10. Molecular Dynamics Simulations

Molecular dynamics simulations

- a technique for computing equilibrium and transport properties of a classical many-body system
- very similar to real experiments in many aspects
- involves preparing a system (a sample), equilibrating it (wait), and computing observables (measurement).
- common mistakes are very similar to those made in real experiments, such as: the system is not prepared correctly, the measurement is too short, or we are not measuring the right thing.

Outline

Initialization

Integrating the equation of motion

Computing observables

Receipe for molecular dynamics simulations

- common workflow of a molecular dynamics simulation:
  1. initialize the system including the positions and velocities of the particles
  2. run the simulation by integrating the equations of motion
  3. compute the observables of interest
- all three steps are crucial for a successful simulation
- the second step is often the most time-consuming part of a simulation

Initialization

- assign the initial positions and velocities for all particles
- initial positions should be compatible with the structure that we want to simulate
- should not have particles overlapping with each other, which is often achieved by initially placing the particles on a cubic lattice
- for molecules such as proteins with known structures, we could use the known structure as the initial configuration

Initialization

- initial velocities need to be compatible with the temperature of the system
- from statistical mechanics, we know that the average kinetic energy per degree of freedom is given by

$$\left\langle \frac{1}{2} m v_\alpha^2 \right\rangle = \frac{1}{2} k_B T$$

- more specifically, the velocity components of a particle follow the Boltzmann distribution

$$p(v_\alpha) = \frac{1}{\sqrt{2\pi k_B T/m}} e^{-m v_\alpha^2 / 2 k_B T}$$

- we can assign the initial velocities by sampling from the Boltzmann distribution
- what is the average speed of a water molecule at room temperature?

Outline

Initialization

Integrating the equation of motion

Computing observables

The equation of motion

- the equation of motion is given by Newton's second law

$$m_i \frac{d^2 \mathbf{r}_i}{dt^2} = -\nabla_i U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$$

- an equivalent form is the Hamiltonian equations of motion

$$\frac{d\mathbf{q}_i}{dt} = \frac{\partial H}{\partial \mathbf{p}_i}, \qquad \frac{d\mathbf{p}_i}{dt} = -\frac{\partial H}{\partial \mathbf{q}_i}$$

where $\mathbf{q}_i = \mathbf{r}_i$, $\mathbf{p}_i = m_i \mathbf{v}_i$, and

$$\begin{aligned} H(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N, \mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N) &= K(\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_N) + U(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N) \\ &= \sum_i \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N) \end{aligned}$$

The Hamiltonian equations of motion

- Hamiltonian equations

$$\frac{d\mathbf{q}_i}{dt} = \frac{\partial H}{\partial \mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i}$$
$$\frac{d\mathbf{p}_i}{dt} = -\frac{\partial H}{\partial \mathbf{q}_i} = -\nabla_i U(\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_N)$$

- in vector form, we have

$$\frac{d}{dt}\begin{pmatrix}\mathbf{q}\\ \mathbf{p}\end{pmatrix} = \begin{pmatrix}0 & \mathbf{I}\\ -\mathbf{I} & 0\end{pmatrix}\begin{pmatrix}\frac{\partial H}{\partial \mathbf{q}}\\ \frac{\partial H}{\partial \mathbf{p}}\end{pmatrix}$$

- a set of first-order differential equations and equivalent to Newton's second law
- provides a more useful framework for understanding the dynamics of a system

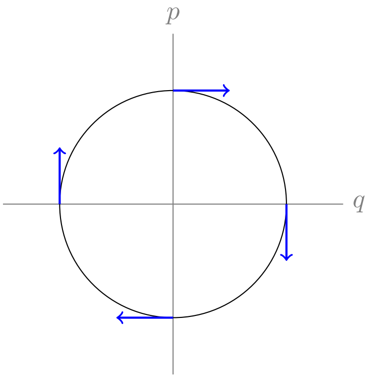
Integrating the Hamiltonian equations of motion

- a solution of the Hamiltonian equations of motion is a trajectory in the phase space of $(\mathbf{q}, \mathbf{p})$
- example: a harmonic oscillator

$$H(q,p) = \frac{p^2}{2m} + \frac{1}{2}kq^2$$

when m = 1 and k = 1

$$\frac{dq}{dt} = \frac{\partial H}{\partial p} = p$$
$$\frac{dp}{dt} = -\frac{\partial H}{\partial q} = -q$$



Integrating the Hamiltonian equations of motion

- most Hamiltonian systems are not solvable analytically
- need numerical methods to integrate the Hamiltonian equations of motion
- the naive method is the Euler method, which is not a good choice
- a much better method is the leap-frog Verlet method

The Euler method

- naive Euler method

$$p(t+\epsilon) = p(t) + \epsilon \frac{dp}{dt} = p(t) - \epsilon \frac{\partial U}{\partial q}(q(t))$$

$$q(t+\epsilon) = q(t) + \epsilon \frac{dq}{dt} = q(t) + \epsilon \frac{p(t)}{m}$$

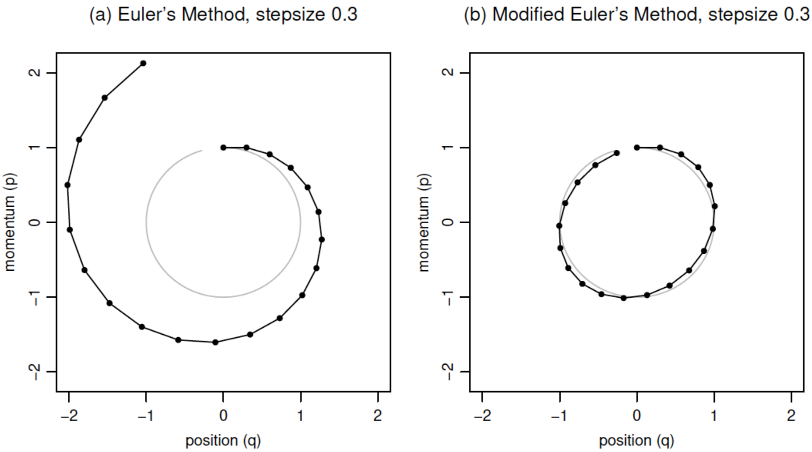
- modified Euler method

$$p(t+\epsilon) = p(t) - \epsilon \frac{\partial U}{\partial q}(q(t))$$

$$q(t+\epsilon) = q(t) + \epsilon \frac{p(t+\epsilon)}{m}$$

The Euler method

- numerical trajectory for the 1-d harmonic oscillator



The leap-frog Verlet method

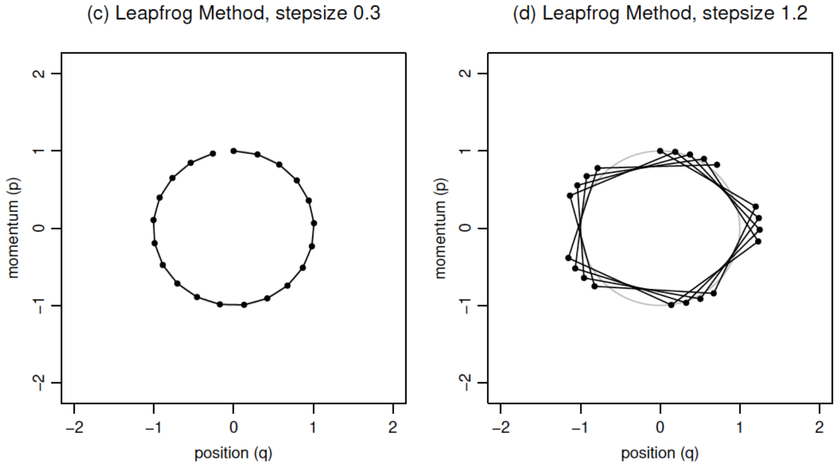
- integration scheme

$$p(t+\epsilon/2) = p(t-\epsilon/2) - 2\frac{\partial U}{\partial q}(q(t))$$
$$q(t+\epsilon) = q(t) + \epsilon\frac{p(t+\epsilon/2)}{m}$$

- is called the leap-frog method because the position and momentum are updated in a staggered fashion
- is a symplectic integrator which preserves the phase space volume
- has better energy conservation properties than the Euler method

The leap-frog Verlet method

- numerical trajectory for the 1-d harmonic oscillator



The Hamiltonian equations are energy conserving

- both the Newtonian and Hamiltonian equations of motion are energy conserving
- trajectories are confined to a constant energy surface in phase space
- conformations of the system are sampled from the *NVE* ensemble, which is also called the *microcanonical ensemble*
- we are mostly interested in the *NVT* ensemble, where the temperature is kept constant
- in the NVT ensemble, the system is contacted with and exchange energy with a heat bath
- in the NVT ensemble, the total energy is not conserved but fluctuates
- integrating the Hamiltonian equations of motion alone will not give us the NVT ensemble

Integrators for the NVT ensemble

- several integrators are available for the NVT ensemble
  - Andersen thermostat
  - Nose-Hoover thermostat
  - Langevin dynamics

- we will only discuss the Langevin dynamics, which is the most widely used method

Langevin dynamics

- the Langevin equation for the *NVT* ensemble is given by

$$m_i \frac{d\mathbf{v}_i}{dt} = -\nabla_i U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) - \gamma m_i \mathbf{v}_i + \sqrt{2 m_i \gamma k_B T} R_i(t)$$

- γ is the friction coefficient
- $R_i(t)$ is an uncorrelated random force with zero mean and unit variance
- $-\gamma m_i \mathbf{v}_i$ is the frictional force modeling the viscous drag
- $\sqrt{2 m_i \gamma k_B T} R_i(t)$ is the random force modeling the interaction with the heat bath

Integrating the Langevin equation

- use the Langevin leap-frog method

$$
\begin{aligned}
\mathbf{v}_i(t+\epsilon/2) &= \mathbf{v}_i(t-\epsilon/2)\alpha - (1-\alpha)/(\gamma m_i)\frac{\partial U}{\partial \mathbf{r}_i}(t) + \sqrt{k_B T(1-\alpha^2)/m_i}R_i(t) \\
\mathbf{r}_i(t+\epsilon) &= \mathbf{r}_i(t) + \epsilon \mathbf{v}_i(t+\epsilon/2) \\
\alpha &= e^{-\gamma\epsilon}
\end{aligned}
$$

- sampled momentums follow the Boltzmann distribution of

$$p(\mathbf{v}_i) = \frac{1}{\sqrt{2\pi k_B T/m_i}} e^{-m_i \mathbf{v}_i^2/2k_B T}$$

- sampled configurations follow the Boltzmann distribution of

$$p(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \frac{1}{Z} e^{-U(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)/k_B T}$$