Statistical Physics and Computer Simulation 529-0483-00L

Exam cheatsheet for the FS24 exam at ETH Zürich by Wu, You

Classical & Quantum Mechanics

Newtonian (Only valid in Cartesian)

System of N particles, $N_{t-t} = 3N$, t-dependent. For each particle $p_t =$ $m_i \dot{r}_i$, where $r^N(t) = \{r_1, ..., r_N\}$

Newton's 2nd law $m_i \ddot{r}_i = \dot{p}_i = F_i$, where $i = 1, 2, ..., N \underline{M} \ddot{r}^N = \dot{p}^N =$ F^N , with $M \in \mathbb{R}^{3N,3N}$ diagonal

Newton's 3rd Law $F_{ii} = -F_{ij}$ with F_{ij} exerted by i

Deterministic Nature Given the EOM and a force expression \rightarrow deterministic. At final state at $t = t_{fin}$, $r^{N}(t_{fin})$ and $p^{N}(t_{fin})$ are predictable.

Conservative Field (CF) Force is field if only *r*-dependent, Curl-free $\mathbb{R} \frac{\partial \mathcal{L}(r^2)}{\partial r^2}$ (Stokes theorem $\frac{\partial}{\partial \mathbf{r}} \times \mathbf{F}_i(\mathbf{r}^N) = 0, \forall i$) implies conservative. Conservative field must fulfill $\oint d\mathbf{r}^N \cdot \mathbf{F}^N(\mathbf{r}^N) = \sum_i^N \oint d\mathbf{r}_i \cdot \mathbf{F}_i(\mathbf{r}^N) = 0$

Potential energy of CF "-" force drives towards lower E_{not} , time-dependency to CF can be brought through $E_{n-t} = \mathcal{V}$

$$\pmb{F}_i(\pmb{r}^N) = -\frac{\partial}{\partial \pmb{r}_i} \mathcal{V}(\pmb{r}^N)$$

$$\begin{array}{l} \underline{\text{Kinetic Energy}} \ \mathcal{K}(\dot{r}^N) = \sum_i^N \frac{m_i r_i^2}{2} = \sum_i^N \frac{m_i^2 \dot{r}_i^2}{2m_i} = \sum_i^N \frac{p_i^2}{2m_i} \\ \underline{\text{Total Energy of CF}} \ \mathcal{V}(r^N) + \mathcal{K}(\dot{r}^N) = E_{\text{tot}} = \text{const} \end{array}$$

Time-dependent
$$\mathcal{V}(\mathbf{r}^N, t) + \mathcal{K}(\dot{\mathbf{r}}^N) = E(t)$$

$$\begin{split} b &= \left(r_{ij} \cdot r_{ij} \right)^{\frac{1}{2}} \\ \theta &= \arccos\left[\frac{r_{ji} \cdot r_{ji}}{r_{ji} \cdot r_{jk}} \right], r_{ji} \coloneqq r_i - r_j \\ F_1 &= \frac{Gm_1m_2}{r^2}, F = -\rho g V e_x, F = q(E + \dot{r} \times B) \end{split}$$

$$F = -k\dot{r}, F = -k(x - x_o)e_s$$

Lagrangian (Valid in any generalized coordinates)

🖒 Enforce constraints, add artificial DOF to dynamics, Generalizing mechanics than just point particles

Generalized Coordinate Systems any set of M scalars (eg. r, a, dihedrals...) sufficient to specify the coordinates of all particles in the **EOM** For m = 1, ..., M there are 2M = 6N 1st-order equations system & exists invertible map to express q_i with $r^N = r^N(q^M) \Leftrightarrow$ $q^M = q^M(r^N)$

Generalized coordinates/Velocities M = 3N for N particles $q^M =$ $\{q_1,...,q_M\}, \dot{\boldsymbol{q}}^M=\{\dot{q}_1,...,\dot{q}_M\}$ s.t. $\dot{\boldsymbol{r}}^N=\dot{\boldsymbol{r}}^N(\boldsymbol{q}^M,\dot{\boldsymbol{q}}^M)$

Energy (Lagrangian CF)

$$\mathcal{L}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M}) = \mathcal{K}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M}) - \mathcal{V}(\boldsymbol{q}^{M})$$

Potential $\mathcal{L}(\mathbf{q}^M, \dot{\mathbf{q}}^M, t) = \mathcal{K}(\mathbf{q}^M, \dot{\mathbf{q}}^M) - \mathcal{V}(\mathbf{q}^M, t), \frac{\partial \mathcal{L}}{\partial t} = -\frac{\partial \mathcal{V}}{\partial t}$ Kinetic (guadratic) $\mathcal{K}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M}) = \frac{1}{2} (\dot{\boldsymbol{q}}^{M})^{T} \boldsymbol{M}(\boldsymbol{q}^{M}) \dot{\boldsymbol{q}}^{M}$ Mass-metric tensor Symmetric but not necessarily diagonal

$$\underline{I}_{kl} = \sum_{i}^{N} m_{i} \frac{\partial r_{i}}{\partial q_{k}} \frac{\partial r_{i}}{\partial q_{l}}$$

EOM (Euler-Lagrange) from principle of least action Physical trajectory between system configurations at two time points minimizes action.

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M})}{\partial \dot{q}_{m}} \right) = \frac{\partial \mathcal{L}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M})}{\partial q_{m}}, \text{with} \ m = 1, 2, ..., M$$

Polar coordinates $(1) \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{a}} \right) = \frac{\partial \mathcal{L}}{\partial \theta} (2) \frac{d}{dt} \left(\frac{\partial \mathcal{L}}{\partial \dot{r}} \right) = \frac{\partial \mathcal{L}}{\partial r}$

Connection to Cartesian

$$\mathcal{L}(oldsymbol{r}^N, \dot{oldsymbol{r}}^N) = \sum_i^N rac{m_i oldsymbol{r}_i^2}{2} - \mathcal{V}(oldsymbol{r}^N)$$

$$\underbrace{\underline{m_i \ddot{r_i}}}_{\mathbb{O}} = \underbrace{\underline{F_i(r^N)}}_{\text{\tiny (B)}}$$

© Use Lagrangian with Cartesian coordinate and with EOM

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}(\mathbf{r}^{N}, \dot{\mathbf{r}}^{N})}{\partial \dot{r}_{i}} \right) = \frac{d}{dt} \left(\frac{\partial}{\partial \dot{r}_{i}} \left(\sum_{i}^{N} \frac{m_{i} \dot{\mathbf{r}_{i}}^{2}}{2} - \mathcal{V}(\mathbf{r}^{N}) \right) \right) = m_{i} \ddot{r}_{i}$$

$$\frac{\partial \mathbf{r}_{i}}{\partial r_{i}} = -\frac{\partial}{\partial r_{i}} \mathcal{V}(\mathbf{r}^{N}) = \mathbf{F}_{i}(\mathbf{r}^{N})$$

Hamiltonian (Valid in any generalized coordinates)

Configuration/conformational space \leftrightarrow coordinates of a system. **Trajectory** $q^{M(t)}$ is the time-dependent evolution of a system in a configuration space, no trajectories can cross \Rightarrow Liouville's theorem (The volume of any closed surface in phase space remains constant as surface moves through phase space) **Phase space** \leftrightarrow coordinates & momenta $x^{2M} = \{q^M, p^M\}$, in Cartesian $x^{2N} = \{r^N, p^N\}$ Orbit $x^{2M}(t)$ **time-dependent** evolution of a system in phase space \bigcirc QM \leftrightarrow CM with Hamiltonian operator

Conjugate momenta
$$p^M = \{p_1, ..., p_M\}$$
 with $p_m := \frac{\partial \mathcal{L}(q^M, \dot{q}^M)}{\partial \dot{q}_m}$
Hamiltonian Function $\mathcal{H} = p \cdot \dot{q} - \mathcal{L}$

Differential of Hamiltonian $d\mathcal{H} = \dot{q}^M \cdot dp^M - \dot{p}^M \cdot dq^M$

$$d\mathcal{H} = d(\boldsymbol{p} \cdot \dot{\boldsymbol{q}}) - \left(\underbrace{rac{\partial \mathcal{L}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M})}{\partial \boldsymbol{q}^{M}}}_{\text{Lagrange } \dot{p}} \cdot d\boldsymbol{q}^{M} + \underbrace{rac{\partial \mathcal{L}(\boldsymbol{q}^{M}, \dot{\boldsymbol{q}}^{M})}{\partial \dot{\boldsymbol{q}}^{M}}}_{\text{Definition of } p} \cdot d\dot{\boldsymbol{q}}^{M}
ight)$$

Energy (Hamiltonian CF) optional time-dependence for \mathcal{V}

$$\mathcal{H}(\boldsymbol{q}^{M},\boldsymbol{p}^{M}) = \underbrace{\mathcal{K}(\boldsymbol{q}^{M},\boldsymbol{p}^{M})}_{\text{kinetic energy}} + \underbrace{\mathcal{V}(\boldsymbol{q}^{M})}_{\text{potential energy}} = E = \text{const}$$

$$\begin{cases} \dot{\boldsymbol{q}}_m \coloneqq \frac{\partial \mathcal{H}(\boldsymbol{q}^M, \boldsymbol{p}^M)}{\partial p_m} \\ \dot{\boldsymbol{p}}_m \coloneqq -\frac{\partial \mathcal{H}(\boldsymbol{q}^M, \boldsymbol{p}^M)}{\partial q_m} \end{cases}$$

$$\begin{split} \underline{\text{Connection to Cartesian}} \mathcal{H}(\mathbf{r}^{N}, \mathbf{p}^{N}, t) &= \underbrace{\sum_{i=2m_{i}}^{N} \frac{p_{i}^{2}}{2m_{i}} + \mathcal{V}(\mathbf{r}^{N}, t)}_{E(t)}}_{E(t)} \\ & \frac{\partial \mathcal{H}(\mathbf{r}^{N}, \mathbf{p}^{N})}{\partial p_{i}} = \underbrace{\frac{p_{i}}{m_{i}} \stackrel{!}{=} \dot{r}_{m}}_{\text{Cartesian momentum}} \\ & \frac{\partial \mathcal{H}(\mathbf{r}^{N}, \mathbf{p}^{N})}{\partial r_{i}} = \frac{\partial \mathcal{V}(\mathbf{r}^{N})}{\partial r_{i}} = \underbrace{\frac{-F_{i} \stackrel{!}{=} -\dot{p}_{i}}_{\text{Newtonian equation}}}_{\text{of motion}} \end{split}$$

EOM in Phase Space with IC $x^{2M}(t_o)$, J Hamiltonian matrix

$$\dot{x}^{2M}(x^{2M}) = \underline{J} \frac{d\mathcal{H}(x^{2M})}{dx^{2M}}, \underline{J} = \begin{pmatrix} 0 & \underline{I}^M \\ -\underline{I}^M & 0 \end{pmatrix}$$

Of a N-particle system

$$\widehat{\mathcal{H}}(m{r}^N,m{p}^N,t) = -\sum_i^N rac{\hbar^2}{2m_i} \underbrace{rac{\partial^2}{\partial r_i^2}}_{ ext{Laplacian } \Delta_i} + \mathcal{V}(m{r}^N,t)$$

Constraints (implicit force law on trajectory)

Active force with explicitly known force law, free, Reactive force

Approach (1) Generalized Coordinates

Replacement for the EOM of the constrained M-th coordinate, with K =M-1 free coordinates.

$$q_M = Q$$

 $\dot{q}_M = 0$

We formulate Lagrangian for the K free coordinates.

$$\mathcal{L}(\boldsymbol{q}^{K}, \dot{\boldsymbol{q}}^{K}; Q) = \mathcal{K}(\boldsymbol{q}^{K}, \dot{\boldsymbol{q}}^{K}; Q) - \mathcal{V}(\boldsymbol{q}^{K}; Q)$$

Formulate K Lagrangian EOM for the free coordinates:

$$\frac{d}{dt} \left(\frac{\partial \mathcal{L}(\boldsymbol{q}^{K}, \dot{\boldsymbol{q}}^{K}; Q)}{\partial \dot{q}_{m}} \right) = \frac{\partial \mathcal{L}(\boldsymbol{q}^{K}, \dot{\boldsymbol{q}}^{K}; Q)}{\partial q_{m}}, \text{with } m = 0, 1, ..., K$$

② Lagrange Multipliers added constraint $C(\mathbf{r}^N) \equiv 0$ explicitly to Cartesian Lagrangian

$$\widehat{\tilde{\mathcal{L}}}(\boldsymbol{r}^{N}, \hat{\boldsymbol{r}}^{N}, \lambda) = \mathcal{L}(\boldsymbol{r}^{N}, \hat{\boldsymbol{r}}^{N}) + \underbrace{\lambda \cdot C(\boldsymbol{r}^{N})}_{\text{does not change Lagrangia}} \underbrace{\lambda \cdot C(\boldsymbol{r}^{N})}_{\text{does not change Lagrangia}}$$

the reactive force. Determine λ with $\frac{d^2}{dt^2}C(\mathbf{r}^N) \stackrel{!}{=} 0$

3 Gauss Principle of Least Constraint (more general, equivalent with the virtual-work principle of D'Alembert)

Quantum Mechanics

As waves (interference patterns in two-slit) • $p = \hbar k$ wave vector k of a particle $k = \frac{2\pi}{3}$ (1D)

As particles (blackbody radia. photo-electric effect)

- $E = \hbar \omega$ energy packed in a photon \sim angular frequency ω .
- $E = h\nu$ with frequency $\nu = \frac{\omega}{2\pi}$

Copenhagen Interpretation System exists in all possible stat simultaneously (superposition). Measurement causes it to random collapse into one state (wave function collapse).

Correspondence principle CM as limit case of QM $(h \rightarrow 0)$. **Observ-** C_{V} = ables mappable to QM Hermitian operators.

$$0 \mathbf{r}_{i} \xrightarrow{q_{i} a_{i}} \mathbf{r}_{i} \otimes \mathbf{p}_{i} \rightarrow -i\hbar \frac{\partial}{\partial \mathbf{r}_{i}} \otimes \mathbf{p}_{i}^{2} \rightarrow -\hbar^{2} \frac{\partial^{2}}{\partial \mathbf{r}^{2}} \otimes E \rightarrow i\hbar \frac{\partial}{\partial t}$$

Particles as waves	Wave as particles	
$oldsymbol{p}=\hbaroldsymbol{k}$	$E = \hbar \omega$	
take position-gradient	take time derivative	
$\rightarrow -i\hbar\frac{\partial}{\partial \boldsymbol{r}}e^{i(\boldsymbol{k}\cdot\boldsymbol{r}-\omega t)}$	$\rightarrow i\hbar\frac{\partial}{\partial t}e^{i({\bf k}\cdot{\bf r}-\omega t)}$	
$= -i\hbar(i{\pmb k})e^{i({\pmb k}\cdot{\pmb r}-\omega t)}$	$= -i\hbar(-i\omega)e^{i({\bf k}\cdot{\bf r}-\omega t)}$	
$= p e^{i(\mathbf{k}\cdot\mathbf{r}-\omega t)}$	$= E e^{i({\pmb k}\cdot {\pmb r} - \omega t)}$	

Operator acts on a wave function and returns another wave function. Hermitian operator has real expectation values.

$$\begin{split} &\int dr \Psi_1^* \hat{\mathcal{A}} \Psi_2 = \left(\int dr \Psi_2^* \hat{\mathcal{A}} \Psi_1\right)^* \\ &\mathbb{E}[\mathcal{A}] \stackrel{\mathrm{at}}{=} \langle \mathcal{A} \rangle \coloneqq \int dr^N \Psi^*(r^N,t) \hat{\mathcal{A}}(t) \Psi(r^N,t) \end{split}$$

TDSE associated with CM through correspondence principle

$$\widehat{\mathcal{H}}(t)\Psi\big(\boldsymbol{r}^{N},t\big)=i\hbar\frac{\partial\Psi\big(\boldsymbol{r}^{N},t\big)}{\partial t}$$

enforces the constraint. Holonomic constraint $C(q^M, t) \equiv 0$ only coor-dinates and time, not on velocities Static holonomic $C(q^M) \equiv 0$ pulsation by rotation in \mathbb{C} -plane. $\frac{dT(t)}{dt} = \frac{E}{i\hbar}T(t) = -i\omega T(t)$

$$\widehat{\mathcal{H}}\Psi_k(\pmb{r}^N)=E\Psi_k(\pmb{r}^N)$$
 with $\int d\pmb{r}^N~|\Psi(\pmb{r}^N)|^2\stackrel{!}{=}1$

Derivation $\widehat{\mathcal{H}}\Psi(\mathbf{r}^N)T(t) = i\hbar\Psi(\mathbf{r}^N)\frac{dT(t)}{dt} = i\hbar\Psi(\mathbf{r}^N)\frac{E}{i\hbar}T(t)$ **TISE solution** for a given $\mathcal{V}(\mathbf{r}^N)$. Defines a **real othonormal basis set**.

$$\Psi(\boldsymbol{r}^N,t) = \sum_k c_k \Psi_k(\boldsymbol{r}^N) \cdot e^{-i\omega_k t}$$

• value pairs $\{E_k, \Psi_k\}$ (EV, associated eigenfunction)

• complex coefficients c_k , $\omega_k = \frac{E_k}{k}$

QMChem atom/molec as many-particle system (nuclei+ e^{-})

Difficulties High cost, limited configurations. Isolated molecules in vacuum only or simple solvation. Approximations No t-dependence, QM. Born-Oppenheimer (motionless nuclei). Neglect electron correlation (partially corrected). Basis set projection to enable for numerical linear algebra sol.

Thermodynamics Basics

with $\lambda = \lambda(t)$ as a function of time which modulates the magnitude of **Extensive** additive, e.g. volume, mass, no. molecules of a species i, U, F, S, heat capacity

> **Intensive** not additive, local only! e.g. P, T, ρ , chemical potential, molar(\div no.)/specific(\div mass) heat capacity, $\kappa_T > 0$, α_P (usually > 0)

Work $W = W_n + W_n$ volume & non-volume (e.g. electrical) • $\delta W_v := -P_{\text{ext}} \cdot dV$ in closed system (deri. from $P_{\text{ext}} := \frac{F_{\text{ext}}}{\Sigma}$)

Chemical Potential Known levels $\{\varepsilon_i\}$ & degeneracies $\{g_i\} \Rightarrow \mu$ determined by the condition $N = \sum n_i$

Thermodynamic Properties (with FD) isothermal compressibility, isobaric expansivity, isochoric heat capacity $\kappa_T = -\frac{1}{V} \left(\frac{\partial V}{\partial P} \right)_{N,T} = \frac{1}{\rho} \left(\frac{\partial \rho}{\partial P} \right)_T = \left(\frac{\partial \ln \rho}{\partial P} \right)_T \stackrel{\text{FD}}{\Rightarrow} \left(\frac{\ln \left(\frac{\rho_2}{\rho_1} \right)}{P_2 - P_1} \right)_T$

$$= \frac{1}{N} \left(\frac{\partial U}{\partial T} \right)_{V} \stackrel{\text{def}}{\to} c_{V} \approx \frac{E_{\text{tot},2} - E_{\text{tot},1}}{T_{2} - T_{1}} + c_{V, \text{ QM-corr}}$$
$$\frac{\partial N}{\partial \mu} \Big)_{V,T} = \frac{N \kappa_{T}}{\nu}$$

Enthalpy H = U + PV, dH = dU + VdP + PdV

Isobaric ($\delta W_{-} = 0, dP = 0$) • $dH = dU + VdP + PdV = (\delta W_v + \delta Q) + PdV = \delta Q$

Heat Capacity δQ needed to raise system temperature by 1 degree. For all system $C_n > C_V > 0$, ideal gas $C_n - C_V > nR$

Туре	Heat capacity	Relation (for closed, uncoupled sys- tem)
Isochoric	$C_V\coloneqq \left(\frac{\partial U}{\partial T}\right)_{n,V}$	$dU = \delta Q = C_V dT$

Isobaric
$$C_P \coloneqq \left(\frac{\partial H}{\partial T}\right)_{n,P}$$
 $dH = \delta Q = C_P dT$

Ideal Gas Assumption

ratio (i.e. low pressure, high temperature) Ideal-gas equation of state barrier heights. PV = nRT

- Ideal Monoatomic Gas $E_{\rm tot} = E_{\rm kin} + E_{\rm pot}$, only kinetic

· Real gas (van der Waals) with corrections for interactions

$$P = \frac{nRT}{\underbrace{V - nb}_{\text{effective volume}}} - a\left(\frac{n}{V}\right)^2$$

without ideal gas: particles interactions via Oapproximate analytical theories (perturb) @MD (sample finite Boltzmann-weighted ensemble numerically) **Particle Insertion** estimate μ_{excess} \Im fails for dense, large particles \leftrightarrow

First Law of Thermodynamics

$$\begin{array}{c} \underbrace{dU}_{\mathrm{independent of path}} = \underbrace{\delta W + \delta Q}_{\mathrm{dependent on path}}\\ \mathcal{V}(r^N) + \mathcal{K}(\dot{r}^N) = E = \mathrm{const}\\ \mathrm{horic} \left(\delta W_n = 0, \, dV = 0\right) \Rightarrow dU = \underbrace{\delta W}_{= 0} + \delta Q = \delta Q \end{array}$$

Second Law of Thermodynamics

Isoc

Inter

$$\underbrace{dS}_{\text{independent of path}} = \frac{\delta Q}{T} + \delta \Sigma$$
nal entropy production $\delta \Sigma = \begin{cases} >0 \text{ spontanenous + irreversible} \\ =0 \text{ reversible} \\ <0 \text{ unnatural} \end{cases}$

Reversible Process $\delta Q = T dS$

Fundamental Equation of Thermodynamics

Integrated form (path-independent)

 $U = -PV + TS + \sum_{i=1}^{N} \mu_{i}n_{i}$

Spontaneity Condition of Processes

$$\delta W + \delta D \geq -PdV + \sum_{i=1}^N \mu_i dn_i$$

Free Energy

Helmholtz F = U - TS• $dF = -PdV - SdT + \sum_{i=1}^{N} \mu_i dn_i$

Gibbs
$$G = F + PV = U + PV - TS = H - FS$$

• $dG = VdP - SdT + \sum_{i}^{N} \mu_{i}dn_{i}$
Derivation $dF = dU - SdT - TdS \stackrel{\odot}{=} \dots$

Derivation $dG = dU + VdP + PdV - SdT - TdS \stackrel{\odot}{=}$

Thermodynamic (Physical & Real, DOF INV)

 $\Delta F \Rightarrow$ change of thermodynamic BC, ΔT , ΔV , Δn

diff T/V, computer $\overline{E}/\overline{P}$ @integrate numerically over T/V

hardly sampled. (can slowly grow particles to cope, TI)

then use $\mathcal{V}(\mathbf{r}_{\text{tost}}) := \mathcal{V}(\mathbf{r}_1, ..., \mathbf{r}_{N+1}) - \mathcal{V}(\mathbf{r}_1, ..., \mathbf{r}_N)$

Conformational (Physical & Virtual, DOF -1)

bation formula above

ing

Calculation methods free energy related to Z, can be formally written Negligible interactions (no medium-range attractions or short-range based on ensemble average. Problem Convergence of ensemble aver-

Temperature/Pressure Integration more robust @NVT simulations at

 $\frac{\partial \left(\frac{T}{T}\right)}{\partial \frac{1}{\pi}} = \langle H \rangle_T \stackrel{\int \cdot d\frac{1}{T}}{\Rightarrow} \frac{F(T_B)}{T_P} - \frac{F(T_A)}{T_A} = \int_{-\infty^{-1}}^{T_B^{-1}} \langle H \rangle_T d(T^{-1})$

no spontaneously formed cavities large enough to give low enough

energy of the inserted particle. Low energy (=good) configuration are

 $\mu_{\text{excess}} = -k_B T \ln \left\langle \exp\left(-\frac{\mathcal{V}(\boldsymbol{r}_{\text{test}})}{k_-T}\right) \right\rangle_{\boldsymbol{r}_1,\dots,\boldsymbol{r}_N}$

For a series of equilibrium configurations of N particles for step i < i

M randomly add 1 test particle r_{test} in step i (state A: N particles; state

 $\Delta F = -k_B T \ln \left[\int \langle \exp\left(-\frac{\mathcal{V}(\boldsymbol{r}_{\text{test}})}{k_B T}\right) \rangle_{\boldsymbol{r}_1,\dots,\boldsymbol{r}_N} d\boldsymbol{r}_{N+1} \right]$

Counting configurations Usable for conformational, alchemical and

appear. Almost never converges within finite trajectories. Insufficient sampling of high energy conformations, insufficient transitions when barrier between states is high. (in eq. f binary assignment function)

 $\Delta F_{BA} = F_B - F_A = -k_B T \ln \left[\frac{\langle f_B(\mathbf{r}) \rangle}{\langle f_A(\mathbf{r}) \rangle} \right] = -k_B T \ln \left[\frac{N_B}{N_A} \right]$

and sufficient transitions occur Phard to design @modify Hamiltonian

to account for bias (2) roll back to unbiased version before direct count-

 $\mathcal{H}_{\text{bias}}(\boldsymbol{r}, \boldsymbol{p}) \coloneqq \mathcal{H}(\boldsymbol{r}, \boldsymbol{p}) + \mathcal{V}_{\text{bias}}(\boldsymbol{r})$

 $w_i = \frac{1}{N} \stackrel{\text{with bias}}{\to} \sum_{j=1}^N w_j = 1, w_i = \underbrace{\frac{\exp\left(+\frac{J_{\text{bias}}(r_j)}{k_B T}\right)}{\sum_{j=0}^N \exp\left(+\frac{V_{\text{bias}}(r_j)}{k_B T}\right)}}$

 $\Delta F_{BA} = F_B - F_A = -k_BT \ln \left[\frac{\langle f_B(\mathbf{r}) \cdot \exp\left(+ \frac{\gamma_{\rm bas}(\mathbf{r})}{k_BT} \right) \rangle_{\rm blas}}{\langle f_A(\mathbf{r}) \cdot \exp\left(+ \frac{\gamma_{\rm bas}(\mathbf{r})}{k_BT} \right) \rangle_{\rm blas}} \right]$

Alchemical (Unphysical, simulation-only! DOF +1)

 $\frac{\partial F}{\partial V} = -P \stackrel{\int \cdot dV}{\Rightarrow} F(V_B) - F(V_A) = -\int_{V}^{V_B} P \cdot dV$

repulsions). Point-like particles only interact via elastic collisions, ran- age in MD Challenge How to sample all relevant parts of phase space, Connected to physical ΔG through thermodynamic cycles, two sepadomizing velocities. Valid when molar density $\frac{n}{v}$ is low, meaning low $\frac{p}{\pi}$ irr. of relative free energies. How to ensure sufficient transitions irr. rate TI calculations for a reaction in different phases, then compare the relative hydration energy.

Δ

Hamiltonian Replica Exchange 1 Do a given number of simulation steps @ MC exchange between neighbouring replicas (λ -points) \bigcirc Smoothen λ curve.

ergy differences. \mathcal{P} fails at orthogonal barrier/ λ insufficiently smooth.

Couple endstates $\mathcal{H}(\boldsymbol{r},\boldsymbol{p}; \lambda) := (1-\lambda)\mathcal{H}_A(\boldsymbol{r},\boldsymbol{p}) + \lambda\mathcal{H}_B(\boldsymbol{r},\boldsymbol{p})$. For

 $\Delta G = \int^{1} d\lambda \left\langle \frac{\partial \mathcal{H}}{\partial \lambda} \right\rangle$

NPT ensemble, and $\langle \cdot \rangle_{\lambda}$ ensemble average at a given λ

$$:= \frac{\left(\mathcal{H}(\mathbf{r}_{j} ; \lambda_{i}) - \mathcal{H}(\mathbf{r}_{i} ; \lambda_{j})\right) - \left(\mathcal{H}(\mathbf{r}_{j} ; \lambda_{j}) - \mathcal{H}(\mathbf{r}_{i} ; \lambda_{i})\right)}{k_{B}T}$$

$$(\lambda_{i}, \dots, \lambda_{i}) = \begin{cases} 1 \text{ for } \Delta \leq 0 \end{cases}$$

Free-Energy Perturbation info from adjacent λ -points \mathcal{P} small enough $\Delta\lambda$ to have overlap between adjacent points

$$\Delta F_{\Delta\lambda} = -k_B T \ln \left[\left\langle \exp \left(-\frac{\mathcal{H}(\boldsymbol{r}, \boldsymbol{p}; \ \lambda + \Delta\lambda) - \mathcal{H}(\boldsymbol{r}, \boldsymbol{p}; \ \lambda)}{k_B T} \right) \right\rangle \right]$$

Combinatorics & Statistics

$$\cdot (p+q)^N = \sum_{n=1}^N {N \choose n} p^n q^{N-n}$$

$$\cdot \underline{\ln N!} = \ln \left(\prod_{n=1}^N n\right) \approx \int_1^N dx \ln x \underline{= N \ln N - N} + O(\ln N)$$

B: N+1 particles) @ estimate $\Delta F = F(N+1) - F(N)$ with the **pertur-•** $\delta(x) = 0, \forall x \neq 0, \int^{\infty} dx \delta(x) \stackrel{!}{=} 1$ (inv unit) or defined as $\delta(x)$

• CLT For large N, the mean of N independent r.v. converges to a normal distribution with mean μ and variance $\frac{\sigma^2}{N}$ (micro property has $\mu, \sigma^2 \Rightarrow$ macro observable peaked at μ)

$$P(X) \coloneqq \frac{1}{(2\pi)^{\frac{1}{2}}\tilde{\sigma}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}, \text{ for } x \in \mathbb{R}, \text{with } \tilde{\sigma}^2 = N^{-1}\sigma^2$$

thermodynamic changes 🖒 simple. 🖓 Both states of interest must • De Moivre (Binomial CLT)

$$\binom{N}{k} p^k q^{N-k} \approx \frac{1}{\sqrt{2\pi Npq}} e^{-\frac{(k-Np)^2}{2Npq}}$$

with $p, q > 0$ and $p+q = 1$.

Combinatorical Problems

$$\begin{split} N_{\rm sel} &= \binom{K}{N} = \frac{K!}{N!(K-N)!}, \ {\rm k \ choose \ n} \\ N_{\rm par} &= \binom{K}{n_1, n_2, \dots, n_M} = \frac{K!}{\prod_m^M n_m!} \end{split}$$

$$N_{\rm ass} = M^K$$

tions $N! (N_{col} = N_{par} \text{ at } M = 2, n_1 = N, n_2 = K - N)$ • Partition a set of K elem into M subsets of $n_1, ..., n_M$ elem, M subsets

1

each with n_i! permutations to divide

• 6 diff fruits into M = 3 subsets of 1, 2 and 3, $N_{\text{par}} = \frac{6!}{1! \cdot 2! \cdot 3!}$

 Assign K objects to M sets. e.g. binary bit assignment, each set m_i as a digit/bit, K objects values $\{0, 1\}$

TI useful for alchemical, conformational and thermodynamic free-en- Indistinguishable distribute K objects into M sets.

 $N_{\rm dis} = {\binom{K+M-1}{K}} = {\frac{(K+M-1)!}{K!(M-1)!}}$

$X \sim$	$p(\pmb{x}) = \mathbb{P}[X = \pmb{x}]$	$\mathbb{E}[X]$	$\operatorname{Var}[X]$
$\mathrm{Ber}(p)$	$p^x(1-p)^{1-x}$	p	p(1-p)
$\mathrm{Bin}(n,p)$	$({^n_k})p^k(1-p)^{n-k}$	np	np(1-p)
$\operatorname{Poisson}(\lambda)$	$e^{-\lambda} rac{\lambda^k}{k!}$	λ	λ
$\operatorname{Geom}(p)$	$p(1-p)^{k-1}$	$\frac{1}{p}$	$\frac{1-p}{p^2}$
$\mathcal{U}([a,b])$	$\begin{cases} \frac{1}{b-a} \text{ if } x \in [a,b] \\ 0 \text{else} \end{cases}$	$\frac{a+b}{2}$	$\tfrac{1}{12}(b-a)^2$
$\mathrm{Exp}(\lambda)$	$\begin{cases} \lambda e^{-\lambda x} \text{ if } x \geq 0 \\ 0 \text{else} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
$\mathcal{N}(\mu,\sigma^2)$	$\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$	μ	σ^2

Stochastic Processes

corresponds to a sequence of values $X_N = \{x_i \mid i = 1, 2, ..., N\}$ (or continuous $X_t = \{x_t \mid t \in \mathbb{R}^+\}$, values in sequence occur randomly and distributed deterministically (i.e. well-defined $p(x_{t+u}, x_t)$)

Uniform Process

Pseudo-random numbers (MCG×congruential, Mersenne Twister) generates non-correlated uniform real numbers over [0, 1). Machine precision involves non-uniformity already. No single necessary and sufficient test!

① Uniformity repeats χ² test for N > 10 · #bin = 10K

$$\chi^{2} = \frac{1}{K} \sum_{k}^{K} \frac{[n_{k} - K^{-1}N]^{2}}{K^{-1}N}$$
$$-\chi^{2} < \frac{2}{\sqrt{K}}$$

• ② Correlation for k points $\langle x_i x_{i+k} \rangle_i := \frac{1}{K-k} \sum_{i=1}^{K-k} x_i x_{i+k}$

$$C(k) = \frac{\langle x_i x_{i+k} \rangle_i - \langle x_i \rangle_i \langle x_{i+k} \rangle_i}{\langle x_i x_i \rangle_i - \langle x_i \rangle_i \langle x_i \rangle_i}$$

No correlation
$$\langle x_i x_{i+k} \rangle_i - \langle x_i \rangle_i \langle x_{i+k} \rangle_i \stackrel{!}{=} 0 \rightarrow C(k) = 0$$

Binomial Process

Bernoulli process independent successive trials and thus independent, observing n successes \Leftrightarrow Binomial distribution

Random Walk direction of diff steps are uncorrelated

In 1D Random walk each step of length z is a Bernoulli process $p \leftrightarrow$ a step to right.

$$E[Z] = N \cdot (2p - 1)z, Var[Z] = 4N \cdot p(1 - p)z^2$$

• Select N elem from set of K, $K \cdot ... (K - N + 1)$, divide by permuta-Probability of a given net move Z over N steps

• Net displacement to right
$$M(n) = n - (N - n)$$
 (right - left)

 $p(\lambda_i \Leftrightarrow \lambda_j) = \begin{cases} e^{-\Delta} & \text{for } \Delta > 0 \end{cases}$

$$\Delta F_{\Delta\lambda} = -k_BT \ln \biggl[\langle \exp\biggl(-\frac{\mathcal{H}(\boldsymbol{r},\boldsymbol{p};\ \lambda + \Delta\lambda) - \mathcal{H}(\boldsymbol{r},\boldsymbol{p};\ \lambda)}{k_BT} \biggr) \rangle_\lambda \biggr] \label{eq:eq:elements}$$

$$\begin{split} (p+q)^N &= \sum_{n=1}^N \binom{N}{n} p^n q^{N-n} \\ \underline{\ln N!} &= \ln \Bigl(\prod_{n=1}^N n \Bigr) \approx \int_1^N dx \ln x \underline{=} N \ln N - N + O(\ln N) \end{split}$$

$$\delta_{x,x_0} = 1, ext{if} \ x = x_0 ext{ otherwise 0 (unitless)}$$

$$f(x) = 0, \forall x \neq 0, f_{-\infty} a dx f(x) = 1$$
 (inv diff) of defined as $\delta(x) = \frac{1}{dx} a^{-1} (dx) = 1$ if $x > 0$, otherwise 0. As limit case of normalized inning function, box gets ∞ -thin/narrow at $\Delta x \to 0$. Property $\sum_{\infty}^{\infty} dx f(x) \delta(x-a) = f(a)$

$$\begin{split} P_b(n;p) &= \binom{N}{n} p^n (1-p)^{N-n} \ \text{for} \ n \in [N] \\ \bar{P}(M) &= \gamma_{M+N} \cdot P_b \Bigl(\frac{m+n}{2};p \Bigr) \\ &= \gamma_{M+N} \cdot \binom{N}{\frac{N+M}{2}} p^n (1-p)^{N-\frac{N+M}{2}} \end{split}$$

De Moivre

$$\begin{split} \bar{P}(M) &\coloneqq \gamma_{M+N} \cdot P_b \Big(\frac{m+n}{2} ; p \Big) \\ &\approx \gamma_{M+N} \cdot \frac{1}{\sqrt{\frac{1}{2} \pi N p(1-p)}} e^{-\frac{\left(\frac{M+N}{2N p(1-p)} - n\right)}{2N p(1-p)}} \\ &\approx \gamma_{M+N} \cdot \frac{1}{\sqrt{\frac{1}{2} \pi N p(1-p)}} e^{-\frac{(M+N(1-2p))}{8N p(1-p)}} \end{split}$$

Symmetric 1D (p = 0.5) mean effec displacement 0

$$\begin{array}{l} \bullet ~~ \tilde{P}(M) = \gamma_{M+N} \cdot {N \choose \frac{N+M}{2}} 2^{-N} \\ \bullet ~~ \mathrm{Var}^{\frac{1}{2}}[Z] = N^{\frac{1}{2}}z ~\mathrm{De} ~\mathrm{Moivre} ~~ \tilde{P}(M) \approx \gamma_{M+N} \cdot \frac{1}{\sqrt{\frac{1}{2}\pi N}} e^{-\frac{M^2}{2N}} \end{array}$$

RW 3D Approach (1) generate Δx_i uniformly & uncorrelatedly \mathcal{P} cubetion) corner (walk of variable size & anisotropy) @ Generate 3D coord for r directly-rescale or discard for fixed steps (inefficient) 3 Use normally distributed Δx_i , rescale to Δr , no rejection needed

• In Polar, $(\Delta x, \Delta y, \Delta z)^T = \Delta r(\cos\theta\cos\varphi, \cos\theta\sin\varphi, \sin\theta)^T$, no rejec-Normalized probability of ption and $\varphi \sim \mathcal{U}([0, 2\pi)), u \sim \mathcal{U}([-1, 1)), \theta = \arccos(u)$ • with weight $W_{m}(Kn) = M^{K}P_{m}(n)$

Gaussian Process (Continuous)

 $\textbf{Cont-Time White-Noise (CTWN)} \ x_t \sim \mathcal{N}(0,\sigma^2), \ \forall t \ \text{needs to be } \oplus \mu = \ \textbf{Peaking of the Distribution Probability } P_K(\textbf{p}) \ \text{of distri$ 0 @ const std dev @ x_t and x_{t+u} uncorrelated $\forall u$.

Wiener Process (Brownian) N i.i.d r.v. $\xi_k \sim \mathcal{N}(0,1) \text{ for } t \in$ $[0, 1], W_t^{(N)} := \frac{1}{\sqrt{N}} \sum_{k=1}^{\lfloor N_t \rfloor} \xi_k, \quad W_t = \lim_{N \to \infty} W_t^{(N)}$ • ①unit of W_t is $\frac{1}{\sqrt{t}}$ ②cont ③not diff-able in t

- (increare Gaussian, $\forall t > 0, u \ge 0, (W_{t+u} W_t) \sim \mathcal{N}(0, u)$, and independent, $\forall s < t, W_{t+u} - W_t$ independent of W_s
- (scale inv, scaled $\frac{1}{2}W_{\alpha^{2}t}$ still Wiener with $\alpha \neq 0$

Statistical Physics

Microstate choice made for all variables (huge amount). Intrinsically equally probable microstates, quantum states $m \in \{0, 1, ..., M-1\}$, Calculation with $\xi^{-1} = h^{3N} N!$ units: $[\xi] = \frac{1}{\operatorname{settion}^{3N}} = [\rho(x^{2N})]$ phase-space volume elements $x^{2N} dx^{2N}$. Macrostate choice made for independent variables eg. $Z = \{n, V, T\}$, projection of huge no. microstates. Compatibility by impose @system constraints: affects

system individually @ensemble constraints: collectively

Assignment vector $m := \{m_1, m_2, ..., m_k\}$

Population vector defines one possible way to distribute n: $\{n_1, n_2, ..., n_M\}$ with $\sum_{m=1}^M n_m \stackrel{!}{=} K$ no. objects assigned to set m is

Map *m* **to** *n* ($N_{\text{dis}} \leq N_{\text{ass}}$ and $N_{\text{dis}} \ll N_{\text{ass}}$ for large *K*, *M*)

$${m}({m}) = \sum_{k=1}^{K} \underbrace{\delta_{m_k,m}}_{\substack{\text{is object k assign} \\ \text{to subset m?}}}$$

Statistical weight W_k of population vector n when distributing K objects into M bins \leftrightarrow no. assignments compatible with the corresponding populations. For *n*, N_{dis} choices and $\sum_{n} W_{K}(n) = N_{\text{ass}}$

$$W_k(n) = \frac{K!}{\prod_{m=1}^M n_m!} \equiv N_{par}$$

• Brute-force $O(M^K)$, $W_k(n) := \sum_{m}^{N_{ass}=M^K} \delta_{n(m),n}$

Example: Given a system with three states M = 3 and an ensemble with six systems K = 6. Gibbs Entropy

How many possible distinct assignment vectors are there?

$$N_{\rm ass} = M^K = 3^6 = 729$$

 What is the number of possible distinct distributions (i.e. number of possible population vectors)?

$$N_{\rm dis} = \frac{(K+M-1)!}{K!(M-1)!} = \frac{(6+3-1)!}{6!(3-1)!} = \frac{8!}{6!2!} = 28$$

 What is the statistical weight of a population that consists of n = $\{3, 2, 1\}$?

$$W_{K=6}(\boldsymbol{n}=\{3,2,1\})=\frac{K!}{\prod_{m=0}^{M-1}n_m!}=\frac{6!}{3!\cdot 2!\cdot 1!}=60$$

Postulate of a priori Equiprobability All microstates equally probable, except when limited by macroscopic constraints \Rightarrow @Without constraints, all $m_{
m assign}$ equally probable @probability of $n_{
m pop}$ in a collection of random ensembles ~ statistical weight of it

Distribution Vector $p = \{p_0, ..., p_{M-1}\} N_{dis}$ choices. (Fractional-Popula-

$$p_m \stackrel{\text{def}}{\coloneqq} K^{-1} n_m \text{ with } \sum_{m=0}^{M-1} p_m =$$

$$P_{K}(\boldsymbol{p}) = \frac{K!}{M^{K} \prod^{M-1} (Kp_{m})!} \text{ with } \sum_{\boldsymbol{p}} P_{K}(\boldsymbol{p}) = 1$$

tributions p becomes more peaked as $K \uparrow$, at limit $K \to \infty$ a single most-probable distribution (the average & the only relevant one). • $\rho = \{\rho_0, \rho_1, ..., \rho_{M-1}\}$ with $\sum_{m=0}^{M-1} \rho_m = 1$

Ensemble Average (E at equilibrium)

For peaked distribution $\langle A \rangle = \sum_{m=0}^{M-1} \rho_m \mathcal{A}_m$

$$\underbrace{\langle A \rangle}_{\text{cro. value}} = \lim_{K \to \infty} \sum_{\substack{p \\ \text{over } \\ \text{distributions}}} P_K(p) \sum_{m=0}^{M-1} p_m \underbrace{\mathcal{A}_m}_{\substack{\text{micro. } v \\ \text{in state}}}$$

$$egin{aligned} &\langle \mathcal{A}
angle = \xi Z^{-1} \int dr dp \mathcal{A} e^{-eta \mathcal{H}} \ &Z = \xi \int dr dp \cdot e^{-eta \mathcal{H}} ext{ with } \xi = \left(h^{3N} N!\right)^{-1} \end{aligned}$$

Phase-Space Probability Density

$$ho(x^{2N})$$
 with $\int dx^{2N}
ho(x^{2N}) =$

Ensemble Average

$$\langle A \rangle = \int d x^{2N} \rho(x^{2N}) \mathcal{A}(x^{2N})$$

Entropy

$$S = -k_B \int d{\pmb x}^{2N} \rho({\pmb x}^{2N}) \ln[\xi^{-1} \rho({\pmb x}^{2N})]$$

Boltzmann & Gibbs Entropy

Boltzmann postulate S as logarithmic measure of $W_K(p)$

Boltzmann Entropy

 $S_K(\mathbf{p}) = k_B K^{-1} \ln W_K(K\mathbf{p})$

$$S(\pmb{p}) = -k_B\sum_{\alpha}^{M-1}p_m\ln(p_m)$$

In the limit of an infinite ensemble

In the limit $K \to \infty$, the entropy of a thermodynamic system is uniquely defined by the given macroscopic constraints S := S(N, V, T).

$$\begin{split} S_{\infty}(\pmb{p}) &= \lim_{K \to \infty} S_{K}(\pmb{p}) \stackrel{\text{fourier}}{=} \lim_{\substack{K \to \infty}} k_{B}K^{-1} \ln W_{K}(\underline{K}\underline{p}) \\ &= k_{B} \lim_{K \to \infty} K^{-1} \ln W_{K}(\underline{K}\underline{p}) \\ \stackrel{\text{Gibbs}}{=} -k_{B} \sum_{m \to 0}^{M-1} \rho_{m} \ln(\rho_{m}) \end{split}$$

Statistical Mechanical Ensembles

hypothetical construct consisting of K i.i.d. copies of the system. Ensemble itself represents 1 macrostate, with system copies in different microstates. Time-independent (Liouville)

- ranked by most common left NPT $> \mu VT \stackrel{?}{=} NVT > NVE$
- isochoric dV, adiabatic dQ, isothermal dT, isobaric dP

Ensemble	Experiment	
microcanoni- cal (NV E_{tot})	isolated (in practice not achievable, in MD no per- fect E _{tot} conservation) • plain MD or MC + "ergostat"	
canonical (NVT) closed + thermostat + rigid container • MD + thermostat or MC or SD		
isothermal- isobaric (NPT/ Gibbs)	closed + thermostat + barostat • MD+thermostat+barostat • MC+barostat • SD+barostat	
grand-canoni- cal (μVT)	open+thermostat+fixed volume • MD+thermostat • MC	

others isoenthalpic-isobaric (NPH), grand-mici isoenthalpic-isobaric, generalized

Microcanonical Ensemble (NVE) (closed, dV, dQ = 0)

Const N, V, E_{tot} (system constraints), N, V define energy levels E = $\{E_0, ..., E_{J-1}\}, E_0 = 0$ and degeneracies $g = \{g_0, ..., g_{J-1}\}$. Value of E some E_i , level denoted as j(E)

Discrete $\rho_m = 1/g_{j(E)}$	Continuous (Classical)
$\tilde{\rho}_m = g_{j(E)}^{-1} \delta_{j_m,j(E)}$	$\rho_m = \Omega^{-1}(E) \delta(E-E_m)$
	$\rho({\boldsymbol x}^{2N}) = \xi \Omega^{-1}(E) \delta\big(E - \mathcal{H}\big({\boldsymbol x}^{2N}\big)\big)$
$S = k_B \ln g_{j(E)}$	$S = k_B \ln \Omega(E)$

Boltzmann Entropy derived from Gibbs Entropy Definition

$$\begin{split} S_{K}(\rho_{m}) & \stackrel{\Delta}{=} k_{B} \sum_{m=0}^{M-1} \Bigl(g_{j(E)}^{-1} \delta_{j_{m},j(E)} \Bigr) \ln\Bigl(g_{j(E)} \delta_{j_{m},j(E)} \Bigr) \\ &= k_{B} g_{j(E)}^{-1} \ln\Bigl(g_{j(E)} \Bigr) \underbrace{\sum_{m=0}^{M-1} \Bigl(\delta_{j_{m},j(E)} \Bigr)}_{g_{j(E)}} \end{split}$$

Density of States (QM & Classical) units: $[\Omega] = \frac{1}{e^{nergy}}$

$$\begin{split} \Omega(E) &= \sum_{m=0}^{M-1} \delta(E-E_m) = \sum_{j=0}^{J-1} g_j \delta(E-E_j) \\ \Omega(E) &= \xi \cdot \underbrace{\int dx^{2N} \delta(E-\mathcal{H}(x^{2N}))}_{\text{area of hypersurface } \mathcal{H} = \mathbf{E}}_{\text{in the phase space}} \end{split}$$

Canonical Ensemble (NVT) (closed, dV, dT = 0)

Const N, V (system) \overline{E} (ensemble), E = E(N, V) limits

(BM) Distribution 2 ensemble constraints,
$$\sum_m p_m = 1$$

• $\Delta E \uparrow \Rightarrow$ exponential population decrease • populated energies p(E) is a peaky region

$$\rho_m=Z^{-1}e^{-\beta E_m}$$
 with $Z=\sum_{m=0}^{M-1}e^{-\beta E_m}$

Lemma, it also holds that

$$\sum_{m=0}^{M-1}\rho_m=1 \text{ and } \overline{E}=\sum_{m=0}^{M-1}\rho_m E_m$$

Discrete	Continuous (Classical)
$\rho_m = Z^{-1} e^{-\beta E_m}$	$\rho(\pmb{x}^{2N}) = \xi Z^{-1} e^{-\beta \mathcal{H}\left(\pmb{x}^{2N}\right)}$
with $Z = \sum_{m=0}^{M-1} e^{-\beta E_m}$	with $Z=\xi\int d\textbf{\textit{x}}^{2N}e^{-\beta\mathcal{H}(\textbf{\textit{x}}^{2N})}$
$S = k_B \ln Z + k_B \beta \overline{E}$	$S = -k_B \int dx^{2N} \ln[\xi^{-1}\rho(x^{2N})]$
$\Rightarrow \underbrace{S = k_B \ln Z + \frac{U - U_o}{T}}_{\star}$	

Helmholtz Free Energy
$$F = -k_B T \ln Z + U_o$$

Reformulate
$$(\star)$$
 to $TS = ..., plug in$

• $F \stackrel{\Delta}{=} U - TS \stackrel{\star}{\Leftrightarrow} F = U - (Tk_B \ln Z + (U - U_o))$

Pressure
$$P = k_B T \left(\frac{\partial \ln Z}{\partial V}\right)_{N,T} - \left(\frac{\partial U}{\partial V}\right)_N$$

• $dF = -PdV - SdT$, plug in F reformulate
• $P = -\left(\frac{\partial F}{\partial V}\right)_{T} = -\left(\frac{\partial (-k_B T \ln Z)}{\partial V} + \frac{\partial U}{\partial V}\right)_{T}$

rocanonical, grand- Internal Energy & two thermo laws (
$$U=\overline{E}+U_{e}$$

$$U = -\left(\frac{\partial \ln Z}{\partial \beta}\right)_{N,V} + U_{c}$$

 (∂U_{α})

for a reversible process, dU in a closed system is reversible

$$dU = \begin{cases} dU_o + \sum_{m=0}^{M-1} \rho_m dE_m + \sum_{m=0}^{M-1} E_m d\rho_m \\ \delta W_{\rm rev} + \delta Q_{\rm rev} = -PdV + TdS \end{cases}$$

Volume Work as Change of Levels not Populations

$$\delta W_{\rm rev} = -PdV = dU_o + \sum_{m=0}^{M-1} \rho_m dE_m$$

Heat Work as Change of Populations not Levels

$$\delta Q_{\rm rev} = T dS = \sum_{m=0}^{M-1} E_m d\rho_m$$

Ideal Monoatomic Gas
$$Z = \frac{V^N}{h^{3N}N!} \cdot \left(\frac{2\pi m}{\beta}\right)^{\frac{-2}{2}}$$

Internal energy

$$U = -\frac{\partial \ln^{-\frac{3N}{2}}}{\partial \beta} = \frac{3N}{2}\beta^{-1} = \frac{3N}{2}k_BT$$

© Pressure $P = \beta^{-1}\frac{\partial \ln V^N}{\partial V} = \beta^{-1}NV^{-1} = \frac{Nk_BT}{V}$

③ Helmholtz Free Energy

$$\begin{split} F &= -\beta^{-1} \ln \left(\frac{V^N}{h^{3N}N!} \cdot \left(\frac{2\pi m}{\beta} \right)^{\frac{3N}{2}} \right) \\ &\stackrel{\text{Stirling}}{=} -\beta^{-1} \ln \left(-N \ln N + N + N \ln V + N \ln \left(\frac{2\pi m}{\beta h^2} \right)^{\frac{3N}{2}} \right) \\ &= -N\beta^{-1} \left(1 + \ln \frac{V}{N} \left(\frac{2\pi m}{\beta h^2} \right)^{\frac{3}{2}} \right) \end{split}$$

④ Entropy

$$\begin{split} S &= k_B \Biggl\{ \frac{3N}{2} + N \Biggl(1 + \ln \frac{V}{N} \Biggl(\frac{2\pi m}{\beta h^2} \Biggr)^{\frac{2}{3}} \Biggr) \Biggr\} \\ &= N k_B \Biggl[\frac{5}{2} + \ln \frac{V}{N} \Biggl(\frac{2\pi m}{\beta h^2} \Biggr)^{\frac{3}{2}} \Biggr] \end{split}$$

Isothermal-isobaric (NPT) (closed, dP, dT = 0)

Const N (system) $\overline{V}, \overline{E}$ (ensemble), V-dependent states. With contin- limit Boltzmann) uous volume variable normalized $\int d\mathcal{V} \sum_m \rho_m(\mathcal{V}) = 1$ and $\varsigma := \beta P$ with unit volume⁻¹

$$egin{aligned} & p_m(\mathcal{V}) = Z_{NPT}^{-1} \cdot \varsigma \cdot e^{-\beta Y_m(\mathcal{V})} \ & Z_{NPT} = \varsigma \int d\mathcal{V} \sum_m e^{-\beta Y_m(\mathcal{V})} \end{aligned}$$

NPT & NVT Partition Function

$$\begin{split} Z_{NPT} &= \varsigma \cdot \int d\mathcal{V} \cdot e^{-\beta P\mathcal{V}} \cdot Z_{NV} \\ Z_{NPT} &= \varsigma \cdot \underline{\xi} \int d\mathcal{V} \underline{\int} dx^{2N} \cdot e^{-\beta Y(x^{2N}, \mathcal{V})} \\ &= \varsigma \cdot \int d\mathcal{V} \cdot e^{-\beta P\mathcal{V}} (\underline{\xi} \cdot \underline{\int} dx^{2N} \cdot e^{-\beta \mathcal{H}(x^{2N}, \mathcal{V})}) \\ &= \varsigma \cdot \int d\mathcal{V} \cdot e^{-\beta P\mathcal{V}} \cdot Z_{NVT} \end{split}$$

Enthalpy $Y_m(\mathcal{V}) = E_m(\mathcal{V}) + P\mathcal{V}$

NPT Ensemble Average

$$A = \langle \mathcal{A} \rangle = \int d\mathcal{V} \sum_m \mathcal{A}_m(\mathcal{V}) \cdot \rho_m(\mathcal{V})$$



Interpretation

1. Compressibility

$$\left(\frac{\partial V}{\partial P}\right)_{N,T} = -V\kappa_T < 0$$

2. Molar Volume

$$\left(\frac{\partial N}{\partial \mu}\right)_{V,T} = \frac{N\kappa_T}{\nu} > 0$$

Ergodic Theorem (Trajectory \Leftrightarrow Ensemble Average)

$$\overline{Q} = \langle Q \rangle$$

All accessible states must be visited multiple times within the observation period. Instead of average over system copies, average over

trajectory of a single system. Holds for $K \to \infty$ and $t \to \infty$ (the latter not in MD).

Equipartition Theorem

 $\overline{E}_{\rm avg} = \frac{1}{2} k_b T \text{ per DOF. Holds only at equilibrium \& classical limit, quan-$ **3D Vector Distribution** $\\ \rho(\nu) = \left(\frac{\beta m}{2\tau}\right)^{\frac{3}{2}} \cdot e^{\frac{-\beta m \nu^2}{2}}$ tum effects (low T) reduce this capacity. Norm Distribution $\int_{0}^{\infty} d\nu \rho(\nu) = 1$

$$\langle \mathcal{K} \rangle = \frac{1}{2} \cdot N_{\text{dof}} k_B T$$

Ideal Gas (No interactions ⇒ Energies additive)

Distinguishable (CM) (atoms in molecule/solid)

$$Z = \prod_{n=1}^{N} Z_n$$

Indistinguishable (QM) (liquid/gas, Fermion/Bosons) neglecting QM symmetry issues (to account, Fermi-Dirac/Bose-Einstein statistics, at

Over states it holds



over levels it holds

$$z_{
m dof} = \sum_{j=0}^{J_{
m dof}-1} g_j^{
m dof} e^-$$

Maxwell-Boltz

evels

$$P_n = \frac{e^{-\frac{E_n}{k_B T}}}{\sum_{i \in \{0,1,2,3\}} e^{-\frac{E_i}{k_B T}}}$$

Equalities for c, μ and σ^2

sound

1.

$$\begin{split} 1 &\stackrel{!}{=} \int_{-\infty}^{\infty} dx p(x) \\ \mu &= \langle x \rangle = \int_{-\infty}^{\infty} dx \cdot x p(x) \\ \sigma^2 &= \langle (x - \overline{x})^2 \rangle = \int_{-\infty}^{\infty} dx (x - \mu)^2 p(x) \end{split}$$

MB Distribution of Velocity (momenta), often for monoatomic gases, **Configuration integral** For ideal gas factorizable into different modes, generally applicable in all phases within classical limit. (speed of For liquids and solids, modes not decouple-able from interactions, no sound, helium voice). Let
$$p(v)dv$$
 be the probability that the velocity possible partition \rightarrow MD necessary! (norm) of a gas particle is between v and $v + dv$, such that $\int_0^{\infty} p(v)dv = \frac{E_{we}}{E_{we}}$

$$Z = \underbrace{Z_{\text{tra}} \cdot Z_{\text{rot}} \cdot \overbrace{Z_{\text{vib}}}^{E_{\text{pot}}} \cdot \underbrace{Z_{\text{elec}} \cdot Z_{\text{nuc}}}_{\text{often only 1 the stark}} \circ \underbrace{Z_{\text{elec}} \cdot Z_{\text{nuc}}}_{\text{optimal best for a constraint of the stark}}$$

Linear rotor (approx valid T sufficient high, else exact)

$$Z_{\rm rot}^{\rm exact} = \sum_{l=0}^\infty (2l+1) e^{-\frac{\beta\hbar^2}{2I} l(l+1)} \overset{x:=l(l+1)}{\approx} \int_{x=0}^\infty e^{-\frac{\beta\hbar^2}{2I} x} dx = \frac{2I}{\beta\hbar^2}$$

Harm Oscillator (CM approx not necessarily valid at room T)

$$Z_{\rm vib}^{\rm exact} = \sum_{n=0}^{\infty} e^{-\beta \hbar \omega n} \stackrel{\star}{=} \frac{1}{1-e^{-\beta \hbar \omega n}}$$

Fluctuations of a quantity $\langle Q \rangle = \sum_{i} P_i Q_i$

$$\langle Q \rangle \coloneqq \int dQ \cdot P(Q) Q \overset{\text{ergodic}}{\Leftrightarrow} \langle Q \rangle = \overline{Q} = \frac{1}{t} \int^t dt Q(t)$$

Standard Deviation Over Ensemble

$$\begin{split} \sigma_Q &= \left(\int dQ \cdot P(Q)(Q - \langle Q \rangle)^2\right)^{\frac{1}{2}} \\ \stackrel{\text{ergodic}}{\Leftrightarrow} \sigma_Q &= \left(\frac{1}{t} \cdot \int_0^t dt \cdot \left(Q(t) - \overline{Q}\right)^2\right)^{\frac{1}{2}} \text{ if } t \to \infty \end{split}$$

Single-Sweep

$$\sigma_Q^2 = \langle Q^2 - 2Q \langle Q \rangle + \langle Q \rangle^2 \rangle = \langle Q^2 \rangle - \langle Q \rangle^2$$

Gaussian $\langle Q \rangle = \mu$ and $\langle (Q - \mu)^2 \rangle = \langle Q^2 \rangle - \mu^2 = \sigma^2$

Canonical/Isothermal-Isobaric/Grand-Canonical the relative fluctuations of all dependent extensive quantities scale with the no. of particles N as $\propto N^{-\frac{1}{2}}$. Fluc. of sum of normal dist scales as $\frac{\sigma(\sum_{i=1}^{N}R_{i})}{\sum^{N}R_{i}} \propto$ $N^{-\frac{1}{2}}$

$$\begin{split} & \sigma_E^2 = k_B T^2 C_V = \beta^{-1} C_V T \stackrel{C_V \propto N}{\Rightarrow} \quad \frac{\sigma_E}{\langle E \rangle} \propto N^{-\frac{1}{2}} \\ & \sigma_V^2 = k_B T V \kappa_T = \beta^{-1} V \kappa_T \stackrel{V \propto N}{\Rightarrow} \quad \frac{\sigma_V}{\langle V \rangle} \propto N^{-\frac{1}{2}} \\ & \sigma_N^2 = k_B T N \nu^{-1} \kappa_T = \beta^{-1} \rho \kappa_T N \Rightarrow \frac{\sigma_N}{\langle N \rangle} \propto N^{-\frac{1}{2}} \end{split}$$

In the **thermodynamic limit** $N \to \infty$, all observables become **fluctu**ation-free in terms of relative fluctuations and a one-component onephase system can be equivalently specified by any combination of three variables (water 1 kg, 1 L, 300 K \equiv water 1 kg, 1 bar, 300 K) De

$$\begin{split} \frac{\partial Z}{\partial \beta} &= \sum_{m} -E_{m}e^{-\beta E_{m}} = -Z\langle E\rangle \\ \frac{\partial^{2} Z}{\partial \beta^{2}} &= \sum_{m} E_{m}^{2}e^{-\beta E_{m}} = Z\langle E^{2}\rangle \\ \tau_{E}^{2} &= \langle E^{2}\rangle - \langle E\rangle^{2} = \frac{1}{Z}\frac{\partial^{2} Z}{\partial \beta^{2}} - \left(\frac{1}{Z^{2}}\left(\frac{\partial Z}{\partial \beta}\right)^{2}\right) = -\frac{\partial}{\partial \beta}\langle E\rangle \\ C_{V} &\triangleq \left(\frac{\partial U}{\partial T}\right)_{N,T} = \left(\frac{\partial U}{\partial \beta}\right)_{N,T}\frac{d\beta}{dT} = -\frac{1}{k_{B}T^{2}}\frac{\partial}{\partial \beta}\langle E\rangle \end{split}$$

Covariance

$$C_{ab} = \langle (a - \overline{a}) \cdot \left(\frac{b}{-} \overline{b} \right) \rangle = \langle ab \rangle - \overline{a}\overline{b}$$

Cross-correlation normalized covariance

$$c_{ab} = \frac{C_{ab}}{\sigma_a \sigma_b} \rightarrow \begin{cases} +1 \text{ perfect correlation} \\ 0 \text{ no correlation} \\ -1 \text{ perfect anticorrelation} \end{cases}$$

Autocorrelation Function $C_O(\tau) = \langle Q(t)Q(\tau+t) \rangle$

Quantum Statistics (i.i.d. & indistinguishable)

Fermi-Dirac/Bose-E/M-Boltzmann (low ρ or high T limit) This plot is not of distribution!

$$p(v) = 4\pi \left(\frac{3m}{4\pi K}\right)^{\frac{3}{2}} v^2 \exp\left(-\frac{3mv^2}{4K}\right)$$
Consider NVT at classical level and in Cartesian coordinate
1D Distribution for velocity component $\alpha = x, y, z$

$$\rho(\nu_{\alpha}) = \left(\frac{\beta m}{2\pi}\right)^{\frac{1}{2}} \cdot e^{\frac{-\beta m \nu_{\alpha}^2}{2}}$$





$$\overline{m}{=}0$$
 $_{
m of}=\sum_{j=1}^{J_{
m dof}}g_{j}^{
m dof}e^{-eta E_{j}^{
m dof}}$

Write down the equation given
$$P_n$$
 for Boltzmann Given energy le $E_n=n\Delta E$ with $n=0,1,2,3$

quation given
$$P_n$$
 for Boltzmann Given end
= $0, 1, 2, 3$

Factorization of Partition Function Z

$$\begin{split} \mathbf{Canonical} \ & Z = Z_{\mathcal{K}} \cdot Z_{\mathcal{V}}, \text{ using } \mathcal{H}(x^{2N}) = \mathcal{K}(p^N) + \mathcal{V}(r^N) \\ & Z_{\text{NVT}} = \xi \int dx^{2N} e^{-\beta \mathcal{H}(x^{2N})} \\ & = \underbrace{\left(\xi \int dp^N e^{-\beta \mathcal{K}(p^N)}\right) \left(\int dr^N e^{-\beta \mathcal{V}(r^N)}\right)}_{Z_{\mathcal{K}} \text{ momentum}} \underbrace{\int \mathcal{L}_{\mathcal{V}} \text{ configuration}}_{Z_{\mathcal{V}} \text{ configuration}} \end{split}$$

average $\langle \nu_{\alpha} \rangle = \int_{-\infty}^{\infty} \nu_{\alpha} \cdot \rho(\nu_{\alpha}) d\nu_{\alpha} \stackrel{\text{symmetric}}{=} 0$

 $\rho(\boldsymbol{\nu}) = 4\pi \left(\frac{\beta m}{2\pi}\right)^{\frac{3}{2}} \cdot \boldsymbol{\nu}^2 \cdot e^{\frac{-\beta m \boldsymbol{\nu}^2}{2}}$

average $\langle \nu \rangle = \left(\frac{8}{\pi^{\beta m}}\right)^{\frac{1}{2}}$

RMS $\langle \nu^2 \rangle^{\frac{1}{2}} = \left(\frac{3}{\beta m}\right)^{\frac{1}{2}}$

Araon : 39.9 a/m

therm

most probable $\nu^* = \left(\frac{2}{\beta m}\right)^{\frac{1}{2}}$

MS $\langle \nu_{\alpha}^2 \rangle = \frac{1}{2}$

Momentum integral depends not on intermolecular interactions.
Translational
$$E_{kin}$$
 only. For canonical ensemble

$$Z_{\rm tra} = Z_{\mathcal{K}} = \xi \left(\int_{-\infty}^{\infty} dp \cdot e^{-\beta \frac{p^2}{2m}} \right)^{3N} \stackrel{\xi = \frac{1}{h^{3N_{\rm H}}}}{=} \frac{1}{N!} \left(\frac{2\pi m}{\beta h^2} \right)^{\frac{3N}{2}}$$



Supra-atomic:

utions with effective pair potentials

taining protein 2nd, tertiary structures.

 $\mathcal{V}^{\text{bond}(\boldsymbol{r}^N)} = \sum_{i=1}^{\text{bonds}} \frac{1}{2} K_i^b [b_i(\boldsymbol{r}^N) - b_i^0]^2$

 $\mathcal{V}^{\text{angle}}(\theta) = \sum_{i}^{\text{angles}} \frac{1}{2} K_i^{\theta} [\theta_i - \theta_i^0]^2$

 $\mathcal{V}^{\text{ele}(\mathbf{r}^N)} = \sum_{\text{nairs } i \leq j} \frac{1}{4\pi\varepsilon_0\varepsilon_r} \frac{q_i q_j}{r_{ii}}$

(Wbv

atoms i and j

Interactions in Classical Atomistic Force Fields (FF)

Classical FF $\mathcal{V}^{\text{phys}}(\mathbf{r}^N)$, with $\mathbf{r}^N = (\mathbf{r}_1, \mathbf{r}_2, ..., \mathbf{r}_N)$

 $\mathcal{V}^{\text{torsion}(\boldsymbol{r}^{N})} = \sum_{i}^{\text{torsions}} K_{i}^{\varphi} \left[1 + \cos(m_{i}\varphi_{i}(\boldsymbol{r}^{N}) - \delta_{i})\right]^{2}$

and N-body polar. energy $\mathcal{V}^{\mathrm{pol}}$, external fields energy $\mathcal{V}^{\mathrm{ext}}$

 $\mathcal{V}^{\mathrm{vdW}(\mathbf{r}^N)} = \sum_{\mathrm{pairs \ i} < j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right]$

Iterative boltzmann inversion reproduces structure, not forces

Force matching reproduces forces, not structure. Compute F_{ave} on

· Fitting to Thermodynamic Properties (MARTINI): Start with atom

· often underdetermined, too many parameter/experimental data

lacks $\mathcal{V}(\mathbf{r}^N) = \mathcal{V}^{\text{phys}}(\mathbf{r}^N) + \mathcal{V}^{\text{special}}(\mathbf{r}^N)$. Special terms from NMR

0.6 0.8

Non-bonded elec, vdW as interactions of atoms independent of net

charge. (Repulsion (short-range forces) of nuclei (Coulomb, Pauli's

tuations in electron clouds (London forces), LJ used to approximate

 $V_{SW(r;r_{1})} = C_{12}(r_{0}-r)^{-12} - C_{6}(r_{0}-r)^{-6}$

pole-dipole interactions, termed London dispersion)

4 0.8 1.2 1.6

$$\begin{split} W_{\rm FD(n)} &= \prod_{j} N_{\rm sel} = \prod_{j} \frac{g_{j}!}{n_{j}!(g_{j} - n_{j})!} \\ W_{\rm BE(n)} &= \prod_{j} N_{\rm dis} = \prod_{j} \frac{(n_{j} + g_{j} - 1)!}{n_{j}!(g_{j} - 1)!} \\ W_{\rm MB(n)} &= \prod_{j} \frac{g^{n_{k}}}{n_{j}!}, \text{where } \frac{g_{j}!}{(g_{j} - n_{j})!} \approx \frac{(g_{j} + n_{j} - 1)!}{(g_{j} - 1)!} \approx g^{n_{k}} \end{split}$$

Distribution Functions ① Use Stirling and differentiate

$$\begin{split} &\frac{\partial}{\partial n_i}\ln W_{\text{FD}(n)} = \ln \frac{g_i - n_i}{n_i} = \ln \left(f_{FD,i}^{-1} - 1\right) \\ &\frac{\partial}{\partial n_i}\ln W_{\text{EE}(n)} = \ln \frac{n_i + g_i - 1}{n_i} = \ln \left(f_{BE,i}^{-1} + 1\right) \\ &\frac{\partial}{\partial n_i}\ln W_{\text{MB}(n)} = \ln \frac{g_i}{n_i} = \ln f_{MB,i}^{-1} \end{split}$$

where $f_{FD,i} \stackrel{\Delta}{=} \frac{n_i}{a}, f_{BE,i} \stackrel{\Delta}{=} \frac{n_i}{a-1}$ and $f_{MB,i} \stackrel{\Delta}{=} \frac{n_i}{a}$ (2) Max. W, Lagrange multipliers $(\beta = \frac{1}{k-T}, \alpha = \frac{-\mu}{k-T} = -\beta\mu)$

$$\begin{split} &\frac{\partial}{\partial n_i} \left(\ln W(n) - \alpha \sum_j n_j - \beta \sum_j n_j \varepsilon_j \right) = 0, \forall i \\ \Rightarrow & \begin{cases} f_{\text{FD},i} = \frac{1}{e^{\alpha + \beta \varepsilon_i} + 1} = \frac{1}{e^{\beta (\varepsilon_i - \mu)} + 1} \\ f_{\text{BE},i} = \frac{1}{e^{\alpha + \beta \varepsilon_i} - 1} = \frac{1}{e^{\beta (\varepsilon_i - \mu)} - 1} \text{ singular at } \varepsilon_i = \mu \\ f_{\text{MR},i} = e^{-(\alpha + \beta \varepsilon_i)} = Z^{-1} \cdot e^{-\beta \varepsilon_i} \text{ with } Z = e^{-\beta \mu} \end{cases} \end{split}$$

Pauli Exclusion Principle There cannot be more than one fermion in a given quantum state, unconstrained for bosons. At limit $T \rightarrow 0$

- FD: all levels ε_0 to $\varepsilon_F := \lim_{T \to 0} \mu$ occupied with 1 Fermion each

BE: all particles at lowest level ε₀ = μ

Biomolecules Amino acids, Carbohydrates (Monosaccharides, Disaccharides), Lipids (amphiphilic), Nucleic acids, DNA/RNA Environment sion) & longer-range attraction by electron correlation (instantaneous ions, water (high C_n, dipole moment, H-bonding capacity) Protein have **secondary structure** (α -helix, π -helix, β -sheets) representation 1) Van der Waals spheres (2) Caroon representation for only backbones (with torsional angles) MD Num integration over the classical EOM $(Newton), Cartesian \ coord. \ Simulating \ biological \ system \leftrightarrow solving \ the \ \ Bonded \ bonds, \ angles \ torsions, \ improper \ dihedrals \ bonds \ bo$ many-particle problem Morse function for covalent bond stretching Cbond-break (dissocia-

Degrees of Freedom - Coarse-grain (CG) model

QM description (with electrons), valid when small systems (100 atoms). Negligible QM effects in large systems.

CG models reduce DOFs & interactions while ① align with research needs @ eliminate mostly decoupled DOFs. @ remained DOF easily rep- Harmonic resentable. PLess transferability, altered physics, unphysical entropy/ energy balance. 🖒 speedup

Atoms: United Atoms. @remove non-essential DOFs (e.g. H-atoms in Effective FF (Parameterized) shall be representative with correct lipids) @replace functional groups (CH, CH₂) with larger atoms 🖓 Not physics, <u>simple</u> with few terms, <u>efficient</u> with no complex derivatives and exponential terms and parameters transferable for a range of for strong hydrogen-bonding atoms (O, N, S, P). molecules

Configuration Generation

mapped atoms for each CG bead-least-square match force distrib-High-dimensional space; global minimum impractical. Goal Generate low-energy, Boltzmann-weighted IC/configurations @Search: Find and minimize low-energy regions. @Sample: Explore configurations (meistic guesses, then systematically scan parameters. 🖓 Can lead to tropolis MC, modified MD with biasing). (Simulate: Use dynamics (MD, loss of essential DOFs and may require elastic networks for main- SD, BD).

Key properties of the required method

EOM: dynamic properties can be calculated



data, biased sampling. Physical terms (if two atoms \rightarrow ignore the \sum): • **Boltzmann**-weighted ensemble: thermodynamic properties can be calculated · Physically-based sequence of configurations through classica

> Init Coordinates Specified Endt.Choose carefully (e.g., X-ray/NMF unless longer equilibration than conformational relaxation time. In Velocities Specified E_{kin} , generally unimportant due to fast relaxatio (ps), useful for initial E. Random from $f_{\rm MB}$ at some T. Bond-Lengt Constraints freeze bond vibrations, valid if bonds weakly coupled, in

> ground state. I freezed bonds allow $\Delta t \rightarrow 2$ fs (before 0.5 - 1 fs for fast vibration)

Integrators good accuracy, conservation of properties compatible with geneous distribution at surface Opporty transferable) thermo/barostating (symplectic, phase-space-volume preserving), reversible in time.

Timestep
$$\Delta t \approx \frac{\tau}{10} \approx \frac{\text{period of fastest motion}}{1000}$$
. **Tradeoff** Too short poor s pling, too long poor energy conservation & overflow

Not time reversible **Euler**, error $O(\Delta^2)$, **RK**, error $O(\Delta^4)$.

Time-reversible **leap-frog**, symplectic, error $O(\Delta^3)$ in both coordi- between minimum images. nates and velocities. Interleaving cancels out leading error term $O(\Delta^2)$

$$v\left(t + \frac{1}{2}\Delta t\right) = v\left(t - \frac{1}{2}\Delta t\right) + a(t)\Delta t$$

 $r(t + \Delta t) = r(t) + v\left(t + \frac{1}{2}\Delta t\right)\Delta t$

principle), Attraction (long-range forces) of induced dipoles from fluc-SHAKE in Cartesian Ceasy geometric constraints, existing iterative

methods Plarge movement within one timestep has no convergence guarantee \Rightarrow indicates bad simulation setup (skip energy minimiza-'vdW* strong short-range repulsion upon atom overlap (Pauli exclu- tion), timestep cannot be too long, lightest atom moves first (Newton's 3rd) SHAKE how iterative for multiple bonds iterate with relative tol 10^{-4} . for a single bond @free-fight step (unconstrained) @coordinate resetting(SHAKE)

Spatial BC

Finite-size (FS) due to miscroscopic nature, lack of solute-solvent tion) possible \mathcal{P} comp. expensive b_{ii} : length of the bond between interactions, cannot model intra-solute vdW/electrostatics. Surface effect (SE) system has a large surface-to-volume ratio, increased surface tension proportionally.



		FS	SE	
	Vacuum			♀compactness, spherical shape, strong electrostatics
	Implicit solvent	-	Large	Cheap ♀exact location of solute- solvent boundary, parameter-sensi- tive (charges, radii)
e al	Finite system (droplet/ solvent layer)	Large		♀P too high, solvent evaporation, sur- face-layer artifacts, strong electrosta- tics
n h	Periodic (most- used)	Large	♡-	artificial anisotropy/periodicity, high effective concentration, expensive, FS still present!

Finite-system BC needs confinement potential (LJ wall to confine sample, prevent evaporation and mimic dispersion), orientation correction potential (prevents solvent preferential orientation & inhomo-

Periodic BC explicit-solvent, mimics infinite lattice of periodic copies of the reference box. Particles exit box through one face translated and am- reenter through opposite. Infinite surface → no SE. Long-range interaction evaluated using lattice-sum methods (Fourier), or Minimum-

image pair, atom with closest periodic replica of another one with cut-off distance. Short-range like covalent easy to compute, act only

Box shape rectangular prism used for long (elongated), if rotates need roto-translational constraint since molecule otherwise interacts with its replica out-of-bound. Hexagon used for DNA, cube isotropic thus no roto-fix needed but requires much solvent, octahedron for spherical molecules, almost isotropic, less solvent. Triclinic for crystal, reshapeable

Minimum solute-to-wall distance $\frac{R}{2}$ no solute atom interacts with solute atom in periodic copy, R no solvent molecule interacts with solute atom in two solute periodic copies.

Thermodynamic BC

Plain MD microcanonical, Newtonian EOM conserves E_{tot} and L_{tot} in vacuum, if periodic rotation in each cell coupled with friction, no conservation of L_{tot} . Quantities T, P dependent, calculated as T = $\langle \mathcal{T} \rangle, P = \langle \mathcal{P} \rangle$, have non-zero fluctuations for finite-systems.

Grand-canonical simulations, no. particles vary (uncommon), µ held constant on average $\bigcirc \bigcirc \square \mu$ is not instantaneous observable, discrete variation \rightarrow jumps in dynamics, equilibration needed after each jump

Thermo/Barostating Overview @Constraining Fixes values exactly 🖒 Hamiltonian form, configuration distribution correct Punphysical, no fluctuations @Weak-coupling (Berendsen) c1-st order exponential relaxation (physical) Pdepends on thermostat coupling time (friction coeff $\zeta_T = \frac{1}{2\pi\pi} \frac{\mathcal{K} - K}{\mathcal{K}}$, unit: time⁻¹), only approx canonical ③<u>Extended</u>

$$\mathcal{V}^{\text{Morse}}\big(\mathbf{r}(t); D_n; K_n^b; b_n^0\big) = \sum_n D_n \{1 - e^{\left(\left\lfloor -\left(\frac{K_n^b}{2}\right)^{\frac{1}{2}}(b_n(t) - b_n^0\right)}\right)}$$

monic approximation \circlearrowright simple, cheap \heartsuit no dissociation

$$\mathcal{V}^{\mathrm{harm}}\big(\pmb{r}(t);K_n^b;b_n^0\big) = \sum \frac{1}{2}K_n^b \Big(b_{n(t)} - b_n^0\Big)^2$$

canonical, thermalization efficient, few artifacts Pnon-deterministic, local thermostating (unphysical dynamics)

$$\begin{array}{l} \textbf{Thermostat} \text{ To ensure } E_{\rm kin} = \mathcal{K}\big(t + \frac{\Delta}{2}\big) = K \text{, we scale the velocity } \lambda \\ \frac{\left(\mathcal{K}(t - \frac{\Delta}{2})\right)}{\mathcal{K}(t + \frac{\Delta}{2})} \\ \end{array} \right)^{\frac{1}{2}} \text{, leapfrog } \boldsymbol{v}\big(t + \frac{\Delta}{2}\big) = \boldsymbol{v}\big(t - \frac{\Delta}{2}\big) + \boldsymbol{M}^{-1}\boldsymbol{F}(t)\Delta \end{array}$$

(a)**Collision** within τ_T period, pick & assign $v \sim f_{\rm MB}(T_{\rm ref})$ of an atom randomly $\Rightarrow \langle T \rangle = T_{ref}$

(a)Langevin EOM $m_i \ddot{r}_i = F_i(r(t)) + R_i(t) - m_i \gamma_i \dot{r}_i(t)$ results in $\langle T \rangle =$ $\frac{\langle \mathbf{R}_i^2 \rangle}{6m_i \gamma_i k_B} = T_{re}$

Barostat $\mathcal{P} = P_{\text{instan}} > P_{\text{target}} \Rightarrow \text{box } V \uparrow$, adjust coord Instan temperature (equipart, at equilibrium $\langle \mathcal{T} \rangle = T$)

> $\mathcal{T} = \frac{2}{\mathcal{N}_{p}k_{p}} \cdot \mathcal{K}$ $\langle \mathcal{K}_{\alpha} \rangle = \left\langle \frac{1}{2} m_{\alpha} \nu_{\alpha}^2 \right\rangle = \frac{k_B T}{2} \stackrel{\text{set of dofs}}{\Rightarrow} \left\langle \mathcal{K} \right\rangle = \frac{\frac{3N}{\mathcal{N}_D} k_B T}{2}$

Instan pressure (virial)

$$\underline{\mathcal{P} = \frac{2(\mathcal{K} - \mathcal{W})}{3\mathcal{V}}} = \frac{Nk_B\mathcal{T}}{\mathcal{V}} - \frac{2\mathcal{W}}{3\mathcal{V}}$$
ideal gas

$$P = \left\langle \frac{2(\mathcal{K} - \mathcal{W})}{3\mathcal{V}} \right\rangle = \underbrace{\overbrace{Nk_B \mathcal{T}}^{\text{marging marging margi$$

instan virial (isotropic) interm forces, correction term

$$\mathcal{W} = -\frac{1}{2}\sum_{i=1}^N \boldsymbol{r}_i\cdot\boldsymbol{F}_i > 0 \Rightarrow P \downarrow (\text{attractive, pull inwards})$$

long tsim for good averaging, P highly fluctuating (1 bar fluc in order 100 bar)

Thermostat	Barostat	
$\dot{N} = 0$	$\dot{N} = 0$	
$\dot{r} = \nabla_p \mathcal{K}$	$\dot{r} = \nabla_{p} \mathcal{K} + \underline{\chi r}$	
$\dot{\boldsymbol{p}} = -\nabla_{\boldsymbol{r}} \mathcal{U} - \zeta_T \boldsymbol{p}$	$\dot{\boldsymbol{p}} = -\nabla_{\boldsymbol{r}} \mathcal{U} - \underline{\chi} \boldsymbol{p}$	
$\dot{\mathcal{V}} = 0$	$\dot{\mathcal{V}} = \underline{3\mathcal{V}\chi}$	

Equipartition Violation (heterogeneous T)

- 1. Hot/Cold-solvent (hetero. spatial regions) solvent subjected to more heating than solute and $E_{\rm kin}$ exchanged slowly, correct \overline{T} but solute simulated to be colder than solvent \rightarrow couple solute and solvent dof to separate thermostats
- modes coupled to the same thermostat should have diff. temp in an ideal gas \rightarrow couple diff modes to separate thermostats or use stochastic approach

Uncoupled DOF Issues: Exclude linear/angular momentum from temperature/pressure calculations to avoid errors (e.g., flving ice-cube effect). Linear Momentum has no effect on system property (Move a @Calculate properties glass of water \Rightarrow T \uparrow), Angular Momentum (centrifugal forces).

Calculation of Properties

MC Sampling

system (Nosé-Hoover) $\zeta_T = \frac{1}{\tau_n^2} (\frac{K}{K} - 1)$ own EOM, K > K, friction \uparrow , \mathring{O} simple, cheap, no derivatives of \mathcal{V} for force needed, applicable No guarantee unless good initial configurations **Best practice:** use brakes particles $\sim \Delta \mathcal{K}_{rel}$ canonical poscil- to discontinuous $\mathcal{V}_{refficient}$ unphysical (but reversible) moves can multiple simulations (too costly), more t_{sim} , sampling enhancement techniques latory(unphysical) @Stochastic 01-st order exponential relaxation, be designed for improved sampling Pnon-deterministic, no dynamic

information, hard to design params (acceptance ratio etc.) Stochastic Dynamics mimics following effects without explicit solvent _ molecules, samples NVT.

① Mean Solvent Effect thermodynamics, dispersive attraction re duced in non-polar solvent, enhanced in polar solvents (hydrophobic) (2) Stochastic Collisions dynamics, random but constrained forces on

solute atoms from solvent.

3 Frictional Drag dynamics, solvent-induced friction reduces solute atom velocities.

$$\textit{Langevin} \; m_i \ddot{r}_i = \underbrace{F_i^\sim}_{\text{freen}} - \underbrace{m_i \gamma_i \dot{r}_i}_{\text{frictional}} + \underbrace{\sigma_i \eta_i}_{\text{stochstic}}$$

• $\eta_i = \eta_i(t) \sim \mathcal{N}$ white-noise in unit of time⁻¹/₂, stochastic force ampli tude σ_i (unit: force \times time¹/₂), γ (unit time⁻¹)

$$\begin{split} m \frac{dv}{dt} &= F - m\gamma v + \sigma\eta \\ v(t) &= e^{-\gamma t} \left[v(0) + m^{-1} \int_0^t dt' \cdot e^{\gamma t'} \left(F(t') + \sigma\eta(t') \right) \right] \end{split}$$

• Zero-friction limit ($\gamma = 0, \sigma = 0$) \leftrightarrow Newton equation of MD High-friction limit ↔ Brownian Dynamics

Brownian Dynamics (BD) valid microscopically with low Re viscosity dominates turbulent flow, derivable from Langevin

$$\begin{split} \underbrace{p_i \dot{r}_i}_{i} &= F_i^\sim - m_i \gamma_i \dot{r}_i + \sigma_i \eta_i \\ \Leftrightarrow m_i \gamma_i \dot{r}_i &= F_i^\sim + \sigma_i \eta_i \end{split}$$

we neglect inertial term by setting the RHS to zero at the limit where $|m_i \ddot{r}_i| \ll |m_i \gamma_i \dot{r}_i|$

Fokker-Plank (FP) with SDE $\dot{x}(t) = A(x,t) + B(x,t)\eta(t)$

$$\dot{p}(x,t) = \left(-\frac{\partial}{\partial x}A(x,t) \times + \frac{1}{2}\frac{\partial^2}{\partial x^2}B^2(x,t) \times\right)p(x,t)$$

FP Equation for BD with SDE $m\gamma \dot{x} = F + \sigma \eta$

$$\dot{p}(x,t) = \left(-\frac{1}{m\gamma} \frac{\partial}{\partial x} \underbrace{ F(x,t)}_{\text{drift}} \times + D \frac{\partial^2}{\partial x^2} }_{\text{diffsion}} \times \right) p(x,t)$$

$$\gamma, \sigma$$
 and T Connection $\sigma^2 = 2\beta^{-1}m\gamma$, $D \coloneqq \frac{1}{2}\left(\frac{\sigma}{m\gamma}\right)$

Calculation of Properties (4-step process)

OPreprocess raw trajectories (periodic gather macromolecule locates) 2. Vibrationally-cold gas (hetero. dof) translation, rotation, vibration at box edges. Gather by pick atom within reference box and follow covalent bonds Q with multiple solutes & roto-translational fitting Q no strict decoupling, x_{rof} -dependent)

> @Time series analysis Statistical Moments (1st-4th) with decreasing accuracy μ , σ^2 , skewness, kurtosis

Fick's Law
$$\boldsymbol{j}(\boldsymbol{r},t) = -D \cdot
abla c(\boldsymbol{r},t)$$

Diffusion Equation $\frac{\partial}{\partial t}c(\mathbf{r},t) = D\nabla^2 c(\mathbf{r},t)$ c concentration of solute. D diffusion constant

Equilibration: Discard initial simulation period to remove non-repre sentative conditions, monitored by stabilization of observables.

Estimation of Statistical Errors Error on μ of N normal distribution, confidence factor $c \ (c = 1 \rightarrow 68\%, 2 \rightarrow 95\%, 3 \rightarrow 99.7\%) \ \varepsilon = c \frac{\sigma}{m^4}$ Effectively the second tive no. samples $\varepsilon = c \frac{\sigma_Q}{N^2}$

Parameter Determinatior

Theoretical OM calculations (reference bond length b_0 , k_1 in $\mathcal{V}^{\text{bond}}$)

Experimental crystallographic structure determination (reference bond length b_0 , infrared/Raman spectroscopy measurements (k_b in 12bond)

Units

•
$$k_B = \frac{R}{N_A} = 1.381 \times 10^{-23} J \cdot K^{-1}$$

• $\beta = \frac{1}{k_B T}, n = \frac{N}{N_A}, N_A \approx 6.02214076 \times 10^{23} \text{mol}^{-1}$
• $h = 6.6 \cdot 10^{-34} J \cdot s \text{ or } h = \frac{h}{2\pi}$
 $E = \frac{h^2 \pi^2}{2\pi} n^2 \text{ for } n = 1.2$

Quantity	Units	
acceleration a	m/s ²	
pressure P	$Pa=N/m^2=J/m^3atm=\!1.013\;bar$	
Moment/work/energy J = Ws = Nm		
Power	W = Nm/s	
Gas constant R	kJ/(mol·K)	
Potential energy \mathcal{V}	kJ/mol	

• $[kJ \cdot mol^{-1} \cdot m^{-3}][m^3] = [mol][kJ \cdot mol^{-1}K^{-1}][K] (PV = nRT)$

Planck constant: [h] = action

• units: $[S] = \frac{\text{energy}}{\text{temperature}}$

diffusion constant with units length2

Maths Tools

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}, \text{for } |x| < 1$$

Gaussian over \mathbb{R}^+ , with $c_n := \begin{cases} \frac{\pi}{2} & \text{if } n \text{ even} \\ 1 & \text{if } n \text{ odd} \end{cases}$

$$I_n(a) = \int_0^\infty x^n e^{-ax^2} dx, a > 0 = \sqrt{c_n} \cdot (n-1)!!(2a)^{-\frac{n+1}{2}}$$

n = 0	n = 2	n = 4
$\frac{1}{2} \left(\frac{\pi}{a}\right)^{\frac{1}{2}}$	$\frac{1}{4} \Big(\frac{\pi}{a^3} \Big)^{\frac{1}{2}}$	$\frac{3}{8} \Big(\frac{\pi}{a^5} \Big)^{\frac{1}{2}}$
n = 1	n = 3	n = 5
$\frac{1}{2a}$	$\frac{1}{2a^2}$	$\frac{1}{a^3}$

ver
$$\mathbb{R}$$
, $\int_{-\infty}^{\infty} x^n e^{-ax^2} dx = \begin{cases} 2I_n(a) \text{ if } n \text{ even} \\ 0 & \text{ if } n \text{ odd} \end{cases}$

Normalized Gaussian
$$2I_{n=0}(a)=rac{1}{\sqrt{2\pi\sigma}}e^{rac{1}{(2\sigma^2)}}$$