MD Ensembles and Thermostats

Konstanze R. Hahn

Department of Physics, University of Cagliari, Italy

4th CP2K Tutorial, Aug 31st – Sep 4th 2015
Goal of MD simulations

- obtain ensemble average from trajectory

\[ \langle A \rangle = \int P(p, t) A(p, r) dp dr = \int A(p(t), r(t)) dt \]

- solve equations of motion

\[ F_i = m_i \ddot{r}_i \]  

\[ \mathcal{H}(r, p) = \frac{|p|^2}{2m} + U(r) \]

*Newton*  
*Hamiltonian*

\[ P: \text{probability} \quad p: \text{momenta} \]

\[ A: \text{property} \quad r: \text{position} \]
Ensembles in MD

- ensemble: all microstates \((r, p)\) that are accessible to the simulation and provide probability of each microstate

<table>
<thead>
<tr>
<th>particle number (N)</th>
<th>chemical potential (\mu)</th>
</tr>
</thead>
<tbody>
<tr>
<td>volume (V)</td>
<td>pressure (P)</td>
</tr>
<tr>
<td>energy (E)</td>
<td>temperature (T)</td>
</tr>
</tbody>
</table>

- \textit{NVE} microcanonical
- \textit{NVT} canonical
- \textit{NPT} isothermal-isobaric
Ensembles in MD

• ideal MD conserves energy and entropy: microcanonical ensemble (NVE)

• realistic systems change energy, volume and particles with external reservoirs \(\rightarrow\) more difficult

• canonical (NVT) most frequently used

\[
P \propto e^{\frac{E(r)}{k_BT}}
\]
NVE – microcanonical ensemble

- system isolated with constant number of particles $N$, volume $V$ and energy $E$
- solving equations of motion without temperature or pressure control

$$F_i = m_i \ddot{r}_i$$

simulation cell
NVE – microcanonical ensemble

- drift in $E$ resulting from rounding and truncation errors
- time reversible
- dynamical variables well defined
- required initial conditions: position and velocity

simulation cell
NVT – canonical ensemble

- constant number of particles $N$ and volume $V$
- system in thermal contact with heat bath
NPT – isothermal-isobaric ensemble

- constant number of particles $N$, pressure $p$ and temperature $T$
- use of thermostat and barostat
Lagrangian equations of motion

- extended ensemble
- difference between kinetic and potential energy

\[ \mathcal{L}(\mathbf{r}^N, \mathbf{v}^N) = K(\mathbf{v}^N) - U(\mathbf{r}^N) \]

- assumption: \( K \) only dependent on \( \mathbf{v} \), \( U \) only on \( \mathbf{r} \)

\[ \frac{d}{dt} \frac{\partial \mathcal{L}}{\partial \mathbf{v}_i} - \frac{\partial \mathcal{L}}{\partial \mathbf{r}_i} = 0 \]

- Hamiltonian

\[ \mathcal{H} = \sum_i p_i \cdot \mathbf{v}_i - \mathcal{L} = \sum_i p_i \cdot \mathbf{v}_i + U \]

\text{kinetic energy} \quad \text{potential energy}
Lagrangian equations of motion

\[ \mathcal{H} = \sum_i p_i \cdot v_i - \mathcal{L} = \sum_i p_i \cdot v_i + U \]
Calculation of temperature

\[ \langle K \rangle = \frac{N_f}{2} k_B T \]

\[ K = \sum_i \frac{p_i^2}{2m_i} \]

\[ T = \sum_{i=1}^{N} \frac{p_i^2}{m_i k_B N_f} = \sum_{i=1}^{N} \sum_{\alpha} \frac{m_i v_{i\alpha}^2(t)}{k_B N_f} \]

\( K \): kinetic energy
Velocity rescaling

- multiply velocities by a factor $\lambda$ to obtain desired temperature $T_0$

$$
\Delta T = \frac{1}{2} \sum_{i=1}^{N} \frac{2m_i (\lambda v_i)^2}{Nk_B} - \frac{1}{2} \sum_{i=1}^{N} \frac{2m_i v_i^2}{Nk_b} = (\lambda^2 - 1)T(t)
$$

$$\lambda = \sqrt{\frac{T_0}{T(t)}}$$
Velocity rescaling

- straightforward
- does not correspond to any ensemble

MD step

compute instantaneous kinetic energy

rescale velocity by $\lambda$

$$\lambda = \sqrt{\frac{T_0}{T(t)}}$$
Berendsen thermostat

- system weakly coupled to heat bath with temperature $T_{\text{bath}}$

\[ \frac{dT(t)}{dt} = \frac{1}{\tau} (T_{\text{bath}} - T(t)) \]

$\tau$ : coupling parameter

- rescaling at each step, with temperature change

\[ \Delta T = \frac{\partial t}{\tau} (T_{\text{bath}} - T(t)) \]

- scaling factor

\[ (\lambda^2 - 1)T(t) = \frac{\partial t}{\tau} (T_{\text{bath}} - T(t)) \]

Berendsen, JCP (1984)
Berendsen thermostat

- smoother than velocity rescaling
- suppresses fluctuations in kinetic energy
  → no ensemble
- global thermostat
- for large systems good approximation
Andersen thermostat

- random collisions of molecules with an imaginary heat bath (randomize velocities)

\[ P(p_1, p_2, \ldots) \propto \exp\left(-\frac{p_1^2}{2k_BTm_1}\right) \times \exp\left(-\frac{p_2^2}{2k_BTm_2}\right) \times \ldots \]

\[ p_{new} = \sqrt{mk_BT}R \]

random Gaussian number

Andersen, JCP (1980)
Andersen thermostat

- canonical ensemble \((NVT)\)
- stochastic
- local thermostat
- destroys momentum transport (true molecular kinetics are not preserved)
  → cannot be used to calculate transport properties (diffusion coefficient)
Langevin thermostat

- velocity corrected by random force and constant friction

\[ m\ddot{\mathbf{r}}_i = -\frac{\partial U}{\partial \mathbf{r}_i} - m\Gamma \dot{\mathbf{r}}_i + \mathbf{W}_i(t) \]

\[ \Gamma : \text{friction coefficient} \]

\[ \mathbf{W} : \text{random force} \]

- relation between magnitude of force and friction

\[ \langle \mathbf{W}_i(t), \mathbf{W}_j(t') \rangle = \partial_{ij} \partial(t - t') 6m \Gamma k_B T \]

Schneider and Stoll, PRB (1978)
Langevin thermostat

- canonical ensemble ($NVT$)
- local thermostat
- ergodic
- allows the use of large time steps
- destroys momentum transport $\rightarrow$ cannot be used to calculate transport properties (diffusion coefficient)
Local and global thermostats

Berendsen (global)

\[ \dot{p}_i = -\gamma p_i = -\left[ \frac{1}{2\tau} \left( \frac{K}{K} - 1 \right) \right] p_i \]

Langevin (local)

\[ \dot{p}_i = -\gamma p_i + \sqrt{2m_i k_B T \gamma \eta_i} \]

if needed: different thermostats for different atoms
Nosé thermostat

- add two additional degrees of freedom the system:
  \( s \) – position of imaginary heat reservoir
  \( p_s \) – conjugate momentum of imaginary heat reservoir

- additional parameter
  \( Q \) – effective mass
  \[ p_s = \frac{\partial L}{\partial \dot{s}} = Q \dot{s} \]

- momenta conjugate to \( r_i \)
  \[ p_i = \frac{\partial L}{\partial \dot{r}_i} = m_i s^2 \dot{r}_i \]

- Hamiltonian
  \[ H_N = \sum_i \frac{p_i^2}{2m_is^2} + U(r^N) + \frac{p_s^2}{2Q} + gk_BT \ln s \]

Nose, JCP (1984)
Nosé thermostat

- canonical ensemble \((NVT)\)
- smooth
- deterministic
- time-reversible

- fluctuation of real time step resulting from scaling factor \(s\)

\[
dt' = s dt
\]
Nosé-Hoover thermostat

- eliminate problem of “real”-time averages

\[ H_{NH} = \sum_i \frac{p_i}{2m_i} + U(r^N) + \frac{1}{2} Q \xi^2 - g k_B T \ln s \]

\[ g = 3N \]

- friction coefficient

\[ \dot{\xi} = \frac{1}{Q} \left( \sum_i \frac{p_i}{m_i} - g k_B T \right) \]

\[ \xi = \frac{\dot{s}}{s} = \frac{d \ln s}{dt} \]

Hoover, PRA (1985)
Nosé-Hoover thermostat

- second order equation on $K$
- proper sampling
- deterministic (can be non-ergodic)
- second order $\rightarrow$ can be oscillating
Nosé-Hoover chains

- canonical ensemble (NVT)

\[ \dot{\xi}_1 = \frac{1}{Q_1} \left( \sum_i \frac{p_i^2}{m_i} gk_B T \right) - \xi_1 \xi_2 \]

\[ \dot{\xi}_j = \frac{1}{Q_j} \left( Q_{j-1} \xi_{j-1}^2 - k_B T \right) - \xi_j \xi_{j+1} \]

\[ \dot{\xi}_M = \frac{1}{Q_M} \left( Q_{M-1} \xi_{M-1}^2 - k_B T \right) \]

Martyna et al., JCP (1992)
Nosé-Hoover chains

- higher order equation on $K$
- canonical
- ergodic
- additional equations for chaotic behavior
Basic input in CP2K

&GLOBAL
   PROJECT  MD_NVE
   RUN_TYPE  MD
   PRINT_LEVEL  LOW
   WALLTIME  600
&END GLOBAL

&MOTION
   &MD
      ENSEMBLE  NVE
      STEPS  1000
      TIMESTEP  0.5
      TEMPERATURE  300.0
   &END MD
&END MOTION
Ensembles in CP2K

&MD
  ENSEMBLE NVE
  STEPS 1000
  TIMESTEP 0.5
  TEMPERATURE 300.0
&END MD
&END MOTION

- microcanonical: NVE
- canonical: NVT
- canonical using Langevin: LANGEVIN
- isobaric-isothermic: NPT_F
- isobaric-isothermic in isotropic cell: NPT_I
- constant pressure: NPE_F
- constant pressure in isotropic cell: NPE_I
- constant kinetic energy: ISOKIN
- HYDROSTATICSHOCK, MSST, MSST_DAMPED, NVT_ADIABATIC, ...
Ensembles in CP2K

```plaintext
&MOTION
 &MD
  ENSEMBLE NVE
  STEPS 1000
  TIMESTEP 0.5
  TEMPERATURE 300.0
  TEMP_TOL 10
 &END MD
 &END MOTION
```

Enable velocity rescaling when \( T < 290 \text{ K} \) or \( T > 310 \text{ K} \)
Velocity rescaling

MD step

compute instantaneous kinetic energy

rescale velocity by $\lambda$

$$\lambda = \sqrt{\frac{T_0}{T(t)}}$$

$T < 290 \text{ K or } T > 310 \text{ K}$
Ensembles in CP2K - CSVR

Kinetic energy

\[ dK = \sum_i \frac{f_i \cdot p_i}{m_i} dt + (\bar{K} - K) \frac{dt}{\tau} + 2 \sqrt{\frac{K\bar{K}}{N_f}} \sqrt{\frac{dW}{\sqrt{\tau}}} \left( \tau = \frac{1}{2\gamma} \right) \]

thermostat part

\[ dK = (\bar{K} - K) \frac{dt}{\tau} + 2 \sqrt{\frac{K\bar{K}}{N_f}} \frac{dW}{\sqrt{\tau}} \]

Berendsen thermostat

noise gives correct fluctuations

Bussi, Donadio and Parrinello, JCP (2007)
Ensembles in CP2K - CSVR

\[ dK = \left( \bar{K} - K \right) \frac{dt}{\tau} + 2 \sqrt{\frac{K\bar{K}}{N_f}} \frac{dW}{\sqrt{\tau}} \]

- stochastic velocity rescaling for \( \tau = 0 \)
- global
- correct fluctuations
- preserves dynamic properties
- recovers Langevin for single degree of freedom
Ensembles in CP2K - CSVR

```plaintext
&MOTION
 &MD
  ENSEMBLE NVT
  STEPS 1000
  TIMESTEP 0.5
  TEMPERATURE 300.0
 &THERMOSTAT
   TYPE CSVR
   REGION GLOBAL
 &CSVR
   TIMECON 50.
 &END CSVR
 &END THERMOSTAT
 &END MD
 &END MOTION
```

ice $I_h$
Ensembles in CP2K - CSVR

Energy fluctuations (in units of $N_f k_b T^2/2$) from a 1.0 ns run.

Berendsen: wrong ensemble

Nosé-Hoover: not-ergodic, especially in solid (NHC solve this)
Ensembles in CP2K - CSVR

\[ \tau = 100 \]

\[ \tau = 1 \]
# Overview of thermostats

<table>
<thead>
<tr>
<th></th>
<th>tune</th>
<th>cont.</th>
<th>L/G</th>
<th>correct</th>
<th>ergodic</th>
<th>cons. q.</th>
<th>determ.</th>
<th>cp2k</th>
</tr>
</thead>
<tbody>
<tr>
<td>Velocity rescaling</td>
<td></td>
<td></td>
<td>G</td>
<td>?</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Andersen</td>
<td>X</td>
<td></td>
<td>L</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Berendsen</td>
<td>X</td>
<td>X</td>
<td>G</td>
<td>?</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
</tr>
<tr>
<td>Nosé-Hoover</td>
<td>X</td>
<td>X</td>
<td>L/G</td>
<td>X</td>
<td></td>
<td>X</td>
<td>X</td>
<td>NOSE*</td>
</tr>
<tr>
<td>Nosé-Hoover chains</td>
<td>X</td>
<td>X</td>
<td>L/G</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>NOSE</td>
</tr>
<tr>
<td>Langevin</td>
<td>X</td>
<td>X</td>
<td>L</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>CSVR**</td>
</tr>
<tr>
<td>Stochastic velocity rescaling</td>
<td>X</td>
<td>X</td>
<td>L/G</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>CSVR</td>
</tr>
</tbody>
</table>

*use “LENGTH=1”

**use “MASSIVE”
Recipes for MD calculations

• equilibration:
  - strong thermostat (small $\tau$)
  - local for ab initio; local or global for classical

• normal production:
  - global, $\tau=1$-100 fs

• difficult temperature control:
  - local, $\tau=1$-100 fs

• two or more “separate” subsystems (solid-liquid, QMMM,…)
  - global, one per subsystem

• always check energy conservation
• always check kinetic energy fluctuations
Analysis of trajectories

\[ \langle A \rangle = \int P(p, t)A(p, r)dpdr = \int A(p(t), r(t))dt \]

from trajectory file \textit{[project]-pos-1.xyz}

property \( A \)
- \( g \) or \( r \)
$g(r)$ of Si

![Graph showing $g(r)$ for Si at different temperatures](image)

- 300K, quenched
- 300K, crystalline
- 1800K, liquid

Sep 2nd, 2015 konstanze.hahn@dsf.unica.it
Analysis of trajectories

\[ \langle A \rangle = \int P(p, t) A(p, r) dp dr = \int A(p(t), r(t)) dt \]

from trajectory file [project]-pos-1.xyz

property A
- \( g \) or \( r \)
- mean square displacement (MSD)
- bond length
- coordination number
- …..
REFTRAJ in CP2K

```
&MOTION
 &MD
  ENSEMBLE REFTRAJ
  STEPS 5
  &REFTRAJ
   TRAJ_FILE_NAME traj.xyz
   FIRST_SNAPSHOT 1
   LAST_SNAPSHOT 5
  &END REFTRAJ
 &END MD
&END MOTION

&SUBSYS
 ...
 &COLVAR
   &COORDINATION
     KINDS_FROM O
     KINDS_TO Si
     R_0 [angstrom] 1.8
   &END COORDINATION
 &END COLVAR
&END SUBSYS
```
Mean square displacement

\[ \text{MSD}(t) = \langle (\Delta \mathbf{r}_i(t))^2 \rangle = \langle (\mathbf{r}_i(t) - \mathbf{r}_i(0))^2 \rangle \]

\&MOTION
\&MD
\ENSEMBLE \ REFTRAJ
\STEPS 5
\&REFTRAJ
\TRAJ\_FILE\_NAME traj.xyz
\FIRST\_SNAPSHOT 1
\LAST\_SNAPSHOT 5
\&MSD
\MSD\_PER\_KIND
\REF0\_FILENAME ref0.xyz
\&END MSD
\&END REFTRAJ
\&END MD
\&END MOTION

reference cell

frames

0 1 2 3 4