Mechanical quantities

notation here:
$U = U(r^N)$ = potential energy
$E = E(T,V)$ = internal energy
f = # of degrees of freedom

- Temperature (NVE MD):

$$T_{\rm kin} = \frac{E_{\rm kin}}{f k_{\rm B}/2}$$

- Internal energy:

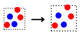
res = residual
see next slide

$$E = \langle E_{\rm kin} + U\rangle \overset{NVT}{=} \frac{f}{2}k_{\rm B}T + \langle U\rangle \equiv E_{\rm id} + E_{\rm res}$$

- Pressure

$\beta = 1/k_{\rm B}T$

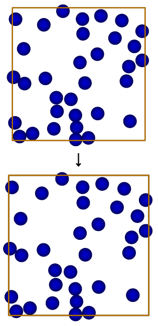
$$p = \frac{N}{V}k_{\rm B}T - \left\langle \left(\frac{\partial U(V^{1/3}\vec\xi^N)}{\partial V}\right)_{\vec\xi^N}\right\rangle \equiv p_{\rm id} + p_{\rm res}$$



- dimensionless (scaled) coordinates $\vec\xi_i$: $\vec r_i = V^{1/3}\vec\xi_i$
- red derivative is calculated at constant $\vec\xi^N$,
  whole configurations is uniformly shrank/swelled
- $p_{\rm id}$ = kinetic contribution (=ideal gas), also $p_{\rm id} = \phi 2E_{\rm kin}/3V$,
  where in periodic b.c. $\phi = N/(N-1)$ takes into acccount 3 zero conserved momenta
- $p_{\rm res}$ = cohesion contribution

Pressure in the *NVT* ensemble: derivation

$$\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V \qquad \langle X\rangle = \frac{1}{Q_N}\int_{V^N} X(\vec r^N)\mathrm{d}\exp[-\beta U(\vec r^N)]\vec r^N$$



$$p = -\left(\frac{\partial F}{\partial V}\right)_T, \quad F = -k_{\rm B}T\ln\frac{Q_N}{N!\Lambda^{3N}}$$

$$Q_N = \int_{V^N}\exp[-\beta U(\vec r^N)]\mathrm{d}\vec r^N \overset{\vec r_i = V^{1/3}\vec\xi_i}{=} \int_{1^{3N}}\exp[-\beta U(V^{1/3}\vec\xi^N)]V^N\mathrm{d}\vec\xi^N$$

$$\begin{aligned} p &= -\left(\frac{\partial F}{\partial V}\right)_T = k_{\rm B}T\left(\frac{\partial \ln Q_N}{\partial V}\right)_{\vec\xi^N} = \frac{k_{\rm B}T}{Q_N}\left(\frac{\partial Q_N}{\partial V}\right)_{\vec\xi^N} \\ &= \frac{k_{\rm B}T}{Q_N}\int_{1^{3N}}\exp[-\beta U(V^{1/3}\vec\xi^N)]NV^{N-1}\mathrm{d}\vec\xi^N \\ &\quad + \frac{k_{\rm B}T}{Q_N}\int_{1^{3N}}\exp[-\beta U(V^{1/3}\vec\xi^N)](-\beta)\left(\frac{\partial U(V^{1/3}\vec\xi^N)}{\partial V}\right)_{\vec\xi^N}V^N\mathrm{d}\vec\xi^N \\ &= \frac{N}{V}k_{\rm B}T - \left\langle\left(\frac{\partial U(V^{1/3}\vec\xi^N)}{\partial V}\right)_{\vec\xi^N}\right\rangle = \boxed{\text{ideal part (kinetic)}} + \boxed{\text{residual part (correction)}} \end{aligned}$$

Residual quantities

= with respect to the standard state of ideal gas at the same temperature, volume (= density), and composition as the given system. Usefull in the canonical (NVT) ensemble.

sometimes called “excess”

For the Helmholtz energy:

ideal gas: $Q_N = V^N$

$$F = -k_{\rm B}T\ln Z_N = -k_{\rm B}T\ln\frac{Q_N}{N!\Lambda^{3N}} = -k_{\rm B}T\ln\frac{V^N}{N!\Lambda^{3N}} - k_{\rm B}T\ln\frac{Q_N}{V^N} \equiv F_{\rm id} + F_{\rm res}$$

Refresh:
de Broglie thermal wavelength:

$$\Lambda = \frac{h}{\sqrt{2\pi m k_{\rm B}T}}$$

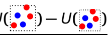
chemical potential of ideal gas:

$$\mu_{\rm id} = \left(\frac{\partial F_{\rm id}}{\partial N}\right)_{T,V} = k_{\rm B}T\ln\frac{N\Lambda^3}{V}$$

Pressure – virtual volume change method

$$p = \frac{N}{V}k_{\rm B}T - \left\langle\left(\frac{\partial U(V^{1/3}\vec\xi^N)}{\partial V}\right)_{\vec\xi^N}\right\rangle$$

Numerical derivative (for a selected series of configurations)



$$\frac{\partial U}{\partial V} = \frac{U(V+\Delta V) - U(V)}{\Delta V} + \mathcal{O}(\Delta V) \equiv \frac{U(\;) - U(\;)}{\Delta V} + \mathcal{O}(\Delta V)$$

$$\frac{\partial U}{\partial V} = \frac{U(V+\Delta V) - U(V-\Delta V)}{2\Delta V} + \mathcal{O}(\Delta V^2)$$

Implementation: $U(V+\Delta V)$ means that the whole configuration (all distances) is swelled by the same ratio; for molecules w.r.t. a reference point (then, N = # of molecules):

$$\left(\frac{V+\Delta V}{V}\right)^{1/3}.$$

The scaled configuration is not included in the trajectory.

For models with a hard core such that swelling the box cannot cause an overlap, shrinking can be used: $P = Nk_{\rm B}T/V + \frac{k_{\rm B}T}{\Delta V}\langle \mathrm{e}^{-[U(V-\Delta V)-U(V)]/k_{\rm B}T}\rangle + \mathcal{O}(\Delta V)$

Pressure from the virial of force

The derivative expanded:

$$\frac{\partial U(V^{1/3}\vec\xi^N)}{\partial V} = \sum_{i=1}^N \frac{1}{3}V^{-2/3}\vec\xi_i\cdot\overset{-\vec f_i}{\frac{\partial U}{\partial \vec r_i}} = \frac{1}{3V}\sum_{i=1}^N \vec r_i\cdot\frac{\partial U}{\partial \vec r_i}$$

The result is

$$p = \frac{N}{V}k_{\rm B}T + \frac{1}{3V}\langle W_f\rangle, \qquad W_f = -\sum_{i=1}^N \vec r_i\cdot\frac{\partial U}{\partial \vec r_i} = \sum_{i=1}^N \vec r_i\cdot\vec f_i \ \ \text{(virial of force)}$$

. . . cannot be directly applied in the periodic boundary conditions.

- Pair additivity in the periodic boundary conditions ⇒

$$p = \frac{N}{V}k_{\rm B}T - \frac{1}{3V}\sum_{i<j}\langle r_{ij}u'(r_{ij})\rangle \equiv p_{\rm id} + p_{\rm res}$$

- For molecular models one can use either the site-frame (N = # of sites/atoms) or molecular frame (N = # of molecules; typically, ref. point = center of mass). Of course, the formulas differ.

Pressure tensor

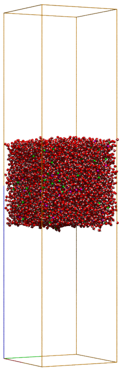
Also called stress tensor (in solids):

$$\overleftrightarrow{P} = \overleftrightarrow{P}_{\rm id} + \overleftrightarrow{P}_{\rm res} = \frac{1}{V}\sum_{i=1}^N(\phi m_i\vec v_i\vec v_i + \vec f_i\vec r_i)$$

- Tensor product $\overleftrightarrow{T} = \vec u\vec v$, also denoted $\overleftrightarrow{T} = \vec u\otimes\vec v$: $T_{ab} = v_a v_b$.
- $\phi = N/(N-1)$ takes into acccount 3 zero conserved momenta in periodic b.c.
- The scalar pressure is 1/3 of its trace, $p = \mathrm{tr}(\overleftrightarrow{P})/3$.
- In models with rigid (constrained) bonds, the virial of constrained forces has to be included (it depends on velocities).
- Diagonal terms are good for calculating surface tension.
- Off-diagonal terms are good for calculating viscosity.*
- Can be also calculated by the virtual box change:
  - elongation in $\hat x \Rightarrow P_{xx}$,
  - simulation cell shape change ⇒ off-diagonal terms.

*Suprisingly, the diagonal terms can be used to calculate viscosity, too.

Surface tension of liquid



Using slab geometry – elongated periodic box (2:1 or more)

$$\gamma = \left(\frac{\partial V}{\partial \mathcal{A}}\right)_{V,T} = \left(\frac{\partial G}{\partial \mathcal{A}}\right)_{p,T} = -\frac{1}{4}L_zP_{\rm t}, \quad \text{where } P_{\rm t} = P_{xx} + P_{yy} - 2P_{zz}.$$

- It is “mechanical quantity”.
- Cannot be used for interfacial energy of crystals (it is “entropic quantity”).
- Cutoff corrections – several variants (cf. simen09).
- P_{zz} = saturated vapor pressure;
  if small enough, P_{xx}, P_{yy} can be replaced by the usual pressure p.
- Alternative: virtual area change method [Gloor et al.: JCP 123, 134703 (2005)]: scaling in $\hat x, \hat y$ and inversely squared in $\hat z$ so that volume is preserved:

$$\gamma = \left\langle\left(\frac{\delta U}{\delta \mathcal{A}}\right)_V\right\rangle$$

Fluctuation quantities

$\langle(\Delta X)^2\rangle = \mathrm{Var}\,X$ = (mean quadratic) fluctuation = variance

$\Delta X = X - \langle X\rangle$

fluctuation = (mech/el/. . . quantity)′ = (thermodynamic potential)′′

Less accurate than mean values!

E.g. (NVT): $P = -\left(\frac{\partial F}{\partial V}\right)_T$, $E = \left(\frac{\partial\beta F}{\partial\beta}\right)_V$, then:

- Heat capacity at constant $[V]$:

$$C_V = \left(\frac{\partial E}{\partial T}\right)_V = \frac{1}{k_{\rm B}T^2}\mathrm{Var}\,E = \frac{1}{k_{\rm B}T^2}\langle(\Delta E_{\rm kin} + \Delta U)^2\rangle$$

$\mathrm{Cov}\,(U, E_{\rm kin}) = \langle\Delta U\Delta E_{\rm kin}\rangle = 0$, $\mathrm{Var}\,E_{\rm kin} = \frac{f}{2}(k_{\rm B}T)^2$ (see exercise) ⇒

$$C_V = \frac{1}{k_{\rm B}T^2}\left\langle\left(\frac{fk_{\rm B}T}{2} + \Delta U\right)^2\right\rangle = \frac{fk_{\rm B}}{2} + \frac{1}{k_{\rm B}T^2}\left\langle(\Delta U)^2\right\rangle \equiv C_{V,\rm id} + C_{V,\rm res}$$

- Isothermal compressibility – in the NPT ensemble

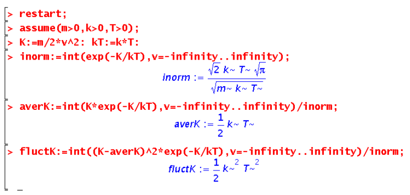
$$\kappa = -\frac{1}{V}\left(\frac{\partial V}{\partial P}\right)_T = \frac{\langle(\Delta V)^2\rangle}{Vk_{\rm B}T}$$

Permittivity can be calculated from the fluctuation of the cell dipole moment, $\langle M^2\rangle$, (more later).

Exercise

Calculate $\langle E_{\rm kin}\rangle$ and $\mathrm{Var}\,(E_{\rm kin})$ for one degree of freedom, $E_{\rm kin} = \frac{1}{2}mv^2$.

$$\begin{aligned} E_{\rm kin} &= \frac{1}{2}mv^2 \\ \pi(v) &= \frac{\exp(-\frac{1}{2}mv^2/k_{\rm B}T)}{\int\exp(-\frac{1}{2}mv^2/k_{\rm B}T)\mathrm{d}v} \\ \langle E_{\rm kin}\rangle &= \int E_{\rm kin}\pi(v)\mathrm{d}v \\ \mathrm{Var}\,(E_{\rm kin}) &= \int(E_{\rm kin} - \langle E_{\rm kin}\rangle)^2\pi(v)\mathrm{d}v \\ &= \frac{1}{2}(k_{\rm B}T)^2 \end{aligned}$$



```
> restart;
> assume(m>0,k>0,T>0);
> K:=m/2*v^2: kT:=k*T:
> inorm:=int(exp(-K/kT),v=-infinity..infinity);
```

$$inorm := \frac{\sqrt{2}\,k\!\sim\,T\!\sim\,\sqrt{\pi}}{\sqrt{m\!\sim\,k\!\sim\,T\!\sim}}$$

```
> averK:=int(K*exp(-K/kT),v=-infinity..infinity)/inorm;
```

$$averK := \frac{1}{2}k\!\sim\,T\!\sim$$

```
> fluctK:=int((K-averK)^2*exp(-K/kT),v=-infinity..infinity)/inorm;
```

$$fluctK := \frac{1}{2}k\!\sim^2\,T\!\sim^2$$

Isochoric heat capacity C_V in the NVE ensemble

In the **canonical** ensemble:

$$\mathrm{Var}\,X = k_{\rm B}T^2\frac{\partial X}{\partial T} \equiv k_{\rm B}T^2X' \quad (X = U,\,E_{\rm kin}) \tag{1}$$

$$\mathrm{Cov}\,(U, E_{\rm kin}) = 0$$

$$C_V = U' + E'_{\rm kin}, \quad E'_{\rm kin} = \frac{fk_{\rm B}}{2}$$

Unnormalized probability distribution for temperature T:

$$w(U, E_{\rm kin}) \sim \exp\left[-\frac{(\delta U)^2}{2\,\mathrm{Var}\,U} - \frac{(\delta E_{\rm kin})^2}{2\,\mathrm{Var}\,E_{\rm kin}}\right]$$

where the deviations are linearized: $(U, E_{\rm kin}) = (U_0 + \delta U, E_{\rm kin,0} + \delta E_{\rm kin})$

Unnormalized probability distribution taking into account change in temperature, $T = T_0 + \delta T$:

$$w(T, U, E_{\rm kin}) \sim \exp\left[-\frac{(\delta U - U'\delta T)^2}{2\,\mathrm{Var}\,U} - \frac{(\delta E_{\rm kin} - E'_{\rm kin}\delta T)^2}{2\,\mathrm{Var}\,E_{\rm kin}}\right]$$

Isochoric heat capacity C_V in the NVE ensemble II

In the **MD NVE** ensemble it holds $U+E_{\rm kin}=E=\text{const}$ or $U=E-E_{\rm kin}$, and of course $\delta E_{\rm kin}=-\delta U$, so that let us choose $\delta E_{\rm kin}$ as the independent variable

$$w(T,E_{\rm kin})\sim\exp\left[-\left(\frac{1}{2k_BT^2U'}+\frac{1}{2k_BT^2E'_{\rm kin}}\right)\delta E_{\rm kin}^2-\left(\frac{U'}{2k_BT^2}+\frac{E'_{\rm kin}}{2k_BT^2}\right)\delta T^2\right]$$

where the variances were replaced by derivatives, see (1).

NB: term $\delta E_{\rm kin}\delta T$ cancels out!

From term at δT^2:

$$\operatorname{Var}T=\frac{k_BT^2}{U'+E'_{\rm kin}}=\frac{k_BT^2}{C_V}$$

which is a known expression for the temperature in the NVE ensemble, however, T here is **not** the kinetic temperature, but T in formula $\mathrm{d}E=T\mathrm{d}S\ [V]$. The average temperature is the same, but not the fluctuation.

Isochoric heat capacity C_V in the NVE ensemble III

From the term at δE^2:

$$\operatorname{Var}E=\operatorname{Var}U=\operatorname{Var}E_{\rm kin}=\frac{k_BT^2}{1/U'+1/E'_{\rm kin}}=\frac{k_BT^2}{1/(C_V-E'_{\rm kin})+1/E'_{\rm kin}}$$

Since $E'_{\rm kin}$ is known:

$$C_V=\frac{fk}{2}\left[\left(\frac{2T^2}{f\operatorname{Var}T_{\rm kin}}-1\right)^{-1}+1\right]$$

where $T_{\rm kin}\equiv E_{\rm kin}/(fk/2)$ is the kinetic temperature.

Final expression applicable in simulations:

$$C_V=\frac{fk}{2}\left[\left(\frac{2\,\langle T_{\rm kin}\rangle^2}{f\,\langle(\Delta T_{\rm kin})^2\rangle}-1\right)^{-1}+1\right]$$

Entropic quantities

These include F, G ($(\Leftarrow$ partition function), S (number of states W), μ, ΔG,...
They cannot be expressed as $\langle\cdot\rangle$.

Formally (for the canonical partition function) but totally useless:

$$Q=\frac{\int e^{-\beta U}d\vec r^N}{1}=\frac{\int e^{-\beta U}d\vec r^N}{V^{-N}\int e^{-\beta U}e^{+\beta U}d\vec r^N}=\frac{V^N}{\langle e^{\beta U}\rangle}$$

Usage: equilibria in general; solubility, ligand–receptor binding, stability of biomolecules, ...

- thermodynamic integration:
  over a real variable (T, V, P) or coupling parameter
- non-Boltzmann sampling:
  gradual insertion, alchemical transmutation;
  umbrella sampling
  multiple histogram reweighting
  metadynamics
- Widom particle insertion method
- reversible work calculated by the integration of force
- local density method

Thermodynamic integration

Remember physical chemistry: $\mathrm{d}F=-S\mathrm{d}T-p\mathrm{d}V$, $\mathrm{d}G=-S\mathrm{d}T+V\mathrm{d}p$

Canonical ensemble ($\beta=1/k_BT$):

$$\left(\frac{\partial F}{\partial V}\right)_T=-p,\quad\left(\frac{\partial(\beta F)}{\partial\beta}\right)_V=E,\quad\text{or}\quad\left(\frac{\partial(\beta F_{\rm res})}{\partial\beta}\right)_V=\langle U\rangle$$

notation here:
$U=U(r^N)$
= potential energy
$E=E(T,V)$
= internal energy

- Numerically integrated: p, E must be determined in many points
- Start from a suitable reference state (known state, ideal gas, harmonic crystal)

Proof # 1 of $\frac{\partial(\beta F)}{\partial\beta}=E$:

$$\frac{\partial(\beta F)}{\partial\beta}=\frac{\partial(F/T)}{\partial(1/T)}=\frac{\partial(F/T)}{\partial T}\Big/\frac{\partial(1/T)}{\partial T}=\frac{-ST-F}{T^2}\Big/\left(\frac{-1}{T^2}\right)=ST+F=E$$

Proof # 2 of $\frac{\partial(\beta F)}{\partial\beta}=E$:

$$\frac{\partial(\beta F)}{\partial\beta}=\frac{-\partial\ln Z}{\partial\beta}=-\frac{1}{Z}\frac{\partial Z}{\partial\beta}=-\frac{\partial\sum_\psi e^{-\beta\mathcal{E}(\psi)}/\partial\beta}{\sum_\psi e^{-\beta\mathcal{E}(\psi)}}=-\frac{\sum_\psi\left[-\mathcal{E}(\psi)e^{-\beta\mathcal{E}(\psi)}\right]}{\sum_\psi e^{-\beta\mathcal{E}(\psi)}}=\langle\mathcal{E}\rangle=E$$

Integration over a coupling parameter

Let us consider any dependence $(\beta U)(\lambda)$, e.g.:

$$(\beta U)(\lambda)=\begin{cases}\beta[U_0+\lambda(U_1-U_0)] & \lambda=\text{coupling parameter, }\lambda\in[0,1]\\ \lambda U & \lambda\equiv\beta\text{: see previous slide}\end{cases}$$

then

$$\frac{\partial\beta F_{\rm res}}{\partial\lambda}=-\frac{\partial\ln Q}{\partial\lambda}=-\frac{1}{Q}\int\frac{\partial e^{-\beta U}}{\partial\lambda}d\vec r^N=\frac{1}{Q}\int\frac{\partial(\beta U)(\lambda)}{\partial\lambda}e^{-\beta U(\lambda)}d\vec r^N=\left\langle\frac{\partial(\beta U)(\lambda)}{\partial\lambda}\right\rangle_\lambda$$

$$(\beta F_{\rm res})(\lambda_1)=(\beta F_{\rm res})(\lambda_0)+\int_{\lambda_0}^{\lambda_1}\left\langle\frac{\partial(\beta U)(\lambda)}{\partial\lambda}\right\rangle_\lambda d\lambda$$

where $\langle\cdot\rangle_\lambda$ = mean value in the ensemble (simulation) with potential $U(\lambda)$

How to integrate:

- Several discrete values of λ_i:
  - fit to a suitable function and integrate it
  - use a numerical quadrature method; e.g., the Simpson formula
    (points with a higher weight should be simulated longer)
- A little change of λ in every MD step (the system is almost in equilibrium) + integration

Examples

- For $\lambda=\beta$ we get as before:

$$\beta_1F_{\rm res}(\beta_1)-\beta_0F_{\rm res}(\beta_0)=\int_{\beta_0}^{\beta_1}\langle U\rangle d\beta$$

- Integration from an Einstein crystal to a real crystal.
  NB: Einstein crystal = independent harmonic oscillators at lattice sites.
  There are minor problems (⇒ correction) when the crystal is detached from the fixed sites.
- Integration from ideal gas around the critical point to liquid. Because of gas singularity at zero density, one of the integrals is (for NPT):

$$\ln\phi=\frac{\mu-\mu^\circ}{RT}=\int_0^p\left(V_m-\frac{RT}{p'}\right)dp'$$

It helps to use a virial equation of state for low densities.

Non-Boltzmann sampling

We want $(\beta U)_1$, but we simulate $(\beta U)_0$ (can change β/U/both)

$$\Delta(\beta U)=(\beta U)_1-(\beta U)_0$$

$$\langle X\rangle_{(\beta U)_1}=\frac{\int Xe^{-(\beta U)_1}d\vec r^N}{\int e^{-(\beta U)_1}d\vec r^N}=\frac{\frac{1}{Q_0}\int Xe^{-(\beta U)_0}e^{-\Delta(\beta U)}d\vec r^N}{\frac{1}{Q_0}\int e^{-(\beta U)_0}e^{-\Delta(\beta U)}d\vec r^N}=\frac{\langle Xe^{-\Delta(\beta U)}\rangle_0}{\langle e^{-\Delta(\beta U)}\rangle_0}$$

Helmholtz energy:

$$\begin{aligned}\Delta(\beta F_{\rm res})&=\beta_1F_{\rm res}((\beta U)_1)-\beta_0F_{\rm res}((\beta U)_0)\\&=-\ln\left(\frac{Q_1}{Q_0}\right)=-\ln\frac{\int e^{-(\beta U)_1}d\vec r^N}{\int e^{-(\beta U)_0}d\vec r^N}\\&=-\ln\frac{\int e^{-(\beta U)_0}e^{-\Delta(\beta U)}d\vec r^N}{\int e^{-(\beta U)_0}d\vec r^N}=-\ln\langle e^{-\Delta(\beta U)}\rangle_0\\&=\ln\langle e^{+\Delta(\beta U)}\rangle_1\end{aligned}$$

where the last equation follows from $0\longleftrightarrow1$ interchange

Non-Boltzmann sampling contd.

$e^x=1+x+\cdots$
$\ln(1+x)=x+\cdots$

- $\Delta(\beta U)$ must not be too large
- the thermodynamic integration is recovered for infinitesimally small $\Delta(\beta U)$:

$$\begin{aligned}\Delta(\beta F_{\rm res})&=-\ln\langle e^{-\Delta(\beta U)}\rangle_0\\&\approx-\ln\langle1-\Delta(\beta U)\rangle_0\\&=-\ln(1-\langle\Delta(\beta U)\rangle_0)\\&\approx\langle\Delta(\beta U)\rangle\end{aligned}$$

$$\Rightarrow\ \partial(\beta F_{\rm res})/\partial\lambda=\langle\partial\Delta(\beta U)/\partial\lambda\rangle_\lambda$$

https://positiveoutlooksblog.com/2015/01/30/surviving-the-storm-together

Umbrella sampling

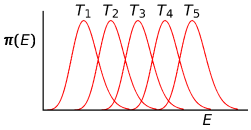
The system in the middle is sampled: mid $=(\beta U)_0+\Delta(\beta U)/2=(\beta_0U_0+\beta_1U_1)/2$:

$$\Delta(\beta F_{\rm res})=\ln\langle e^{+\Delta(\beta U)/2}\rangle_{\rm mid}-\ln\langle e^{-\Delta(\beta U)/2}\rangle_{\rm mid}$$

Often, term “umbrella sampling” is extended to using many steps of similar kind

Multiple histogram reweighting I

Building the **density of states** as a funtion of energy in a wide range of temperatures from overlapping histograms of energies obtained in a number of simulations at different temperatures.



$$\boldsymbol{\pi}(E)=D(E)e^{-\beta E}$$

$$D(E)\propto E^{kN}$$

where k is a constant

Configurational integral and residual Helmholtz energy at temperature T_i, $\beta_i=1/k_BT_i$:

$$Q_i=e^{-\beta_iF_i}=\int e^{-\beta_iU(\vec r^N)}d\vec r^N=\int D(E)e^{-\beta_iE}dE$$

where $D(E)$ is the density of states:

$$\begin{aligned}D(E)\quad&=\quad\int_{U(\vec r^N)=E}1\,d\vec r^N\equiv\int\delta(U(\vec r^N)-E)\,d\vec r^N\\&\overset{\text{rectangle}}{\approx}\int_{U(\vec r^N)\in(E-\Delta E/2,E+\Delta E/2)}1d\vec r^N\end{aligned}$$

[Ghoufi et al. (2008)]

Detour: Density of states for a particle in a box

One particle in 1D in a box of length L, energies of eigenstates according to the Schrödinger equation:

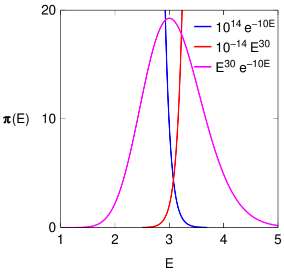
$$E_n=\frac{n^2h^2}{8L^2}\propto n^2,\ n=1,2,\ldots$$

Let’s have $f=3N$ such degrees of freedom. The number of states $\#(E)$ with energy less than E satisfies the equation

$$n_1^2+n_2^2+\cdots+n_f^2<E=[E^{1/2}]^2$$

For large E, $\#(E)=1/2^f$ of the volume of the $E^{1/2}$-ball in f-dimendional space:

$$\#(E)=\frac{\pi^{f/2}E^{f/2}}{\Gamma(f/2+1)}\propto E^{f/2}\quad\Rightarrow\quad D(E)=\frac{d\#(E)}{dE}\propto E^{f/2-1}$$



$$\boldsymbol{\pi}(E)=D(E)e^{-\beta E}$$

The Boltzmann factor $e^{-\beta E}$ eventually wins!
For large N, the product converges to a Gaussian with $\operatorname{Var}(E)\propto f\propto N$ ($\sigma=N^{1/2}$)

Multiple histogram reweighting II

In the simulation, we calculate the histogram $h_i(E)$ for a set of (equidistant) energies E, or some equivalent Gaussian-based δ-function approximation. We will use the normalized histogram and the $\int$-form. (To repeat, subscript $_i$ refers to T_i.)

$$\sum_E h_i(E) = \int h_i(E)\mathrm{d}E = 1$$

$$h_i(E) = \frac{D(E)\mathrm{e}^{-\beta_i E}}{\int D(E)\mathrm{e}^{-\beta_i E}\mathrm{d}E} = D(E)\mathrm{e}^{-\beta_i(E-F_i)}$$

Using one temperature only but F_i is not known (yet):

$$D(E) = h_i(E)\mathrm{e}^{\beta_i(E-F_i)} \qquad (2)$$

- We will average $D(E)$ from several simulations at different temperatures.
- $D(E)$ does not depend on T_i, but our calculation does ± statistical errors.
- At different T_i different ranges of E are sampled.

⇒ We compose the total $D(E)$ from all data:

$$D(E) = \sum_i w_i(E)h_i(E)\mathrm{e}^{\beta_i(E-F_i)}, \qquad \sum_i w_i(E) = 1$$

Multiple histogram reweighting III

Determining the weights: minimization of the error $\delta D(E)$ (or some estimate). Using reasonable assumptions, we get

$$w_i(E) = \frac{N_i h_i(E)}{\sum_j N_j h_j(E)} = \frac{N_i \mathrm{e}^{-\beta_i(E-F_i)}}{\sum_j N_j \mathrm{e}^{-\beta_j(E-F_j)}}$$

where N_i is the number of measurements at temperature β_i. ⇒

$$\mathrm{e}^{-\beta_i F_i} = \int D(E)\mathrm{e}^{-\beta_i E}\,\mathrm{d}E = \int \sum_j w_j(E)h_j(E)\mathrm{e}^{\beta_j(E-F_j)}\mathrm{e}^{-\beta_i E}\,\mathrm{d}E$$

$$= \int \frac{\sum_j N_j \mathrm{e}^{-\beta_j(E-F_j)} h_j(E)\mathrm{e}^{\beta_j(E-F_j)}}{\sum_j N_j \mathrm{e}^{-\beta_j(E-F_j)}}\mathrm{e}^{-\beta_i E}\,\mathrm{d}E = \int \frac{\sum_j N_j h_j(E)\mathrm{e}^{-\beta_j E}}{\sum_j N_j \mathrm{e}^{-\beta_j(E-F_j)}}\mathrm{d}E$$

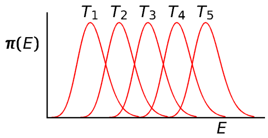
can be solved by iterations (self-consistent solution).
F_i are determined but an additive constant, $D(E)$ but a multiplicative factor

Multiple histogram reweighting IV

Expectation value at temperature β:

$$\langle X\rangle_\beta = \frac{\int X(E)D(E)\mathrm{e}^{-\beta E}\mathrm{d}E}{\int D(E)\mathrm{e}^{-\beta E}\mathrm{d}E} = \frac{\displaystyle\int X(E)\frac{\sum_i h_i(E)}{\sum_j N_j \mathrm{e}^{-\beta_j(E-F_j)}}\mathrm{e}^{-\beta E}\mathrm{d}E}{\displaystyle\int \frac{\sum_i h_i(E)}{\sum_j N_j \mathrm{e}^{-\beta_j(E-F_j)}}\mathrm{e}^{-\beta E}\mathrm{d}E}$$

- $\int \mathrm{d}E$ is over histograms of width ΔE
- if ΔE is very short, all calculated energies are stored and $\int X(E)h_i(E)\mathrm{d}E$ is replaced by a sum of E over $h_i(E) = 1/N_i$
- $\sqrt{\mathrm{Var}E}/\langle E\rangle \approx N^{-1/2} \Rightarrow \beta_i/\beta_{i+1} \approx 1 \pm N^{-1/2}$ (overlapping distributions)



Parallel tempering (replica exchange)

k simulations at temperatures $\beta_1 < \beta_2 \ldots$ are run in parallel.

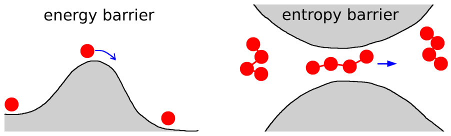
Once a while, 2 subsystems β_i, β_j (normally $|i-j| = 1$) are interchanged, the acceptance probability is

$$\min\left\{1, \frac{\exp(-\beta_i E_j - \beta_j E_i)}{\exp(-\beta_i E_i - \beta_j E_j)}\right\}$$

- Pros: easier barrier crossing, improved ergodicity, faster convergence at low temperatures
- Cons: correlations between subsystems, difficult error estimation

Conformational flooding, metadynamics, Wang–Landau

Method suitable for fast **barrier crossing** (bad ergodicity – “bottleneck”), incl. calculation of the free energy profile (ΔF or ΔG), based on decreasing the energy barrier.



- (several predecessors)
- Conformational flooding H. Grubmüller (1995)
- Wang–Landau (MC) F. Wang, D.P. Landau (2001)
- Metadynamics (MD) A. Laio, M. Parrinello (2002)
- Formal equivalence C. Junghans, D. Perez, T. Vogel (2014)

Conformational flooding, metadynamics, Wang–Landau

- Our system is defined by the potential $U_0(\vec{r}^N)$
- **Collective variable** $\lambda = \lambda(\vec{r}^N)$ describes the studied process, e.g.: reaction coordinate, distance ligand–receptor (optionally + orientation), etc.
  There may be 2 or even more collective variables
- We simulate system with potential $U(\vec{r}^N) = U_0(\vec{r}^N) + \Delta U(\lambda)$, where $\Delta U = 0$ at start
- $\Delta U(\lambda)$ is periodically updated:

$$\Delta U := \Delta U + \omega\frac{\delta(\lambda)}{h(\lambda)}, \quad \lambda = \lambda(\vec{r}^N)$$

δ = aproximation of δ-function (MC: histogram bin, better and MD: Gauss)
ω = small enough relaxation parameter, $\omega \ll k_\mathrm{B}T$
$h(\lambda)$ = density of Cartesian points on hypersurface λ; e.g., $h(\lambda) = 4\pi\lambda^2$ for $\lambda = r_{12}$
⇒ the probability of visiting the same λ again decreases

Conformational flooding, metadynamics, Wang–Landau

Free energy profile along λ:

- strictly, updating should be turned off in the final run (MC: microreversibility violation, MD: heating)
- then, the residual almost-uniform $p(\lambda)$ is determined and:

$$F(\lambda) = \mathrm{const} - \Delta U(\lambda) - k_\mathrm{B}T\ln[p(\lambda)/h(\lambda)]$$

- in practice, with small enough ω and continued updating, we can assume $p(\lambda)/h(\lambda) = \mathrm{const}^\dagger$ so that (in the sampled region of λ)

$$F(\lambda) = \mathrm{const} - \Delta U(\lambda)$$

Free energy of a well-define “basin” of states (e.g., a bond state) is:

$$F = -k_\mathrm{B}T\ln\int_{\lambda_1}^{\lambda_2}\mathrm{e}^{-F(\lambda)/k_\mathrm{B}T}h(\lambda)\mathrm{d}\lambda$$

integration over region for more collective variables

†symbols const are different

[plot/metadynamics.sh 1 2]

Conformational flooding, metadynamics, Wang–Landau: Case study

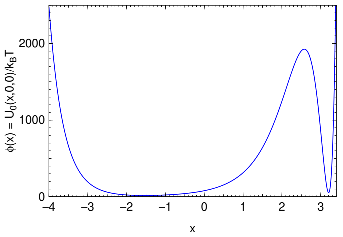
3D system, interaction energy: online simulation is for T=1 K and T=2 K, type $\boxed{r}$ to reread

$$\frac{U_0(x,y,z)}{k_\mathrm{B}\,\mathrm{K}} = \phi(x)(y^2+z^2+1), \text{ where } \phi(x) = \Big[(x+1)(x+2)(x-2)(x-4)+21\Big]\mathrm{e}^{(x+1.5)^2/3}$$

- The barrier is $E^*/k_\mathrm{B} = 1910\,\mathrm{K}$, $\mathrm{e}^{-E^*/k_\mathrm{B}1\mathrm{K}} = 10^{-929}$
- Collective variable $\lambda = x$
- Histogram: triangular $\delta(\lambda)$, grid = 1/100
- On purpose short MC step in x (≈ MD)
- Initial $\omega = 0.25k_\mathrm{B}T$, decrease to finish
  fine + stop ($\omega \equiv a < 0$): aaaV in the plot window
- graphs shown:
  – running: $\Delta U - F_\mathrm{exact}$, where

$$F_\mathrm{exact}(\lambda) = \phi(\lambda) + k_\mathrm{B}T\ln(\phi(\lambda))$$

  – final: ΔU and $\phi(x) = U_0(x,0,0)$
  – F and F_exact comparison



Note: in real systems, the second term will be a result of many degrees of freedom, not just 2 (y, z)

[xmaple ../maple/metadynamics-case-study.mw]

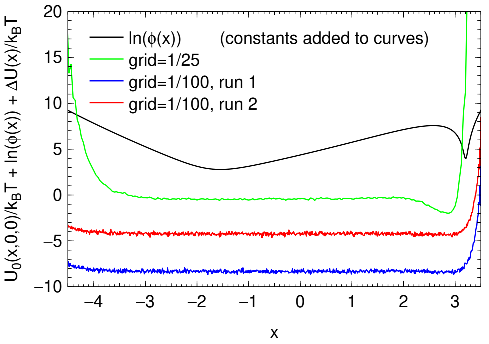
Conformational flooding, metadynamics, Wang–Landau: Case study

The final graph shows $F(x) + \Delta U(x)$ (two independent simulations $T = 1\,\mathrm{K}$, different start)

1/grid	$\Delta F/k_\mathrm{B}T$
50	37.85(8)
100	40.52(5)
200	40.98(5)
400	41.08(3)
exact*	41.13

where ΔF is the difference between the right and left basins.

* see Maple



Widom particle insertion method I

Open system

$$\mathrm{d}F = -S\mathrm{d}T - p\mathrm{d}V + \mu\mathrm{d}N$$

$$\beta\mu = \left(\frac{\partial(\beta F)}{\partial N}\right)_{V,T} = -\left(\frac{\partial\ln Z_N}{\partial N}\right)_{V,T}$$

$$\beta\mu_\mathrm{res} = \left(\frac{\partial(\beta F_\mathrm{res})}{\partial N}\right)_{V,T} = -\left(\frac{\partial\ln(Q_\mathrm{N}/V^N)}{\partial N}\right)_{V,T} \approx -\left(\ln\frac{Q_{N+1}}{V^{N+1}} - \ln\frac{Q_N}{V^N}\right)$$

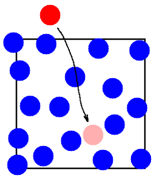
$$\exp(-\beta\mu_\mathrm{res}) = \frac{1}{V}\frac{Q_{N+1}}{Q_N}$$

Or for the full chemical potential:

$$\mathrm{e}^{-\beta\mu} = \frac{Z_{N+1}}{Z_N} = \frac{Q_{N+1}/[(N+1)!\Lambda^{3(N+1)}]}{Q_N/[N!\Lambda^{3N}]} = \frac{1}{(N+1)\Lambda^3}\frac{Q_{N+1}}{Q_N} \approx \frac{1}{N\Lambda^3}\frac{Q_{N+1}}{Q_N}$$

then by subtracting $\mu^\mathrm{id} = k_\mathrm{B}T\ln\left(\frac{N\Lambda^3}{V}\right)$ we get the same $\mu_\mathrm{res} = \mu - \mu^\mathrm{id}$

Widom particle insertion method II



$$\exp(-\beta\mu_{\rm res}) = \frac{1}{V}\frac{Q_{N+1}}{Q_N}$$

$$N \to N+1 \qquad U_{N+1} = U_N + \Psi(N)$$

$$\frac{1}{V}\frac{Q_{N+1}}{Q_N} = \frac{1}{VQ_N}\int \exp(-\beta U_N - \beta\Psi)\mathrm{d}\vec r_1 \ldots \mathrm{d}\vec r_{N+1} = \frac{1}{V}\int \langle \mathrm{e}^{-\beta\Psi}\rangle_N \mathrm{d}\vec r_{N+1}$$

$$\exp(-\beta\mu_{\rm res}) = \frac{1}{V}\int \langle \mathrm{e}^{-\beta\Psi}\rangle_N \mathrm{d}\vec r_{N+1} = \Big\langle \langle \mathrm{e}^{-\beta\Psi}\rangle_N \Big\rangle_{\text{random } \vec r_{N+1}}$$

where $\frac{1}{V}\int X \mathrm{d}\vec r_{N+1} = \langle X\rangle_{\text{random } \vec r_{N+1}}$ = mean value of X over positions of the $(N+1)$-th particle in volume V, calculated by MC integration (inserting a particle at rancom place)

- $(N+1)$-th particle does not influence the system – it is virtual (fictitious, ghost)
- Problem: dense systems, large solutes
  Remedy: gradual insertion (thermodynamic integration or by finite steps)

Similar: solute insertion ⇒ solubility, Henry constant

Example

[cd ../maple; xmaple simul07+18+Widom.mw]

We have simulated $N = 500$ Ar atoms (Lennard-Jones: $\sigma = 3.405$ Å, $\epsilon/k_B = 119.8$ K) in a box of volume $V = 15.791\,\mathrm{nm}^3$ at temperature $T = 150$ K. By the Widom method, we found that[‡]

$$\exp(-\mu_{\rm res}/k_BT) = 0.749(3)$$

Calculate μ°, the chemical potential of Ar with respect to the standard state of ideal gas at pressure $p^{\rm st} = 1$ bar and temperature T.

$$\mu^{\rm id} = k_BT\ln\frac{N\Lambda^3}{V}$$

Hints: $\mu_{\rm res} = \mu - \mu^{\rm id}(T,V)$ (N particles in volume V)

$\mu^\circ = \mu - \mu^{\rm id}(T,V^{\rm id})$ ($V^{\rm id}$ = volume of ideal gas of N particles at $T, p^{\rm st}$)

$$\begin{aligned}\mu^\circ &= \mu_{\rm res} + \mu^{\rm id}(T,V) - \mu^{\rm id}(T,V^{\rm id})\\ &= \mu_{\rm res} + k_BT\ln\frac{Nk_BT}{p^{\rm st}V}\end{aligned}$$

$$\begin{aligned}\mu_{\rm res} &= 5.985\times10^{-22}\,\mathrm{J}\ \text{(per molecule)}\\ \sigma(\mu_{\rm res}) &= k_BT\times(0.003/0.749) = 8.3\times10^{-24}\,\mathrm{J}\ \text{(std. error)}\\ V_{\rm id} &= Nk_BT/p^{\rm st} = 10355\,\mathrm{nm}^3\\ \Delta\mu &\equiv \mu(15.791\,\mathrm{nm}^3) - \mu(10355\,\mathrm{nm}^3) = -T\Delta S_m\\ &= -k_BT\ln(15.791/10355) = 1.3432\times10^{-20}\,\mathrm{J}\\ \mu &= \mu_{\rm res} + \Delta\mu = 1.4030(8)\times10^{-20}\,\mathrm{J}\\ \mu &= 8.449(5)\,\mathrm{kJ\,mol^{-1}}\ \text{(per mole)}\end{aligned}$$

[‡] The value in parentheses is the estimated standard uncertainty in the unit of the least significant digit.

Henry constant

Gas (2) dissolved in liquid (1), mole fraction in liquid $x_2 = N_2/(N_1+N_2)$, $x_2 \ll 1$.
One form of the Henry law for partial pressure p_2 of (2) in equilibrium with solution:

$$p_2 = K_H x_2$$

By virtual insertion of one molecule ($N_2 = 1$) of gas (2) to pure liquid (1), we get the residual chemical potential of (2) at $x_2 = 1/(N_1+1) \approx 1/N_1$,

$$\mu^{\rm id}(V) = k_BT\ln\frac{N\Lambda^3}{V}$$

$$\left.\begin{aligned}&\mu_{\rm res,2} = \mu_2(x_2) - \mu_2^{\rm id}(V)\\ &\text{In equilibrium:}\\ &\mu_2(x_2) \overset{\rm eql.}{=} \mu^{\rm id}(V_2) = k_BT\ln\frac{\Lambda_2^3}{V_2}\end{aligned}\right\}\ \mu_{\rm res,2} = k_BT\ln\frac{N_2\Lambda^3}{V_2} - k_BT\ln\frac{N_2\Lambda^3}{V} = k_BT\ln\frac{V}{V_2}$$

where V_2 is the volume corresponding to pressure p_2 (don't forget that $N_2 = 1$),

$$V_2 = \frac{1k_BT}{p_2} = \frac{k_BT}{x_2K_H} = \frac{k_BTN_1}{K_H}$$

where $\rho_1 = N_1/V$ = number density of liquid (1). Finally,

$$K_H = \frac{\rho_1k_BT}{\mathrm{e}^{-\beta\mu_{\rm res,2}}} = \frac{\rho_1k_BT}{\big\langle\langle \mathrm{e}^{-\beta\Psi}\rangle_{N_1}\big\rangle_{\text{random insertion of (2)}}}$$

Reversible work by integrating the mean force

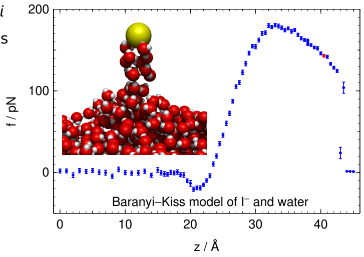
From thermodynamics:

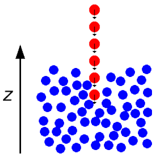
$$\Delta G = W_{\text{other than pressure–volume}} \quad [p,T]$$

$$\Delta\mu_i = -\int_{\vec r_i(1)}^{\vec r_i(2)} \langle\vec f_i\rangle\cdot\mathrm{d}\vec r_i$$

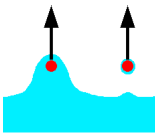
where $\vec f_i = -\partial U/\partial\vec r_i$ is the force acting on particle i

Molecules: the force applies to the center of mass or other reference point.





Problem: hysteresis



Mean force and its potential

Let us define the **single-particle density** as (also denoted as n_1)

$$\rho_1(\vec r_1) = \frac{N}{Q_N}\int \mathrm{e}^{-\beta U(\vec r_1,\ldots,\vec r_N)}\mathrm{d}\vec r_2\ldots\mathrm{d}\vec r_N$$

– $\rho_1(\vec r_1)\mathrm{d}\vec r_1$ is the probability of finding a particle (any one) in $\mathrm{d}\vec r_1$
– extension to mixtures: $N/Q_N \to N_{\rm species}/Q_N$

The **potential of mean force** is defined by

$$\bar U_1(\vec r_1) = -k_BT\ln[V\rho_1(\vec r_1)]$$

The corresponding force is

$$\bar f_i = -\left(\frac{\partial\bar U_1}{\partial\vec r_1}\right) = k_BT\frac{\partial\rho_1/\partial\vec r_1}{\rho_1} = \left\langle -\left(\frac{\partial U}{\partial\vec r_1}\right)\right\rangle_{\vec r_2,\ldots,\vec r_N} = \langle\vec f_1\rangle_{\vec r_2,\ldots,\vec r_N}$$

i.e,. it is indeed the mean force on particle 1 held at fixed position $\vec r_1$.

Note: Similarly, for a 2-particle distribution, which for a pair of particles in an isotropic fluid is $g(r)$, one defines the potential of mean force as:

$$\bar U_2(r) = -k_BT\ln[g(r)]$$

Interfacial (surface) energy of solids

Cleaving [Davidchack, Laird: JCP 118, 7651 (2003)]: Thermodynamic integration over a gradually growing “knife” (e.g., Gaussian potential) inserted between crystalographic planes.

Molding [Espinosa, Vega, Sanz: JCP 141, 134709 (2014)]: Thermodynamic integration over a gradually growing “mold” (potential wells) to keep a crystal growing in a part of the system.

- General problem of both methods: hysteresis

Gibbs energy of crystals

- Einstein crystal[§] at given T as a reference, integration over a coupling parameter [Frenkel, Ladd: JCP 81 3188 (1984), Frenkel, Mulder: Mol. Phys. 55, 1171 (1985)].
- Classical method: harmonic vibrations as a reference, thermodynamic integration of the difference $0 \to T$ [Kolafa JCTC 15, 68 (2019) and references therein]

[§] independent harmonic oscillators, here treated classically

---

Local density/concentration method

Let a solute i be a subject of external potential $U_i^{\rm ext}(\vec r)$ (e.g., “gravity”). In equilibrium:

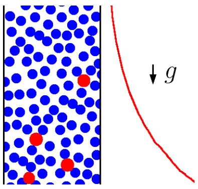
$$\mu_i(\vec r) + U_i^{\rm ext}(\vec r) = \text{const}$$

or

$$\mu_i(\vec r_1) - \mu_i(\vec r_2) = -[U_i^{\rm ext}(\vec r_1) - U_i^{\rm ext}(\vec r_2)]$$

We determine the concentration or density at $\vec r_1$ and know the chemical potential (with respect to a certain reference)

Example:



- reference = infinite dilution approximation ($\gamma = 1$) in the region of small concentration
- the activity coefficient γ in the region of high concentration can be calculated