

# 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

- 1 Episode 2: Running simulations
  - Key parameters (OpenMM User Guide §3.7)
  - Time discretization
  - Classical integrators
  - Leapfrog scheme
  - Conservación y NVE
  - Termostatos
  - Barostatos
  - Restricciones
  - Condiciones periódicas
  - Minimum image
  - Cutoffs y suavizado
  - Electrostática de largo alcance
  - Salida y reproducibilidad

# 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

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

- `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  $\gamma$ .
- 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

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

- $\Delta t$  must resolve the fastest vibrations.

# Local and global error

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

- Order- $p$  integrators control the accumulation of error.

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

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

# Practical guideline

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

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

- Simple but unstable for molecular dynamics.

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

- Symmetric and time-reversible.

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

- Combines stability with direct access to velocities.

# Symmetry and stability

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



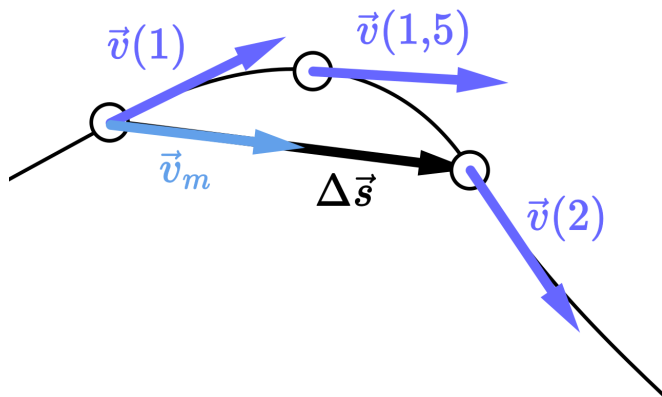
# Practical choice

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

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

- 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.

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

- En NVE ideal,  $E(t)$  es constante.

# Deriva energética

- Errores numéricos producen deriva secular.
- Se controla reduciendo  $\Delta t$ .



# Temperatura instantánea

$$T = \frac{2\langle K \rangle}{3Nk_B}.$$

- En NVE,  $T$  fluctúa alrededor de un valor medio.

# Comprobaciones rápidas

- Estabilidad de  $E(t)$  y distribución de velocidades.
- Comparar  $\langle K \rangle$  con el valor teórico.

# Tiempo de equilibración

- Un pre-equilibrado estabiliza la energía antes de producir datos.

# Objetivo del termostato

- Imponer una distribución canónica a temperatura  $T$ .
- Extraer o inyectar energía de forma controlada.

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

- Termostato estocástico con fricción y ruido.

- Reasigna velocidades aleatoriamente con frecuencia  $\nu$ .
- Bueno para muestreo, menos realista dinámicamente.

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

- Termostato determinista con variable extendida.

- Langevin: robusto y estable.
- Nosé-Hoover: dinámica más realista si está bien parametrizado.



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

- La presión depende de energía cinética y fuerzas internas.

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

- Rápido para equilibrar, no reproduce fluctuaciones exactas.

- Escala la celda de simulación con variables dinámicas.
- Permite cambios anisotrópicos del volumen.

- Ensemble realista para condiciones experimentales.
- Combina termostato y barostato.

- Equilibración: Berendsen + Langevin.
- Producción: barostato más riguroso (Monte Carlo barostat).

- Eliminar vibraciones rápidas permite aumentar  $\Delta t$ .

# Restricciones holónomas

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

- Fijan distancias o ángulos internos.

- SHAKE corrige posiciones; RATTLE corrige posiciones y velocidades.



- Aumenta estabilidad y reduce coste de simulación.
- Puede alterar modos de alta frecuencia.

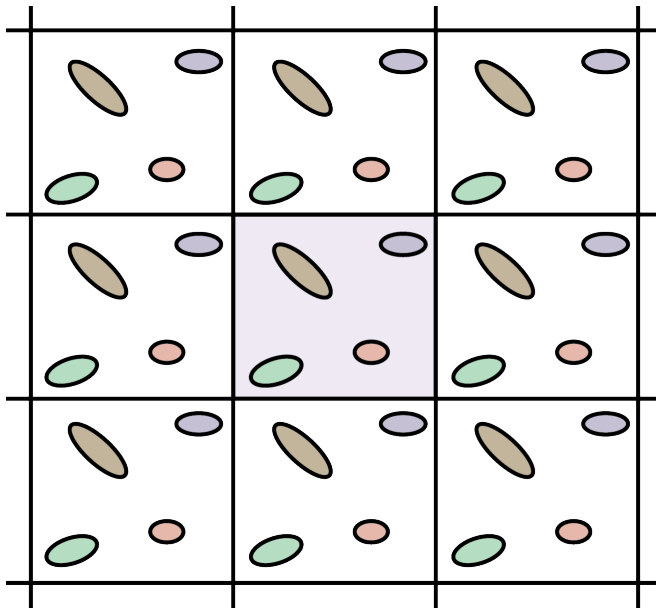
- Usar restricciones en enlaces con H para biomoléculas.

- Evita efectos de superficie en sistemas finitos.
- Emula un sistema infinito por replicación de la celda.

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

- $\mathbf{a}, \mathbf{b}, \mathbf{c}$  definen la caja de simulación.

## Ejemplo 2D



# Implicaciones en dinámica

- Las partículas que salen reingresan por la cara opuesta.
- Se preserva la densidad en el volumen simulado.

# Cajas no ortogonales

- Celdas triclinicas para cristales o membranas inclinadas.

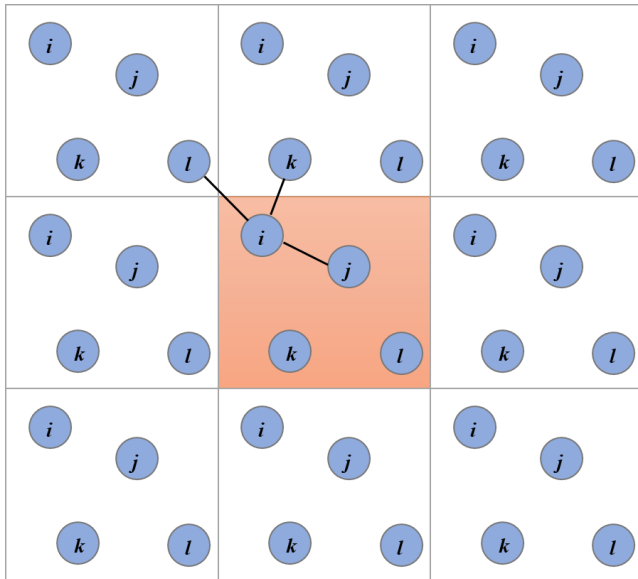
# Convención de imagen mínima

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

- Usa la imagen periódica más cercana para calcular interacciones.



# Ejemplo visual



- La imagen mínima requiere  $r_c < L/2$  para evitar dobles conteos.

- Errores si el corte es grande o la densidad baja.

- Ajustar  $r_c$  y tamaño de caja según la densidad del sistema.

# Interacciones de corto alcance

- LJ se trunca a  $r_c$  con o sin función de suavizado.

# Función switch

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

- Evita discontinuidades en fuerzas.

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

- Corrige energía y presión al truncar LJ.

# Escogiendo $r_c$

- Mayor  $r_c$  mejora precisión, pero aumenta coste.



# Implicaciones en estabilidad

- Un corte abrupto puede introducir ruido en la dinámica.

# Problema de Coulomb

$$u_C(r) \propto \frac{1}{r} \quad (\text{largo alcance}).$$

- El truncamiento directo produce grandes errores.

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

- Divide Coulomb en suma real y recíproca.

- Particle Mesh Ewald usa FFT para acelerar el cálculo.

- Controla el error con parámetros de malla y  $\alpha$  de Ewald.

- Para biomoléculas en solvente, PME es el estándar.

- Guardar coordenadas y velocidades en intervalos regulares.
- Balance entre almacenamiento y resolución temporal.

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

- Autocorrelation determines the effective number of samples.



- 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](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](https://commons.wikimedia.org/wiki/File:Periodic_Boundary_Conditions_in_2D.png) (visited on 01/12/2026).