Molecular Dynamics

Part 2: MD algorithms

Tadeusz Andruniów

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!

Wrocław University of Technology

EUROPEAN SOCIAL FUND

Project co-financed from the EU European Social Fund

Characteristics of MD

- Molecular dynamics allows for simulating the motion of each particle as a function of time by using equations of motion and functional form of potential energy (force field)
- Equations of motion are integrated using discreet time steps (total time length is divided into small time steps)
- To use the finite time-step method the potential must be a continous function of particle positions (particles must change positions smoothly with time)
  - If the potential varies sharply there will be a discontinuity in velocities

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Characteristics of MD

- MD is (formally) deterministic - given a state of the system at the present time t_0 it is possible to determine the state of the system in the future
- What about chaos?
  - Lyapunov instability - any integration error, will cause the generated trajectory to diverge from the 'correct' trajectory
- MD integrators assume that velocites and accelerations are constant over a given time step

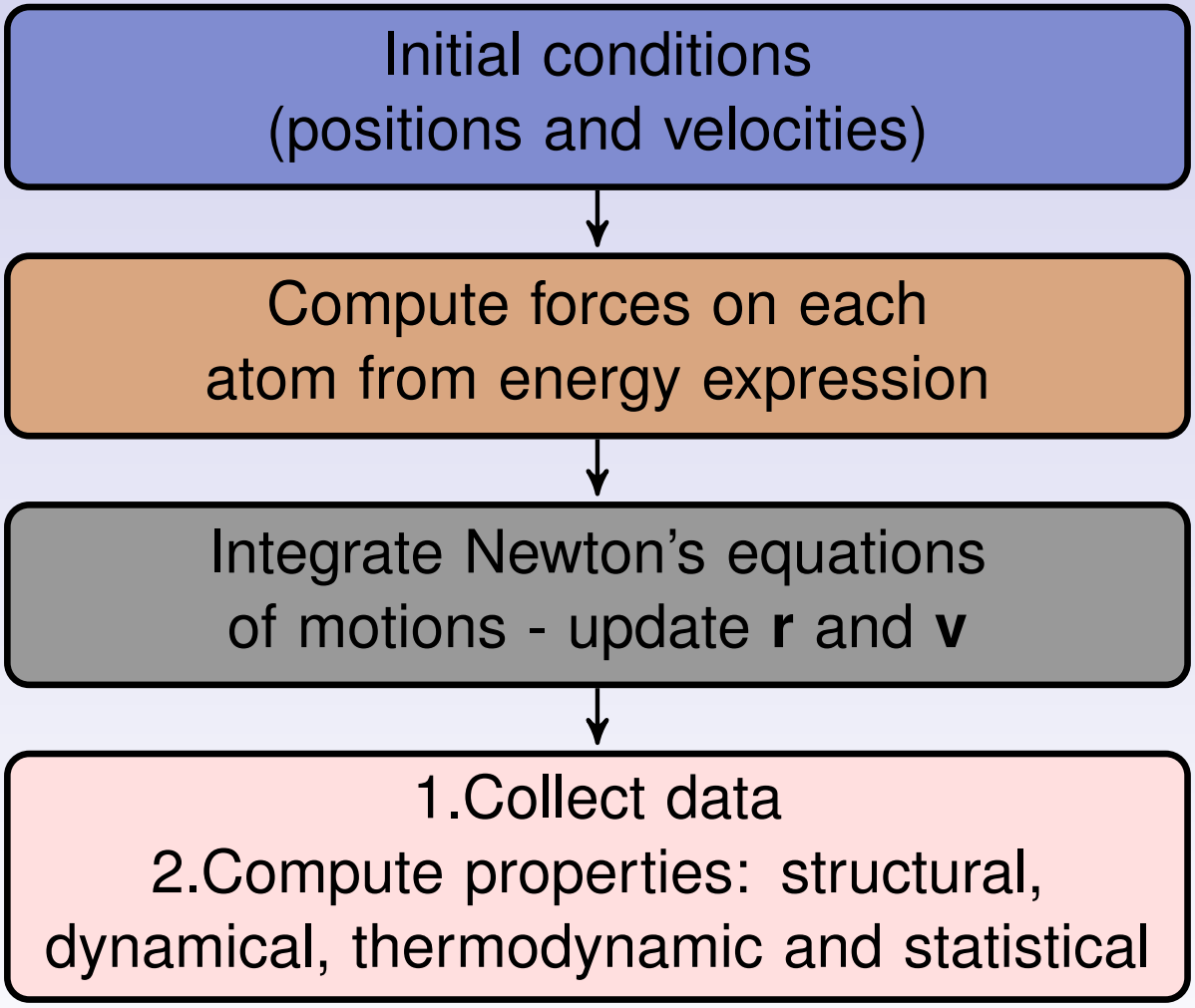
HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Characteristics of MD

- All particles move in time according to the laws of Newtonian mechanics
  - Classical mechanical treatment of nuclei movement
  - Quantum effects are neglected
- Random initial distribution of particles
- Every new distribution is derived from an old one
- A sequence of distributions of particles formed as time advances = trajectory

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Steps in MD simulations



HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Newton's formalism

System of N particles can be described by 3N independent generalized coordinates q_i and 3N velocities $\dot{q}_i$

$$\mathbf{F}_i = m_i \mathbf{a}_i = m_i \ddot{\mathbf{r}}_i$$

$$\mathbf{F}_i = -\frac{\partial}{\partial \mathbf{r}_i} U = \nabla U(\mathbf{r}_i)$$

$$\begin{cases} \dot{\mathbf{r}}_i = \dfrac{\mathbf{p}_i}{m_i} \\ \dot{\mathbf{p}}_i = \mathbf{F}_i \end{cases}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Lagrange formulation

- Lagrange function is the difference between kinetic and potential energies:

$$\mathcal{L}(\mathbf{r}, \dot{\mathbf{r}}) = T(\mathbf{r}, \dot{\mathbf{r}}) - U(\mathbf{r})$$

$$\mathcal{L} = \frac{1}{2} \sum_{i=1}^{3N} m_i \dot{\mathbf{r}}_i^2 - U(\mathbf{r}_i)$$

- Lagrange function is independent of the coordinate system

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Lagrange formulation

- Equations of motion of a system in the Lagrangan form are:

$$\frac{d}{dt}\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}} - \frac{d\mathcal{L}}{d\mathbf{r}} = 0$$

- Equation of motion has the form of a single-order differential equation

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Hamilton formulation

Hamilton function is the sum of kinetic and potential energies:

$$\mathcal{H} = T + U$$

Equations of motion can be derived from Hamiltonian:

$$\begin{cases} -\dfrac{\partial \mathcal{H}}{\partial \mathbf{r}_i} = \dot{\mathbf{p}}_i \\ \dfrac{\partial \mathcal{H}}{\partial \mathbf{p}_i} = \dot{\mathbf{r}}_i \end{cases}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Comparison of different formalisms

$$\mathcal{H} = \sum_i \dot{\mathbf{r}}_i \frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i} - \mathcal{L}$$

- Total energy, linear (and angular) momentum are conserved.
- In Hamilton formulation a coupled set of first-order differential equations is solved wheras in Lagrange formalism it is a single second-order differential equation.
- In both Hamilton and Lagrange formulations any set of nonredundant variables can be used (in Newtonian only Cartesian spatial coordinates and corresponding velocities)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Integration schemes

In all integration schemes (if we assume that a classical trajectory is continuous) positions, velocities, accelerations and higher order derivatives are expanded in Taylor series (for any continuous and differentiable funtion):

$$\mathbf{r}_i(t+\delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{1}{2!}\mathbf{a}_i(t)\delta t^2 + \frac{1}{3!}\mathbf{b}_i(t)\delta t^3 + \frac{1}{4!}\mathbf{c}_i(t)\delta t^4 + \ldots$$

$$\mathbf{v}_i(t+\delta t) = \mathbf{v}_i(t) + \mathbf{a}_i(t)\delta t + \frac{1}{2!}\mathbf{b}_i(t)\delta t^2 + \frac{1}{3!}\mathbf{c}_i(t)\delta t^4 + \ldots$$

$$\mathbf{a}_i(t+\delta t) = \mathbf{a}_i(t) + \mathbf{b}_i(t)\delta t + \frac{1}{2!}\mathbf{c}_i(t)\delta t^2 + \ldots$$

$$\mathbf{b}_i(t+\delta t) = \mathbf{b}_i(t) + \mathbf{c}_i(t)\delta t + \ldots$$

$$\underbrace{\mathbf{v} = \frac{d\mathbf{r}}{dt}}_{\textit{velocity}};\ \underbrace{\mathbf{a} = \frac{d^2\mathbf{r}}{dt^2}}_{\textit{acceleration}};\ \underbrace{\mathbf{b} = \frac{d^3\mathbf{r}}{dt^3}}_{\textit{jerk}};\ \underbrace{\mathbf{c} = \frac{d^4\mathbf{r}}{dt^4}}_{\textit{snap (jounce)}}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Integration schemes

- From Taylor-based equations it is possible to estimate new values of positions, velocities and accelerations from one time step to another
- Taylor-based equations will not generate correct trajectories as time advances!
- To do that one has to correct predicted quantities
- Equations of motion have to be introduced, e.g. through a predictor-corrector algorithm
- From the predicted new positions r(t+δt), having the equation of motion, the forces and then correct accelerations are calculated at time t+δt

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Desired features of good algorithms

- $\Rightarrow$ fast
- $\Rightarrow$ small memory and storage requirements
- $\Rightarrow$ easy to implement
- $\Rightarrow$ time-reversible
- $\Rightarrow$ area-preserving (symplectic)
- $\Rightarrow$ reveal small divergence from exact classical trajectory
- $\Rightarrow$ conserve energy and momentum
- $\Rightarrow$ good stability when using large time steps δt

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Euler

Euler algorithm uses positions, velocities and accelerations at time t to compute new positions at time t+δt

$$\mathbf{r}_i(t+\delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{1}{2}\dot{\mathbf{v}}_i\delta t^2$$

$$\mathbf{v}_i(t+\delta t) = \mathbf{v}_i(t) + \dot{\mathbf{v}}_i(t) + \dot{\mathbf{v}}_i(t)\delta t^2$$

- Recently Euler is NEVER used due to energy drifts

- It is no area-preserving and not time reversible

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Verlet algorithm

- Verlet uses positions and acceleration at time step t and previous step t-δt to update positions at t+δt
- Taylor expansion around r_i(t) can be written as:

$$\mathbf{r}_i(t+\delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{1}{2}\dot{\mathbf{v}}_i(t)\delta t^2 + \frac{1}{6}\ddot{\mathbf{v}}_i(t)\delta t^3 + O(\delta t^4)$$

$$\mathbf{r}_i(t-\delta t) = \mathbf{r}_i(t) - \mathbf{v}_i(t)\delta t - \frac{1}{2}\dot{\mathbf{v}}_i(t)\delta t^2 + \frac{1}{6}\ddot{\mathbf{v}}_i(t)\delta t^3 + O(\delta t^4)$$

- Add equations for r_i(t+δt) and r_i(t-δt) and rearrange:

$$\mathbf{r}_i(t+\delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\delta t) + \ddot{\mathbf{r}}_i(t)\delta t^2 + O(\delta t^4)$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Verlet algorithm

⇒ Velocity is not used for propagation but can be calculated as:

$$\mathbf{v}_i(t + \frac{1}{2}\delta t) = \frac{\left[\mathbf{r}_i(t + \delta t) - \mathbf{r}_i(t)\right]}{\delta t} \quad or$$

$$\mathbf{v}_i(t) = \frac{\left[\mathbf{r}_i(t + \delta t) - \mathbf{r}_i(t - \delta t)\right]}{2\delta t}$$

⇒ Velocities are not needed to generate trajectories but are useful to calculate kinetic energy

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Features of Verlet integrator

- $\Rightarrow$ The truncation error is of order $O(\delta t^4)$ for positions update while velocities are subject to errors $O(\delta t^2)$
- $\Rightarrow$ time-reversible
- $\Rightarrow$ good energy conserving properties even for large time steps
- $\Rightarrow$ numerically imprecise as small term $O(\delta t^2)$ is added to much larger term $O(\delta t^0)$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Leap-frog

First calculate the velocities at time t+1/2δt:

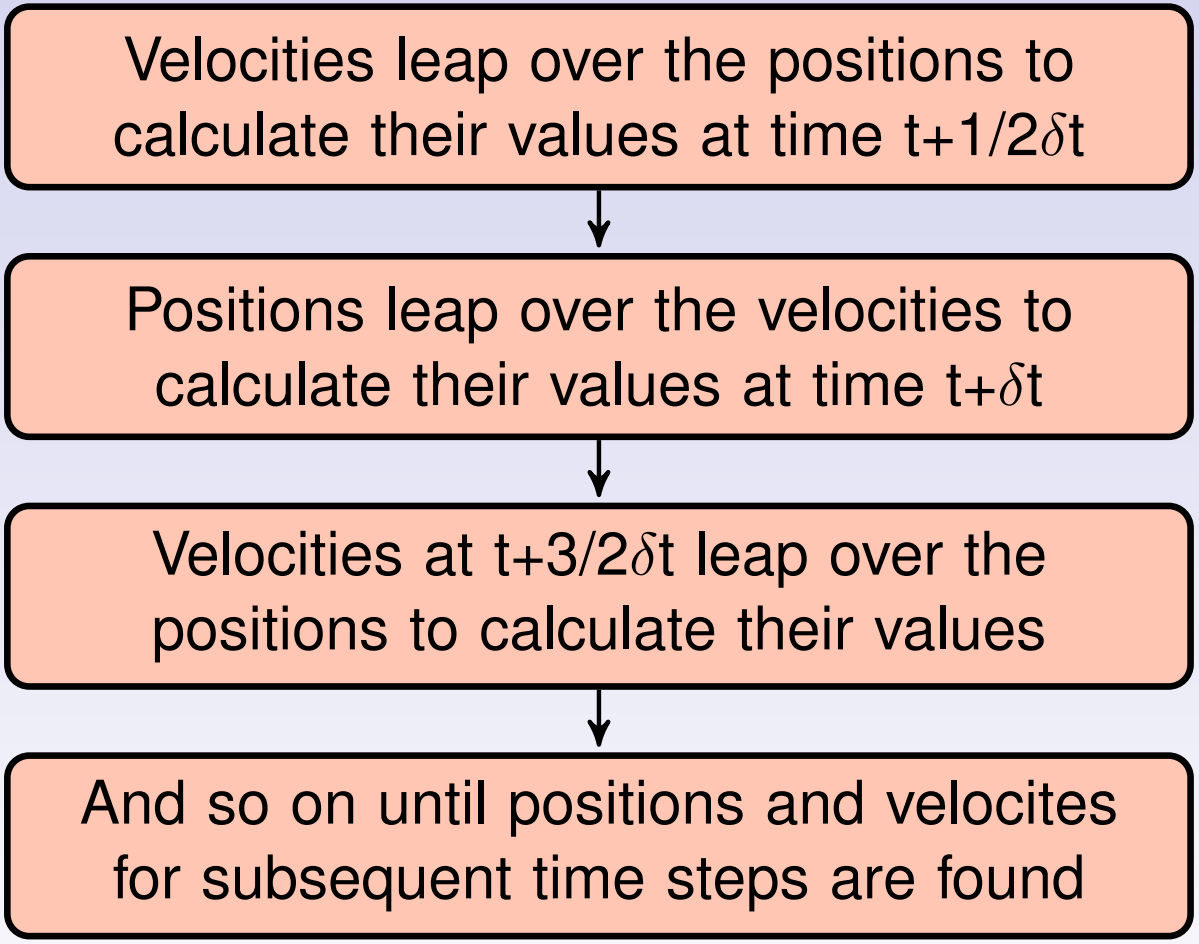
$$\mathbf{v}_i(t + \frac{1}{2}\delta t) = \mathbf{v}_i(t - \frac{1}{2}\delta t)\delta t + \dot{\mathbf{v}}_i(t)\delta t$$

use velocities at time t+1/2δt to calculate positions at time t+δt:

$$\mathbf{r}_i(t + \delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t + \frac{1}{2}\delta t)$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Leap-frog scheme



number of time steps=time length of simulations divided by time step δt

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Leap-frog

- Velocities are not calulated at the same time as positions (velocities are known accurately at t+1/2δt while positions at t+δt)
- However, velocities at time t can be estimated from:

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

- Improvements over standard Verlet:
  - $\Rightarrow$ more accurate (no calculation of differences between large and small numbers)
  - $\Rightarrow$ the same memory requirements as Verlet
  - $\Rightarrow$ velocites are explicitly calculated in leap-frog scheme

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Velocity Verlet

Velocity Verlet yields positions, velocities and accelerations at the same time t, thus minimizing round-off errors.
Calculate positions at t+δt:

$$\mathbf{r}_i(t+\delta t) = \mathbf{r}_i(t) + \mathbf{v}_i(t)\delta t + \frac{1}{2}\dot{\mathbf{v}}_i(t)\delta t^2$$

Calculate velocities at t+1/2δt:

$$\mathbf{v}_i(t+\frac{1}{2}\delta t) = \mathbf{v}_i(t) + \frac{1}{2}\dot{\mathbf{v}}_i\delta t$$

Calculate accelerations at t+1/2δt using previously calculated values of $\mathbf{r}_i$(t+δt) and $\mathbf{v}_i$(t+1/2δt)
Calculate velocities at t+δt:

$$\mathbf{v}_i(t+\delta t) = \mathbf{v}_i(t) + \frac{1}{2}\left[\dot{\mathbf{v}}_i(t) + \dot{\mathbf{v}}_i(t+\delta t)\right]\delta t$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Velocity Verlet

$\Rightarrow$ Truncation errors are the same as in standard Verlet

$\Rightarrow$ More accurate expressions for velocities are used by Beeman algorithm ($O(\delta t^3)$) and Velocity-corrected Verlet algorithm ($O(\delta t^4)$)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Predictor-Corrector

1. Predict $\mathbf{r}_i$, $\mathbf{v}_i$ and $\mathbf{a}_i$ at a time t+δt using the current values of these quantities at a time t (from Taylor expressions)
2. Calculate forces, and via equations of motion accelerations, from new positions $\mathbf{r}$(t+δt)
3. Compare values of $\mathbf{a}$(t+δt) obtained from both steps:

$$\Delta\mathbf{a}(t+\delta t) = \mathbf{a}^{C}(t+\delta t) - \mathbf{a}^{Taylor}(t+\delta t)$$

4. Correct positions, velocities, accelerations and higher order terms using new accelerations $\mathbf{a}$(t+δt)
5. return to point 1 for the next step

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Predictor-corrector

$$\mathbf{r}^c(t+\delta t) = \mathbf{r}^{Taylor}(t+\delta t) + c_0\,\Delta\mathbf{a}(t+\delta t)$$
$$\mathbf{v}^c(t+\delta t) = \mathbf{v}^{Taylor}(t+\delta t) + c_1\,\Delta\mathbf{a}(t+\delta t)$$
$$\mathbf{a}^c\frac{(t+\delta t)}{2} = \mathbf{a}^{Taylor}\frac{(t+\delta t)}{2} + c_2\,\Delta\mathbf{a}(t+\delta t)$$
$$\mathbf{b}^c\frac{(t+\delta t)}{6} = \mathbf{b}^{Taylor}\frac{(t+\delta t)}{6} + c_3\,\Delta\mathbf{a}(t+\delta t)$$

coefficients determined to maximize the stability of algorithm
c_0=1/6, c_1=5/6, c_2=1 and c_3=1/3 (Gear)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Predictor-corrector

- both positions and velocities corrected to δt^4
- accurate for small δt (Verlet is more accurate for large δt)
- not time reversible and not area preserving
- time consuming (two force evaluations per step vs. single force evaluation per step for Verlet)
- high memory requirements(15N vs. 9N for Verlet)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Time step in MD

To simulate physical and/or biological phenomena occuring on very long time scales one needs to use time step as large as possible.

Dilemma?

- $\Rightarrow$ a too small time step increses the cpu time of simulations and prohibits investigations of many long-time processes
- $\Rightarrow$ a too large time step results in less accurate (large forces) and less stable (large forces=small interparticles distances) integration procedures

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Time step in MD

- There is always a compromise between the accuracy and the value of the time step.
- The largest value of δt is determined by the fastest process taking place in the system.
  - $\Rightarrow$ The fastest motion in most molecular systems is due to vibrational stretching of C-H bonds which has a period of 10 fs (ca. 3000 cm^{-1})
- For molecular systems δt should be at least 10-fold smaller than the time of the fastest motion δt<1 fs.
- Time for sampling should be ca. 10 times longer than the timescale of the investigated process.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Accuracy vs. cpu time

1. reduce truncation errors by using higher order integration schemes resulting in the possibility of choosing a larger time step to achieve the same accuracy
2. freeze high-frequency modes (for example C-H stretching modes) which allows to increase time step (SHAKE and RATTLE algorithms)
3. use different time steps for different parts of the system oscillating on different time scales (one can also split forces into short and long-range to which different time step may be applied)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Multiple time-step

- The state of the system at a time t is given by:

$$\Gamma(t) = e^{iLt}\Gamma(0)$$

  - $\Gamma(\mathbf{r}, \mathbf{p}, t)$ is a classical mechanical distribution function
  - e^{iLt} is a classical time evolution operator

- Liouville operator for a molecular system consisted of N atoms in Cartesian coordinates (3N coordinates) is defined as:

$$iL = \sum_{i=1}^{3N} \left[\dot{\mathbf{r}}_i \frac{\partial}{\partial \mathbf{r}_i} + \mathbf{F}_i \frac{\partial}{\partial \mathbf{p}_i} \right]$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Multiple time-step

- F(x) can be defined as:

$$F(\mathbf{r}) = F_{stretch}(\mathbf{r}) + F_{bend}(\mathbf{r}) + F_{dihedral}(\mathbf{r}) + F_{vdw}(\mathbf{r}) + F_{electrostatic}(\mathbf{r})$$

- nonbonded F(**r**) (F_{nb}) can be divided into (e.g. through switching function):

$$F_{nb}(\mathbf{r}) = F_{nb}^{short} + F_{nb}^{long}$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Multiple time-step

- Liouville operator can be decomposed into parts:

$$L = L_1 + L_2 + ...$$

- and then each Liouville term is associated with specific force, e.g. F_{nb}^{short} with L_1 and F_{nb}^{long} with L_2
- let δt_1 be the time step employed for calculation of short-range non-bonded forces, then δt_2 which is the time step for long-range non-bonded forces can be written as:

$$\delta t_2 = n_1 \delta t_1$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Constrained dynamics

If the highest frequency motions are removed with holonomic constraints (SHAKE, RATTLE) the time step can be increased by a factor of 2
Constraints are introduced through Lagrangian (or Hamiltonian):

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}}{\partial \dot{\mathbf{r}}_i}\right) - \frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = \mathbf{g}_i$$

where the constraint force is given by:

$$\mathbf{g}_i = \sum_{\alpha=1}^{M} \Lambda_\alpha \left(\frac{\partial \chi_\alpha}{\partial \dot{\mathbf{r}}_i}\right)$$

Λ_α - undetermined Lagrange multiplier associated with α constraint

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Constrained dynamics

Constraints need to satisfy the following condition:

$$\begin{cases} \chi_1(\mathbf{r}) = 0 \\ \quad . \\ \quad . \\ \quad . \\ \chi_M(\mathbf{r}) = 0 \end{cases}$$

in case a bond has a fixed length χ will be definded as:

$$\chi(\mathbf{r}_1, \mathbf{r}_2) = (\mathbf{r}_1 - \mathbf{r}_2)^2 - b^2 = 0$$

where b is the fixed bond length between atoms 1 and 2

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

SHAKE algorithm

SHAKE uses the Verlet (leap-frog) algorithm for unconstrained particles:

$$\mathbf{r}_i'(t+\delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\delta t) + \frac{\delta t^2}{m_i}\mathbf{f}_i$$

after imposing constraints:

$$\mathbf{r}_i(t+\delta t) = 2\mathbf{r}_i(t) - \mathbf{r}_i(t-\delta t) + \frac{\delta t^2}{m_i}(\mathbf{f}_i - \mathbf{g}_i)$$

nonlinear equation for Λ multiplier has to be solved:

$$\mathbf{r}_i(t+\delta t) = \mathbf{r}_i'(t+\delta t) - \frac{\delta t^2}{m_i}\sum_{\alpha=1}^{M}\Lambda_\alpha\left(\frac{\partial\chi_\alpha(r(t))}{\partial\mathbf{r}_i}\right)$$

choosing Λ such that $\chi(\mathbf{r}(t+\delta t)) = 0$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

SHAKE vs. RATTLE

SHAKE cycles through constraints adjusting positions iterativelly until all constrained equations are satisfed (to within a specified tolerance)

RATTLE uses the Velocity Verlet algorithm and two sets of Lagrange multipliers

In SHAKE only positions are restrained, in RATTLE both positions and velocities

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Periodic boundary conditions

- In small systems surface effects can dominate
- The number of particles close to the surface is ca. $6N^{2/3}$ and the ratio of surface particles to bulk particles is $6N^{-1/3}$
  - for a system conisisting of 1000 water molecules ca. 60% of them will be at the surface (even for 10^6 water molecules surface effects are still non-negligible with a surface fraction of 6%)
- This will introduce artifacts

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Periodic boundary conditions

- It is not possible to simulate the number of atoms contained in macroscopic samples of substances (of the order of 10^{23} particles)
- In simulations the ratio between the boundary atoms and the total number of atoms will always be much larger than in reality resulting in much more pronounced surface effects
- A solution to the boundary effects caused by finite-size is to use periodic boundary conditions

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Issues with periodic boundary conditions

- Upon introducing periodic boundary conditions particles are placed in a box
- The box is replicated throughout space by translations in all directions
- The central box is surrounded by 26 copies of the original box
- When a particle leaves the central box its image (from the neighboring box) will simultaneously enter the central box from the other side
- The number of particles in the central box and thus in the entire system is conserved

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Issues with periodic boundary conditions

- Arbitrary shift of the cell boundaries has no effect on the system
- PBC eliminates surface effects - no boundaries=no surface problem!
- A particle from central box will interact not only with other particles in this box but also with their images in neighboring boxes
- As a result the number of pair-wise interactions vastly increases

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Issues with periodic boundary conditions

- PBC does not eliminate artifacts caused by a small number of particles
  - the simulation box should be large enough not to generate additional artifacts
- for some physical processes, e.g. adsorption onto a surface (motion perpendicular to the surface), standard PBC cannot be used in all directions
- PBC will not work properly if the cell size is smaller than the distance over which the interactions act (e.g. liquid-gas critical point)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

What should be the shape of the box?

- Any space-filling shape:
  - cube
  - hexagonal prism
  - truncated octahedron
  - rhombic dodecahedron
  - elongasted dodecahedron

It is suggested to choose the periodic cell that reflects most closely the shape of the system.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Minimum image convention

- only the closest image (each atom interacts with at most one image of every atom in the system) is considered in calculations of energy and forces
- interactions between all pairs of atoms (including images) that are further apart than the cut-off distance are set to zero
- the largest value of the cut-off may not exceed half the length of the box (when using a cubic box; for a rectangular box it should not be greater than half the length of the shortest side)
  - the particle cannot see its own image otherwise duplicate forces are calculated and simulations are erroneous

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Problems with minimum image convention

- the nearest image is not always energetically favourable
- splitting charged parts of molecules
- suppress long-range correlations
- new artificial correlations

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Look-up tables

- precompute the values of non-bonded potential for different distances and store them in the form of a table
- these values are then used during MD simulations to evaluate the potential through interpolation schemes

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Neighbor list methods

- Finding all interacting pairs is a time-consuming process for biomolecules ($O(N^2)$)
- Not all of these interactions are relevant and cut-off techniques and swtiching/shifting potentials are utilized
- However, it still may be a daunting task to calculate the list of neighbors
- Efficient algorithms, such as cell list and Verlet neighbor list, are utilized

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Verlet list

- define neighbor cut-off radius as potential cut-off radius and 'skin' radius ($r_L = r_c + r_{skin}$)
- update list whenever atoms move more than r_{skin}
- the list update is $O(N^2)$ so where is the computational-cost gain?
- particles are assumed to move slowly so the list will have to be updated only occasionally
- for solids the list may not need to be updated ever while for liquids very seldom
- large memory requirements (size of the list is $(4\pi(r_c+r_{skin})\cdot\rho\cdot N/3)$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Cell list

- divide the volume into a set of cells of equal size with an edge length of at least r_c
- prepare the list of atoms in each cell
- interaction is computed between particles in the same and adjacent cells
- as the interaction between particle i and j is the same as between particle j and i only half of the neighboring cells needs to be considered

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Cell list

- computational cost is reduced from $O(N^2)$ to $O(N)$
- with cells having an edge length of r_c the alogorithm is highly inefficient (84% of all calculated pair-wise interactions are unnecessary)
- choosing the cell edge-length of $r_c/2$ improves the efficiency (63% of spurious distance evaluations)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Ensembles in MD

In MD one is interested to study properties of systems as a function of temperature and pressure rather than volume and energy.

- microcanonical (NVE)
- grand-canonical (μVT)
- canonical (NVT)
- isothermal-isobaric (NPT)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Temperature in MD

Equipartition energy theorem relates the temperature to the average kinetic energy of the system:

$$< K >= \frac{3}{2} N k_B T$$

Instantenous temperature is expressed by:

$$K(t) = \frac{3}{2} N k_B T(t)$$

In some cases the number of degrees of freedom for N particles will be 3N-6 as center of mass velocity and angular momentum are constant

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Pressure in MD

For pair-wise potentials the pressure is related to the sum of the products of the particles positions and forces acting on them
The calculation of the pressure in MD is based on virial theorem of Clausius.

$$PV = \overbrace{Nk_BT}^{ideal\ gas} - \frac{1}{3}\left\langle \sum_{i}^{N}\sum_{j>i}^{N} \mathbf{r}_{ij}\mathbf{f}_{ij} \right\rangle$$

$\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$
and $\mathbf{f}_{ij}$ is the force on particle i due to particle j.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Controlling temperature

The velocity is simply scaled at each time step:

$$\mathbf{v}_i^{new} = c_t \mathbf{v}_i^{old}$$

where

$$c_t = \sqrt{\frac{T_0}{T}}$$

This procedure leads to rapid energy transfer between different degrees of freedom of the particle. Resetting of veloticites may be required from time to time.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Berendsen thermostat

To maintain the temperaure the system is coupled weakly to an external heat bath with a fixed temperature (T_0).

The equation of motion is modified to obtain the desired temperature

$$\dot{\mathbf{r}}_i(t) = \dot{\mathbf{v}}_i(t)$$

$$m_i\dot{\mathbf{v}}_i(t) = \dot{\mathbf{F}}_i(t) - \gamma(t)m_i\dot{\mathbf{v}}_i(t)$$

$$\gamma(t) = \frac{1}{2\tau}\left(1 - \frac{T_0}{T(t)}\right)$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Berendsen thermostat

The velocities are scaled at each step by the scaling factor c_t:

$$c_t = \sqrt{1 - \frac{\delta t}{\tau}\left(1 - \frac{T_0}{T(t)}\right)}$$

where
γ - heat flow variable
T - instantanous kinetic temperature
T_0 - thermostat temperature
τ - time constant of the coupling to the heat bath

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Berendsen thermostat

The value of τ reveals the strength of coupling between the system and the bath. To achieve good temperature control this value has to be chosen with care.

- when $\tau \rightarrow \infty$ Berendsen thermostat does not work (sampling microcanonical ensemble)
- when $\tau = \delta$t - Berendsen thermostat turns into a simple scaling scheme
- when τ is very small the coupling to the heat bath is large and there is a significant energy exchange between the system and the thermal bath
- efficient
- does not lead to the canonical distribution

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Berendsen barostat

To obtain constant pressure (under assumptions of PBC, finite volume, isotropic system) Berendsen proposed:

- scaling of Cartesian coordinates:

$$\mathbf{r}_i^{new} = c_p \mathbf{r}_i^{old}$$

- scaling of volume

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Berendsen barostat

The scaling factor is expressed as:

$$c_p = \left[1 - \frac{\beta \Delta t}{\tau}(P_0 - P)\right]^{1/3}$$

where

β - isothermal compressibility

P - instantanous pressure

P_0 - target pressure

τ - pressure coupling time

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Berendsen barostat

The equation of motion can be written as:

$$\dot{\mathbf{r}}_i(t) = \mathbf{v}_i(t) - \left(\frac{\beta(P_0 - P(t)}{3\tau}\right)\mathbf{r}_i(t)$$

Pressure is kept constant while the volume of the system is allowed to fluctuate by changing the cell size uniformly but not its shape (for isotropic system)

does not lead to canonical distribution

For this reason the Langevin piston method is mostly used in MD simulations!

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Andersen thermostat

The Andersen approach couples velocity of a particle to a heat bath via stochastic collisions
Procedure of Andersen algorithm:

- starting from intial conditions integrate the equations of motion for δt
- particles that undergo collisions with the heat bath are selected

  the probability of selecting a praticle during the integration time step δt is vδt, where v is the probability of collision
- new velocites for selected particles are chosen at random from Maxwell- Boltzmann distribution at the desired temperature; all other particles remain unaffected

generate canonical ensemble
less realistic dynamics and thus not used for computing dynamic quantities (e.g. diffusion coefficients, lifetime of water hydrogen bonds)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Nose-Hoover thermostat

Nose-Hoover proposed that the bath is an integral part of the system by adding a fictitious degree of freedom to the physical system.
As a result the system does possess 6N+1 degrees of freedom in Γ space

$$\dot{\mathbf{r}}_i(t) = \dot{\mathbf{v}}_i(t)$$

$$m_i\dot{\mathbf{v}}_i(t) = \mathbf{F}_i(t) - \zeta(t)m_i\dot{\mathbf{v}}_i(t)$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Nose-Hoover thermostat

The heat flow variable has its own equation of motion enabling fluctuations of temperature:

$$m_t \dot{\zeta}_t(t) = 2v^T mv - gk_B T_0$$

where
ζ_t - heat flow variable (thermodynamic friction coefficient)
m_t - fictitious mass
g - number of degrees of freedom in the system

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Nose-Hoover thermostat

System Hamiltonian is conserved:

$$H(\mathbf{r}_i, \mathbf{p}_i, \zeta) = K(\mathbf{p}_i) + U(\mathbf{r}_i) + \frac{1}{2}(m_t \zeta_t^2) + \frac{3}{2} N k_b T_0 x_t$$

Nose-Hoover:

- allows the temperature to fluctuate about the average value and controls the oscillations of temperature

- may oscillate for a system far from equilibrium thus it is not recommended for the initial stages of simulations

- leads to canonical distribution

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Nose-Hoover barostat

- volume of the system is not constant but instead is allowed to fluctuate under constant pressure
- an approach analogous to Nose-Hoover thermostat for NVT ensemble is used for describing Nose-Hoover barostat (NPT ensemble)
- two sets of variables are used: (x_p,m_p,ζ_p) associated with thermostat and (x_t,m_t,ζ_t) associated with barostat
- the algorithm drives the system to the state where the average internal pressure is equal to the applied external pressure.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

In experiments an extremely large number of particles are sampling conformational space very effectively (the value of a measured property is an average value over time of the measurement).

$$\overline{A} = \lim_{\tau \to \infty} \frac{1}{\tau} \int_{t=0}^{\tau} A(\mathbf{p}^N(t), \mathbf{r}^N(t))\, dt$$

In molecular simulations the value of A is obtained as time average on the trajectory:

$$\overline{A} = \lim_{n_t \to \infty} \frac{1}{n_t} \sum_{t=0}^{n_t} A(\mathbf{p}^N(t), \mathbf{r}^N(t))$$

where n_t is the number of time-steps

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Can MD predict thermodynamic quantities?

- Ergodic hypothesis links conformations of the particle sampled by MD over simultation time with the conformations of the particle in a thermodynamic ensemble.
- In other words, the ergodic hypothesis assumes that temporal average value of A ($\overline{A}$) of an equilibrium system is equal to the ensemble average value of A ($\langle A \rangle$).
- Therefore, in MD simulations:

$$\overline{A} = \lim_{n_t \to \infty} \frac{1}{n_t} \sum_{t=0}^{n_t} A(\mathbf{p}^N(t), \mathbf{r}^N(t)) = \lim_{M \to \infty} \frac{1}{M} \sum_{i=1}^{M} A_i(\mathbf{p}^N, \mathbf{r}^N) = \langle A \rangle$$

M is the number of configurations

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Is the ergodic hypothesis true?

- For most cases no, as the sampling time does not approach infinity.
- However, it is true for special cases where barriers between states are on the order of RT.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Examples of time averages in MD

Average potential energy:

$$\overline{U} = \frac{1}{M} \sum_{i=1}^{M} U_i$$

Average kinetic energy:

$$\overline{K} = \frac{1}{M} \sum_{j=1}^{M} \left[\sum_{i=1}^{N} \frac{m_i}{2} v_i v_j \right]$$

where M is the # of configurations and N is the # of atoms

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Autocorrelation functions

Time-dependent autocorrelation functions in MD measure the correlation between particle quantity at time t and time zero.
A general expression for autocorrelation function:

$$C_{AA}(t) = \frac{1}{N}\sum_{i=1}^{N}\frac{\langle A_i(t)A_i(0)\rangle - \langle A_i(0)\rangle^2}{\langle A_i(0)A_i(0)\rangle - \langle A_i(0)\rangle^2}$$

Autocorrelation functions (ACFs) can be defined and calculated for any particle quantity (e.g. v_i) or any system quantity (e.g. U, T, P, ρ)
The initial value of ACF at t=0 is 1 and it approaches 0 as $t \rightarrow \infty$ (usually the decay of ACFs is exponential and the relaxation time is the time required for the function to approach zero).

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Autocorrelation functions

- To calculate an autocorrelation function of any property the relaxation time of this property must be much shorter than the simulation time
- In most cases this condition is fulfilled and many sets of data can be extracted from trajectory (# of independent data sets to be used in computation of ACF is equal to the # of time steps comprising the total simulation time minus the # of time steps needed to reach the relaxation time) and average over to calculate the correlation function.
- The more data sets we use in calculations of ACF the more statistically precise value of ACF can be determined.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Velocity autocorrelation function

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

sum over M goes through different time origins (t_j)

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Dipole autocorrelation function

Fourier transform of the dipole autocorrelation function is related to the far-IR spectrum of the solvent.

$$I(\omega) = \int_{-\infty}^{+\infty} \langle \mu(t)\mu(0) \rangle e^{j\omega t} dt$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Dynamical properties

Diffusion coefficient:

$$D = \lim_{t\to\infty} \frac{1}{6Nt} \left\langle \sum_{i=1}^{N} \left[\mathbf{r}_i(t) - \mathbf{r}_i(0)\right]^2 \right\rangle$$

Formula based on velocity autocorrelation function:

$$D = \frac{1}{3N} \int_0^\infty \left\langle \sum_{i=0}^{N} \mathbf{v}_i(t)\mathbf{v}_i(0) \right\rangle dt$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Dynamical properties

Thermal conductivity

$$\lambda = \lim_{t\to\infty}\left(\frac{1}{6kT^2Vt}\right)\left\langle \sum_{\alpha}\left[\sum_{i} r_{\alpha i}(t)e_i(t) - \sum r_{\alpha i}(0)e_i(0)\right]^2\right\rangle$$

e_i is the instantenous excess energy of particle i

α stands for spatial coordinates

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Dynamical properties

Formula based on autocorrelation function:

$$\lambda = \frac{V}{3kT^2} \int_0^\infty \langle \mathbf{S}(t)\mathbf{S}(0) \rangle dt$$

where

$$\mathbf{S} = \frac{1}{V} \left[\sum_i e_i \mathbf{v}_i + \frac{1}{2} \sum_{ij} \mathbf{r}_{ij} (\mathbf{f}_{ij} \mathbf{v}_{ij}) \right]$$

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Radial distribution function

Radial distribution function (radial pair correlation function) descibes the probability of finding a particle within a distance range of r and r+dr from another particle relative to the ideal gas distribution.
Radial distribution function g(r) can be written as:

$$g(r) = \frac{1}{N} \frac{dn(r)}{4\pi r^2 dr \rho}$$

$4\pi r^2 dr$ is the volume of a spherical shell of thickness dr at a distance r from a chosen atom

ρ is an ideal gas density

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

Radial distribution function

dn(r) is the number of particles between r and r+dr from another particle:

$$dn(r) = \sum_{i}^{N} \sum_{i \neq j}^{N} \delta(r - r_{ij})$$

g(r) is larger/smaller than 1 when the probability of finding a particle within a distance range of r and r+dr from another particle is larger/smaller than that expected from a completely uniform distribution (e.g. an ideal gas).

In the ideal gas g(r)=1.

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund

g(r) is normalized in the following way:

$$\frac{1}{N}\int_0^\infty \rho g(r) 4\pi r^2 dr = 1$$

It is important to choose the right bin size otherwise the results will be meaningless:

- too large dr leads to low resolution
- too small dr leads to large statistical errors

g(r) can be obtained from experimental studies:

- in a crystal directly from X-ray
- in solution from neutron diffraction by Fourier transform of structure factor

HUMAN CAPITAL
HUMAN – BEST INVESTMENT!
Wrocław University of Technology
EUROPEAN SOCIAL FUND
Project co-financed from the EU European Social Fund