C7790 Introduction to Molecular Modelling
TSM Modelling Molecular Structures

Lesson 26

Molecular Dynamics I

PS/2020 Distant Form of Teaching: Rev1

Petr Kulhánek

kulhanek@chemi.muni.cz

National Centre for Biomolecular Research, Faculty of Science
Masaryk University, Kamenice 5, CZ-62500 Brno

Context

macroworld **microworld**

states

(thermodynamic properties, G, T,…)

phenomenological thermodynamics

equilibrium (equilibrium constant)

kinetics (rate constant)

free energy

(Gibbs/Helmholtz)

partition function

statistical thermodynamics

microstates

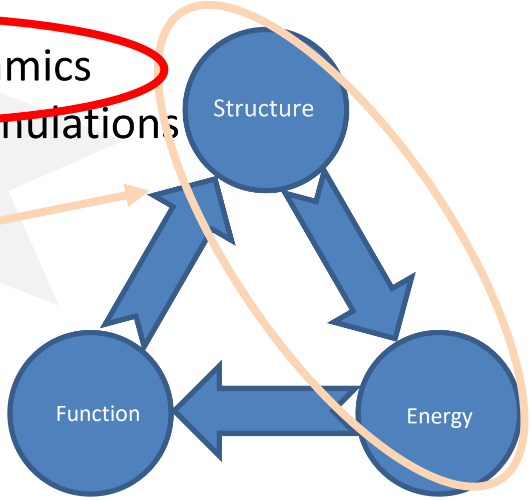
(mechanical properties, E)

Description levels (model chemistry):

- quantum mechanics
  - semiempirical methods
  - ab initio methods
  - post-HF methods
  - DFT methods
- molecular mechanics
- coarse-grained mechanics

Simulations:

- molecular dynamics
- Monte Carlo simulations
- docking
- …



microstate ≠ microworld

System Evolution in Time

How to simulate time evolution of the system?

$$\overline{M} = \frac{1}{t_{tot}} \int_{o}^{t_{tot}} M(t)dt$$

snapshots of the system are a microstates

🟦 the system is composed of N atoms

Mechanical Description (classical physics)*:

- Newton's laws of motion
  - **First law** states that an object at rest will stay at rest, and an object in motion will stay in motion unless acted on by a net external force.
  - **Second law** states that the acceleration of a body over time is directly proportional to the force applied and occurs in the same direction as the applied force.

$$\boldsymbol{F}_i = m_i \boldsymbol{a}_i$$

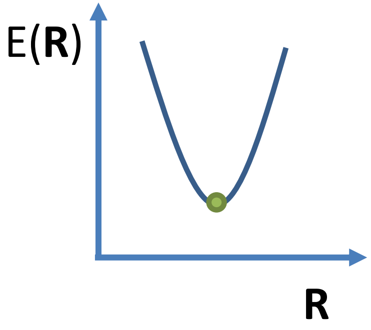
  - **Third law** states that all forces between two objects exist in equal magnitude and opposite direction.

* time evolution can also be described by QM but at cost of theoretical and computational complexity

Forces

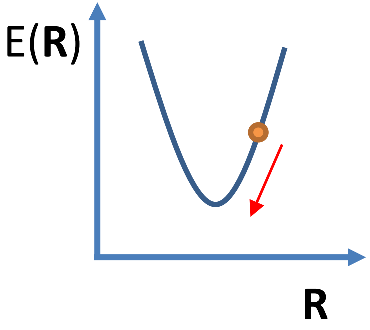
The only forces that can act on atoms in the system are from **interatomic interactions**.

Origin of interatomic forces:



- the most stable local configuration
- the system at rest

$$\frac{\partial E(\boldsymbol{R})}{\partial \boldsymbol{R}} = 0$$



- the system tends to reach more energy favorable configuration

$$\frac{\partial E(\boldsymbol{R})}{\partial \boldsymbol{R}} \neq 0$$

Interatomic forces:

total potential energy

$$\boldsymbol{F}_i = -\frac{\partial E(\boldsymbol{R})}{\partial \boldsymbol{r}_i}$$

negative value of potential energy gradient is force

force acting on atom *i*

position of atom *i*

Equation of Motions

Second Newton's Law

$$\boldsymbol{F}_i = m_i \boldsymbol{a}_i$$

Forces in molecular systems

$$\boldsymbol{F}_i = -\frac{\partial E(\boldsymbol{R})}{\partial \boldsymbol{r}_i}$$

Final equations of motions (EM):

$$m_i \boldsymbol{a}_i = -\frac{\partial E(\boldsymbol{R})}{\partial \boldsymbol{r}_i}$$

$$m_i \frac{d^2 \boldsymbol{r}_i}{dt^2} = -\frac{\partial E(\boldsymbol{R})}{\partial \boldsymbol{r}_i}$$

To describe evolution of the system in time, it is necessary to solve system of N (number of atoms) second order differential equations or motions.

Result: position of atoms in time (trajectory)

$$\boldsymbol{R}(\boldsymbol{t}) = \{\boldsymbol{r}_1(t), \boldsymbol{r}_2(t), \dots, \boldsymbol{r}_N(t)\}$$

Numerical Integration

The solution of EM can be obtained by integration of differential equations. Unfortunately, the analytical solution is not feasible even for small systems (three and more atoms).

Numerical integrations

- **Finite difference methods**
  - **leap-frog algorithm (a variant of Verlet algorithm)**
  - **velocity Verlet algorithm**
- **Gear corrector-predictor methods**
- **Runge-Kutta methods**

most often used algorithms in MD simulations of (bio)chemical systems

Leap-frog algorithm

t ← time

dt ← time step (integration step)

1) Initial conditions:

$$\boldsymbol{r}(t);\ \boldsymbol{v}(t - dt/2)$$

2) Molecular dynamics (MD loop)
    1) Calculation of forces and accelerations

$$\boldsymbol{a}(t) = \frac{1}{m}\frac{\partial E(\boldsymbol{R}(t))}{\partial \boldsymbol{r}}$$

    2) Update velocities

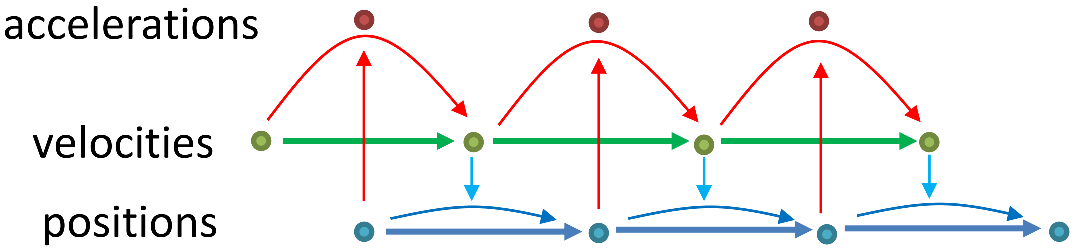
$$\boldsymbol{v}(t + dt/2) = \boldsymbol{v}(t - dt/2) + \boldsymbol{a}(t) \cdot dt$$

$$\boldsymbol{a}(t) = \frac{\boldsymbol{v}(t + dt/2) - \boldsymbol{v}(t - dt/2)}{dt}$$

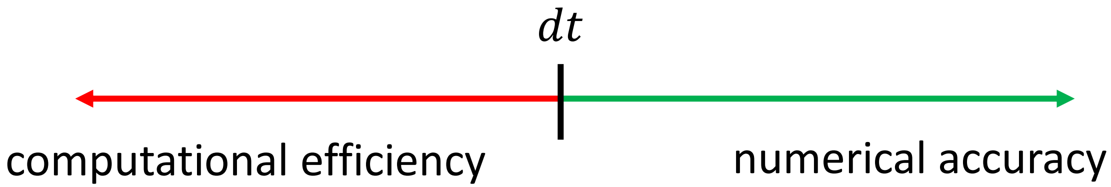
    1) Update positions

$$\boldsymbol{r}(t + dt) = \boldsymbol{r}(t) + \boldsymbol{v}(t + dt/2) \cdot dt$$

$$\boldsymbol{v}(t + dt/2) = \frac{\boldsymbol{r}(t + dt) - \boldsymbol{r}(t)}{dt}$$



Time Step



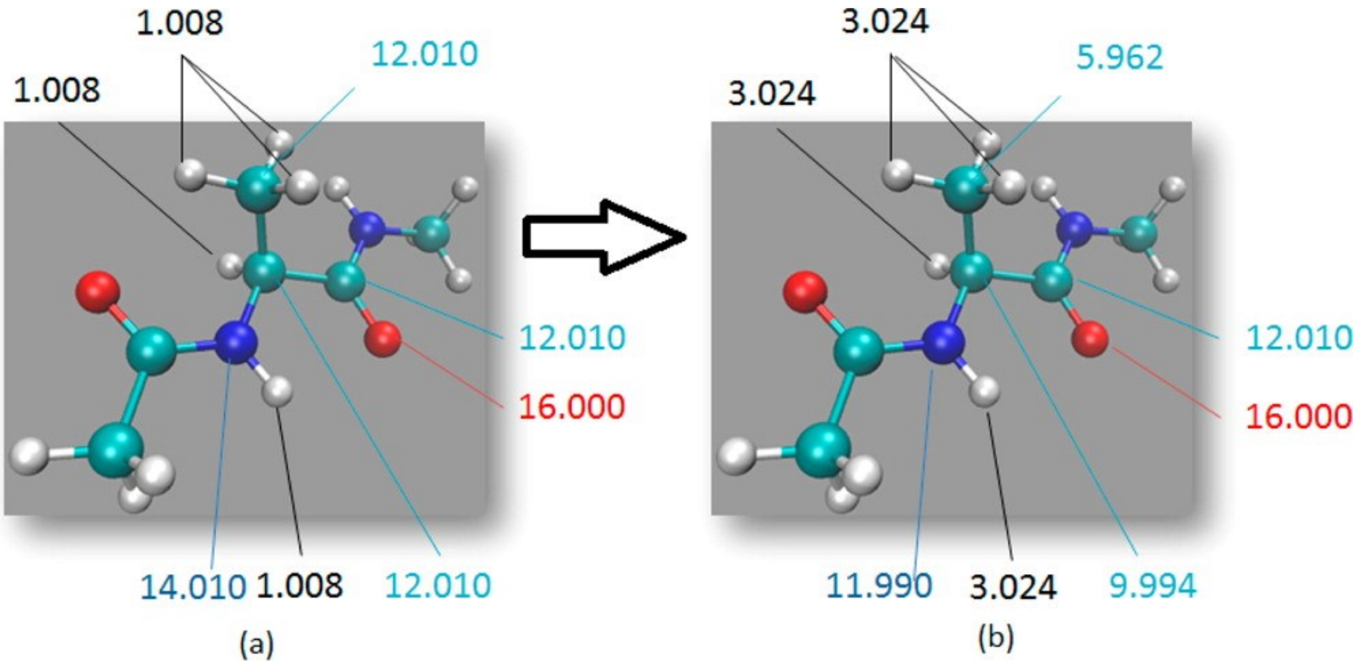
- The time step size is usually taken as 1/10 of the fastest motions.
- The fastest motions are X-H vibrations (higher PES curvature, light atom (hydrogen)).
- Then, the typical size of the integration step is 1 fs (10^{-15} s)

Strategies how to increase the integration time step:

- remove the fastest motions by the constraining X-H distances, which allows a 2-fs step size
    - SHAKE, RATTLE, SETTLE, LINCS algorithms
- in addition, constrain valence angles (mathematically too complex, not use)
- hydrogen mass repartitioning (up to 4 fs)
- multiple time-step integrators
    - computationally cheap short-range forces (shorter integration time step)
    - computationally expensive long-range forces (longer integration time step)

HMR - Hydrogen Mass Repartitioning

Since the molecular dynamics uses the classical physics, each degree of freedom is thermalized to $\frac{1}{2}k_B T$, which is independent to atom masses.



Hopkins, C. W.; Le Grand, S.; Walker, R. C.; Roitberg, A. E. Long-Time-Step Molecular Dynamics through Hydrogen Mass Repartitioning. *J. Chem. Theory Comput.* **2015**, *11* (4), 1864–1874. https://doi.org/10.1021/ct5010406.

Initial Conditions, Equilibration

For integration of EM, we need **initial geometry and velocities**:

$$\boldsymbol{r}(t);\ \boldsymbol{v}(t - dt/2)$$

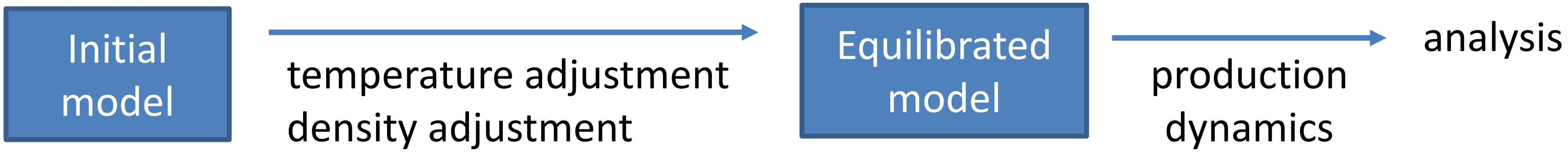
velocities can be generated randomly to satisfy Maxwell-Boltzmann distribution for given temperature

initial geometry (structure) of model

The initial geometry of models derived from experimental structures (X-RAY, NMR, CryoEM, etc.) is usually of low quality.

Equilibration:

- The aim of the equilibration is to bring model to desired thermodynamic state (temperature, pressure, density, etc.).



Initial model → temperature adjustment, density adjustment → Equilibrated model → production dynamics → analysis

Thermostats

Equipartition principle:

mean kinetic energy

$$\frac{3N-c}{2}k_BT = \langle E_k \rangle = \left\langle \sum_{i=1}^{N} \frac{1}{2} m_i \boldsymbol{v}_i^2 \right\rangle$$

degrees of freedom (DOF)
(c - constrained DOF)

temperature

The temperature can be controlled by a thermostat:

- weak coupling thermostat, Berendsen thermostat
  - simple, incorrect ensemble
  - temperature is controlled by velocity scaling
  - dangerous to use for simulations in vacuum
  - susceptible to various artefacts (flying ice cube, etc.)
- Langevin thermostat (stochastic, correct ensemble)
  - it thermalizes each degree of freedom by random collisions
- Nosé-Hoover barostat (correct ensemble)

Barostats

Virial theorem (Clausius 1870):

$$2\langle E_k\rangle = -\sum_{i=N}^{N}\langle \boldsymbol{F}_i\boldsymbol{r}_i\rangle$$

time average of kinetic energy

the virial (it reflects potential energy)

NpT ensemble (ideal gas model):

$$2\langle E_k\rangle = -3PV - \sum_{i=N}^{N}\langle \boldsymbol{F}_i\boldsymbol{r}_i\rangle \longrightarrow P = \frac{1}{V}\left[Nk_BT - \frac{1}{3}\sum_{i=N}^{N}\langle \boldsymbol{F}_i\boldsymbol{r}_i\rangle\right]$$

pressure

from equipartition principle

The pressure can be controlled by a barostat:

- weak coupling barostat (simple, incorrect ensemble)
- Monte-Carlo barostat (stochastic, correct ensemble)
- Nosé-Hoover barostat (correct ensemble)

the pressure change is achieved by changing the size of the simulation box.

Output from MD

actual temperature (K) and pressure (atm)
(they DO NOT represent thermodynamical properties)

```
NSTEP =    60000   TIME(PS) = 3970719.913  TEMP(K) =   300.74  PRESS =  -209.4
Etot   =    -68239.6682  EKtot   =     14401.7354  EPtot      =    -82641.4035
BOND   =       225.2198  ANGLE   =       498.4282  DIHED      =       712.6121
1-4 NB =       259.0139  1-4 EEL =     -4287.9441  VDWAALS    =     10154.5260
EELEC  =    -90203.2595  EHBOND  =         0.0000  RESTRAINT  =         0.0000
EKCMT  =      6899.5998  VIRIAL  =      7977.8787  VOLUME     =    238531.9588
                                                   Density    =         1.0214
```

the same property, which is the actual kinetic energy,
expressed in different units

Output from MD, cont.

Thermostat: T = 300 K (weak coupling)
Barostat: p = 1 atm (weak coupling)

thermodynamic temperature (K) and pressure (atm)

```
A V E R A G E S   O V E R
NSTEP =   5000000   TIME(PS) = 3990599.911  TEMP(K) =   299.90  PRESS =      2.0
Etot   =    -68232.2253  EKtot   =     14361.2461  EPtot      =    -82593.4714
BOND   =       248.6326  ANGLE   =       517.2225  DIHED      =       724.2102
1-4 NB =       253.7846  1-4 EEL =     -4299.0145  VDWAALS    =     10259.0679
EELEC  =    -90297.3769  EHBOND  =         0.0000  RESTRAINT  =         0.0023
EAMBER (non-restraint)  =    -82593.4736
EKCMT  =      6859.7668  VIRIAL  =      6849.2815  VOLUME     =    238422.2094
                                                   Density    =         1.0219

R M S  F L U C T U A T I O N S
NSTEP =   5000000   TIME(PS) = 3990599.911  TEMP(K) =     1.53  PRESS =    153.2
Etot   =        17.0060  EKtot   =        73.4906  EPtot      =        75.4230
BOND   =        12.9980  ANGLE   =        16.9666  DIHED      =        10.9737
1-4 NB =         5.9275  1-4 EEL =        19.2799  VDWAALS    =       126.9605
EELEC  =       159.5253  EHBOND  =         0.0000  RESTRAINT  =         0.0359
EAMBER (non-restraint)  =        75.3871
EKCMT  =        57.0484  VIRIAL  =       788.8336  VOLUME     =       174.8435
                                                   Density    =         0.0007
```