

# Practical Course on Molecular Dynamics and Trajectory Analysis

## Episode 1: Introduction (environment and data)

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 Episodio 1: Introducción (entorno y datos)
  - Capa de aplicación de OpenMM
  - Panorama y escalas
  - Variables y estados
  - Entropía y energía libre
  - Ensembles y partición
  - Hamiltoniano y dinámica
  - Fuerzas y gradientes
  - Partición y probabilidades
  - Distribución de Boltzmann
  - Espacio de fases y ergodicidad
  - Superficie de energía potencial
  - Campos de fuerza
  - Entorno y datos

# Flujo general de OpenMM

- Cargar estructuras PDB/GROMACS/AMBER/CHARMM/TINKER desde los scripts oficiales (p.ej., `simulatePdb.py`, `simulateGromacs.py`, `simulateAmber.py` y `simulateCharmm.py`).
- Generar objetos 'Topology', 'System' e 'Integrator' y sincronizar 'StateReporter'/'TrajectoryReporter' para registrar energías, fuerzas y coordenadas.
- Ejecutar integraciones con pasos definidos desde la capa de aplicación y guardar checkpoints con la rutina de OpenMM-Setup (User Guide §3.5-3.15).

# Entradas externas

- `simulateAmber.py` acepta `'prmtop' + '.inpcrd'`, mientras que `simulateCharmm.py` y `simulateGromacs.py` respetan sus formatos nativos.
- `simulatePdb.py` es el punto de partida tradicional para el estudio de alanina o proteínas-ligando sin parametrizar.
- Todos incluyen funciones para modificar 'reporters' y acomodar formatos de salida que luego procesaremos con scripts Python de análisis.

# Objetivos del episodio

- Conectar termodinámica, mecánica estadística y simulación.
- Definir variables, funciones de estado y observables.
- Establecer el flujo de datos y el entorno de trabajo.

# Escalas de tiempo y longitud

$$\Delta t_{\text{vibraciones}} \sim 10^{-15} \text{ s},$$

$$\Delta t_{\text{conformaciones}} \sim 10^{-9} \text{ s a } 10^{-3} \text{ s},$$

$$L_{\text{atomos}} \sim 10^{-10} \text{ m}, \quad L_{\text{proteinas}} \sim 10^{-8} \text{ m}.$$

- La simulación atómica resuelve escalas ultrarrápidas.
- La estadística conecta microestados con promedios macroscópicos.

$$\text{Estado} = (\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{p}_1, \dots, \mathbf{p}_N),$$
$$\mathbf{p}_i = m_i \dot{\mathbf{r}}_i.$$

- El sistema está definido por  $6N$  coordenadas en el espacio de fases.
- La energía potencial codifica la interacción entre partículas.

$$\langle A \rangle = \int A(\mathbf{r}, \mathbf{p}) \rho(\mathbf{r}, \mathbf{p}) d\mathbf{r} d\mathbf{p},$$

$$\bar{A} = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(t) dt.$$

- La ergodicidad conecta promedios temporales y de conjunto.



# Pipeline de simulación

- 1 Preparación del sistema (topología, coordenadas, parámetros).
- 2 Definición del Hamiltoniano y condiciones de contorno.
- 3 Integración temporal y generación de trayectoria.
- 4 Análisis estadístico de observables.

$$P, V, T, U, H = U + PV, G = H - TS, A = U - TS.$$

- Funciones de estado: dependen solo del estado macroscópico.
- Funciones de trayectoria: trabajo  $W$  y calor  $Q$ .

# Primera ley y trabajo

$$dU = \delta Q + \delta W,$$
$$\delta W = -P dV + \sum_i f_i dx_i.$$

- La energía interna cambia por intercambio de calor o trabajo.
- En simulación,  $\delta W$  proviene de fuerzas sobre partículas.

## Segunda ley y entropía

$$dS \geq \frac{\delta Q}{T},$$
$$S = k_B \ln \Omega.$$

- $\Omega$  es el número de microestados compatibles con el macroestado.

# Potenciales termodinámicos

$$G = U + PV - TS,$$
$$A = U - TS.$$

- $G$  minimiza a  $T, P$  constantes;  $A$  minimiza a  $T, V$  constantes.

$$dG = -S dT + V dP + \sum_i \mu_i dN_i,$$
$$dA = -S dT - P dV + \sum_i \mu_i dN_i.$$

- Las derivadas parciales dan observables medibles.

$$S = k_B \ln \Omega, \quad \Omega = \sum_{\text{micro}} 1.$$

- Incrementos de entropía favorecen estados con más configuraciones.

$$\Delta G = \Delta H - T\Delta S,$$
$$\Delta A = \Delta U - T\Delta S.$$

- En equilibrio,  $\Delta G = 0$  para procesos a  $T, P$  fijos.



$$\mu_i = \left( \frac{\partial G}{\partial N_i} \right)_{T, P, N_{j \neq i}} .$$

- Controla intercambios de materia y equilibrio de fases.

$$P(\text{estado}) \propto e^{-\beta G}, \quad \beta = 1/(k_B T).$$

- La energía libre gobierna el peso estadístico de macroestados.

## Ejemplo: equilibrio $A \rightleftharpoons B$

$$K = e^{-\beta \Delta G},$$
$$\Delta G = -k_B T \ln K.$$

- La estabilidad relativa se expresa en términos de  $\Delta G$ .

# Ensembles clásicos

- Microcanónico (NVE):  $N, V, E$  fijos.
- Canónico (NVT):  $N, V, T$  fijos.
- Isothermo-isobárico (NPT):  $N, P, T$  fijos.

# Función de partición canónica

$$Z = \frac{1}{h^{3N} N!} \int e^{-\beta H(\mathbf{r}, \mathbf{p})} d\mathbf{r} d\mathbf{p}.$$

- Central para derivar energías libres y promedios.

$$\langle A \rangle = \frac{1}{Z} \int A e^{-\beta H} d\mathbf{r} d\mathbf{p}.$$

- En simulación, se estiman con promedios temporales.

$$\Delta = \int dV e^{-\beta P V} Z(N, V, T),$$
$$G = -k_B T \ln \Delta.$$

- Incluye fluctuaciones de volumen controladas por el barostato.

$$C_V = \frac{\langle E^2 \rangle - \langle E \rangle^2}{k_B T^2},$$
$$\kappa_T = \frac{\langle V^2 \rangle - \langle V \rangle^2}{k_B T \langle V \rangle}.$$

- Las fluctuaciones conectan estadística y respuesta macroscópica.



$$H(\mathbf{r}, \mathbf{p}) = \sum_{i=1}^N \frac{\mathbf{p}_i^2}{2m_i} + U(\mathbf{r}).$$

- Se separa en energía cinética y potencial.

$$\dot{\mathbf{r}}_i = \frac{\partial H}{\partial \mathbf{p}_i},$$
$$\dot{\mathbf{p}}_i = -\frac{\partial H}{\partial \mathbf{r}_i}.$$

- Equivalentes a la segunda ley de Newton.

$$m_i \ddot{\mathbf{r}}_i = -\nabla_i U(\mathbf{r}).$$

- La fuerza es el gradiente negativo del potencial.

$$\frac{dH}{dt} = 0 \quad (\text{NVE ideal}).$$

- La integración numérica introduce errores controlables.

# Flujo en el espacio de fases

$$\frac{d\rho}{dt} = \{\rho, H\} = 0 \quad (\text{ecuación de Liouville}).$$

- La densidad de probabilidad se conserva en el flujo hamiltoniano.

# Fuerzas a partir del potencial

$$\mathbf{F}_i = -\nabla_i U(\mathbf{r}).$$

- En MD, el coste dominante es evaluar  $U$  y  $\mathbf{F}_i$ .

$$U = \sum_{i < j} u(r_{ij}), \quad r_{ij} = \|\mathbf{r}_i - \mathbf{r}_j\|.$$

- Simplifican la energía total con interacciones binarias.

## Ejemplo: Lennard-Jones

$$u_{\text{LJ}}(r) = 4\epsilon \left[ \left( \frac{\sigma}{r} \right)^{12} - \left( \frac{\sigma}{r} \right)^6 \right].$$

- Repulsión a corta distancia, atracción a media distancia.



$$u_C(r) = \frac{1}{4\pi\epsilon_0} \frac{q_i q_j}{r_{ij}}.$$

- Dominante en sistemas biomoleculares.

$$\nabla U \equiv \left( \frac{\partial U}{\partial r_1}, \dots, \frac{\partial U}{\partial r_{3N}} \right),$$

$$\text{Hess} = \nabla \nabla U.$$

- El Hessiano describe curvaturas locales (modos normales).

# Separación cinética/potencial

$$Z = Z_{\text{kin}} Z_{\text{conf}},$$
$$Z_{\text{kin}} = \prod_i \left( \frac{2\pi m_i}{\beta h^2} \right)^{3/2}.$$

- $Z_{\text{conf}}$  depende solo de  $U(\mathbf{r})$ .

# Distribución de configuraciones

$$P(\mathbf{r}) = \frac{e^{-\beta U(\mathbf{r})}}{Z_{\text{conf}}}.$$

- Base del muestreo Monte Carlo y MD con termostato.

# Energía libre configuracional

$$A = -k_B T \ln Z_{\text{conf}} + \text{const.}$$

- Relaciona el muestreo con  $\Delta A$  entre estados.

$$P(\mathbf{v}) \propto e^{-\beta \sum_i \frac{1}{2} m_i v_i^2}.$$

- Permite inicializar velocidades a temperatura  $T$ .

# Teorema de equipartición

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

- Cada grado de libertad cuadrático aporta  $\frac{1}{2} k_B T$ .

$$P(E) = \frac{1}{Z} e^{-\beta E}.$$

- El peso exponencial penaliza estados de alta energía.



# Distribución de Maxwell-Boltzmann

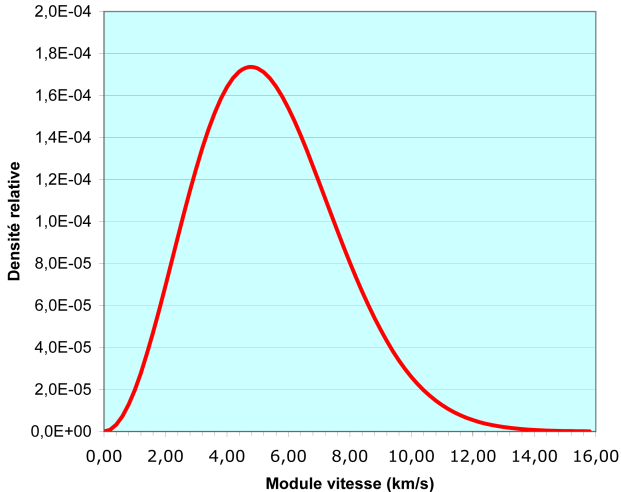
$$f(v) = 4\pi \left( \frac{m}{2\pi k_B T} \right)^{3/2} v^2 e^{-\frac{mv^2}{2k_B T}}.$$

- Caracteriza la distribución de velocidades en equilibrio.

- La mayoría de partículas se agrupan alrededor de  $v_m$ .
- El ensanchamiento crece con  $T$ .

## Ejemplo gráfico

**Distribution des vitesses  
des molécules d'hélium à 5500 K  
(Distribution de Maxwell-Boltzmann)**



- Inicializar velocidades según  $f(v)$  evita sesgos térmicos.
- Permite comprobar que el termostato reproduce el equilibrio.

$$\Gamma = (\mathbf{r}, \mathbf{p}) \in \mathbb{R}^{6N}.$$

- Cada punto representa un microestado completo.

# Ecuación de Liouville

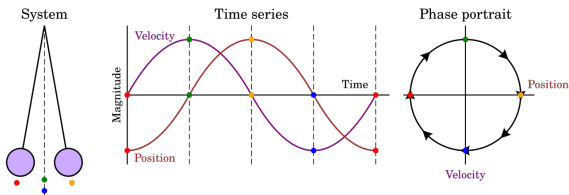
$$\frac{\partial \rho}{\partial t} + \{\rho, H\} = 0.$$

- El volumen en  $\Gamma$  se conserva (teorema de Liouville).

$$\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T A(t) dt = \langle A \rangle_{\text{ensemble}}.$$

- Supuesto clave para sustituir promedios de conjunto por tiempo.

# Trayectorias en $\Gamma$



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



- Órbitas cerradas: movimiento periódico.
- Órbitas abiertas: difusión en regiones de alta energía.

$$U(\mathbf{r}) : \mathbb{R}^{3N} \rightarrow \mathbb{R}.$$

- Describe el paisaje energético donde se mueve el sistema.

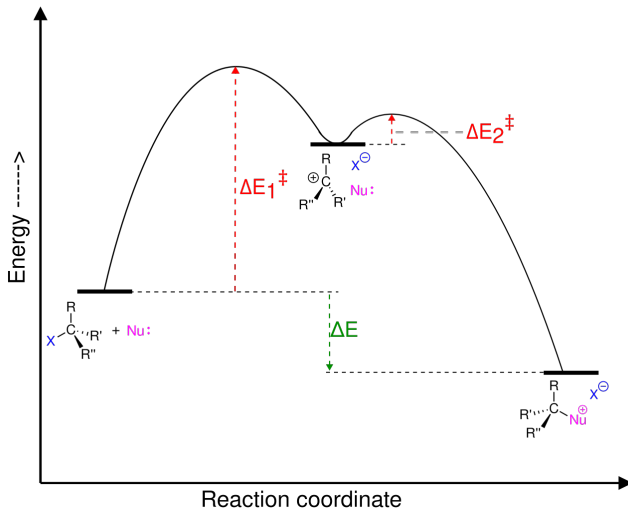
$$\xi = \xi(\mathbf{r}), \quad F(\xi) = -k_B T \ln P(\xi).$$

- Proyección que resume la dinámica en una coordenada clave.

# Barreras y estados metaestables

- Mínimos locales: conformaciones metaestables.
- Barreras: controlan las escalas de transición.

# Ejemplo de superficie



Fuente: Wikimedia Commons (CC BY-SA 4.0). [4]

# Dinámica sobre la superficie

- Integración clásica produce trayectorias en  $U(\mathbf{r})$ .
- Los termostatos aseguran muestreo canónico.

# Descomposición del potencial

$$U = U_{\text{bond}} + U_{\text{angle}} + U_{\text{dihedral}} + U_{\text{nonbond}}.$$

- Separación entre términos enlazados y no enlazados.

$$U_{\text{bond}} = \sum_b k_b (r_b - r_b^0)^2,$$

$$U_{\text{angle}} = \sum_a k_a (\theta_a - \theta_a^0)^2.$$

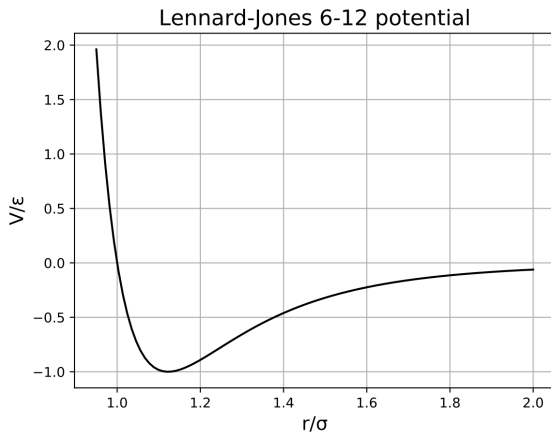
- Aproximación armónica alrededor de geometrías de equilibrio.



$$U_{\text{dihedral}} = \sum_d \frac{V_d}{2} [1 + \cos(n_d \phi_d - \gamma_d)].$$

- Controlan barreras rotacionales y preferencias conformacionales.

# Ejemplo LJ



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

# Parámetros y validación

- Los campos de fuerza se ajustan a datos cuánticos y experimentales.
- La validación incluye densidades, energías y conformaciones.

- Variable de entorno `COURSE_DIR`.
- Estructura: `data/` (entradas) y `results/` (salidas).

- Sistema simple: alanina.
- Sistema complejo: proteína + ligando.

- 1 Descarga de datos base.
- 2 Preparación de topología y coordenadas.
- 3 Validación del entorno de cálculo.

- Controlar versiones de librerías y semillas aleatorias.
- Guardar archivos de entrada y scripts.

# Resumen del episodio

- La base termodinámica se conecta con el muestreo estadístico.
- El potencial define fuerzas y probabilidades.
- El entorno de trabajo garantiza reproducibilidad.



# References I

- [1] AdiPhysics. *Lennard Jones potential graph*. CC BY-SA 4.0. URL: [https://commons.wikimedia.org/wiki/File:Lennard\\_Jones\\_potential\\_graph.svg](https://commons.wikimedia.org/wiki/File:Lennard_Jones_potential_graph.svg) (visited on 01/12/2026).
- [2] Krishnavedala. *Pendulum phase portrait illustration*. CC BY-SA 4.0. URL: [https://commons.wikimedia.org/wiki/File:Pendulum\\_phase\\_portrait\\_illustration.svg](https://commons.wikimedia.org/wiki/File:Pendulum_phase_portrait_illustration.svg) (visited on 01/12/2026).
- [3] Gilles MAIRET. *Distribution de Maxwell-Boltzmann. Distribution des vitesses de l'hélium a 5500 K*. CC BY-SA 3.0. URL: [https://commons.wikimedia.org/wiki/File:Distribution\\_de\\_Maxwell-Boltzmann.\\_Distribution\\_des\\_vitesses\\_de\\_l%27h%C3%A9lium\\_%C3%A0\\_5500\\_K.png](https://commons.wikimedia.org/wiki/File:Distribution_de_Maxwell-Boltzmann._Distribution_des_vitesses_de_l%27h%C3%A9lium_%C3%A0_5500_K.png) (visited on 01/12/2026).

- [4] Unknown author. *Potential Energy Surface for a generic SN1 reaction*. CC BY-SA 4.0. Source: LibreTexts (Organic Chemistry, Morsch et al.)  
URL: [https://commons.wikimedia.org/wiki/File:Potential\\_Energy\\_Surface\\_for\\_a\\_generic\\_SN1\\_reaction.svg](https://commons.wikimedia.org/wiki/File:Potential_Energy_Surface_for_a_generic_SN1_reaction.svg)  
(visited on 01/12/2026).