Practical Course on Molecular Dynamics and Trajectory Analysis

Episode 2: Running simulations

Jordi Villà i Freixa

Universitat de Vic - Universitat Central de Catalunya
Facultat de Ciències, Tecnologia i Enginyeries (FCTE)

jordi.villa@uvic.cat

MD Course and Trajectory Analysis
Concepcion, January 2026

FACULTAT DE CIÈNCIES, TECNOLOGIA I ENGINYERIES
UVIC | UVIC-UCC

Contents

Speed vs. accuracy trade-offs

- Speed: large time steps and reduced reporting.
- Numerical accuracy: symplectic integrators, constraints, and a small timestep.
- OpenMM exposes knobs for `dt`, `constraints`, `hydrogenMass`, and thermostats (§3.7.6–3.7.8).

Time step and constraints

$$\Delta t \approx \frac{1}{10\,\omega_{\max}}, \quad \omega_{\max} \sim \sqrt{\frac{k}{\mu}}.$$

- With `constraints=HBonds`, the timestep doubles.
- `constraints=AllBonds/HAngles` allow bigger jumps but reduce flexibility.

Heavy hydrogens

$$m_H^{\text{new}} = \alpha\, m_H, \quad m_{heavy}^{\text{new}} = m_{heavy} - (\alpha - 1) m_H.$$

- `hydrogenMass=1.5*amu` slows oscillations without changing the total mass.
- Used in integration scripts, e.g., `argon-chemical-potential.py` adopts this trick.

Thermal and pressure coupling

- `LangevinIntegrator` simulates a heat bath with friction γ.
- Nosé-Hoover (chain) controls the canonical distribution while preserving momentum.
- Monte Carlo or Parrinello-Rahman barostats appear in the `simulateAmber.py` scripts.

Reporting and checkpoints

- ‘StateDataReporter‘ tracks energy, temperature, volume, and RMSD.
- ‘DCDReporter‘ / ‘NetCDFReporter‘ save trajectories (OpenMM App §3.13–3.14).
- Use ‘CheckpointReporter‘ to resume long simulations (OpenMM §3.15).

Continuous and discrete equations

$$\dot{\mathbf{r}}(t) = \mathbf{v}(t), \quad \dot{\mathbf{v}}(t) = \frac{\mathbf{F}(t)}{m}.$$

- Simulation replaces derivatives with finite increments.

Time step

$$\mathbf{r}(t + \Delta t) = \mathbf{r}(t) + \mathbf{v}(t)\Delta t + \mathcal{O}(\Delta t^2).$$

- Δt must resolve the fastest vibrations.

Local and global error

$$\text{error local} = \mathcal{O}(\Delta t^{p+1}),$$
$$\text{error global} = \mathcal{O}(\Delta t^{p}).$$

- Order-p integrators control the accumulation of error.

Stability scale

$$\Delta t \lesssim \frac{1}{10\,\omega_{\max}}.$$

- $\omega_{\max}$ comes from the fastest vibrational modes.

Practical guideline

- Without constraints: $\Delta t \approx 1$ fs.
- With SHAKE/RATTLE: $\Delta t \approx 2$ fs.

Explicit Euler

$$\mathbf{v}_{t+\Delta t} = \mathbf{v}_t + \frac{\mathbf{F}_t}{m}\Delta t,$$
$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}_t \Delta t.$$

- Simple but unstable for molecular dynamics.

Classical Verlet

$$\mathbf{r}_{t+\Delta t} = 2\mathbf{r}_t - \mathbf{r}_{t-\Delta t} + \frac{\mathbf{F}_t}{m}\Delta t^2.$$

- Symmetric and time-reversible.

Velocity Verlet

$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}_t \Delta t + \frac{\mathbf{F}_t}{2m} \Delta t^2,$$
$$\mathbf{v}_{t+\Delta t} = \mathbf{v}_t + \frac{\mathbf{F}_t + \mathbf{F}_{t+\Delta t}}{2m} \Delta t.$$

- Combines stability with direct access to velocities.

Symmetry and stability

- Symplectic integrators conserve volume in phase space.
- They prevent energy drift in NVE.

Practical choice

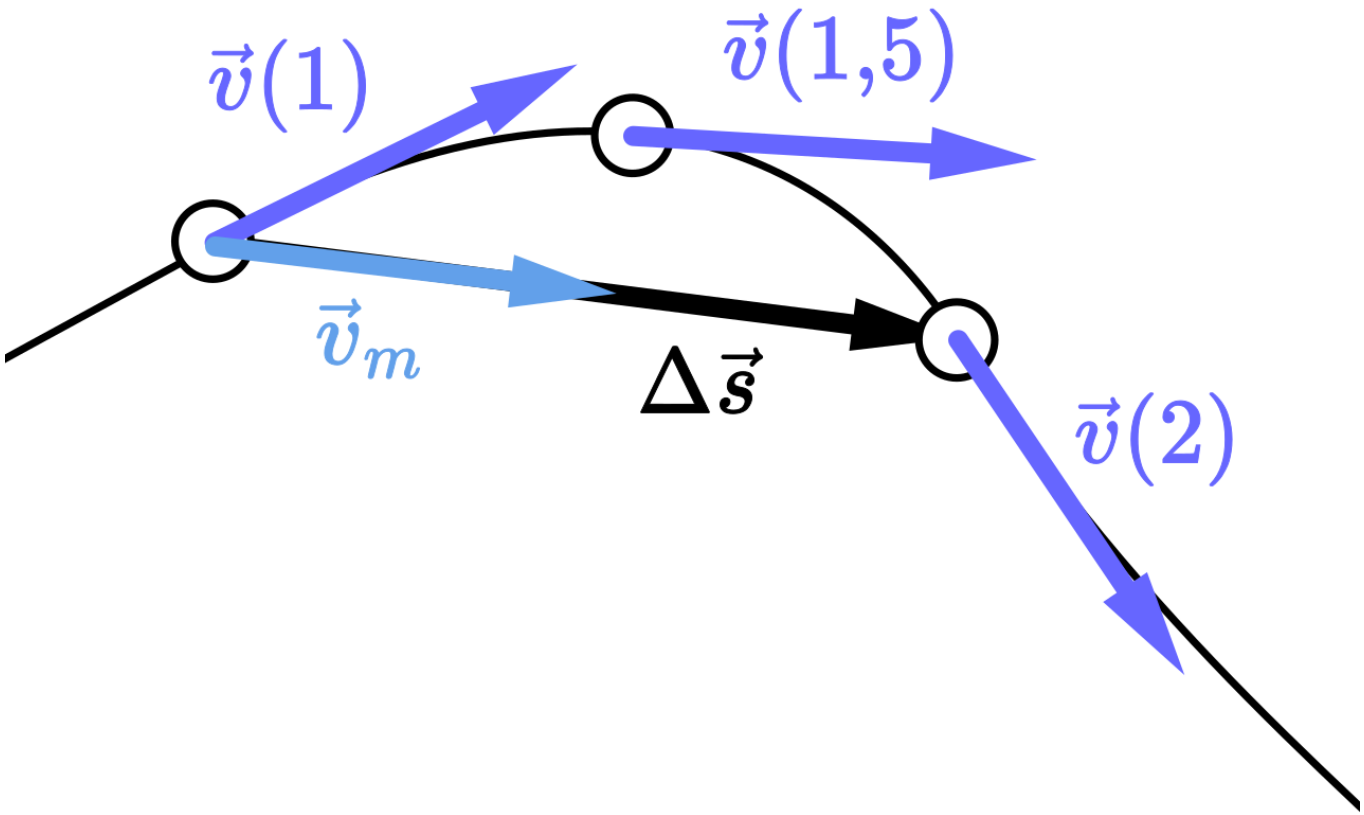
- OpenMM uses Verlet- or Langevin-type integrators.
- The choice depends on the desired ensemble.

Leapfrog

$$\mathbf{v}_{t+\Delta t/2} = \mathbf{v}_{t-\Delta t/2} + \frac{\mathbf{F}_t}{m}\Delta t,$$
$$\mathbf{r}_{t+\Delta t} = \mathbf{r}_t + \mathbf{v}_{t+\Delta t/2}\Delta t.$$

- Velocities and positions "leapfrog" through time.

Graphical interpretation



Source: Wikimedia Commons (CC BY-SA 4.0). [2]

Advantages

- Symmetric, reversible, and inexpensive.
- Preserves energy better than Euler.

Relation to Verlet

- Leapfrog and velocity Verlet are algebraically equivalent.
- Differences appear in how velocities are stored.

Integrator choice

- NVE: Verlet or Leapfrog.
- NVT: Langevin or Nosé-Hoover.

Total energy

$$E(t) = K(t) + U(t).$$

- In ideal NVE, $E(t)$ is constant.

Energy drift

- Numerical errors produce secular drift.
- Controlled by reducing Δt.

Instantaneous temperature

$$T = \frac{2\langle K \rangle}{3Nk_{\mathrm{B}}}.$$

- In NVE, T fluctuates around a mean value.

Quick checks

- Check stability of $E(t)$ and the velocity distribution.
- Compare $\langle K \rangle$ with the theoretical value.

Equilibration time

- A pre-equilibration stabilizes the energy before collecting data.

Thermostat goal

- Impose a canonical distribution at temperature T.
- Extract or inject energy in a controlled way.

Langevin dynamics

$$m\ddot{\mathbf{r}} = -\nabla U - \gamma m\dot{\mathbf{r}} + \mathbf{R}(t),$$
$$\langle \mathbf{R}(t)\mathbf{R}(t')\rangle = 2\gamma m k_{\mathrm{B}} T\, \delta(t - t').$$

- Stochastic thermostat with friction and noise.

Andersen

- Reassigns velocities randomly with frequency ν.
- Good for sampling, less realistic dynamically.

Nosé-Hoover

$$\dot{\mathbf{p}}_i = \mathbf{F}_i - \xi \mathbf{p}_i, \quad \dot{\xi} = \frac{1}{Q}\left(\sum_i \frac{\mathbf{p}_i^2}{m_i} - 3Nk_{\mathrm{B}}T\right).$$

- Deterministic thermostat with an extended variable.

Thermostat choice

- Langevin: robust and stable.
- Nosé-Hoover: more realistic dynamics if well tuned.

Microscopic pressure

$$P = \frac{N k_B T}{V} + \frac{1}{3V} \sum_{i<j} \mathbf{r}_{ij} \cdot \mathbf{F}_{ij}.$$

- Pressure depends on kinetic energy and internal forces.

Berendsen barostat

$$\frac{dV}{dt} = \frac{1}{\tau_P}(P_0 - P)V.$$

- Fast to equilibrate, does not reproduce exact fluctuations.

Parrinello-Rahman

- Scales the simulation cell with dynamic variables.
- Allows anisotropic volume changes.

NPT

- Realistic ensemble for experimental conditions.
- Combines thermostat and barostat.

Practical choice

- Equilibration: Berendsen + Langevin.
- Production: more rigorous barostat (Monte Carlo barostat).

Motivation

- Eliminating fast vibrations allows increasing Δt.

Holonomic constraints

$$g_k(\mathbf{r}) = 0, \quad k = 1, \ldots, M.$$

- Fix distances or internal angles.

SHAKE/RATTLE

- SHAKE corrects positions; RATTLE corrects positions and velocities.

Numerical impact

- Increases stability and reduces simulation cost.
- May alter high-frequency modes.

Recommendation

- Use constraints on H bonds for biomolecules.

Motivation for PBC

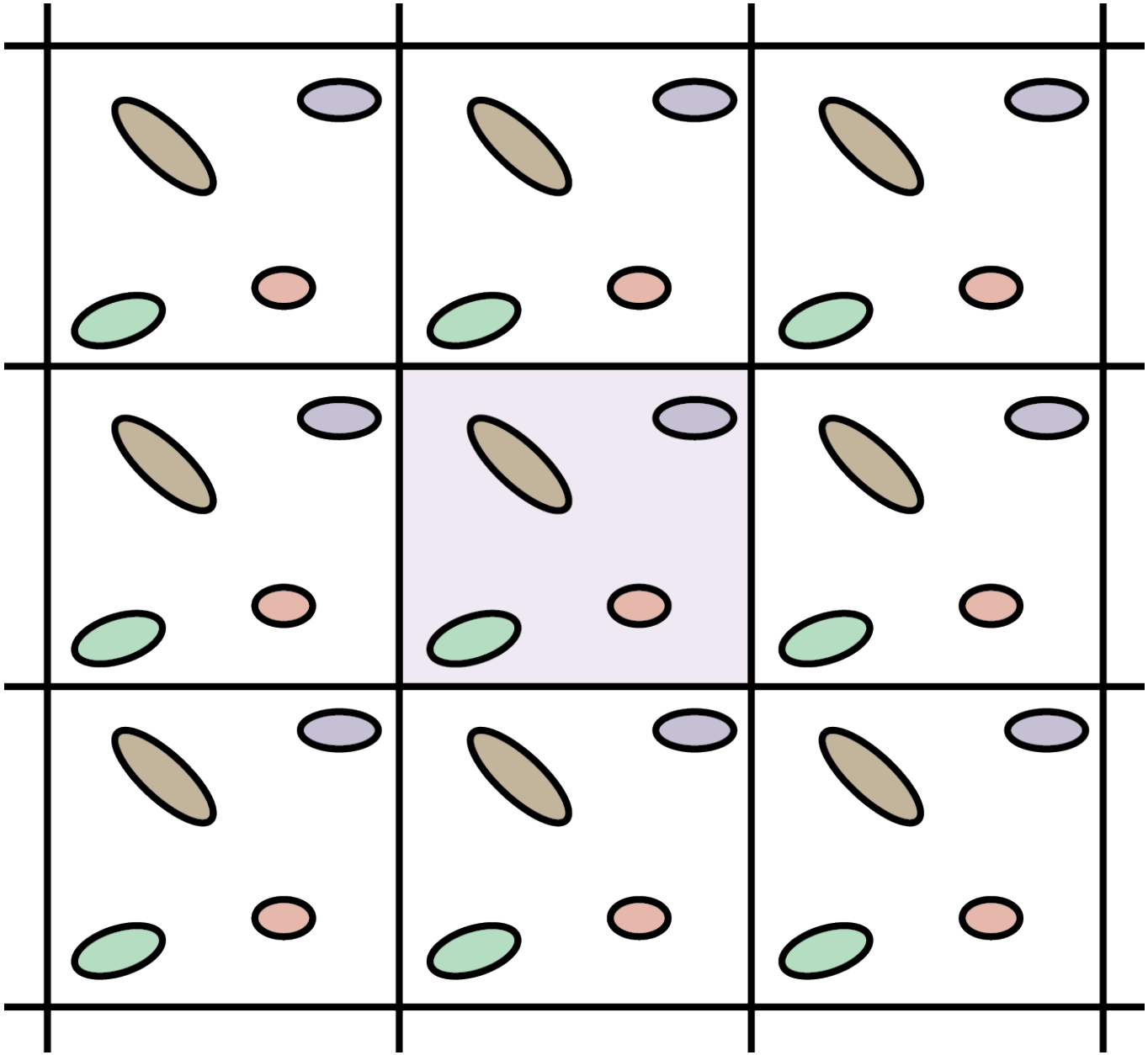
- Avoids surface effects in finite systems.
- Mimics an infinite system by replicating the cell.

Periodic cell

$$\mathbf{r} \rightarrow \mathbf{r} + n_x\mathbf{a} + n_y\mathbf{b} + n_z\mathbf{c}.$$

- $\mathbf{a}, \mathbf{b}, \mathbf{c}$ define the simulation box.

2D example



Source: Wikimedia Commons (CC BY SA 4.0) [2]

Implications for dynamics

- Particles leaving re-enter through the opposite face.
- The density in the simulated volume stays constant.

Non-orthogonal boxes

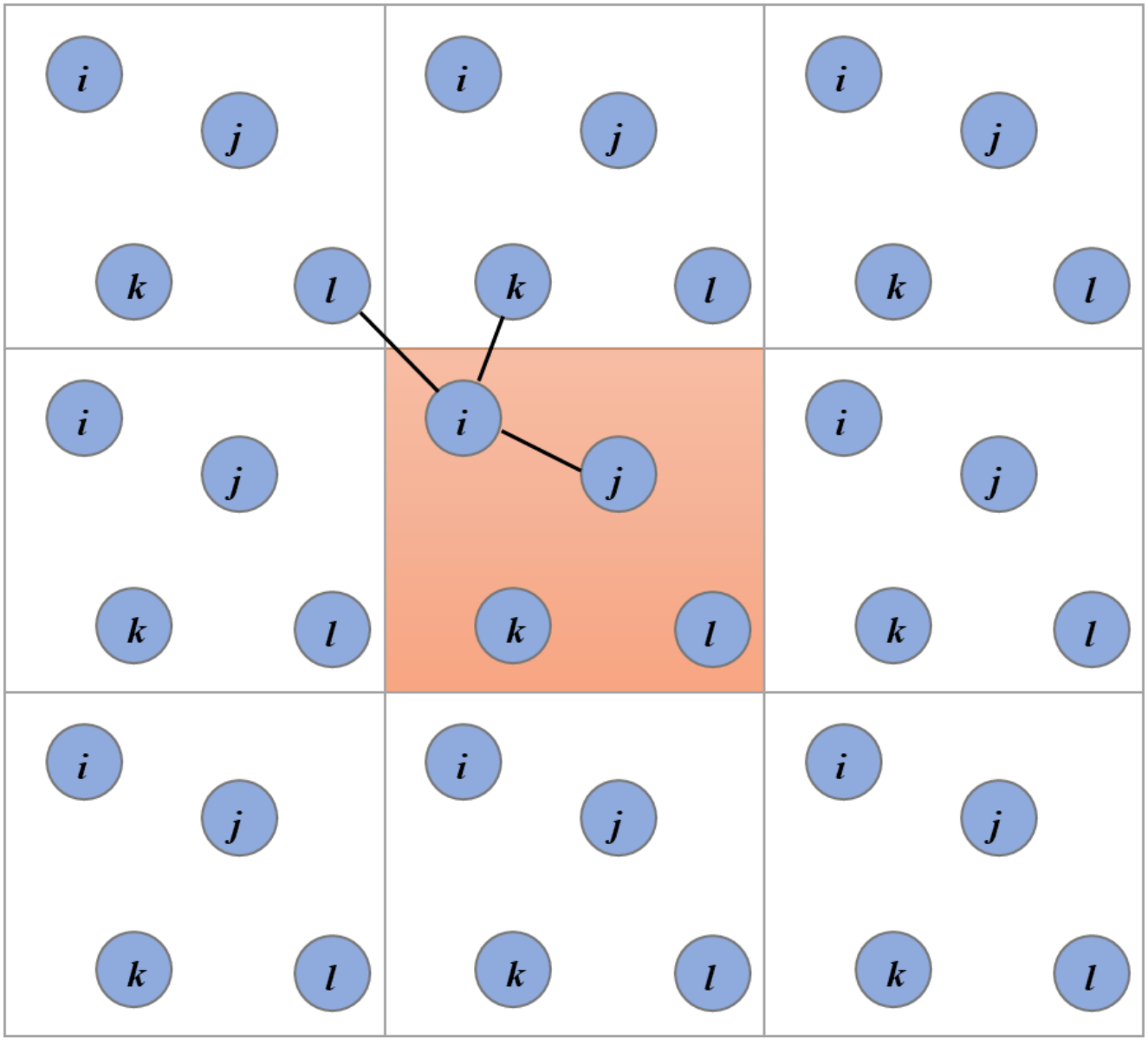
- Triclinic cells for crystals or tilted membranes.

Minimum image convention

$$r_{ij} = \min_{\mathbf{n}} \|\mathbf{r}_i - \mathbf{r}_j + \mathbf{n}\|.$$

- Use the nearest periodic image to compute interactions.

Visual example



Source: Wikimedia Commons (CC BY-SA 4.0). [1]

Cutoff radius

- The minimum image requires $r_c < L/2$ to avoid double counting.

Energy impact

- Errors arise if the cutoff is large or the density is low.

Best practices

- Adjust r_c and box size according to the system density.

Short-range interactions

- LJ is truncated at r_c with or without a smoothing function.

Switch function

$$u_{\text{switch}}(r) = s(r)\, u(r), \quad s(r) \in [0, 1].$$

- Avoids discontinuities in forces.

Tail correction

$$U_{\text{tail}} \approx 2\pi\rho \int_{r_c}^{\infty} r^2 u(r)\, dr.$$

- Corrects energy and pressure when truncating LJ.

Choosing r_c

- Larger r_c improves accuracy but increases cost.

Stability implications

- An abrupt cutoff can introduce noise in the dynamics.

Coulomb problem

$$u_C(r) \propto \frac{1}{r} \quad \text{(long range)}.$$

- Direct truncation produces large errors.

Ewald

$$U = U_{\text{real}} + U_{\text{rec}} + U_{\text{self}}.$$

- Split Coulomb into real and reciprocal sums.

PME

- Particle Mesh Ewald uses FFT to speed up the calculation.

Accuracy

- Control the error with mesh parameters and the Ewald alpha.

Practical choice

- For solvated biomolecules, PME is the standard.

Trajectories

- Save coordinates and velocities at regular intervals.
- Balance between storage and temporal resolution.

Observables

$$\langle A \rangle \approx \frac{1}{M} \sum_{k=1}^{M} A_k .$$

- Autocorrelation determines the effective number of samples.

Random seeds

- Document seeds to reproduce stochastic thermostats.

Checkpoints

- Save states to restart long simulations.

Episode summary

- Numerical integration controls stability and accuracy.
- Thermostats and barostats set the ensemble.
- PBC and long-range electrostatics are critical in biomolecules.

References I

[1] Abusaleh AA. *Minimum Image Convention*. CC BY-SA 4.0. URL: https://commons.wikimedia.org/wiki/File:Minimum_Image_Convention.png (visited on 01/12/2026).

[2] MikeRun. *Leapfrog-method-argument*. CC BY-SA 4.0. URL: https://commons.wikimedia.org/wiki/File:Leapfrog-method-argument.svg (visited on 01/12/2026).

[3] Christopher Rowley. *Periodic Boundary Conditions in 2D*. CC BY-SA 4.0. URL: https://commons.wikimedia.org/wiki/File:Periodic_Boundary_Conditions_in_2D.png (visited on 01/12/2026).