Computational Statistical Physics

Part II: Interacting particles and molecular dynamics

Marina Marinkovic

April 13, 2022 - updated

ETH Zürich
Institute for Theoretical Physics
HIT G 41.5
Wolfgang-Pauli-Strasse 27
8093 Zürich

402-0812-00L
FS 2022

Molecular dynamics

Molecular dynamics

To model interacting particle systems, we use generalized coordinates

$$\mathbf{q}_i = \left(q_i^1, \ldots, q_i^d\right) \quad \text{and} \quad \mathbf{p}_i = \left(p_i^1, \ldots, p_i^d\right). \tag{1}$$

in a system where each particle has d degrees of freedom.

Molecular dynamics

The system of N particles is then described by

$$Q = (\mathbf{q}_1, \ldots, \mathbf{q}_N) \quad \text{and} \quad P = (\mathbf{p}_1, \ldots, \mathbf{p}_N), \tag{2}$$

using the Hamiltonian

$$\mathcal{H}(P, Q) = K(P) + V(Q) \tag{3}$$

with $K(P) = \sum_{i,k} \frac{\left(p_i^k\right)^2}{2m_i}$ being the kinetic energy, m_i the mass of the i^{th} particle and $V(Q)$ the potential energy. The sum over $k \in \{1, ..., d\}$ accounts for the d degrees of freedom.

Molecular dynamics

The potential (e.g., an attractive or repulsive electromagnetic potential) determines the mutual interactions of all particles and therefore their dynamics. An expansion of the potential energy yields:

$$V(Q) = \sum_i v_1(q_i) + \sum_i \sum_{j>i} v_2(q_i, q_j) + \sum_i \sum_{j>i} \sum_{k>j} v_3(q_i, q_j, q_k) + \dots . \tag{4}$$

Molecular dynamics

Typically three or more body interactions are neglected and their effect is considered in an effective two body interaction described by

$$v_2^{\text{eff}}(q_i, q_j) = v^{\text{attr}}(r) + v^{\text{rep}}(r) \quad \text{with} \quad r = |\mathbf{q}_i - \mathbf{q}_j|, \qquad (5)$$

where $v^{\text{attr}}(r)$ and $v^{\text{rep}}(r)$ represent attractive and repulsive part of the effective potential, respectively.

Molecular dynamics

For now, we only consider potentials that depend on distance, not particle orientation. Analytically, the simplest potential is the hard sphere interaction potential

$$v^{\text{rep}}(r) = \begin{cases} \infty & \text{if } r < \sigma, \\ 0 & \text{if } r \geq \sigma. \end{cases} \tag{6}$$

Molecular dynamics

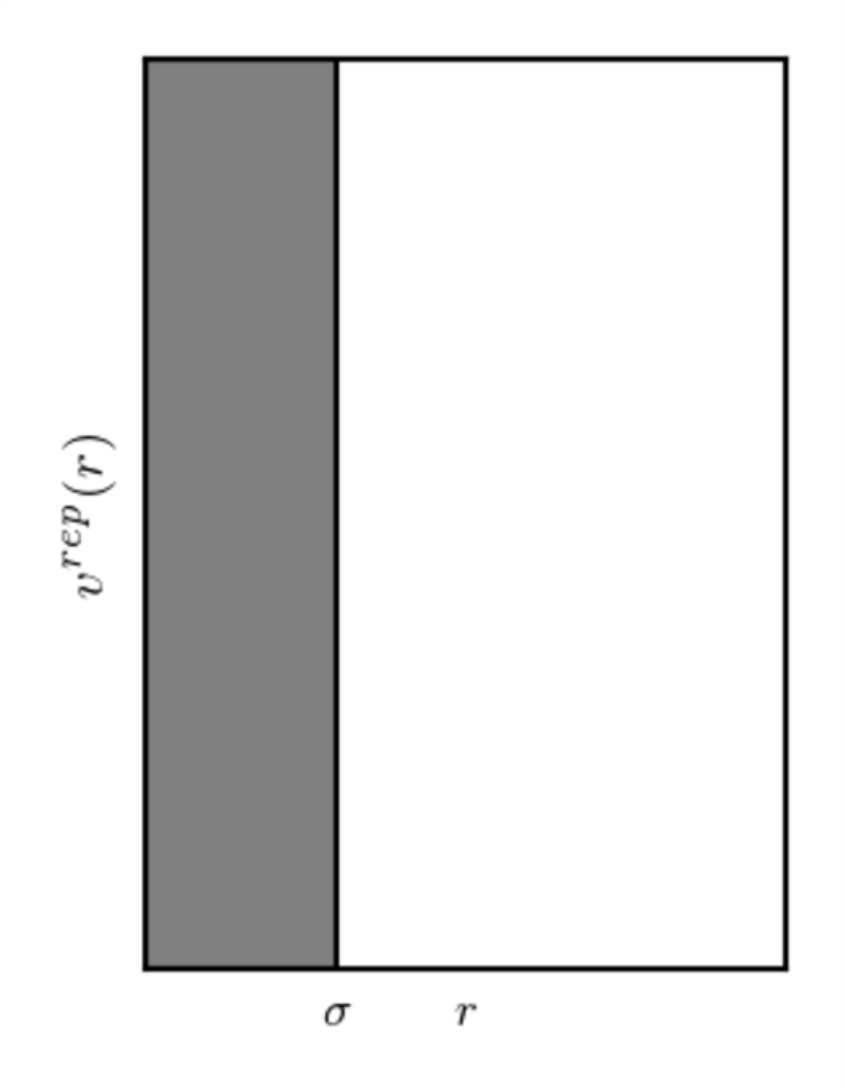

Figure 1: An example of a hard sphere potential.

Equations of motion

The first order Taylor approximation of a symmetric attractive or repulsive potential is given by an elastic potential. For two particles with radii R_1 and R_2, the potential is given by

$$v^{\text{rep}}(r) = \begin{cases} \frac{k}{2}(R-r)^2 & \text{if } r < R \\ 0 & \text{if } r > R \end{cases} \quad \text{with} \quad R = R_1 + R_2, \qquad (7)$$

where k is the elastic spring constant.

Equations of motion

Another very important form of potential typically used to describe the interaction between molecules is the *Lennard-Jones* potential

$$v^{\mathrm{LJ}}(r) = 4\epsilon \left[\left(\frac{\sigma}{r} \right)^{12} - \left(\frac{\sigma}{r} \right)^{6} \right], \tag{8}$$

where ϵ is the attractive energy and σ the interaction range.

LJ potential approximates the spherical symmetric interaction between a pair of neutral atoms or molecules.

Equations of motion

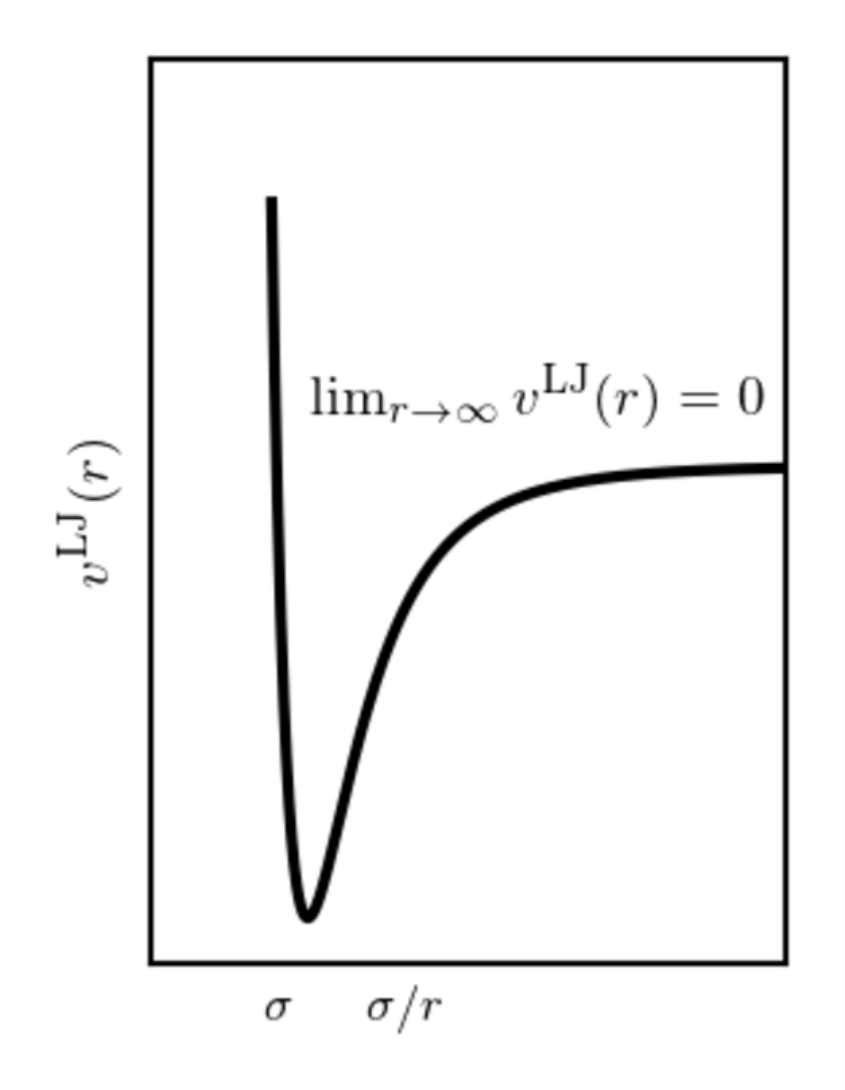

Figure 2: An example of a Lennard-Jones potential, cf. `http://www.atomsinmotion.com/`.

Equations of motion

Once the interaction potential has been defined, we can easily derive the equations of motion using the Hamilton equations

$$\dot{q}_i^k = \frac{\partial \mathcal{H}}{\partial p_i^k}, \quad \dot{p}_i^k = -\frac{\partial \mathcal{H}}{\partial q_i^k}, \tag{9}$$

where $k \in \{1, \ldots, d\}$ and $i \in \{1, \ldots, N\}$.

Equations of motion

For every particle, we identify q_i with the position vector x_i and $\dot{\mathbf{q}}_i = \dot{\mathbf{x}}_i$ with the velocity vector $\dot{\mathbf{v}}_i$. Due to $\dot{\mathbf{x}}_i = \mathbf{v}_i = \frac{\mathbf{p}_i}{m}$ and $\dot{\mathbf{p}}_i = -\nabla V(Q) = \mathbf{f}_i$, the equations of motion are:

$$m_i \ddot{\mathbf{x}}_i = \mathbf{f}_i = \sum_j \mathbf{f}_{ij}, \tag{10}$$

where $\mathbf{f}_{ij}$ is the force exerted by particle j on particle i.

Contact time

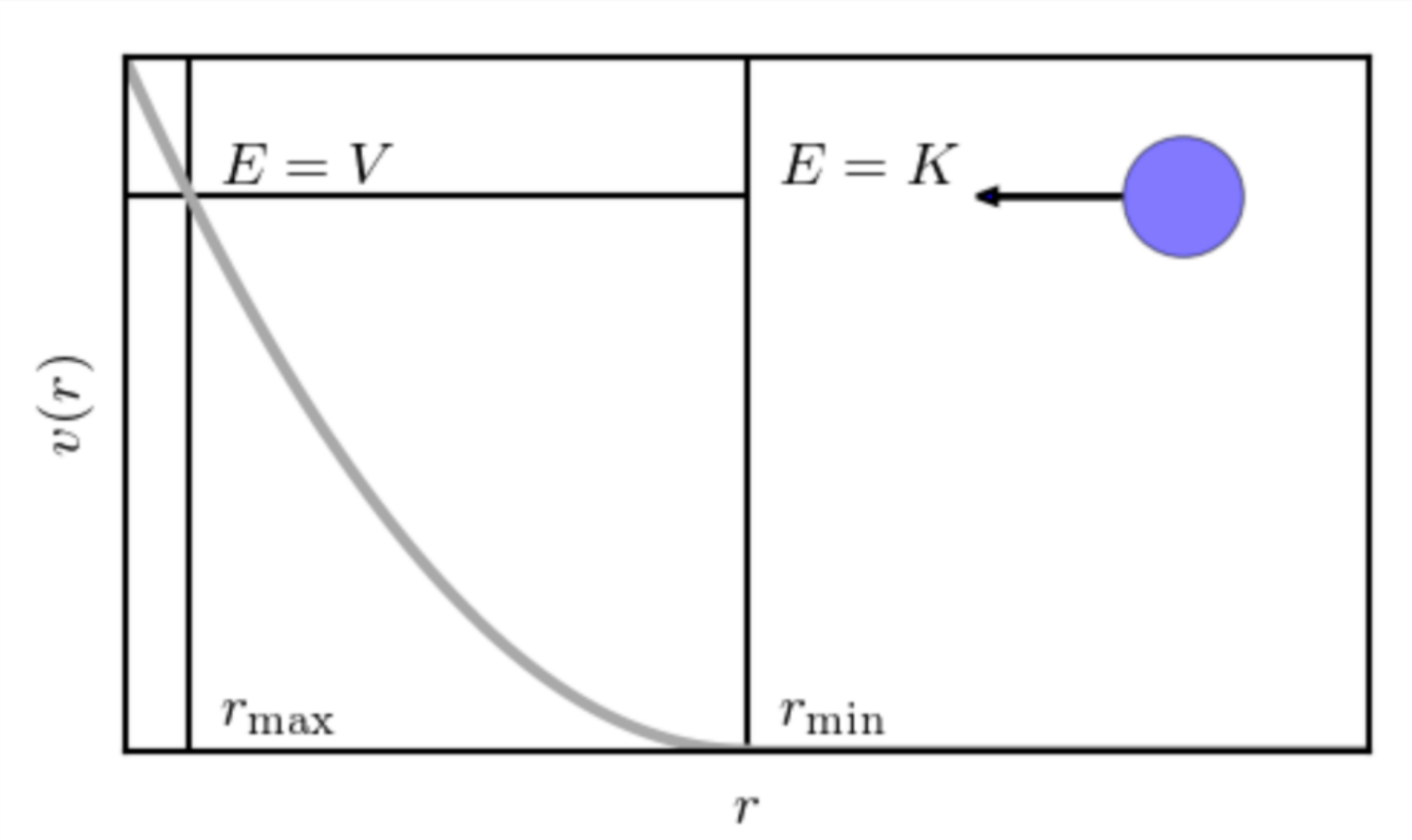

Figure 3: Derivation of the contact time.

Contact time

Using the equations for energy

$$E = \frac{1}{2} m \dot{r}^2 + V(r) = \text{const.} \tag{11}$$

and radial velocity

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \left[\frac{2}{m} \left(E - V(r) \right) \right]^{\frac{1}{2}}, \tag{12}$$

we derive the contact time

$$t_c = 2 \int_0^{\frac{1}{2} t_c} \mathrm{d}t = 2 \int_{r_{\min}}^{r_{\max}} \frac{\mathrm{d}t}{\mathrm{d}r} \mathrm{d}r = 2 \int_{r_{\min}}^{r_{\max}} \left[\frac{2}{m} \left(E - V(r) \right) \right]^{-\frac{1}{2}} \mathrm{d}r, \tag{13}$$

where $r_{\min}$ and $r_{\max}$ are the range of the potential and the turning point of a colliding particle, respectively.

Contact time

We expect reasonable results only if the time step is not larger than the smallest contact time. The time integration of the equations of motion is then possible using an integration method such as

- Euler's method,
- Runge-Kutta methods,
- Predictor-corrector methods,
- Verlet methods,
- Leap-frog methods.

Verlet method [Verlet '67]

We begin with a Taylor expansion of $x(t+\Delta t)$ for sufficiently small time steps Δt so that

$$\begin{aligned} \mathbf{x}\,(t+\Delta t) &= \mathbf{x}\,(t) + \Delta t\,\mathbf{v}\,(t) + \frac{1}{2}\Delta t^2\,\dot{\mathbf{v}} + \mathcal{O}(\Delta t^3), \\ \mathbf{x}\,(t-\Delta t) &= \mathbf{x}\,(t) - \Delta t\,\mathbf{v}\,(t) + \frac{1}{2}\Delta t^2\,\dot{\mathbf{v}} - \mathcal{O}(\Delta t^3). \end{aligned} \tag{14}$$

Adding the latter two expressions yields

$$\mathbf{x}\,(t+\Delta t) = 2\mathbf{x}\,(t) - \mathbf{x}\,(t-\Delta t) + \Delta t^2\,\ddot{\mathbf{x}}\,(t) + \mathcal{O}(\Delta t^4). \tag{15}$$

Verlet method

Newton's second law enables us to express $\ddot{\mathbf{x}}(t)$ as

$$\ddot{\mathbf{x}}_i(t) = \frac{1}{m_i} \sum_j \mathbf{f}_{ij}(t) \quad \text{with} \quad \mathbf{f}_{ij}(t) = -\nabla V(r_{ij}(t)). \tag{16}$$

The particle trajectories are then computed by plugging in the latter results in Eq. (15). Typically, we use a time step of approximately $\Delta t \approx t_c/20$, with t_c a contact time defined in Eq. (13).

Verlet method

Some general remarks about the Verlet method:

- Two time steps need to be stored (t and $t - \Delta t$).
- Velocities can be computed with $\mathbf{v}(t) = \frac{\mathbf{x}(t+\Delta t)-\mathbf{x}(t-\Delta t)}{2\Delta t}$.
- The local numerical error is of order $\mathcal{O}\left(\Delta t^4\right)$, i.e. it is globally a third order algorithm.
- The numbers which are added are of order $\mathcal{O}\left(\Delta t^0\right)$ and $\mathcal{O}\left(\Delta t^2\right)$.
- Improvable by systematical inclusion of higher orders (very inefficient).
- The method is time reversible, which allows to estimate the error accumulation by reversing the process and comparing it to the initial conditions.

Leapfrog Method

For the derivation of the Leapfrog method, we consider velocities at intermediate steps:

$$\mathbf{v}\left(t+\frac{1}{2}\Delta t\right) = v(t) + \frac{1}{2}\Delta t\,\dot{\mathbf{v}}(t) + \mathcal{O}(\Delta t^2), \tag{17}$$

$$\mathbf{v}\left(t-\frac{1}{2}\Delta t\right) = v(t) - \frac{1}{2}\Delta t\,\dot{\mathbf{v}}(t) + \mathcal{O}(\Delta t^2). \tag{18}$$

Taking the difference of the two equations leads to

$$\mathbf{v}\left(t+\frac{1}{2}\Delta t\right) = \mathbf{v}\left(t-\frac{1}{2}\Delta t\right) + \Delta t\,\ddot{\mathbf{x}}\,(t) + \mathcal{O}(\Delta t^3) \tag{19}$$

and we then update the positions according to

$$\mathbf{x}\,(t+\Delta t) = \mathbf{x}\,(t) + \Delta t\,\mathbf{v}\left(t+\frac{1}{2}\Delta t\right) + \mathcal{O}(\Delta t^4). \tag{20}$$

Leapfrog Method

The analogies and differences between the Leapfrog method

$$\begin{aligned} \dot{\mathbf{v}}\left(t+\Delta t\right) &= \frac{f\left(\mathbf{x}\left(t\right)\right)}{m}, \\ \mathbf{v}\left(t+\Delta t\right) &= \mathbf{v}\left(t\right) + \Delta t\, \dot{\mathbf{v}}\left(t+\Delta t\right), \\ \mathbf{x}\left(t+\Delta t\right) &= \mathbf{x}\left(t\right) + \Delta t\, \mathbf{v}\left(t+\Delta t\right) \end{aligned} \tag{21}$$

and the forward Euler integration

$$\begin{aligned} \dot{\mathbf{v}}\left(t+\Delta t\right) &= \frac{f\left(\mathbf{x}\left(t\right)\right)}{m}, \\ \mathbf{x}\left(t+\Delta t\right) &= \mathbf{x}\left(t\right) + \Delta t\, \mathbf{v}\left(t\right), \\ \mathbf{v}\left(t+\Delta t\right) &= \mathbf{v}\left(t\right) + \Delta t\, \dot{\mathbf{v}}\left(t+\Delta t\right) \end{aligned} \tag{22}$$

are the following: The update of the variables is done in a different order (both methods rely on explicit forward integration). In the case of the Leapfrog method, the position is not updated using the previous velocity, as it is done in the usual Euler method.

Leapfrog Method

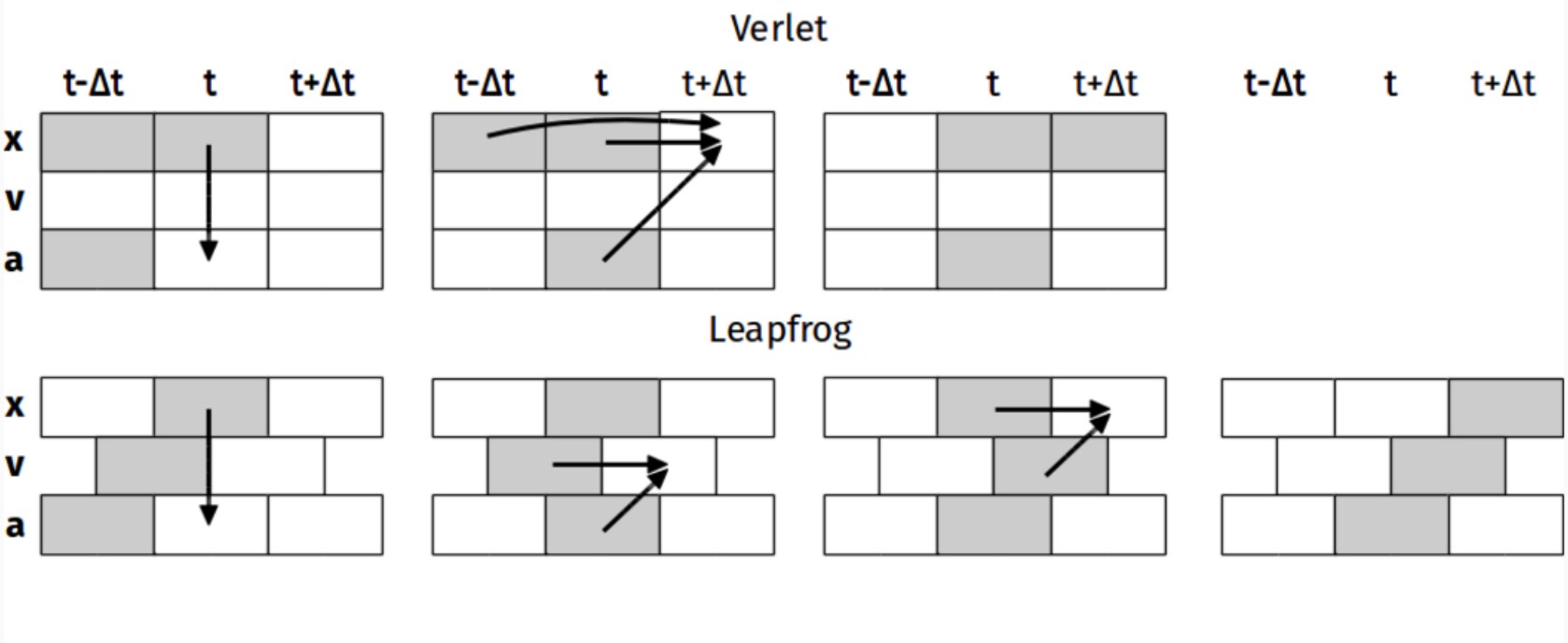

Figure 4: A comparison between Verlet and Leapfrog update schemes.

Optimization

Optimization

For MD simulations of N particles – we have an operation of complexity $\mathcal{O}\left(N^2\right)$.

Alternatively, let our potential be a function $v(r) \sim r^{-2n}$ with $n \geq 1$. We can then omit the computation of the square root in

$$r_{ij} = \sqrt{\sum_{\alpha=1}^{d} \left(x_i^\alpha - x_j^\alpha\right)^2} \tag{23}$$

since for the chosen potential $\mathbf{f} = -\nabla r^{-2n} \propto r^{-2(n-1)}\mathbf{r}$ and $\mathbf{f}_i = f\left(r^{-2(n-1)}\right)\mathbf{r}_i$.

Optimization

If the potential is not a simple function and its calculation would imply a lot of tedious calculations, discretizing the potential and storing its values in a lookup table might be helpful. For short range potentials, we define a cutoff r_c and discretize the interval $(0, r_c^2)$ in K pieces, i.e.,

$$l_k = \frac{k}{K} r_c^2. \tag{24}$$

The force values stored in a lookup table are $f_k = f\left(\sqrt{l_k}\right)$ and the corresponding index k is given by

$$k = \left\lfloor S \sum_{\alpha=1}^{d} \left(x_i^\alpha - x_j^\alpha\right)^2 \right\rfloor + 1, \tag{25}$$

where $\lfloor \cdot \rfloor$ denotes the floor function and $S = K/r_c^2$.

Optimization

The definition of a cutoff makes it necessary that we introduce a cutoff potential $\tilde{v}(r)$ according to

$$\tilde{v}\left(r\right) = \begin{cases} v(r) - v(r_c) - \left.\frac{\partial v}{\partial r}\right|_{r=r_c}(r - r_c) & \text{if } r \leq r_c, \\ 0 & \text{if } r > r_c, \end{cases} \tag{26}$$

where $v(r_c)$ is the value of the original potential at r_c. Without adding the derivative term to the potential $\tilde{v}(r)$, there would be a discontinuity in the corresponding force.

In the case of the Lennard-Jones potential, a value of $r_c = 2.5\sigma$ is typically used. Care must be taken for potentials decaying as r^{-1}, since forces at large distances are not negligible.

Verlet tables

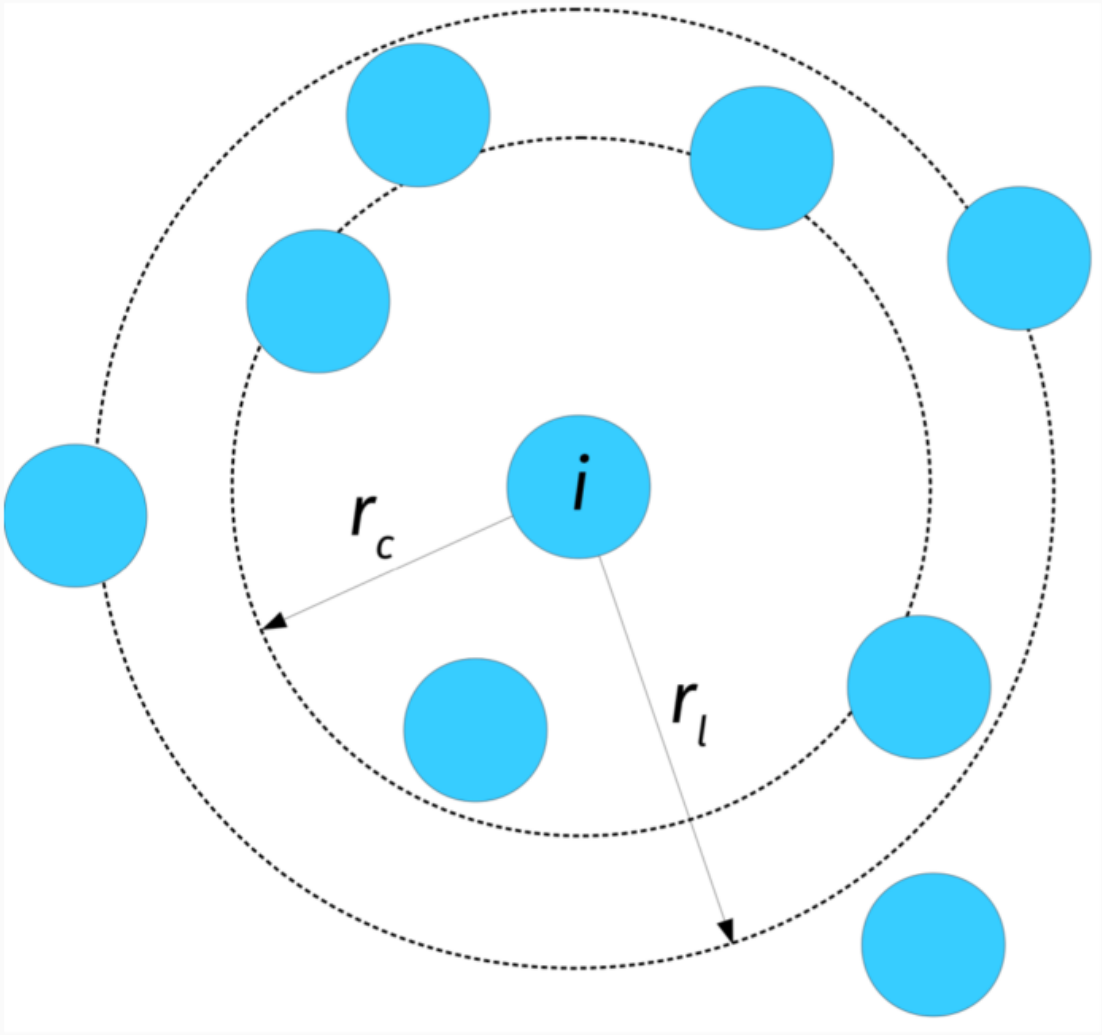

Figure 5: An illustration of the Verlet table method. Only particles within a distance of $r_l > r_c$ from particle i are considered

Verlet tables

In order to reduce the amount of necessary computations, we wish to omit the computation of particle-particle interactions when they are negligible. Therefore, we only consider particles in a certain range $r_l > r_c$.

For every particle, we then store the coordinates of the neighbouring particles in a list which is referred to as *Verlet table*. As the particles move with time, the table has to be updated after

$$n = \frac{r_l - r_c}{\Delta t v_{\max}} \tag{27}$$

time steps, where $v_{\max}$ denotes the maximal velocity. Updating the whole list is still an operation of complexity $\mathcal{O}\left(N^2\right)$.

Linked-Cell Method

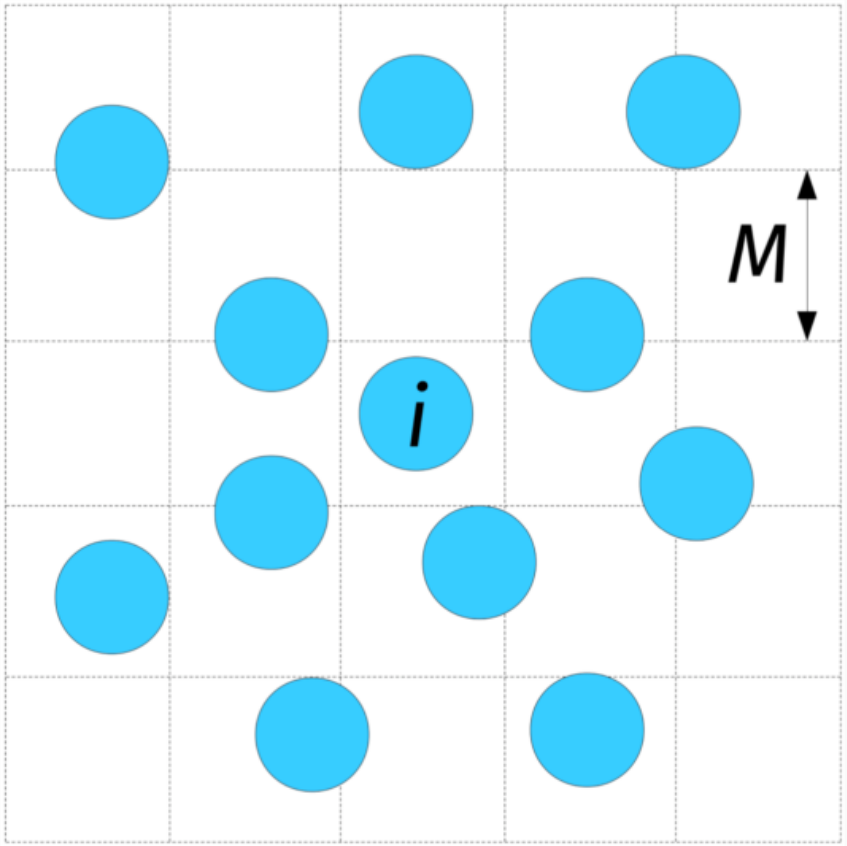

Figure 6: An illustration of the linked-cell method. A grid with grid spacing $M(\frac{r_c}{2} < M < r_c)$ is placed on top of the MD simulation geometry. Only interactions between particles in a certain cell neighborhood have to be considered [D. Knuth, The Art of Computer Programming, '68].

Linked-Cell Method

In d dimensions there are 3^d cells of interest. On average, we thus have to compute the interactions of $N3^dN/M^d$ particles. To keep track of the locations of all particles, we define a vector FIRST of length $N_M = M^d$ to store the index of a particle located in cell j in $\mathrm{FIRST}\,[j]$. If cell j is empty, then $\mathrm{FIRST}\,[j] = 0$. In a second vector LIST of length N, the indices of the remaining particles located in the same cell are stored. If the particle i is the last one in a cell, then $\mathrm{LIST}\,[i] = 0$.

Linked-Cell Method

The following code shows an example of how to extract the particles located in cell $i = 2$.

```
i=2;
A[1]=FIRST[i];
while(M[i-1]!=0)
{
        A[j]=LIST[M[j-1]];
}
```

When a particle changes the cell, FIRST and LIST are updated locally to avoid loops over all particles. The algorithm is thus of order $\mathcal{O}(N)$. In addition, this method is well suited for parallelization (domain composition).

Dynamics of composed particles

Dynamics of composed particles

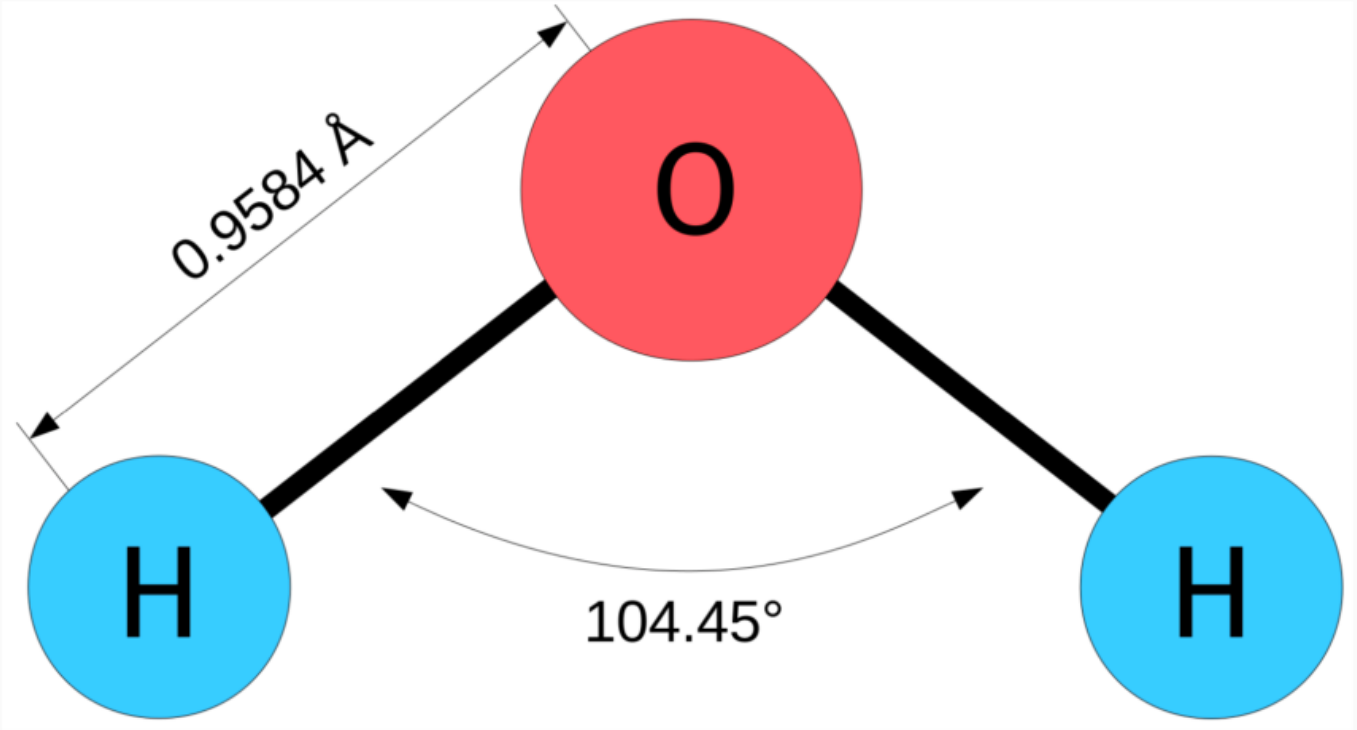

Figure 7: A water molecule as a composed particle system consisting of two hydrogen and one oxygen atom.

Lagrange multipliers

One of the first description of composed particle systems based on an additional force term in the equations of motions has been suggested in [Ryckaert, Ciccotti and Berendsen, 1977]. The idea is to rewrite the equation of motion for each particle as

$$m_i \ddot{\mathbf{x}}_i = \underbrace{\mathbf{f}_i}_{\text{external interaction}} + \underbrace{\mathbf{g}_i}_{\text{internal constraints}} , \tag{28}$$

where the first term accounts for interactions between different composed particles and the second one describes the constraint forces.

Lagrange multipliers

We now impose such constraints to enforce the geometric arrangement of the molecules, e.g., certain distances d_{12} and d_{23} between atoms. Therefore, we define a potential such that the constraint forces g_i are proportional to the difference of the actual and the desired distance of the particles. Considering a water molecule consisting of three particles, the two distance measures

$$\chi_{12} = r_{12}^2 - d_{12}^2, \tag{29}$$

$$\chi_{23} = r_{23}^2 - d_{23}^2, \tag{30}$$

are zero if the particles have the desired distance.

Lagrange multipliers

With $r_{ij} = \|\mathbf{r}_{ij}\|$ and $\mathbf{r}_{ij} = \mathbf{x}_i - \mathbf{x}_j$ we obtain

$$\mathbf{g}_k = \frac{\lambda_{12}}{2} \nabla_{\mathbf{x}_k} \chi_{12} + \frac{\lambda_{23}}{2} \nabla_{\mathbf{x}_k} \chi_{23}, \tag{31}$$

for $k \in \{1, 2, 3\}$.

Lagrange multipliers

The yet undetermined Lagrange multipliers are defined by λ_{12} and λ_{23}. We compute these multipliers by imposing the constraints. According to Eq. (31), the constraint forces are

$$\mathbf{g}_1 = \lambda_{12}\mathbf{r}_{12}, \quad \mathbf{g}_2 = \lambda_{23}\mathbf{r}_{23} - \lambda_{12}\mathbf{r}_{12}, \quad \mathbf{g}_3 = -\lambda_{23}\mathbf{r}_{23}. \qquad (32)$$

Lagrange multipliers

The previous equations describe nothing but a linear spring with a yet to be determined spring constant $\lambda_{(\cdot)}$. To obtain the values of the Lagrange multipliers $\lambda_{(\cdot)}$, the Verlet algorithm is executed in two steps. We first compute the Verlet update without constraint to obtain

$$\tilde{\mathbf{x}}_i\left(t+\Delta t\right) = 2\mathbf{x}_i - \mathbf{x}_i\left(t-\Delta t\right) + \Delta t^2 \frac{\mathbf{f}_i}{m_i}. \tag{33}$$

Lagrange multipliers

Then we correct the value using the constraints according to

$$\mathbf{x}_i\left(t+\Delta t\right)=\tilde{\mathbf{x}}_i\left(t+\Delta t\right)+\Delta t^2\frac{\mathbf{g}_i}{m_i}. \tag{34}$$

Lagrange multipliers

By combining Eqs. (34) and (31), the updated positions are given by

$$\mathbf{x}_1\left(t+\Delta t\right)=\tilde{\mathbf{x}}_1\left(t+\Delta t\right)+\Delta t^2\frac{\lambda_{12}}{m_1}\mathbf{r}_{12}\left(t\right), \tag{35}$$

$$\mathbf{x}_2\left(t+\Delta t\right)=\tilde{\mathbf{x}}_2\left(t+\Delta t\right)+\Delta t^2\frac{\lambda_{23}}{m_2}\mathbf{r}_{23}\left(t\right)-\Delta t^2\frac{\lambda_{12}}{m_2}\mathbf{r}_{12}\left(t\right), \tag{36}$$

$$\mathbf{x}_3\left(t+\Delta t\right)=\tilde{\mathbf{x}}_3\left(t+\Delta t\right)-\Delta t^2\frac{\lambda_{23}}{m_3}\mathbf{r}_{23}\left(t\right). \tag{37}$$

Lagrange multipliers

With these expressions, we now obtain λ_{12} and λ_{23} by inserting (35), (36) and (37) into the constraint condition, i.e.,

$$\begin{aligned} \left|\mathbf{x}_1\left(t+\Delta t\right)-\mathbf{x}_2\left(t+\Delta t\right)\right|^2 &= d_{12}^2, \\ \left|\mathbf{x}_2\left(t+\Delta t\right)-\mathbf{x}_3\left(t+\Delta t\right)\right|^2 &= d_{23}^2, \end{aligned} \tag{38}$$

and finally

$$\begin{aligned} \left|\tilde{\mathbf{r}}_{12}\left(t+\Delta t\right)+\Delta t^2\lambda_{12}\left(\frac{1}{m_1}+\frac{1}{m_2}\right)\mathbf{r}_{12}\left(t\right)-\Delta t^2\frac{\lambda_{23}}{m_2}\mathbf{r}_{23}\left(t\right)\right|^2 &= d_{12}^2, \\ \left|\tilde{\mathbf{r}}_{23}\left(t+\Delta t\right)+\Delta t^2\lambda_{23}\left(\frac{1}{m_2}+\frac{1}{m_3}\right)\mathbf{r}_{23}\left(t\right)-\Delta t^2\frac{\lambda_{12}}{m_2}\mathbf{r}_{12}\left(t\right)\right|^2 &= d_{23}^2, \end{aligned} \tag{39}$$

where $\tilde{\mathbf{r}}_{ij}=\tilde{\mathbf{x}}_i-\tilde{\mathbf{x}}_j$. These coupled quadratic equations are solved in practice perturbatively by linearising in t.