space.
 - Use in X-ray crystallographic and NMR data refinement.
- 

9.1. X-ray Crystallographic Refinement

- Obtain a structure that fits the better with experimental data.
- Traditionally, least-square methods.
- By gradually changing the structure .
- The degree of agreement is quantified by the value of the cristallographic R factor :

$$R = \frac{\sum ||F_{\text{obs}}| - |F_{\text{calc}}||}{\sum |F_{\text{obs}}|}$$

- Where F is the structure factor :

$$F = |F|e^{i\Phi}$$

- In restrained MD approach, the total 'potential energy' is written as the sum of the usual potential energy and the penalty term E_{sf} .

$$E_{sf} = S \sum ||F_{obs}| - |F_{calc}||^2$$

- S is a scaling factor chosen so that the gradient of E_{sf} is comparable to the gradient of the potential energy part of the function.
- The conformational space is explored with simulated annealing and the temperature is gradually reduced as the structure settles into a conformation which has a low energy and a low R factor.

9.2. Molecular Dynamics Refinement of NMR Data

- Different ways to penalise a structure :
 - *Harmonic restraints* terms of the form $k(d-d_0)^2$ where k is a force constant (determine the tightness of the restraint) ;
 - *Torsional restraints* terms

9.2.1. Harmonic restraints

a.

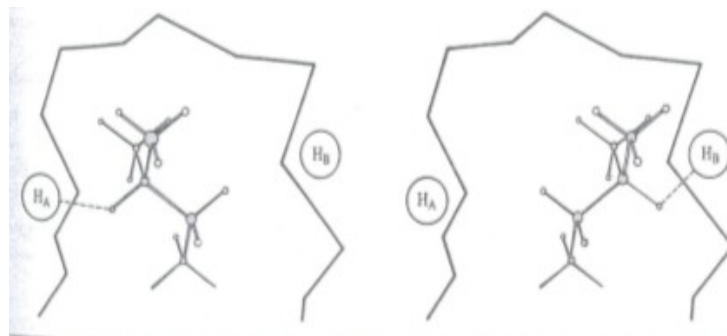
$$\begin{aligned} v(d) &= k(d - d_0)^2 & d > d_0 \\ v(d) &= 0 & d \leq d_0 \end{aligned}$$

b.

$$\begin{aligned} v(d) &= k_l(d - d_l)^2 & d < d_l \\ v(d) &= 0 & d_l \leq d \leq d_u \\ v(d) &= k_u(d - d_u)^2 & d_u < d \end{aligned}$$

9.3. Time-averaged NMR Refinement

- If the molecule swap between two conformations on a timescale more rapid than the chemical shift timescale, one obtain an average signal.
- Two sets of restraints can be derived and a standard refinement procedure would attempt to satisfy both sets of restraints simultaneously
- \Rightarrow lead to a conformation positioned at the top of barrier between the two minima.
- Time-averaged restraints overcomes this problem : it uses a value averaged over time rather than instantaneous value of a distance.



- One drawback of any restraint method : additional penalty terms represent an unnatural perturbation of the forces within the molecule.
- When using 'static' restraints the size of the force constants for the restraint terms can be large
- \Rightarrow can cause the conformations to have rather high energies

