Molecular Dynamics: A Mathematical Introduction

Gabriel STOLTZ

stoltz@cermics.enpc.fr

(CERMICS, Ecole des Ponts & MICMAC team, INRIA Rocquencourt)

Workshop “Modèles Stochastiques en Temps Long“

Outline

- **Statistical physics: some elements** [Lecture 1]
  - Microscopic description of physical systems
  - Macroscopic description: thermodynamic ensembles
- **Sampling the microcanonical ensemble** [Lecture 1]
  - Hamiltonian dynamics and ergodic assumption
  - Longtime numerical integration of the Hamiltonian dynamics
- **Sampling the canonical ensemble** [Lectures 1-2]
  - Markov chain approaches (Metropolis-Hastings)
  - SDEs: Langevin dynamics
  - Deterministic methods
- **Computation of free energy differences** [Lectures 2-3]
- **Computation of transport coefficients** [Lecture 3]

General references (1)

- Statistical physics: theoretical presentations
  - R. Balian, *From Microphysics to Macrophysics. Methods and Applications of Statistical Physics*, volume I - II (Springer, 2007).
  - many other books: Chandler, Ma, Phillies, Zwanzig, ...

- Computational Statistical Physics
  - D. Frenkel and B. Smit, *Understanding Molecular Simulation, From Algorithms to Applications* (Academic Press, 2002)
  - M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford, 2010)
  - M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (Oxford University Press, 1987)
  - D. C. Rapaport, *The Art of Molecular Dynamics Simulations* (Cambridge University Press, 1995)
  - T. Schlick, *Molecular Modeling and Simulation* (Springer, 2002)

General references (2)

- Longtime integration of the Hamiltonian dynamics
  - E. Hairer, C. Lubich and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for ODEs* (Springer, 2006)
  - B. J. Leimkuhler and S. Reich, *Simulating Hamiltonian dynamics*, (Cambridge University Press, 2005)
  - E. Hairer, C. Lubich and G. Wanner, Geometric numerical integration illustrated by the Störmer-Verlet method, *Acta Numerica* **12** (2003) 399–450

- Sampling the canonical measure
  - L. Rey-Bellet, Ergodic properties of Markov processes, *Lecture Notes in Mathematics*, **1881** 1–39 (2006)
  - E. Cancès, F. Legoll and G. Stoltz, Theoretical and numerical comparison of some sampling methods, *Math. Model. Numer. Anal.* **41**(2) (2007) 351-390
  - T. Lelièvre, M. Rousset and G. Stoltz, *Free Energy Computations: A Mathematical Perspective* (Imperial College Press, 2010)

- J.N. Roux, S. Rodts and G. Stoltz, *Introduction à la physique statistique et à la physique quantique*, cours Ecole des Ponts (2009)
`http://cermics.enpc.fr/~stoltz/poly_phys_stat_quantique.pdf`

Some elements of statistical physics

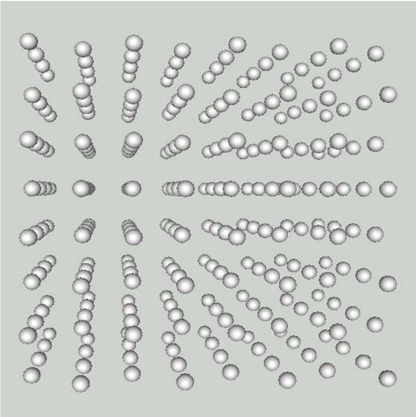
General perspective (1)

- Aims of computational statistical physics:
  - numerical microscope
  - computation of average properties, static or dynamic

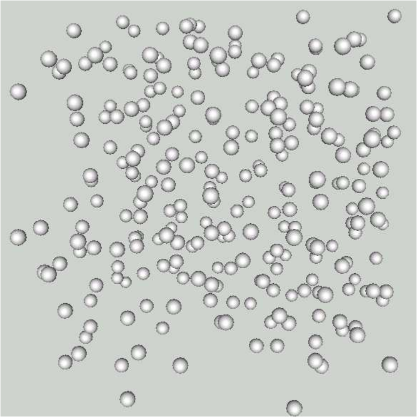
- Orders of magnitude
  - distances $\sim 1\ \mathring{A} = 10^{-10}$ m
  - energy per particle $\sim k_{\rm B}T \sim 4 \times 10^{-21}$ J at room temperature
  - atomic masses $\sim 10^{-26}$ kg
  - time $\sim 10^{-15}$ s
  - number of particles $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

- "Standard" simulations
  - 10^6 particles ["world records": around 10^9 particles]
  - integration time: (fraction of) ns ["world records": (fraction of) μs]

What is the melting temperature of argon?



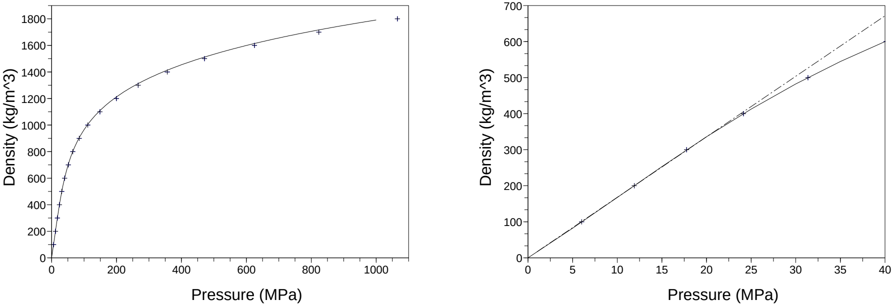
(a) Solid argon (low temperature)



(b) Liquid argon (high temperature)

General perspective (3)

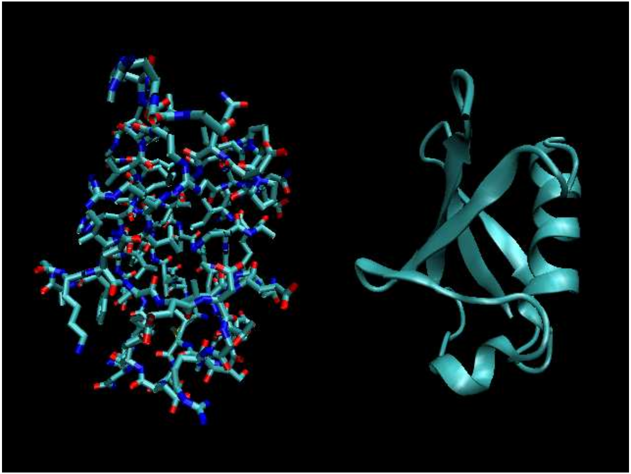
"Given the structure and the laws of interaction of the particles, what are the macroscopic properties of the matter composed of these particles?"



Equation of state (pressure/density diagram) for argon at $T = 300$ K

General perspective (4)

What is the structure of the protein? What are its typical conformations, and what are the transition pathways from one conformation to another?



Microscopic description of physical systems: unknowns

• Microstate of a classical system of N particles:

$$(q,p) = (q_1, \dots, q_N,\ p_1, \dots, p_N) \in \mathcal{E}$$

Positions q (configuration), momenta p (to be thought of as $M\dot{q}$)

• In the simplest cases, $\mathcal{E} = \mathcal{D} \times \mathbb{R}^{3N}$ with $\mathcal{D} = \mathbb{R}^{3N}$ or $\mathbb{T}^{3N}$

• More complicated situations can be considered: molecular constraints defining submanifolds of the phase space

• Hamiltonian $H(q,p) = E_{\mathrm{kin}}(p) + V(q)$, where the kinetic energy is

$$E_{\mathrm{kin}}(p) = \frac{1}{2}\, p^T M^{-1} p, \qquad M = \begin{pmatrix} m_1 \,\mathrm{Id}_3 & & 0 \\ & \ddots & \\ 0 & & m_N \,\mathrm{Id}_3 \end{pmatrix}.$$

Microscopic description: interaction laws

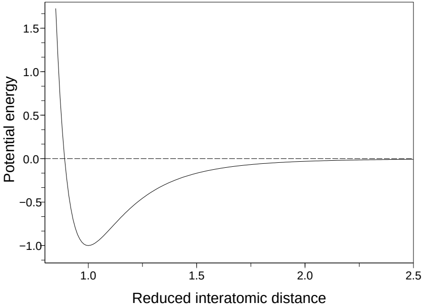
- All the physics is contained in V
  - ideally derived from quantum mechanical computations
  - in practice, empirical potentials for large scale calculations

- An example: Lennard-Jones pair interactions to describe noble gases

$$V(q_1,\dots,q_N) = \sum_{1 \leqslant i<j \leqslant N} v(|q_j - q_i|)$$

$$v(r) = 4\varepsilon \left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6} \right]$$

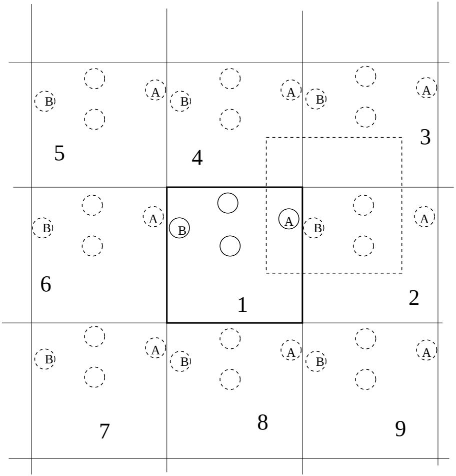
$$\text{Argon:} \quad \begin{cases} \sigma = 3.405 \times 10^{-10} \ \text{m} \\ \varepsilon / k_{\mathrm{B}} = 119.8 \ \ \text{K} \end{cases}$$



Microscopic description: boundary conditions

Various types of boundary conditions:

- Periodic boundary conditions: easiest way to mimick bulk conditions
- Systems *in vacuo* ($\mathcal{D} = \mathbb{R}^3$)
- Confined systems (specular reflection): large surface effects
- Stochastic boundary conditions (inflow/outflow of particles, energy, ...)



Thermodynamic ensembles (1)

- Macrostate of the system described by a probability measure

Equilibrium thermodynamic properties (pressure,...)

$$\langle A \rangle_\mu = \mathbb{E}_\mu(A) = \int_{\mathcal{E}} A(q,p)\,\mu(dq\,dp)$$

- Choice of thermodynamic ensemble
  - least biased measure compatible with the observed macroscopic data
  - Volume, energy, number of particles, ... fixed exactly or in average
  - Equivalence of ensembles (as $N \to +\infty$)

- Constraints satisfied in average: constrained maximisation of entropy

$$S(\rho) = -k_{\mathrm{B}} \int \rho \ln \rho \, d\lambda,$$

(λ reference measure), conditions $\rho \geqslant 0, \quad \int \rho\, d\lambda = 1, \quad \int A_i\, \rho\, d\lambda = \mathcal{A}_i$

Two examples: NVT, NPT ensembles

- Canonical ensemble = measure on (q,p), average energy fixed $A_0 = H$

$$\mu_{\rm NVT}(dq\,dp) = Z_{\rm NVT}^{-1}\,\mathrm{e}^{-\beta H(q,p)}\,dq\,dp$$

with β the Lagrange multiplier of the constraint $\displaystyle\int_{\mathcal{E}} H\,\rho\,dq\,dp = E_0$

- NPT ensemble = measure on (q,p,x) with $x \in (-1,+\infty)$
  - x indexes volume changes (fixed geometry): $\mathcal{D}_x = \Big((1+x)L\mathbb{T}\Big)^{3N}$
  - Fixed average energy and volume $\displaystyle\int (1+x)^3 L^3\,\rho\,\lambda(dq\,dp\,dx)$
  - Lagrange multiplier of the volume constraint: βP (pressure)

$$\mu_{\rm NPT}(dx\,dq\,dp) = Z_{\rm NPT}^{-1}\,\mathrm{e}^{-\beta P L^3(1+x)^3}\,\mathrm{e}^{-\beta H(q,p)}\,\mathbf{1}_{\{q\in[L(1+x)\mathbb{T}]^{3N}\}}\,dx\,dq\,dp$$

Observables

- May depend on the chosen ensemble! Given by physicists, by some analogy with macrosocpic, continuum thermodynamics
  - Pressure (derivative of the free energy with respect to volume)

$$A(q,p) = \frac{1}{3|\mathcal{D}|} \sum_{i=1}^{N} \left(\frac{p_i^2}{m_i} - q_i \cdot \nabla_{q_i} V(q) \right)$$

  - Kinetic temperature $A(q,p) = \frac{1}{3Nk_{\rm B}} \sum_{i=1}^{N} \frac{p_i^2}{m_i}$
  - Specific heat at constant volume: canonical average

$$C_V = \frac{\mathcal{N}_{\rm a}}{Nk_{\rm B}T^2} \left(\langle H^2 \rangle_{\rm NVT} - \langle H \rangle_{\rm NVT}^2 \right)$$

Main issue

Computation of high-dimensional integrals... Ergodic averages

- Also techniques to compute interesting trajectories (not presented here)

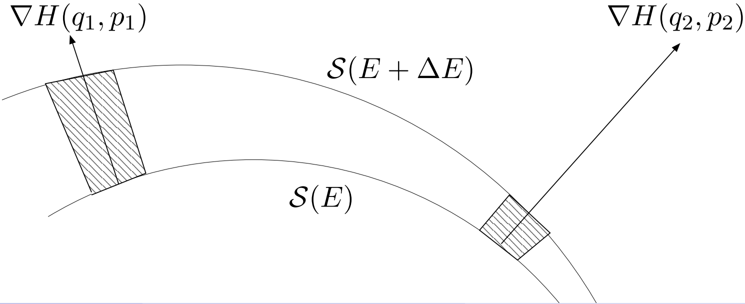
Sampling the microcanonical ensemble

The microcanonical measure

Lebesgue measure conditioned to $\mathcal{S}(E) = \left\{ (q,p) \in \mathcal{E} \,\middle|\, H(q,p) = E \right\}$ (co-area formula)

Microcanonical measure

$$\mu_{\mathrm{mc},E}(dq\,dp) = Z_E^{-1} \delta_{H(q,p)-E}(dq\,dp) = Z_E^{-1} \frac{\sigma_{\mathcal{S}(E)}(dq\,dp)}{|\nabla H(q,p)|}$$



The Hamiltonian dynamics

Hamiltonian dynamics

$$\begin{cases} \dfrac{dq(t)}{dt} = \nabla_p H(q(t),p(t)) = M^{-1}p(t) \\ \dfrac{dp(t)}{dt} = -\nabla_q H(q(t),p(t)) = -\nabla V(q(t)) \end{cases}$$

Assumed to be well-posed (*e.g.* when the energy is a Lyapunov function)

- Some simple properties (with ϕ_t the flow of the dynamics)
  - Preservation of energy $H \circ \phi_t = H$
  - Time-reversibility $\phi_{-t} = S \circ \phi_t \circ S$ where $S(q,p) = (q,-p)$
  - Symmetry $\phi_{-t} = \phi_t^{-1}$
  - Volume preservation $\displaystyle\int_{\phi_t(B)} dq\,dp = \int_B dq\,dp$

Invariance of the microcanonical measure

- Invariance by the Hamiltonian flow: proof using the co-area

$$
\begin{aligned}
\int_{\mathbb{R}} g(E) \int_{\mathcal{S}(E)} f(\phi_t(q,p))\, \delta_{H(q,p)-E}(dq\, dp)\, dE
&= \int_{\mathcal{E}} g(H(q,p))\, f(\phi_t(q,p))\, dq\, dp \\
&= \int_{\mathcal{E}} g(H(Q,P))\, f(Q,P))\, dQ\, dP \\
&= \int_{\mathbb{R}} g(E) \int_{\mathcal{S}(E)} f(q,p)\, \delta_{H(q,p)-E}(dq\, dp)\, dE
\end{aligned}
$$

- More intuitively with the limiting procedure $\Delta E \to 0$

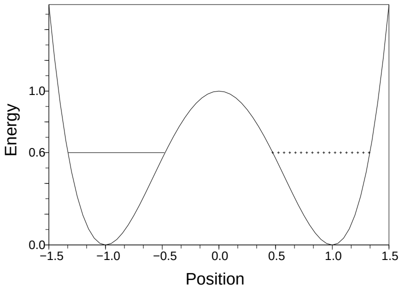
$$
\frac{1}{\Delta E} \int_{E \leqslant H \leqslant E+\Delta E} f = \frac{1}{\Delta E} \int_{E \leqslant H \leqslant E+\Delta E} f \circ \phi_t
$$

Ergodicity of the Hamiltonian dynamics

Ergodic **assumption**

$$\langle A \rangle_{\mathrm{NVE}} = \int_{\mathcal{S}(E)} A(q,p)\, \mu_{\mathrm{mc},E}(dq\, dp) = \lim_{T \to +\infty} \frac{1}{T} \int_0^T A(\phi_t(q,p))\, dt$$

- Wrong when spurious invariants are known, such as $\sum_{i=1}^{N} p_i$



Numerical approximation

- The ergodic assumption is true...
  - for completely integrable systems and perturbations thereof (KAM), upon conditioning the microcanonical measure by all invariants
  - if stochastic perturbations are considered[1]

$\rightarrow$ Although questionable, ergodic averages are the only realistic option

- Requires trajectories with good energy preservation over very long times

$\rightarrow$ disqualifies default schemes (Explicit/Implicit Euler, RK4, ...)

- Standard (simplest) estimator: integrator $(q^{n+1}, p^{n+1}) = \Phi_{\Delta t}(q^n, p^n)$

$$\langle A \rangle_{\mathrm{NVE}} \simeq \frac{1}{N_{\mathrm{iter}}} \sum_{n=1}^{N_{\mathrm{iter}}} A(q^n, p^n)$$

or refined estimators using some filtering strategy[2]

[1] E. Faou and T. Lelièvre, *Math. Comput.* **78**, 2047–2074 (2009)

[2] Cancès *et. al*, *J. Chem. Phys.*, 2004 and *Numer. Math.*, 2005

Longtime integration: failure of default schemes

Hamiltonian dynamics as a first-order differential equation

$$y = (q,p), \qquad \dot{y} = J\nabla H(y), \qquad J = \begin{pmatrix} 0 & I_{dN} \\ -I_{dN} & 0 \end{pmatrix}$$

- Analytical study of $\Phi_{\Delta t}$ for 1D harmonic potential $V(q) = \dfrac{1}{2}\omega^2 q^2$

$$\left\{ \begin{aligned} q^{n+1} &= q^n + \Delta t M^{-1} p^n, \\ p^{n+1} &= p^n - \Delta t \nabla V(q^n), \end{aligned} \right. \quad \text{so that } y^{n+1} = \begin{pmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 \end{pmatrix} y^n$$

Modulus of eigenvalues $|\lambda_\pm| = \sqrt{1+\omega^2\Delta t^2} > 1$, hence exponential increase of the energy

- For implicit Euler and Runge-Kutta 4 (for Δt small enough), exponential decrease of the energy

- Numerical confirmation for general (anharmonic) potentials

Longtime integration: symplecticity

- A mapping $g \; : \; U \text{ open} \to \mathbb{R}^{2dN}$ is symplectic when

$$[g'(q,p)]^T \cdot J \cdot g'(q,p) = J$$

- A mapping is symplectic if and only if it is (locally) Hamiltonian

Approximate longtime energy conservation

For an analytic Hamiltonian H and a symplectic method $\Phi_{\Delta t}$ of order p, and if the numerical trajectory remains in a compact subset, then there exists $h > 0$ and $\Delta t^* > 0$ such that, for $\Delta t \leqslant \Delta t^*$,

$$H(q^n, p^n) = H(q^0, p^0) + \mathrm{O}(\Delta t^p)$$

for exponentially long times $n\Delta t \leqslant \mathrm{e}^{h/\Delta t}$.

Weaker results under weaker assumptions[3]

[3] Hairer/Lubich/Wanner, Springer, 2006 and *Acta Numerica*, 2003

Longtime integration: construction of symplectic schemes

- Splitting strategy: decompose as $\left\{\begin{array}{ll}\dot{q} = & M^{-1}\,p,\\ \dot{p} = & 0,\end{array}\right.$ and $\left\{\begin{array}{ll}\dot{q} = & 0,\\ \dot{p} = & -\nabla V(q).\end{array}\right.$

- Flows $\phi_t^1(q,p) = (q + t\,M^{-1}p, p)$ and $\phi_t^2(q,p) = (q, p - t\nabla V(q))$

- Symplectic Euler A: first order scheme $\Phi_{\Delta t} = \phi_{\Delta t}^2 \circ \phi_{\Delta t}^1$

$$\left\{\begin{array}{l} q^{n+1} = q^n + \Delta t M^{-1}\,p^n \\ p^{n+1} = p^n - \Delta t \nabla V(q^{n+1}) \end{array}\right.$$

Composition of Hamiltonian flows hence symplectic

- Linear stability: harmonic potential $A(\Delta t) = \begin{pmatrix} 1 & \Delta t \\ -\omega^2 \Delta t & 1 - (\omega \Delta t)^2 \end{pmatrix}$

- Eigenvalues $|\lambda_\pm| = 1$ provided $\omega \Delta t < 2$

$\rightarrow$ time-step limited by the highest frequencies

Longtime integration: symmetrization of schemes[4]

- Strang splitting $\Phi_{\Delta t} = \phi^2_{\Delta t/2} \circ \phi^1_{\Delta t} \circ \phi^2_{\Delta t/2}$, second order scheme

Störmer-Verlet scheme

$$\begin{cases} p^{n+1/2} = p^n - \dfrac{\Delta t}{2} \nabla V(q^n) \\ q^{n+1} = q^n + \Delta t \ M^{-1} p^{n+1/2} \\ p^{n+1} = p^{n+1/2} - \dfrac{\Delta t}{2} \nabla V(q^{n+1}) \end{cases}$$

- Properties:
  - Symplectic, symmetric, time-reversible
  - One force evaluation per time-step, linear stability condition $\omega \Delta t < 2$
  - In fact, $M \dfrac{q^{n+1} - 2q^n + q^{n-1}}{\Delta t^2} = -\nabla V(q^n)$

[4] L. Verlet, *Phys. Rev.* **159**(1) (1967) 98-105

Some elements of backward error analysis

- Philosophy of backward analysis for EDOs: the numerical solution is...
  - an approximate solution of the exact dynamics $\dot{y} = f(y)$
  - the exact solution of a modified dynamics : $y^n = z(t_n)$

$\rightarrow$ properties of numerical scheme deduced from properties of $\dot{z} = f_{\Delta t}(z)$

Modified dynamics

$$\dot{z} = f_{\Delta t}(z) = f(z) + \Delta t F_1(z) + \Delta t^2 F_2(z) + ..., \qquad z(0) = y^0$$

- For Hamiltonian systems $(f(y) = J\nabla H(y))$ and symplectic scheme:

  Exact conservation of an approximate Hamiltonian $H_{\Delta t}$, hence approximate conservation of the exact Hamiltonian

- Harmonic oscillator: $H_{\Delta t}(q,p) = H(q,p) - \dfrac{(\omega \Delta t)^2 q^2}{4}$ for Verlet

General construction of the modified dynamics

- Iterative procedure (carried out up to an arbitrary truncation order)
- Taylor expansion of the solution of the modified dynamics

$$z(\Delta t) = z(0) + \Delta t \dot{z}(0) + \frac{\Delta t^2}{2} \ddot{z}(0) + \dots$$

with $\begin{cases} \dot{z}(0) = f(z(0)) + \Delta t F_1(z(0)) + \mathrm{O}(\Delta t^2) \\ \ddot{z}(0) = \partial_z f(z(0)) \cdot f(z(0)) + \mathrm{O}(\Delta t) \end{cases}$

Modified dynamics: first order correction

$$z(\Delta t) = y^0 + \Delta t\, f(y^0) + \Delta t^2 \left(F_1(y^0) + \frac{1}{2} \partial_z f(y^0) f(y^0) \right) + \mathrm{O}(\Delta t^3)$$

- To be compared to $y^1 = \Phi_{\Delta t}(y^0) = y^0 + \Delta t f(y^0) + \dots$

Some examples

- Explicit Euler $y^1 = y^0 + \Delta t\, f(y^0)$: the correction is not Hamiltonian

$$F_1(z) = -\frac{1}{2}\partial_z f(z) f(z) = \frac{1}{2}\begin{pmatrix} M^{-1}\nabla_q V(q) \\ \nabla_q^2 V(q)\cdot M^{-1}p \end{pmatrix} \neq \begin{pmatrix} \nabla_p H_1 \\ -\nabla_q H_1 \end{pmatrix}$$

- Symplectic Euler A

$$\left\{ \begin{aligned} q^{n+1} &= q^n + \Delta t M^{-1}\, p^n, \\ p^{n+1} &= p^n - \Delta t \nabla_q V(q^n) - \Delta t \nabla_q^2 V(q^n) M^{-1} p^n + \mathrm{O}(\Delta t^3) \end{aligned} \right.$$

The correction derives from the Hamiltonian $H_1(q,p) = \frac{1}{2} p^T M^{-1} \nabla_q V(q)$

$$F_1(q,p) = \frac{1}{2}\begin{pmatrix} M^{-1}\nabla_q V(q) \\ -\nabla_q^2 V(q)\cdot M^{-1}p \end{pmatrix} = \begin{pmatrix} \nabla_p H_1(q,p) \\ -\nabla_q H_1(q,p) \end{pmatrix}$$

Energy $H + \Delta t H_1$ preserved at order 2, while H preserved only at order 1

Sampling the canonical ensemble

Classification of the methods

- Computation of $\langle A \rangle = \int_{\mathcal{E}} A(q,p)\,\mu(dq\,dp)$ with

$$\mu(dq\,dp) = Z_\mu^{-1} \mathrm{e}^{-\beta H(q,p)}\,dq\,dp, \qquad \beta = \frac{1}{k_{\mathrm{B}}T}$$

- Actual issue: sampling canonical measure on configurational space

$$\nu(dq) = Z_\nu^{-1} \mathrm{e}^{-\beta V(q)}\,dq$$

- Several strategies (theoretical and numerical comparison[5])
  - Purely stochastic methods (i.i.d sample) $\to$ impossible...
  - Markov chain methods
  - Stochastic differential equations
  - Deterministic methods *à la* Nosé-Hoover

In practice, no clear-cut distinction due to blending...

[5] E. Cancès, F. Legoll and G. Stoltz, *M2AN*, 2007

Outline

- **Markov chain methods**
  - Metropolis-Hastings algorithm
  - (Generalized) Hybrid Monte Carlo

- **Stochastic differential approaches**
  - General perspective (convergence results, ...)
  - Overdamped Langevin dynamics (Einstein-Schmolukowski)
  - Langevin dynamics
  - Extensions: DPD, Generalized Langevin

- **Deterministic methods**
  - Nosé-Hoover and the like
  - Nosé-Hoover Langevin

- **Sampling constraints in average**
  - A first example of a nonlinear dynamics

Metropolis-Hastings algorithm (1)

- Markov chain method[6,7], on position space

  - Given q^n, propose $\tilde{q}^{n+1}$ according to transition probability $T(q^n, \tilde{q})$
  - Accept the proposition with probability

$$\min\left(1, \frac{T(\tilde{q}^{n+1}, q^n)\,\nu(\tilde{q}^{n+1})}{T(q^n, \tilde{q}^{n+1})\,\nu(q^n)}\right),$$

  and set in this case $q^{n+1} = \tilde{q}^{n+1}$; otherwise, set $q^{n+1} = q^n$.

- Example of proposals
  - Gaussian displacement $\tilde{q}^{n+1} = q^n + \sigma\, G^n$ with $G^n \sim \mathcal{N}(0, \mathrm{Id})$
  - Biased random walk[8,9] $\tilde{q}^{n+1} = q^n - \alpha \nabla V(q^n) + \sqrt{\frac{2\alpha}{\beta}}\, G^n$

---

[6] Metropolis, Rosenbluth ($\times 2$), Teller ($\times 2$), *J. Chem. Phys.* (1953)
[7] W. K. Hastings, *Biometrika* (1970)
[8] G. Roberts and R.L. Tweedie, *Bernoulli* (1996)
[9] P.J. Rossky, J.D. Doll and H.L. Friedman, *J. Chem. Phys.* (1978)

Metropolis-Hastings algorithm (2)

- Transition kernel

$$P(q, dq') = \min\Big(1, r(q,q')\Big) T(q,q')\, dq' + \Big(1 - \alpha(q)\Big)\delta_q(dq'),$$

where $\alpha(q) \in [0,1]$ is the probability to accept a move starting from q:

$$\alpha(q) = \int_{\mathcal{D}} \min\Big(1, r(q,q')\Big) T(q,q')\, dq'.$$

- The canonical measure is reversible with respect to ν, hence invariant:

$$P(q, dq')\nu(dq) = P(q', dq)\nu(dq')$$

- Irreducibility: for almost all q_0 and any set A of positive measure, there exists n_0 such that, for $n \geqslant n_0$,

$$P^n(q_0, A) = \int_{x \in \mathcal{D}} P(q_0, dx)\, P^{n-1}(x, A) > 0$$

- Pathwise ergodicity[10] $\displaystyle \lim_{N \to +\infty} \frac{1}{N} \sum_{n=1}^{N} A(q^n) = \int_{\mathcal{D}} A(q)\, \nu(dq)$

[10] S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability* (1993)

Metropolis-Hastings algorithm (3)

- Central limit theorem for Markov chains under additional assumptions:

$$\sqrt{N}\left|\frac{1}{N}\sum_{n=1}^{N} A(q^n) - \int_{\mathcal{D}} A(q)\,\nu(dq)\right| \xrightarrow[N\to+\infty]{\text{law}} \mathcal{N}(0,\sigma^2)$$

- The asymptotic variance σ^2 takes into account the correlations:

$$\sigma^2 = \mathrm{Var}_\nu(A) + 2\sum_{n=1}^{+\infty} \mathbb{E}_\nu\Big[\big(A(q^0)-\mathbb{E}_\nu(A)\big)\big(A(q^n)-\mathbb{E}_\nu(A)\big)\Big]$$

- Numerical efficiency: trade-off between acceptance and sufficiently large moves in space to reduce autocorrelation (rejection rate around[11] 0.5)

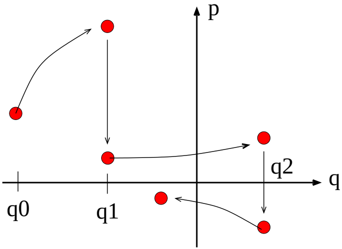
- Refined Monte Carlo moves such as parallel tempering/replica exchanges

- A way to stabilize discretization schemes for SDEs

[11] See B. Jourdain's talk...

(Generalized) Hybrid Monte Carlo (1)

- Markov chain in the configuration space[12,13], parameters: τ and Δt
  - generate momenta p^n according to $Z_p^{-1}\, \mathrm{e}^{-\beta p^2/2m} dp$
  - compute (an approximation of) the flow $\Phi_\tau(q^n, p^n) = (\tilde{q}^{n+1}, \tilde{p}^{n+1})$ of the Hamiltonian dynamics
  - accept $\tilde{q}^{n+1}$ and set $q^{n+1} = \tilde{q}^{n+1}$ with probability $\min\left(1, \mathrm{e}^{-\beta(\widetilde{E}^{n+1} - E_n)}\right)$; otherwise set $q^{n+1} = q^n$.



- Extensions: correlated momenta, random times τ, constraints, ...

- Ergodicity is an issue (harmonic case with $\tau =$ period): can be proved for potentials bounded above and ∇V globally Lipschitz[14]

---

[12] S. Duane, A. Kennedy, B. Pendleton and D. Roweth, *Phys. Lett. B* (1987)

[13] Ch. Schütte, *Habilitation Thesis* (1999)

[14] E. Cancès, F. Legoll et G. Stoltz, *M2AN* (2007)

(Generalized) Hybrid Monte Carlo (2)

- Transformation $S = S^{-1}$ leaving $\pi(dx)$ invariant, *e.g.* $S(q,p) = (q,-p)$
- Assume that $r(x,x') = \dfrac{T(S(x'),S(dx))\,\pi(dx')}{T(x,dx')\,\pi(dx)}$ is defined and positive

Generalized Hybrid Monte Carlo

- given x^n, propose a new state $\tilde{x}^{n+1}$ from x^n according to $T(x^n,\cdot)$;
- accept the move with probability $\min\left(1, r(x^n,\tilde{x}^{n+1})\right)$, and set in this case $x^{n+1} = \tilde{x}^{n+1}$; otherwise, set $x^{n+1} = S(x^n)$.

- Reversibility up to S, *i.e.* $P(x,dx')\,\pi(dx) = P(S(x'),S(dx))\,\pi(dx')$
- Standard HMC: $T(q,dq') = \delta_{\Phi_\tau(q)}(dq')$, momentum reversal upon rejection (not important since momenta are resampled, but is important when momenta are partially resampled)

Generalities on SDEs (1)

• Consider $dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$, smooth drift and diffusion (not true in practice hence many open problems...)

• Configuration space $\mathcal{X}$, law $\psi(t,x)$ of X_t

• Generator $\mathcal{A} = b(x)\cdot\nabla + \frac{1}{2}\sigma\sigma^T(x) : \nabla^2$

• Fokker-Planck equation $\partial_t \psi = \mathcal{A}^* \psi$ (adjoint on $L^2(\mathcal{X})$)

• Invariant measure $\psi_\infty(x)\,dx$ solution of $\mathcal{A}^*\psi_\infty = 0$

• Define $f = \psi/\psi_\infty$, then Fokker-Planck equation

$$\partial_t f = \mathcal{A}^* f$$

with adjoints on $L^2(\psi_\infty)$ defined as $\int_{\mathcal{X}} f\,(\mathcal{A}g)\,\psi_\infty = \int_{\mathcal{X}} (\mathcal{A}^* f)\,g\,\psi_\infty$

• Reversibility: the paths $(x_t)_{t\in[0,T]}$ and $(x_{T-t})_{t\in[0,T]}$ have the same laws when $x_0 \sim \psi_\infty$, equivalent to $\mathcal{A}^* = \mathcal{A}$

Generalities on SDEs (2)

- Irreducibility: show that $P_t(x,A) = \mathbb{E}_x(X_t \in A) > 0$ when A is open (support theorem Stroock-Varadhan), proof based on controlled ODE

$$\dot{x}(t) = b(x(t)) + \sigma(x(t))\, u(t)$$

- Smoothness of the transition probabilities: Hypoellipticity[15]
  - Operator rewritten as $\mathcal{A} = X_0 + \sum_{i=1}^{M} X_i^* X_i$
  - Commutators $[S,T] = ST - TS$
  - If $\{X_i\}_{i=0,\dots,M}$, $\{[X_i,X_j]\}_{i,j=0,\dots,M}$, $\{[[X_i,X_j],X_k]\}_{i,j,k=0,\dots,M}$, ... has full rank at every point, then $\mathcal{A}$ is hypoelliptic on $\mathcal{X}$
  - If $\{X_i\}_{i=1,\dots,M}$, $\{[X_i,X_j]\}_{i,j=0,\dots,M}$, ... has full rank at every point, then $\partial_t - \mathcal{A}$ is hypoelliptic on $\mathbb{R} \times \mathcal{X}$

[15] L. Hörmander, *Acta Mathematica* (1967)

Generalities on SDEs (3)

- When $\partial_t - \mathcal{A}$ hypoelliptic: smooth transition probability $p(t,x,y)\,dy$

- Hypoellipticity is a local property: it does not imply uniqueness of the invariant measure[16] (requires irreducibility = global)

- Irreducibility and existence of invariant measure with density ψ_∞ gives uniqueness and

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \varphi(X_t)\,dt = \int \varphi(x)\,\psi_\infty(x)\,dx \qquad \text{a.s.}$$

- Rate of convergence given by Central Limit Theorem: $\widetilde{\varphi} = \varphi - \int \varphi\,\psi_\infty$

$$\sqrt{T}\left(\frac{1}{T}\int_0^T \varphi(X_t)\,dt - \int \varphi\,\psi_\infty\right) \xrightarrow[T\to+\infty]{\text{law}} \mathcal{N}(0,\sigma_\varphi^2)$$

with $\sigma_\varphi^2 = 2\,\mathbb{E}\left[\int_0^{+\infty} \widetilde{\varphi}(X_t)\widetilde{\varphi}(X_0)dt\right]$ (decay estimates/resolvent bounds)

[16] K. Ichihara and H. Kunita, *Z. Wahrscheinlichkeit* (1974)

Generalities on SDEs (4)

- Existence and uniqueness of ψ_∞: irreducibility, hypoellipticity and

Lyapunov condition

Function W with values in $[1,+\infty)$ such that

$$W(x) \xrightarrow[|x|\to+\infty]{} +\infty, \qquad \mathcal{A}W \leqslant -cW + b\,\mathbf{1}_K \quad (c>0,\ K\,\text{compact})$$

Useful when the invariant measure is not known (*e.g.* discretization)

$$\|\psi(t)-\psi_\infty\|_W \leqslant C\|\psi(0)-\psi_\infty\|_W\,\mathrm{e}^{-\lambda t}, \qquad \|\varphi\|_W = \sup_{x\in\mathcal{X}} \frac{|\varphi(x)|}{W(x)}$$

Proof via coupling argument[17] or spectral method[18]

- Rate of convergence not very explicit...
- More explicit rates: functional setting (ISL, hypocoercivity, ...)

---

[17] M. Hairer and J. Mattingly, *Progr. Probab.* (2011)
[18] L. Rey-Bellet, *Lecture Notes in Mathematics* (2006)

Generalities on SDEs: numerics (1)

- Numerical discretization: various schemes (Markov chains)

$$x^{n+1} = x^n + \Delta t\, b(x^n) + \sqrt{2\Delta t\, \sigma(x^n)}\, G^n, \qquad G^n \sim \mathcal{N}(0, \mathrm{Id})$$

- Ergodic for the probability measure $\psi_{\infty,\Delta t}$

- Estimator $\Phi_{N_{\rm iter}} = \dfrac{1}{N_{\rm iter}} \displaystyle\sum_{n=1}^{N_{\rm iter}} \varphi(x^n)$

- Errors $\sqrt{N_{\rm iter}} \left(\Phi_{N_{\rm iter}} - \displaystyle\int \varphi\, \psi_{\infty,\Delta t} \right) \xrightarrow[N_{\rm iter} \to +\infty]{\rm law} \mathcal{N}(0, \sigma^2_{\Delta t,\varphi})$
  - Statistical error: using a Central Limit Theorem
  - Systematic errors: perfect sampling bias and finite sampling bias

$$\left| \int \varphi\, \psi_{\infty,\Delta t} - \int \varphi\, \psi_\infty \right| \leqslant C_\varphi\, \Delta t^p$$

Numerical analysis of perfect sampling bias: Talay-Tubaro[19]

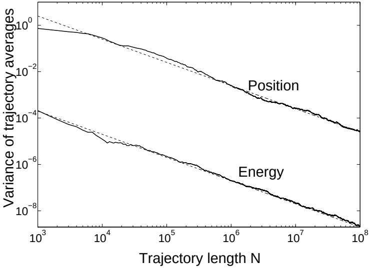
[19] D. Talay and L. Tubaro, *Stoch. Anal. Appl.* (1990)

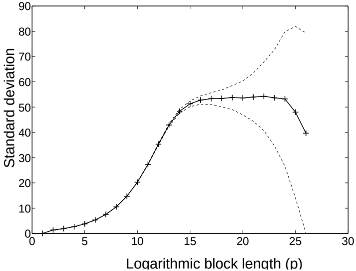
- Expression of the asymptotic variance: using $\widetilde{\varphi} = \varphi - \int \varphi\, \psi_{\infty,\Delta t}$

$$\sigma^2_{\Delta t,\varphi} = \mathrm{Var}(\varphi) + 2\sum_{n=1}^{+\infty} \mathbb{E}\Big(\widetilde{\varphi}(q^0,p^0)\widetilde{\varphi}(q^n,p^n)\Big) \sim \frac{2}{\Delta t}\mathbb{E}\left[\int_0^{+\infty} \widetilde{\varphi}(X_t)\widetilde{\varphi}(X_0)\,dt\right]$$

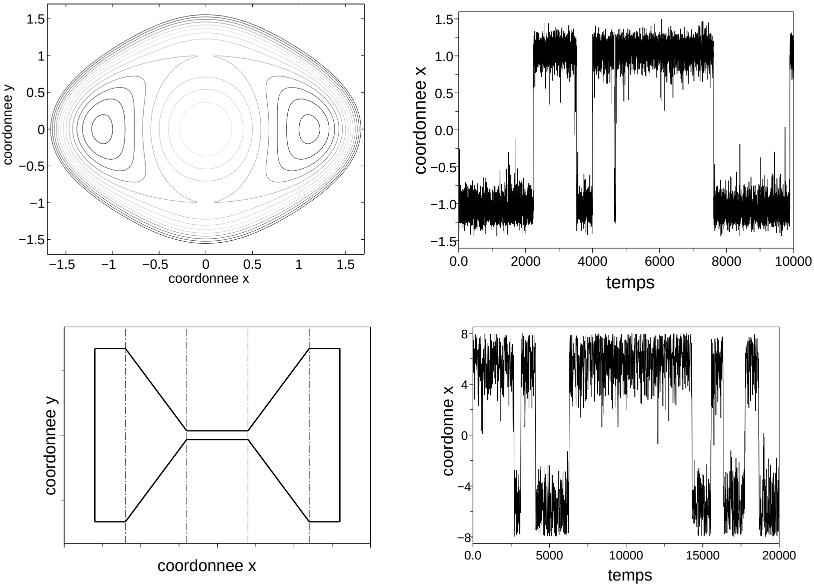
- Estimation of $\sigma_{\Delta t,\varphi}$ by block averaging (batch means)

$$\sigma^2_{\Delta t,\varphi} = \lim_{N,M\to+\infty} \frac{N}{M}\sum_{k=1}^{M}\left(\Phi^k_N - \Phi_{NM}\right)^2, \qquad \Phi^k_N = \frac{1}{N}\sum_{i=(k-1)N+1}^{kN} \varphi(q^i,p^i)$$





Metastability: large variances...



Overdamped Langevin dynamics

- SDE on the configurational part only (momenta trivial to sample)

$$dq_t = -\nabla V(q_t)\, dt + \sqrt{\frac{2}{\beta}} dW_t$$

- Invariance of the canonical measure $\nu(dq) = \psi_0(q)\, dq$

$$\psi_0(q) = Z^{-1}\, \mathrm{e}^{-\beta V(q)}, \qquad Z = \int_{\mathcal{D}} \mathrm{e}^{-\beta V(q)}\, dq$$

- Generator $\mathcal{A}_0 = -\nabla V(q) \cdot \nabla + \frac{1}{\beta}\Delta = \mathrm{div}\left(\psi_0 \nabla \left(\frac{\cdot}{\psi_0}\right)\right)$
  - self-adjoint on $L^2(\psi_0)$, hence reversibility
  - elliptic generator hence irreducibility and ergodicity
- Discretization $q^{n+1} = q^n - \Delta t\, \nabla V(q^n) + \sqrt{\frac{2\Delta t}{\beta}}\, G^n$ (+ Metropolization)

Overdamped Langevin dynamics: convergence

- Convergence of the law: $\|\psi(t,\cdot) - \psi_0\|_{\mathrm{TV}} \leqslant \sqrt{2\mathcal{H}(\psi(t,\cdot)\,|\,\psi_0)}$

$$\mathcal{H}(\psi(t,\cdot)\,|\,\psi_0) = \int_{\mathcal{D}} \ln\left(\frac{\psi(t,\cdot)}{\psi_0}\right)\psi(t,\cdot) \qquad \text{(relative entropy)}$$

- Decay in time $\frac{d}{dt}\mathcal{H}(\psi(t,\cdot)\,|\,\psi_0) = -\frac{1}{\beta} I(\psi(t,\cdot)\,|\,\psi_0)$ with

$$I(\psi(t,\cdot)\,|\,\psi_0) = \int_{\mathcal{D}} \left|\nabla \ln\left(\frac{\psi(t,\cdot)}{\psi_0}\right)\right|^2 \psi(t,\cdot) \qquad \text{(Fisher information)}$$

Logarithmic Sobolev Inequality for ψ_0 (metastability: small R)

$$\mathcal{H}(\phi\,|\,\psi_0) \leqslant \frac{1}{2R} I(\phi\,|\,\psi_0)$$

Gronwall: $\mathcal{H}(\psi(t)\,|\,\psi_0) \leqslant \mathcal{H}(\psi(0)\,|\,\psi_0)\exp(-2Rt/\beta)$

- Obtaining LSI? Bakry-Emery criterion (convexity), Gross (tensorization), Holley-Stroock's perturbation result

Langevin dynamics (1)

- Stochastic perturbation of the Hamiltonian dynamics

$$\begin{cases} dq_t = M^{-1} p_t \, dt \\ dp_t = -\nabla V(q_t)\, dt - \gamma M^{-1} p_t \, dt + \sigma \, dW_t \end{cases}$$

- Fluctuation/dissipation relation $\sigma \sigma^T = \frac{2}{\beta}\gamma$
- Reference space $L^2(\psi_0)$ where $\psi_0(q,p) = \mathrm{e}^{-\beta H(q,p)}$
- Generator $\mathcal{A}_0 = \mathcal{A}_{\mathrm{ham}} + \mathcal{A}_{\mathrm{thm}}$ with $\mathcal{A}_{\mathrm{ham}}^* = -\mathcal{A}_{\mathrm{ham}}$ and $\mathcal{A}_{\mathrm{thm}}^* = \mathcal{A}_{\mathrm{thm}}$

$$\mathcal{A}_{\mathrm{ham}} = \frac{p}{m} \cdot \nabla_q - \nabla V(q) \cdot \nabla_p,$$

$$\mathcal{A}_{\mathrm{thm}} = \gamma \left(-\frac{p}{m} \cdot \nabla_p + \frac{1}{\beta} \Delta_p \right) = -\frac{\gamma}{\beta} \sum_{i=1}^{N} (\partial_{p_i})^* \, \partial_{p_i}$$

- Invariance of the canonical measure: $\mathcal{A}_0^* \mathbf{1} = 0$

Langevin dynamics (2)

- Reversibility $\int_{\mathcal{E}} \mathcal{A}_0 f\, g\, \psi_0 = \int_{\mathcal{E}} (f \circ S)\mathcal{A}_0 (g \circ S)\, \psi_0$ for $S(q,p) = (q,-p)$

- Hypoellipticity: $[\partial_{p_{\alpha i}}, \mathcal{A}_{\rm ham}] = \frac{1}{m}\partial_{q_{\alpha i}}$

- Irreducibility: for given initial conditions $(q_{\rm i}, p_{\rm i})$ and final condition $(q_{\rm f}, p_{\rm f})$, consider any (smooth) path $\{Q(s)\}_{0 \leqslant s \leqslant t}$ such that

$$\Big(Q(0), Q'(0)\Big) = \Big(q_{\rm i}, M^{-1}p_{\rm i}\Big), \qquad \Big(Q(t), Q'(t)\Big) = \Big(q_{\rm f}, M^{-1}p_{\rm f}\Big)$$

and $u(s) = \sqrt{\frac{\beta}{2\gamma}}\left(\ddot{Q}(s) + \nabla V(Q(s)) + \gamma M^{-1}\dot{Q}(s)\right)$

- Conclusion: ψ_0 is the unique invariant probability measure and

$$\lim_{T\to+\infty} \frac{1}{T}\int_0^T \varphi(q_t,p_t)\,dt = \int_{\mathcal{E}} \varphi(q,p)\,\psi_0(q,p)\,dq\,dp \qquad \text{a.s.}$$

Langevin dynamics (3)

- Rate of convergence?

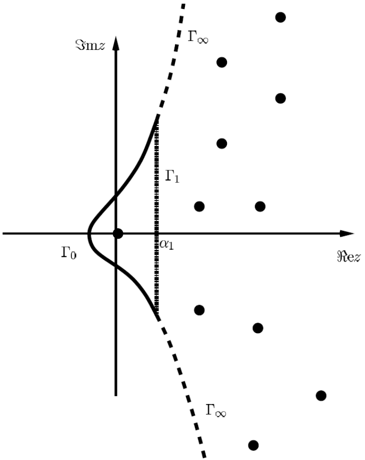
Hypocoercivity[a,b,c,d,e] results on

$$\mathcal{H} = \left\{ f \in L^2(\psi_0) \,\middle|\, \int_{\mathcal{E}} f\psi_0 = 0 \right\}$$
$$= L^2(\psi_0) \cap \mathrm{Ker}\,(\mathcal{A}_0)^\perp$$

- Operator $\mathcal{A}_0 = X_0 - \sum_{i=1}^{M} X_i^* X_i$

with $X_0 = \mathcal{A}_{\mathrm{ham}}$, $X_i = \sqrt{\dfrac{\gamma}{\beta}}\,\partial_{p_i}$

- $\mathcal{A}_0^{-1}$ compact on $\mathcal{H}$



[a] D. Talay, *Markov Proc. Rel. Fields*, **8** (2002)
[b] J.-P. Eckmann and M. Hairer, *Commun. Math. Phys.*, **235** (2003)
[c] F. Hérau and F. Nier, *Arch. Ration. Mech. Anal.*, **171** (2004)
[d] C. Villani, *Trans. AMS* **950** (2009)
[e] G. Pavliotis and M. Hairer, *J. Stat. Phys.* **131** (2008)

Langevin dynamics (4)

- Basic hypocoercivity result: $C_i = [X_i, X_0]$ $(1 \leqslant i \leqslant M)$, assume
  - $X_0^* = -X_0$
  - (for $i,j \geqslant 1$) X_i and X_i^* commute with C_j, X_i commutes with X_j
  - appropriate commutator bounds
  - $\displaystyle\sum_{i=1}^M X_i^* X_i + \sum_{i=1}^M C_i^* C_i$ is coercive

Then time-decay of the semigroup $\left\| \mathrm{e}^{t\mathcal{A}_0} \right\|_{\mathcal{B}(H^1(\psi_0)\cap\mathcal{H})} \leqslant C\mathrm{e}^{-\lambda t}$

- The proof uses a scalar product involving mixed derivatives $(a \gg b \gg 1)$

$$\langle\langle u, v \rangle\rangle = a\,\langle u, v\rangle + \sum_{i=1}^M b\,\langle X_i u, X_i v\rangle + \langle X_i u, C_i v\rangle + \langle C_i u, X_i v\rangle + b\langle C_i u, C_i v\rangle$$

- Langevin: $C_i = \dfrac{1}{m}\partial_{q_i}$, coercivity by Poincaré inequality

Overdamped limit of the Langevin dynamics

- Either $M = \varepsilon \to 0$ (for $\gamma = 1$) or $\gamma = \dfrac{1}{\varepsilon} \to +\infty$ (for $m = 1$ and an appropriate time-rescaling $t \to t/\varepsilon$)

$$\begin{cases} dq_t^\varepsilon = v_t^\varepsilon \, dt \\ \varepsilon \, dv_t^\varepsilon = -\nabla V(q_t^\varepsilon) \, dt - v_t^\varepsilon \, dt + \sqrt{\dfrac{2}{\beta}} \, dW_t \end{cases}$$

- Limiting dynamics $dq_t^0 = -\nabla V(q_t^0) \, dt + \sqrt{\dfrac{2}{\beta}} \, dW_t$

- Convergence result: $\lim_{\varepsilon \to 0} \left(\sup_{0 \leqslant s \leqslant t} \| q_s^\varepsilon - q_s^0 \| \right) = 0$ (a.s.), relying on

$$q_t^\varepsilon - q_t^0 = v_0 \, \varepsilon \left(1 - \mathrm{e}^{-t/\varepsilon} \right) - \int_0^t \left(1 - \mathrm{e}^{-(t-r)/\varepsilon} \right) \left(\nabla V(q_r^\varepsilon) - \nabla V(q_r^0) \right) \, dr$$
$$+ \int_0^t \mathrm{e}^{-(t-r)/\varepsilon} \, \nabla V(q_r^0) \, dr - \sqrt{2} \int_0^t \mathrm{e}^{-(t-r)/\varepsilon} \, dW_r$$

Numerical integration of the Langevin dynamics (1)

- Many possible schemes... Some implicitness helps for convergence results on non-compact configuration spaces
- Splitting: Hamiltonian *vs.* fluctuation/dissipation ($\alpha_{\Delta t} = \mathrm{e}^{-\gamma M^{-1}\Delta t}$)

$$\begin{cases} \widetilde{p}^{n+1/2} = \alpha_{\Delta t/2} p^n + \sqrt{\dfrac{1-\alpha_{\Delta t}}{\beta} M}\, G^n, \\ p^{n+1/2} = \widetilde{p}^{n+1/2} - \dfrac{\Delta t}{2}\, \nabla V(q^n), \\ q^{n+1} = q^n + \Delta t\, M^{-1} p^{n+1/2}, \\ \widetilde{p}^{n+1} = p^{n+1/2} - \dfrac{\Delta t}{2}\, \nabla V(q^{n+1}), \\ p^{n+1} = \alpha_{\Delta t/2} \widetilde{p}^{n+1} + \sqrt{\dfrac{1-\alpha_{\Delta t}}{\beta} M}\, G^{n+1/2}, \end{cases}$$

- Compact state spaces: Lyapunov function $W(q,p) = 1 + |p|^s$ ($s \geqslant 2$)
- Metropolization using Generalized HMC (Verlet part): flip momenta!

Numerical integration of the Langevin dynamics (3)

- Evolution operator $P_{\Delta t} = \mathrm{e}^{\Delta t\, C/2}\mathrm{e}^{\Delta t\, B/2}\mathrm{e}^{\Delta t\, A}\mathrm{e}^{\Delta t\, B/2}\mathrm{e}^{\Delta t\, C/2}$ with

$$A = M^{-1}p \cdot \nabla_q, \quad B = -\nabla V(q) \cdot \nabla_p, \quad C = \gamma\left(-M^{-1}p \cdot \nabla_p + \frac{1}{\beta}\Delta_p\right)$$

- Existence of a unique invariant measure $\mu_{\Delta t}$ for compact position spaces

- Exact remainders for the expansion of the evolution operator

$$P_{\Delta t} = I + \Delta t \mathcal{A}_0 + \frac{\Delta t^2}{2}\mathcal{A}_0^2 + \Delta t^3 S_2 + \Delta t^4 R_{\Delta t,2} = I + \Delta t \mathcal{A}_0 + \Delta t^2 \widetilde{R}_{\Delta t,2}$$

Error estimates

For a smooth observable ψ,

$$\int_{\mathcal{E}} \psi\, d\mu_{\Delta t} = \int_{\mathcal{E}} \psi\, d\mu + \Delta t^2 \int_{\mathcal{E}} \psi\, f\, d\mu + \mathrm{O}_{\psi}(\Delta t^3)$$

with $f = -\left(\mathcal{A}_0^{-1}\right)^* S_2^* \mathbf{1}$ (use BCH formula)

Numerical integration of the Langevin dynamics (2)

- Elements of the proof: use $\displaystyle\int_{\mathcal{E}} (I - P_{\Delta t})\varphi \, d\mu_{\Delta t} = 0$,

$$\int_{\mathcal{E}} (I - P_{\Delta t})\varphi \cdot \left(1 + \Delta t^2 f\right) d\mu = -\Delta t^3 \int_{\mathcal{E}} \left[\mathcal{A}_0 \varphi \cdot f + S_2 \varphi\right] d\mu - \Delta t^4 \int_{\mathcal{E}} \left[R_{\Delta t,2}\varphi + \left(\widetilde{R}_{\Delta t,2}\varphi\right) f\right] d\mu$$

and consider $\varphi = Q_{\Delta t,2}\psi$ with $\dfrac{\mathrm{Id} - P_{\Delta t}}{\Delta t} Q_{\Delta t,2} = \mathrm{Id} + \Delta t^3 Z_{\Delta t,2}$

- The correction term can be numerically approximated as ($g = S_2^* \mathbf{1}$)

$$\int_{\mathcal{E}} \psi \, \left(\mathcal{A}_0^{-1}\right)^* g \, d\mu = -\int_0^{+\infty} \mathbb{E}\Big(\psi(q_t, p_t) g(q_0, p_0)\Big) dt \simeq \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t}\Big(\psi\left(q^{n+1}, p^{n+1}\right) g\left(q^0, p^0\right)\Big)$$

- Rate of convergence? ("Numerical" hypocoercivity?)

Some extensions (1)

- The Langevin dynamics is not Galilean invariant, hence not consistent with hydrodynamics $\to$ friction forces depending on relative velocities

Dissipative Particle Dynamics

$$\left\{ \begin{aligned} dq &= M^{-1} p_t \, dt \\ dp_{i,t} &= -\nabla_{q_i} V(q_t)\, dt + \sum_{i \neq j} \left(-\gamma \chi^2(r_{ij,t}) v_{ij,t} + \sqrt{\frac{2\gamma}{\beta}} \chi(r_{ij,t})\, dW_{ij} \right) \end{aligned} \right.$$

with $\gamma > 0$, $r_{ij} = |q_i - q_j|$, $v_{ij} = \frac{p_i}{m_i} - \frac{p_j}{m_j}$, $\chi \geqslant 0$, and $W_{ij} = -W_{ji}$

- Invariance of the canonical measure, preservation of $\displaystyle\sum_{i=1}^N p_i$
- Ergodicity is an issue[20]
- Numerical scheme: splitting strategy[21]

[20] T. Shardlow and Y. Yan, *Stoch. Dynam.* (2006)

[21] T. Shardlow, *SIAM J. Sci. Comput.* (2003)

Some extensions (2)

• Mori-Zwanzig derivation[22] from a generalized Hamiltonian system: particle coupled to harmonic oscillators with a distribution of frequencies

Generalized Langevin equation ($M = \mathrm{Id}$)

$$\begin{cases} dq = p_t \, dt \\ dp_t = -\nabla V(q_t) \, dt + R_t \, dt \\ \varepsilon \, dR_t = -R_t \, dt - \gamma p_t \, dt + \sqrt{\dfrac{2\gamma}{\beta}} \, dW_t \end{cases}$$

• Invariant measure $\Pi(q,p,R) = Z_{\gamma,\varepsilon}^{-1} \exp\left(-\beta\left[H(q,p) + \dfrac{\varepsilon}{2\gamma} R^2\right]\right)$

• Langevin equation recovered in the limit $\varepsilon \to 0$

• Ergodicity proofs (hypocoercivity): as for the Langevin equation[23]

---

[22] R. Kupferman, A. Stuart, J. Terry and P. Tupper, *Stoch. Dyn.* (2002)

[23] M. Ottobre and G. Pavliotis, *Nonlinearity* (2011)

Deterministic methods: Nosé-Hoover and the like (1)

EDO on extended phase space, additional parameter $Q > 0$

$$\begin{cases} \dot{q} = M^{-1}p \\ \dot{p} = -\nabla V(q) - \xi p \\ \dot{\xi} = Q^{-1}\left(p^T M^{-1} p - N k_B T\right) \end{cases}$$

- Invariant measure $\pi(dq\,dp\,d\xi) = Z_Q^{-1} \mathrm{e}^{-\beta H(q,p)} \mathrm{e}^{-\beta Q \xi^2/2}$
- Discretization: reversible schemes, or resort to Hamiltonian reformulation
- It converges fast (as $1/N_{\mathrm{iter}}$)... but maybe not to the correct value!
- Ergodicity is an issue!
  - Proofs of non-ergodicity in limiting regimes (KAM tori)[24]
  - Practical difficulties when heterogeneities (*e.g.* very different masses)

[24] F. Legoll, M. Luskin and R. Moeckel, *ARMA* (2007), *Nonlinearity* (2009)

Deterministic methods: Nosé-Hoover and the like (2)

• Various (unsatisfactory) remedies: Nosé-Hoover chains, massive Nosé-Hoover thermostatting, etc[25]

• A more serious remedy: add some stochasticity[26]

Langevin Nosé-Hoover

$$\begin{cases} dq_t = M^{-1} p_t \, dt \\ dp_t = (-\nabla V(q_t) - \xi_t p_t) \, dt \\ d\xi_t = \left[Q^{-1} \left(p_t^T M^{-1} p_t - \dfrac{N}{\beta} \right) - \gamma \right] dt + \sqrt{\dfrac{2\gamma}{\beta Q}} \, dW_t \end{cases}$$

Ergodic for the measure π (hypoellipticity + existence of invariant probability measure)

[25] M. Tuckerman, *Statistical Mechanics:...* (2010)

[26] B. Leimkuhler, N. Noorizadeh and F. Theil, *J. Stat. Phys.* (2009)

Sampling constraints in average (1)

- Set some external parameter (temperature, pressure/volume) to obtain the correct value of a given thermodynamic property

- Example of external parameter: temperature T in the canonical ensemble $\mu_T(dq\,dp) = Z^{-1}\mathrm{e}^{-H(q,p)/(k_{\mathrm{B}}T)}$

Formulation of the problem

Given an observable A and $\mathscr{A} \in \mathbb{R}$, find T such that

$$\langle A \rangle_T = \mathbb{E}_{\mu_T}(A) = \mathscr{A}$$

- Momenta are straightforward to sample: consider $A \equiv A(q)$

- Possible strategies
  - Newton method on T (accurate approximation of derivatives?)
  - New thermodynamic ensembles (physical meaning?)
  - Temperature as an additional variable + feedback mechanism[27]

[27] J.-B. Maillet and G. Stoltz, *Appl. Math. Res. Express* (2009)

Sampling constraints in average (2)

- Motivation: computation of Hugoniot curve = all admissible shocks

$$\mathscr{E} - \mathscr{E}_0 - \frac{1}{2}(\mathscr{P} + \mathscr{P}_0)(\mathscr{V}_0 - \mathscr{V}) = 0$$

- Statistical physics reformulation?
  - simulation cell $\mathcal{D}_c = \left(cL\mathbb{T} \times (L\mathbb{T})^2\right)^N$
  - Pole: reference temperature T_0 and volume with $c = 1$
  - vary the compression rate $c = |\mathcal{D}|/|\mathcal{D}_0|$

For a given compression $c_{\mathsf{max}} \leqslant c \leqslant 1$, find $T \equiv T(c)$ such that

$$\langle A_c \rangle_{|\mathcal{D}_c|,T} = 0$$

with $A_c(q,p) = H(q,p) - \langle H \rangle_{|\mathcal{D}_0|,T_0} + \frac{1}{2}\left(P_{xx,c}(q,p) + \langle P \rangle_{|\mathcal{D}_0|,T_0}\right)(1-c)|\mathcal{D}_0|$

where $P_{xx,c}(q,p) = \dfrac{1}{|\mathcal{D}_c|} \displaystyle\sum_{i=1}^{N} \frac{p_{i,x}^2}{m_i} - q_{i,x}\partial_{q_{i,x}} V(q)$

Sampling constraints in average (3)

- Assume that $\langle A \rangle_{T^*} = 0$ and locally $\alpha \leqslant \dfrac{\langle A \rangle_T - \langle A \rangle_{T^*}}{T - T^*} \leqslant a$
- The (deterministic) dynamics $T'(t) = -\gamma \langle A \rangle_{T(t)}$ is such that $T(t) \to T^*$
- Approximate the equilibrium canonical expectation by the current one:

$$\begin{cases} dq_t = -\nabla V(q_t)\, dt + \sqrt{2 k_{\rm B} T(t)}\, dW_t \\ T'(t) = -\gamma\, \mathbb{E}(A(q_t)) \end{cases}$$

- Consistency: (T^*, ν_{T^*}) is invariant (with $\nu_T(q) = Z_T^{-1} \mathrm{e}^{-V(q)/(k_{\rm B} T)}$)

Nonlinear PDE on the law $\psi(t,q)$ of the process q_t

$$\begin{cases} \partial_t \psi = k_{\rm B} T(t)\, \nabla \cdot \left[\nu_{T(t)} \nabla \left(\dfrac{\psi}{\nu_{T(t)}} \right) \right] = k_{\rm B} T(t)\, \Delta \psi + \nabla \cdot (\psi \nabla V), \\ T'(t) = -\gamma \displaystyle\int_{\mathcal{D}} A(q)\, \psi(t,q)\, dq \end{cases}$$

Sampling constraints in average (4)

Well-posedness (short time)

Assume A, V smooth enough, $T^0 > 0$ and $\psi^0 \in \mathrm{H}^2(\mathcal{D})$. Then there exists a unique solution $(T,\psi) \in \mathrm{C}^1([0,\tau],\mathbb{R}) \times \mathrm{C}^0([0,\tau],\mathrm{H}^2(\mathcal{D}))$ for a time

$$\tau \geqslant \frac{T^0}{2\gamma\|A\|_\infty} > 0$$

In particular, the temperature remains positive

Proof = Schauder fixed-point theorem using a mapping $T \mapsto \psi_T \mapsto g(T)$

- Longtime behavior? Convergence results for initial conditions close to the fixed-point

- Total entropy $\mathcal{E}(t) = E(t) + \frac{1}{2}(T(t) - T^*)^2$, where the reference measure in the spatial entropy is time-dependent:

$$E(t) = \int_{\mathcal{D}} \ln\left(\frac{\psi}{\nu_{T(t)}}\right) \psi$$

Sampling constraints in average (5)

- If $\mathcal{E}(t) \to 0$ then $T(t) \to T^*$ and $\psi \to \mu_{T^*}$
- It holds $E'(t) = -k_{\rm B} T(t) \displaystyle\int_{\mathcal{D}} \left| \nabla \ln \left(\frac{\psi}{\nu_{T(t)}} \right) \right|^2 \psi + \frac{T'(t)}{k_{\rm B} T(t)^2} \int_{\mathcal{D}} \dots \, \nu_{T(t)}$
- First term bounded by $-\rho E(t)$ using some LSI, remainder small when γ small enough (since $T'(t) \propto \gamma$)

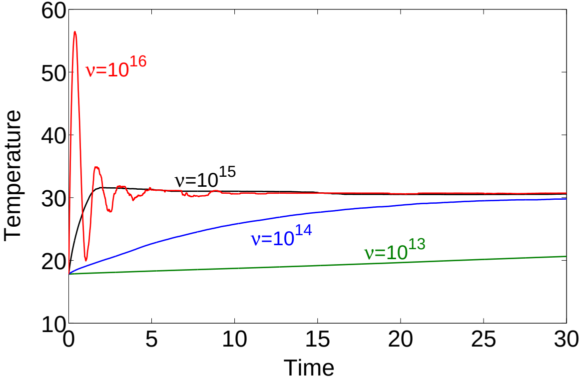
Convergence result

Consider (T^0, ψ^0) with $\psi^0 \in \mathrm{H}^2(\mathcal{D})$ such that $\mathcal{E}(0) \leqslant \mathcal{E}^*$ (depends on range of temperatures where LSI holds uniformly).
Then, for $\gamma \leqslant \gamma^*$, the solution is global in time and $\mathcal{E}(t) \leqslant \mathcal{E}(0) \exp(-\kappa t)$ for some $\kappa > 0$.
In particular, the temperature remains positive at all times, and it converges exponentially fast to T^*.

Rate of convergence larger when ρ larger (relaxation of the spatial distribution at a fixed temperature happens faster)

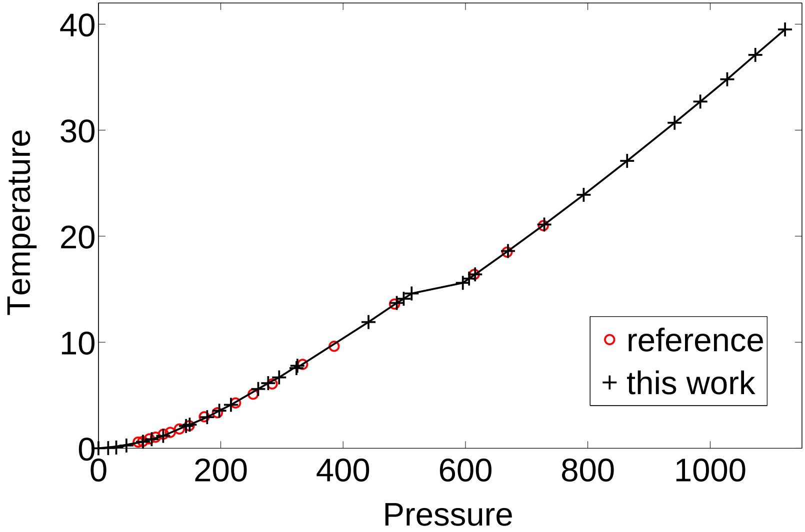
Sampling constraints in average (6)

- Time averages $dT_t = -\gamma \left(\dfrac{\displaystyle\int_0^t A(q_s)\, \delta_{T_t - T_s}\, ds}{\displaystyle\int_0^t \delta_{T_t - T_s}\, ds} \right) dt$



Hugoniot problem: fixed compression $c = 0.62$, pole $\rho_0 = 1.806 \times 10^3$ kg/m^3, $T_0 = 10$ K

Sampling constraints in average (7)



Hugoniot curve (reduced units)

Computation of free energy differences

Outline

- **Definition of (relative) free energies**
  - Thermodynamic definitions
  - Alchemical transitions *vs* reaction coordinates
  - Relation to metastability
- **Computational methods: based on...**
  - simple sampling methods (histogram methods, free energy perturbation)
  - constrained dynamics (thermodynamic integration)
  - nonequilibrium dynamics (Jarzynski equality)
  - adaptive biasing techniques (adaptive biasing force, Wang-Landau, ...)

What is free energy?

- A quantity of physical/chemical interest

Absolute free energy

$$F = -\frac{1}{\beta} \ln Z, \qquad Z = \int_{\mathcal{E}} \mathrm{e}^{-\beta H(q,p)} \, dq \, dp$$

- Motivation (Gibbs, 1902): Analogy with macroscopic thermodynamics

$$F = U - TS$$

energy $U = \int_{\mathcal{E}} H\psi$, entropy $S = -k_{\mathrm{B}} \int_{\mathcal{E}} \psi \ln \psi$ with $\psi = Z^{-1} \mathrm{e}^{-\beta H}$

- Can be analytically computed for ideal gases $(V = 0)$, and solids at low temperature

- Usually only free energy differences matter! (relative likelihood)

Free energy differences: The alchemical case

- Alchemical transition: indexed by an external parameter λ (force field parameter, magnetic field,...)

Alchemical free energy difference

$$F(1) - F(0) = -\beta^{-1} \ln \left(\frac{\displaystyle\int_{\mathcal{E}} \mathrm{e}^{-\beta H_1(q,p)} \, dq \, dp}{\displaystyle\int_{\mathcal{E}} \mathrm{e}^{-\beta H_0(q,p)} \, dq \, dp} \right)$$

- Typically, $H_\lambda = (1-\lambda) H_0 + \lambda H_1$

- Example: Widom insertion $\to$ chemical potential $\mu = F(1) - F(0)$

$$V_\lambda(q) = \sum_{1 \leqslant i < j \leqslant N} v(|q^i - q^j|) + \lambda \sum_{1 \leqslant i \leqslant N} v(|q^i - q^{N+1}|)$$

Free energy differences: The reaction coordinate case

- Reaction coordinate $\xi : \mathbb{R}^{3N} \to \mathbb{R}^m$ (angle, length,...)

- Foliation of the configurational space using level sets of ξ

$$\mathcal{D} = \bigcup_{z \in \mathbb{R}^m} \Sigma(z), \qquad \Sigma(z) = \left\{ q \in \mathcal{D} \,\middle|\, \xi(q) = z \right\}$$

Free energy difference: relative likelihood of marginals in ξ

$$F(z_1) - F(z_0) = -\beta^{-1} \ln \left(\frac{\displaystyle\int_{\Sigma(z)\times\mathbb{R}^{3N}} \mathrm{e}^{-\beta H(q,p)} \, \delta_{\xi(q)-z_1}(dq) \, dp}{\displaystyle\int_{\Sigma(z)\times\mathbb{R}^{3N}} \mathrm{e}^{-\beta H(q,p)} \, \delta_{\xi(q)-z_0}(dq) \, dp} \right).$$

with (as in the microcanonical case) $\delta_{\xi(q)-z}(dq) = \dfrac{\sigma_{\Sigma(z)}(dq)}{|\nabla\xi(q)|}$

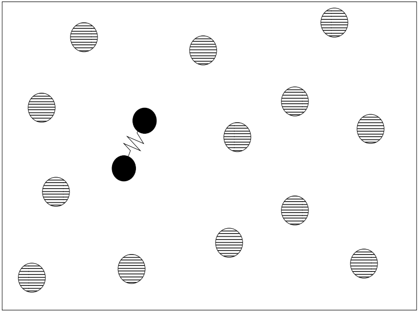
- Depends on the choice of ξ and not only on the foliation

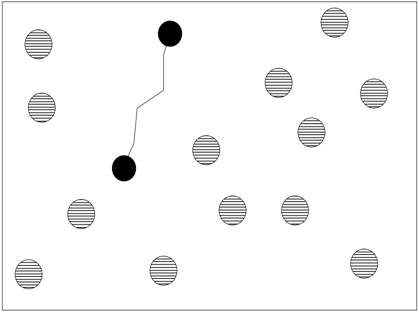
Free energy differences: The reaction coordinate case (2)

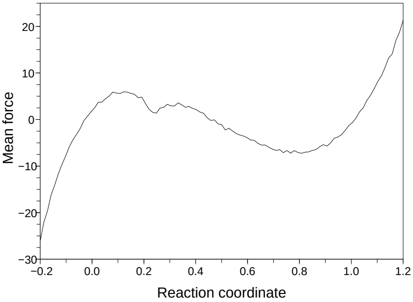
- Two particules (q_1,q_2), interaction $V_{\mathrm{S}}(r) = h\left[1 - \frac{(r-r_0-w)^2}{w^2}\right]^2$
- Solvent: purely repulsive potential $V_{\mathrm{WCA}}(r) = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^{6}\right] + \varepsilon$

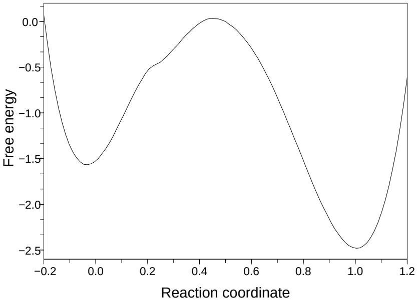
if $r \leqslant r_0$, and 0 for $r > r_0$

- Choose $\xi(q) = \frac{|q_1 - q_2| - r_0}{2w}$ (0 for compact, 1 for stretched)









Left: Estimated mean force $F'(z)$.
Right: Corresponding potential of mean force $F(z)$.

Parameters: $\beta = 1$, $N = 100$ particles, solvent density $\rho = 0.436$, WCA interactions $\sigma = 1$ and $\varepsilon = 1$, dimer $w = 2$ and $h = 2$.

Another view on free energy: Remove metastability (1)

- Remove metastability: uniform distribution of ξ under $\propto \mathrm{e}^{-\beta(V-F\circ\xi)}$

$\rightarrow$ Application to other fields, such as Bayesian statistics

- Data set $\{y_n\}_{n=1,\dots,N_{\mathrm{data}}}$ approximated by mixture of K Gaussians

$$f(y\,|\,\theta) = \sum_{i=1}^{K} q_i \sqrt{\frac{\lambda_i}{2\pi}} \exp\left(-\frac{\lambda_i}{2}(y-\mu_i)^2\right)$$

- Parameters $\theta = (q_1,\dots,q_{K-1},\mu_1,\dots,\mu_K,\lambda_1,\dots,\lambda_K)$ with

$$\mu_i \in \mathbb{R}, \qquad \lambda_i \geqslant 0, \quad 0 \leqslant q_i \leqslant 1, \qquad \sum_{i=1}^{K-1} q_i \leqslant 1$$

- Prior distribution $p(\theta)$: Random beta model[28,29]

Aim

Find the values of the parameters (namely θ, and possibly K as well) describing correctly the data

[28] S. Richardson and P. J. Green. *J. Roy. Stat. Soc. B*, 1997

[29] A. Jasra, C. Holmes and D. Stephens, *Statist. Science*, 2005

Another view on free energy: Remove metastability (2)

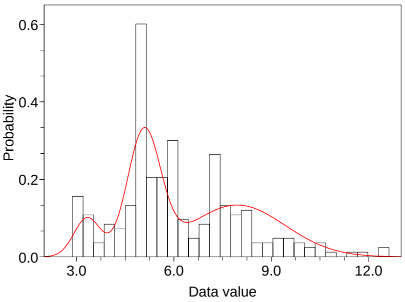
Prior distribution: additional variable $\beta \sim \Gamma(g,h)$

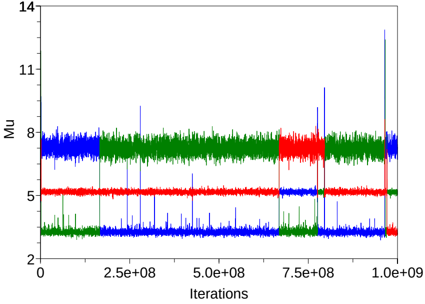
- uniform distribution of the weights q_i
- $\mu_k \sim \mathcal{N}\left(M, \dfrac{R^2}{4}\right)$ with M = mean of data, $R = \max - \min$
- $\lambda_k \sim \Gamma(\alpha,\beta)$ with $g = 0.2$ and $h = 100g/\alpha R^2$

Posterior density $\pi(\theta) = \dfrac{1}{Z_K} p(\theta) \displaystyle\prod_{n=1}^{N_{\rm data}} f(y_n \,|\, \theta)$

- Initial conditions: equal weights, means and variances for the Gaussians
- Metropolis random walk with (anisotropic) Gaussian proposals
- Metastability: at least $K! - 1$ symmetric replicates of any mode, but there may be additional metastable states
- Metastability increased when $N_{\rm data}$ increases

Another view on free energy: Remove metastability (3)



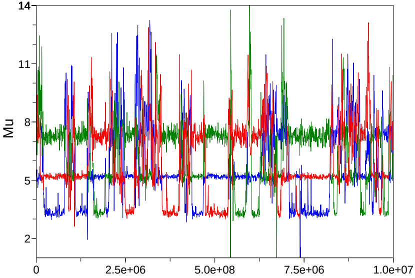


Left: Lengths of snappers ("Fish data"), $N_{\text{data}} = 256$, and a possible fit for $K = 3$ (last configuration from the trajectory)

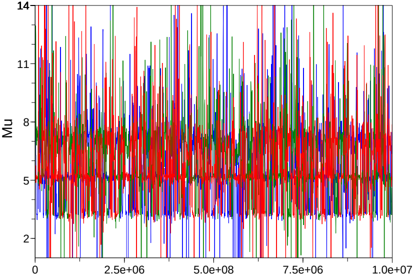
Right: Typical sampling trajectory, gaussian random walk with $(\sigma_q, \sigma_\mu, \sigma_v, \sigma_\beta) = (0.0005, 0.025, 0.05, 0.005)$.

[IS88] A. J. Izenman and C. J. Sommer, *J. Am. Stat. Assoc.*, 1988.
[BMY97] K. Basford *et al.*, *J. Appl. Stat.*, 1997

Another view on free energy: Remove metastability (4)



$\xi = \beta$



$\xi = V$

- Sampling of $\pi_F(\theta) \propto \pi(\theta)\, \mathrm{e}^{F(\xi(\theta))}$ with F free energy associated with ξ
- Choice of ξ? Computation of F? Efficiency of the reweighting?[30]

$$\mathbb{E}_\pi(\varphi) = \frac{\mathbb{E}_{\pi_F}\Big(\varphi \exp\left\{-F \circ \xi\right\}\Big)}{\mathbb{E}_{\pi_F}\Big(\exp\left\{-F \circ \xi\right\}\Big)}$$

[30] N. Chopin, T. Lelièvre and G. Stoltz, *Statist. Comput.*, 2012

Classification of available methods

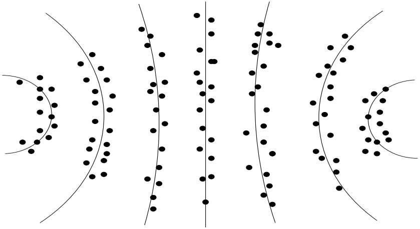
- Increasing order of mathematical complexity

Free energy perturbation	$\rightarrow$	Homogeneous MCs and SDEs
Histogram methods	$\rightarrow$	Homogeneous MCs and SDEs
Thermodynamic integration	$\rightarrow$	Projected MCs and SDEs
Nonequilibrium dynamics	$\rightarrow$	Nonhomogenous MCs and SDEs
Adaptive dynamics	$\rightarrow$	Nonlinear SDEs and MCs

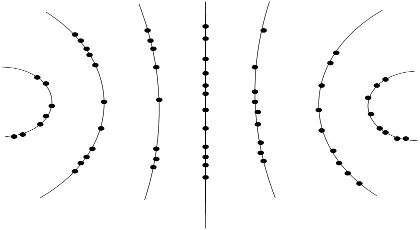
- On top of that: selection procedures can be added $\rightarrow$ particle systems and jump processes

- Questions:
  - Consistency (convergence)
  - Efficiency (error estimates = rate of convergence)

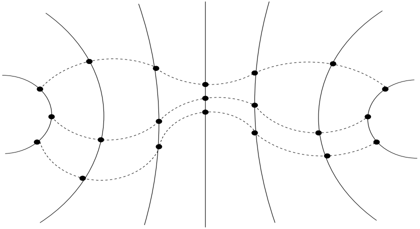
A cartoon comparison of available methods



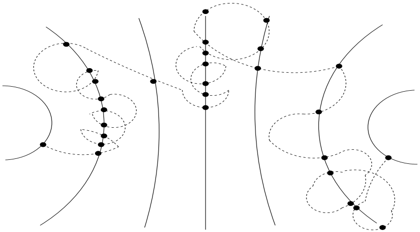
(a) Histogram methods



(b) Thermodynamic integration



(c) Nonequilibrium dynamics



(d) Adaptive dynamics

Free energy perturbation (1)

- Alchemical case only! Express ΔF as an average[31]

$$F(\lambda) - F(0) = -\beta^{-1} \ln \frac{\displaystyle\int_{\mathcal{E}} \mathrm{e}^{-\beta(H_\lambda(q,p) - H_0(q,p))} \, \mu_0(dq\,dp)}{\displaystyle\int_{\mathcal{E}} \mu_0(dq\,dp)}$$

with $\mu_0(dq\,dp) = Z^{-1} \mathrm{e}^{-\beta H_0(q,p)} \, dq\,dp$

- All usual sampling techniques can be used to sample from μ_0

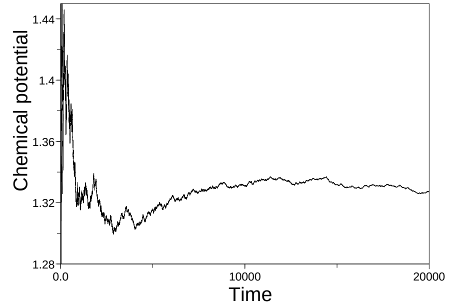
- Simplest estimator

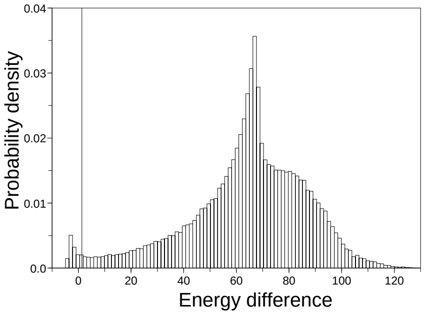
$$\widehat{\Delta F}_M = -\frac{1}{\beta} \ln \left(\frac{1}{M} \sum_{i=1}^{M} \mathrm{e}^{-\beta(H_1(q^i,p^i) - H_0(q^i,p^i))} \right), \qquad (q^i, p^i) \sim \mu_0$$

[31] Zwanzig, *J. Chem. Phys.* **22**, 1420 (1954)

Free energy perturbation (2)

Widom insertion. Left: Estimate of the chemical potential. Right: Distribution $P_0(dU)$ of insertion energies $U = H_1 - H_0$.





- The convergence is plagued by a very large variance... Remedies?

- Staging (stratification): $F(1) - F(0) = \sum_{i=1}^{I} F(\lambda_{i+1}) - F(\lambda_i)$

Free energy perturbation (3)

- Umbrella sampling[32] (importance sampling)

$$F(\lambda) - F(0) = -\beta^{-1} \ln \frac{\displaystyle\int_{\mathcal{E}} \mathrm{e}^{-\beta(H_\lambda - W)} d\mu_W}{\displaystyle\int_{\mathcal{E}} \mathrm{e}^{-\beta(H_0 - W)} d\mu_W}, \qquad \mu_W \propto \mu_0 \mathrm{e}^{-\beta W}$$

- Bridge sampling[33]: sample from the two distributions μ_0, μ_1 and optimize α to reduce the (asymptotic) variance

$$\frac{Z_1}{Z_0} = \frac{\displaystyle\int_{\mathcal{E}} \alpha\, \mathrm{e}^{-\beta H_1}\, d\mu_0}{\displaystyle\int_{\mathcal{E}} \alpha\, \mathrm{e}^{-\beta H_0}\, d\mu_1}, \quad \widehat{r}^{n_1,n_2} = \frac{\displaystyle\frac{1}{n_2}\sum_{j=1}^{n_2} \frac{f_1(x^{2,j})}{n_1 f_1(x^{2,j}) + n_2\, \widehat{r}^{n_1,n_2}\, f_2(x^{2,j})}}{\displaystyle\frac{1}{n_1}\sum_{j=1}^{n_1} \frac{f_2(x^{1,j})}{n_1 f_1(x^{1,j}) + n_2\, \widehat{r}^{n_1,n_2}\, f_2(x^{1,j})}}$$

[32] G.M. Torrie and J.P. Valleau, *J. Comp. Phys.* **23**, 187 (1977)

[33] C. Bennett, *J. Comput. Phys.* **22**, pp. 245–268 (1976)

Thermodynamic integration: Alchemical case

- Free energy = integral of an average force[34]

$$F(1) - F(0) = \int_0^1 F'(\lambda)\, d\lambda \simeq \sum_{i=1}^{M} (\lambda_i - \lambda_{i-1})\, F'(\lambda_i)$$

- Average force: computed by any method sampling the canonical measure

$$F'(\lambda) = \mathbb{E}_{\mu_\lambda}\left(\frac{\partial H_\lambda}{\partial \lambda}\right), \qquad \mu_\lambda(dq\, dp) = Z_\lambda^{-1} \mathrm{e}^{-\beta H_\lambda(q,p)}\, dq\, dp$$

- Optimization of the quadrature points to minimize the variance

- Extension to the case of reaction coordinates using projected SDEs, mean force = average Lagrange multiplier of the constraint[35]

[34] Kirkwood, *J. Chem. Phys.* **3**, 300 (1935)

[35] Ciccotti, Lelièvre, Vanden-Eijnden, Comm. Pure Appl. Math. (2008)

Thermodynamic integration: Constrained overdamped (1)

- Constrained configuration space $\Sigma(z) = \left\{ q \in \mathcal{D} \;\middle|\; \xi(q) = z \right\}$

Constrained overdamped Langevin dynamics

$$\begin{cases} dq_t = -\nabla V(q_t)\, dt + \sqrt{\frac{2}{\beta}} dW_t + \nabla \xi(q_t)\, d\lambda_t, \\ \xi(q_t) = z \end{cases}$$

- Ergodic and reversible for $\nu_{\Sigma(z)}(dq) = Z^{-1} \mathrm{e}^{-\beta V(q)}\, \sigma_{\Sigma(z)}(dq)$

$$F(z) = F_{\mathrm{rgd}}(z) - \beta^{-1} \ln \left(\int_{\Sigma(z)} (\det G)^{-1/2} d\nu_{\Sigma(z)} \right) + \mathrm{C},$$

with $\nabla F_{\mathrm{rgd}}(z) = \dfrac{\displaystyle\int_{\Sigma(z)} f_{\mathrm{rgd}} \exp(-\beta V)\, d\sigma_{\Sigma(z)}}{\displaystyle\int_{\Sigma(z)} \exp(-\beta V)\, d\sigma_{\Sigma(z)}}$ (complicated expression...)

Thermodynamic integration: Constrained overdamped (2)

- Numerical scheme (well-posed for Δt sufficiently small)

$$\begin{cases} q^{n+1} = q^n - \nabla V(q^n)\,\Delta t + \sqrt{\dfrac{2\Delta t}{\beta}}\,G^n + \lambda \nabla \xi(q^{n(+1)}), \\ \xi(q^{n+1}) = 0, \end{cases}$$

- Invariant measure $d\nu^{\Delta t}_{\Sigma(z)}(dq)$ with[36] $\left| \int_{\Sigma(z)} \varphi \, d\nu^{\Delta t}_{\Sigma(z)} - \int_{\Sigma(z)} \varphi \, d\nu_{\Sigma(z)} \right| \leqslant C \Delta t$

- Estimation of $\nabla F_{\rm rgd}$ using the Lagrange multipliers

$$\lim_{T\to\infty} \lim_{\Delta t \to 0} \frac{1}{M \Delta t} \sum_{n=1}^{M} \lambda^n = \nabla F_{\rm rgd}(z)$$

- Variance reduction (antithetic variables): use G^n and $-G^n$ and average Lagrange multipliers $\to$ removes the martingale part

[36] E. Faou and T. Lelièvre, *Math. Comput.* (2009)

Thermodynamic integration: Constrained Langevin (1)

Constrained Langevin dynamics

$$\begin{cases} dq_t = M^{-1} p_t \, dt, \\ dp_t = -\nabla V(q_t) \, dt - \gamma(q_t) M^{-1} p_t \, dt + \sigma(q_t) \, dW_t + \nabla \xi(q_t) \, d\lambda_t, \\ \xi(q_t) = z \end{cases}$$

- Standard fluctuation/dissipation relation $\sigma \sigma^T = \frac{2}{\beta} \gamma$
- Hidden velocity constraint: $\frac{d\xi(q_t)}{dt} = v_\xi(q_t, p_t) = \nabla \xi(q_t)^T M^{-1} p_t = 0$
- The corresponding phase-space is $\Sigma_{\xi, v_\xi}(z, 0)$ where

$$\Sigma_{\xi, v_\xi}(z, v_z) = \left\{ (q,p) \in \mathbb{R}^{6N} \,\middle|\, \xi(q) = z, \; v_\xi(q,p) = v_z \right\}$$

- An explicit expression of the Lagrange multiplier can be found by computing the second derivative in time of the constraint

Thermodynamic integration: Constrained Langevin (2)

Invariant measure

$$\mu_{\Sigma_{\xi,v_\xi}(z,0)}(dq\,dp) = Z_{z,0}^{-1}\,\mathrm{e}^{-\beta H(q,p)}\,\sigma_{\Sigma_{\xi,v_\xi}(z,0)}(dq\,dp)$$

with $\sigma_{\Sigma_{\xi,v_\xi}(z,v_z)}(dq\,dp)$ phase space Liouville measure induced by J

- Reversibility and detailed balance up to momentum reversal, ergodicity
- The free energy can be estimated from constrained samplings as

$$F(z) = F_{\rm rgd}^M(z) - \frac{1}{\beta}\ln\int_{\Sigma_{\xi,v_\xi}(z,0)} (\det \nabla\xi^T M^{-1}\nabla\xi)^{-1/2} d\mu_{\Sigma_{\xi,v_\xi}(z,0)} + \mathrm{C}$$

with rigid free energy $F_{\rm rgd}^M(z) = -\frac{1}{\beta}\ln\int_{\Sigma_{\xi,v_\xi}(z,0)} \mathrm{e}^{-\beta H(q,p)} d\mu_{\Sigma_{\xi,v_\xi}(z,0)}$

- Thermodynamic integration through the computation of the mean force

$$\nabla_z F_{\rm rgd}^M(z) = \int_{\Sigma_{\xi,v_\xi}(z,0)} f_{\rm rgd}^M(q,p)\,\mu_{\Sigma_{\xi,v_\xi}(z,0)}(dq\,dp)$$

Thermodynamic integration: Constrained Langevin (3)

- Splitting into Hamiltonian & constrained Ornstein-Uhlenbeck
- Midpoint scheme for momenta (reversible for constrained measure)

$$p^{n+1/4} = p^n - \frac{\Delta t}{4}\gamma\, M^{-1}(p^n + p^{n+1/4}) + \sqrt{\frac{\Delta t}{2}}\,\sigma\, G^n + \nabla\xi(q^n)\,\lambda^{n+1/4},$$

with the constraint $\nabla\xi(q^n)^T M^{-1} p^{n+1/4} = 0$

- RATTLE scheme (symplectic)

$$\begin{cases} p^{n+1/2} = p^{n+1/4} - \dfrac{\Delta t}{2}\nabla V(q^n) + \nabla\xi(q^n)\,\lambda^{n+1/2}, \\ q^{n+1} = q^n + \Delta t\, M^{-1}\, p^{n+1/2}, \\ p^{n+3/4} = p^{n+1/2} - \dfrac{\Delta t}{2}\nabla V(q^{n+1}) + \nabla\xi(q^{n+1})\,\lambda^{n+3/4}, \end{cases}$$

with $\xi(q^{n+1}) = z$ and $\nabla\xi(q^{n+1})^T M^{-1} p^{n+3/4} = 0$

- Overdamped limit obtained when $\dfrac{\Delta t}{4}\gamma = M \propto \mathrm{Id}$

Thermodynamic integration: Constrained Langevin (4)

- Metropolization of the RATTLE part to eliminate the time-step error in the sampled measure

- Longtime (a.s.) convergence (No second order derivatives of ξ needed)

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T d\lambda_t = \nabla_z F_{\mathrm{rgd}}^M(z)$$

- Variance reduction: keep only the Hamiltonian part of λ_t

- Numerical discretization: only Lagrange multipliers from RATTLE:

$$\nabla_z F_{\mathrm{rgd}}^M(z) \simeq \frac{1}{N} \sum_{n=0}^{N-1} f_{\mathrm{rgd}}^M(q^n, p^n) \simeq \frac{1}{N\Delta t} \sum_{n=0}^{N-1} (\lambda^{n+1/2} + \lambda^{n+3/4})$$

- Consistency result

$$\lambda^{n+1/2} + \lambda^{n+3/4} = \frac{\Delta t}{2} \Big(f_{\mathrm{rgd}}^M(q^n, p^{n+1/4}) + f_{\mathrm{rgd}}^M(q^{n+1}, p^{n+3/4}) \Big) + \mathrm{O}(\Delta t^3)$$

Nonequilibrium dynamics (1)

- Basic idea: switch from the initial to the final state in a finite time, starting from equilibrium, and reweight trajectories appropriately[37]

- Simplest possible setting: schedule $\Lambda(0) = 0, \Lambda(T) = 1$

$$\begin{cases} \dot{q}(t) = \nabla_p H_{\Lambda(t)}(q(t), p(t)) \\ \dot{p}(t) = -\nabla_q H_{\Lambda(t)}(q(t), p(t)) \end{cases}$$

- Work $\mathcal{W}(q,p) = \displaystyle\int_0^T \frac{\partial H_{\Lambda(t)}}{\partial \lambda}\big(\phi_t^\Lambda(q,p)\big)\Lambda'(t)\,dt = H_1\big(\phi_T^\Lambda(q,p)\big) - H_0(q,p)$

Jarzynski equality: exponential reweighting of the works

$$\mathbb{E}_{\mu_0}\left(\mathrm{e}^{-\beta\mathcal{W}}\right) = Z_0^{-1}\int_{\mathcal{E}} \mathrm{e}^{-\beta H_1(\phi_T^\Lambda(q,p))}\,dq\,dp = \frac{Z_1}{Z_0} = \mathrm{e}^{-\beta(F(1)-F(0))}$$

[37] C. Jarzynski, *Phys. Rev. Lett.* & *Phys. Rev. E* (1997)

Nonequilibrium dynamics (2)

- Generalization: $x = q$ or (q,p), invariant measure $\pi_t = \nu_{\Lambda(t)}$ or $\mu_{\Lambda(t)}$

$$\mathcal{L}_t = p^T M^{-1}\nabla_q - \nabla V_{\Lambda(t)} \cdot \nabla_p - \gamma p^T M^{-1}\nabla_p + \frac{\gamma}{\beta}\Delta_p \quad \text{(Langevin)}$$

- Work $\mathcal{W}_t\big(\{X_s\}_{0 \leqslant s \leqslant t}\big) = \displaystyle\int_0^t \frac{\partial E_{\Lambda(s)}}{\partial \lambda}(X_s)\dot{\Lambda}(s)\,ds$ (with $E_\lambda = V_\lambda$ or H_λ)
- Stochastic dynamics in the alchemical case: Feynman-Kac formula

$$P^w_{s,t}\varphi(x) = \mathbb{E}\left(\varphi(X_t)\,\mathrm{e}^{-\beta(\mathcal{W}_t - \mathcal{W}_s)}\,\middle|\, X_s = x\right)$$

satisfies the following backward Kolmogorov evolution

$$\partial_s P^w_{s,t} = -\mathcal{L}_s P^w_{s,t} + \beta \frac{\partial E_{\Lambda(s)}}{\partial \lambda}\,\dot{\Lambda}(s)\,P^w_{s,t}$$

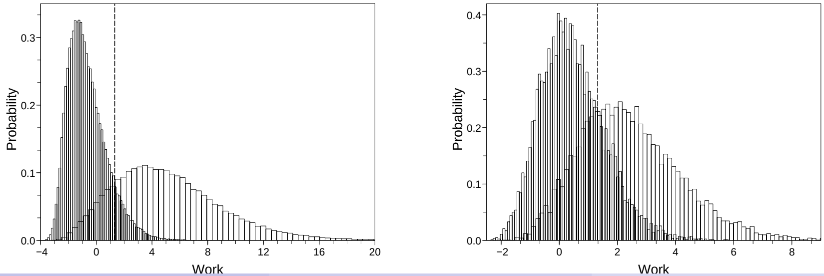
and recall that $X_0 \sim \pi_0$ (equilibrium initial conditions)

$$\frac{Z_t}{Z_0}\int \varphi\,d\pi_t = \mathbb{E}\Big(\varphi(X_t)\,\mathrm{e}^{-\beta\mathcal{W}_t}\Big)$$

Nonequilibrium dynamics (3)

- Mostly of theoretical interest: weight degeneracies (same as FEP)
- Free energy inequality $\mathbb{E}(\mathcal{W}_t) \geqslant F(\Lambda(t)) - F(0)$ (Jensen)
- Extensions...
  - Metropolis dynamics
  - Forward/backward versions (Crooks), path sampling, bridge estimators

$$\frac{Z_T}{Z_0}\,\mathbb{E}\left(\varphi^{\mathrm{r}}_{[0,T]}(X^{\mathrm{b}})\,\mathrm{e}^{-\beta\theta\mathcal{W}^{\mathrm{b}}_{0,T}}\right) = \mathbb{E}\left(\varphi_{[0,T]}(X^{\mathrm{f}})\,\mathrm{e}^{-\beta(1-\theta)\mathcal{W}^{\mathrm{f}}_{0,T}}\right)$$



Nonequilibrium dynamics (4)

- Reaction coordinate case: driven constrained processes[38]

$$\begin{cases} \quad dq_t = M^{-1} p_t \, dt \\ \quad dp_t = -\nabla V(q_t)\, dt - \gamma_P(q_t) M^{-1} p_t \, dt + \sigma_P(q_t)\, dW_t + \nabla \xi(q_t)\, d\lambda_t \\ \xi(q_t) = z(t) \end{cases}$$

with equilibrium initial conditions $(q_0, p_0) \sim \mu_{\Sigma_{\xi, v_\xi}(z(0), \dot{z}(0))}(dq\, dp)$

- Projected fluctuation/dissipation relation $(\sigma_P, \gamma_P) := (P_M\, \sigma, P_M\, \gamma\, P_M^T)$ so that the noise act only in the direction orthogonal to $\nabla \xi$
- Several expressions for work, *e.g.* $\mathcal{W}_{0,T}\left(\{q_t, p_t\}_{0 \leqslant t \leqslant T}\right) = \displaystyle\int_0^T \dot{z}(t)^T d\lambda_t$
- Free energy identity (corrector C to account for velocity constraints)

$$F(z(T)) - F(z(0)) = -\frac{1}{\beta} \ln \frac{\mathbb{E}\left(\mathrm{e}^{-\beta\left[\mathcal{W}_{0,T}\left(\{q_t,p_t\}_{t\in[0,T]}\right) + C(T,q_T)\right]}\right)}{\mathbb{E}\left(\mathrm{e}^{-\beta C(0,q_0)}\right)}$$

- Many extensions (path functionals, Crooks, discrete versions, ...)

[38] T. Lelièvre, M. Rousset and G. Stoltz, *Math. Comput.* (2012)

Adaptive biasing force (1)

- Simplified setting: $q = (x, y)$ and $\xi(q) = x \in \mathbb{R}$ so that

$$F(x_2) - F(x_1) = -\beta^{-1} \ln \left(\frac{\overline{\nu}(x_2)}{\overline{\nu}(x_1)} \right), \qquad \overline{\nu}(x) = \int \mathrm{e}^{-\beta V(x,y)} \, dy$$

- The mean force is $F'(x) = \dfrac{\displaystyle\int \partial_x V(x,y) \, \mathrm{e}^{-\beta V(x,y)} \, dy}{\displaystyle\int \mathrm{e}^{-\beta V(x,y)} \, dy}$

- The dynamics $dq_t = -\nabla V(q_t) \, dt + \sqrt{\frac{2}{\beta}} \, dW_t$ is metastable, contrarily to

$$\left\{ \begin{aligned} & dq_t = -\nabla \Big(V(q_t) - F(\xi(q_t)) \Big) \, dt + \sqrt{\frac{2}{\beta}} \, dW_t \\ & F'(x) = \mathbb{E}_\nu \Big(\partial_x V(q) \, \Big| \, \xi(q) = x \Big) = \mathbb{E}_{\widetilde{\nu}} \Big(\partial_x V(q) \, \Big| \, \xi(q) = x \Big) \end{aligned} \right.$$

where the last equality holds for any $\widetilde{\nu}(dq) \propto \nu(dq) \, g(x)$ (with $g \geqslant 0$)

Adaptive biasing force (2)

- Bias the dynamics by an approximation of F' computed on-the-fly

$\rightarrow$ Replace equilibrium expectations by $F'(t,x) = \mathbb{E}\Big(\partial_x V(q_t) \,\Big|\, \xi(q_t) = x\Big)$

ABF dynamics

$$\left\{ \begin{aligned} dq_t &= -\nabla\Big(V(q_t) - F_t(\xi(q_t))\Big)\, dt + \sqrt{\frac{2}{\beta}}\, dW_t \\ F_t'(x) &= \mathbb{E}\Big(\partial_x V(q) \,\Big|\, \xi(q_t) = x\Big) \end{aligned} \right.$$

- Reformulation as a nonlinear PDE on the law $\psi(t,q)$

$$\left\{ \begin{aligned} \partial_t \psi &= \mathrm{div}\Big[\nabla\big(V - F_{\mathrm{bias}}(t,x)\big)\,\psi + \beta^{-1}\nabla\psi\Big], \\ F_{\mathrm{bias}}'(t,x) &= \frac{\displaystyle\int \partial_x V(x,y)\,\psi(t,x,y)\,dy}{\displaystyle\int \psi(t,x,y)\,dy}. \end{aligned} \right.$$

Adaptive biasing force (3)

- Stationary solution $\psi_\infty \propto \mathrm{e}^{-\beta(V-F\circ\xi)}$

Convergence rate of ABF (the spirit of it)

Assume that

- the conditioned measures $\dfrac{\nu(x,y)}{\overline{\nu}(x)}\,dy$ satisfy LSI(ρ) for all x
- there is a bounded coupling $\|\partial_x\partial_y V\|_{L^\infty} < +\infty$

Then $\|\psi(t)-\psi_\infty\|_{L^1} \leqslant C\mathrm{e}^{-\beta\rho t}$.

- Improvement in the convergence rate when ρ (LSI for conditioned measures) is much larger than R (LSI for ψ_∞) $\to$ choice of ξ

- Elements of the proof
  - Marginals $\overline{\psi}(t,x) = \displaystyle\int \psi(t,x,y)\,dy$: simple diffusion $\partial_t\overline{\psi} = \partial_{xx}\overline{\psi}$
  - Decomposition of the total relative entropy $E(t) = \mathcal{H}(\psi\,|\,\psi_\infty)$ into a macroscopic contribution E_M (marginals in x) and a microscopic one E_m (conditioned measures)

Adaptive Biasing Potential techniques

- Self-Healing Umbrella Sampling[39]: unbiasing on-the-fly the occupation measure

$$\left\{ \begin{aligned} dq_t &= -\nabla(V - F_t \circ \xi)(q_t)\, dt + \sqrt{\frac{2}{\beta}}\, dW_t, \\ \mathrm{e}^{-\beta F_t(z)} &= \frac{1}{Z_t} \left(1 + \int_0^t \delta^\varepsilon(\xi(q_s) - z)\, \mathrm{e}^{-\beta F_s(\xi(q_s))}\, ds \right), \end{aligned} \right.$$

- If instantaneous equilibrium $q_t \sim \psi^{\mathrm{eq}}(t) \propto \mathrm{e}^{-\beta(V - F_t \circ \xi)}$ (consistency)

$$\lim_{\varepsilon \to 0} \mathbb{E}_{\psi^{\mathrm{eq}}(t)} \left[\delta^\varepsilon(\xi(q_t) - z)\, \mathrm{e}^{-\beta F_t(\xi(q_t))} \right] = \int_{\Sigma(z)} \mathrm{e}^{-\beta V} \delta_{\xi(q)-z}(dq) = \mathrm{e}^{-\beta F(z)}$$

- Metadynamics and its many versions/extensions/modifications[40]...

---

[39] S. Marsili *et al.*, *J. Phys. Chem. B* (2006)

[40] G. Bussi, A. Laio and M. Parinello, *Phys. Rev. Lett.* (2006)

The Wang-Landau algorithm (1)

- Partitioning of the configuration space $\mathcal{D}$ intro subsets $\mathcal{D}_i$ with weights

$$\theta_\star(i) \stackrel{\text{def}}{=} \int_{\mathcal{D}_i} \nu(q)\,dq, \qquad \nu(q) = Z^{-1}\mathrm{e}^{-\beta V(q)}$$

- Typically, $\mathcal{D}_i = \xi^{-1}\Big([\alpha_{i-1}, \alpha_i)\Big)$, originally[41] $\xi = V$

- Importance sampling to reduce metastability issues: biased measure

$$\nu_\theta(q) = \left(\sum_{i=1}^d \frac{\theta_\star(i)}{\theta(i)}\right)^{-1} \sum_{i=1}^d \frac{\nu(q)}{\theta(i)}\ \mathbb{1}_{\mathcal{D}_i}(q)$$

for any $\theta \in \Theta = \left\{\theta = (\theta(1), \cdots, \theta(d)) \ \middle|\ 0 < \theta(i) < 1,\ \sum_{i=1}^d \theta(i) = 1\right\}$

[41] F. Wang and D. Landau, *Phys. Rev. Lett.* & *Phys. Rev. E* (2001)

The Wang-Landau algorithm (2)

Linearized WL in the stochastic approximation setting

Given $q^0 \in \mathcal{D}$ and weights $\theta_0 \in \Theta$ (typically $\theta_0(i) = 1/d$),

(1) draw q^{n+1} from conditional distribution $P_{\theta_n}(q^n, \cdot)$ (Metropolis);

(2) assume that $q^{n+1} \in \mathcal{D}_i$. The weights are then updated as

$$\begin{cases} \theta_{n+1}(i) = \theta_n(i) + \gamma_{n+1}\ \theta_n(i)\,(1 - \theta_n(i)) & \\ \theta_{n+1}(k) = \theta_n(k) - \gamma_{n+1}\ \theta_n(k)\ \theta_n(i) & \text{for } k \neq i. \end{cases} \tag{1}$$

- Comparison with original Wang-Landau algorithm[42,43]
  - deterministic step-sizes γ_n, to be chosen appropriately
  - no "flat histogram" criterion
  - linearized weight update $\theta_{n+1}(i) = \theta_n(i) \dfrac{1 + \gamma_{n+1} \mathbb{1}_{I(X_{n+1})=i}}{1 + \sum_{j=1}^{d} \gamma_{n+1}\theta_n(j) \mathbb{1}_{I(X_{n+1})=i}}$

[42] Y. Atchade and J. Liu, *Stat. Sinica* (2010)

[43] F. Liang, *J. Am. Stat. Assoc.* (2005)

The Wang-Landau algorithm (3)

Stochastic approximation reformulation

Define $\eta_{n+1} = H(q^{n+1}, \theta_n) - h(\theta_n)$ and $h(\theta) = \int_{\mathcal{D}} H(q,\theta)\, \nu_\theta(q)\, dq$.

Then,

$$\theta_{n+1} = \theta_n + \gamma_{n+1}\, h(\theta_n) + \gamma_{n+1}\eta_{n+1}.$$

with $H_i(x,\theta) = \theta(i)\Big[\mathbb{1}_{\mathcal{D}_i}(x) - \theta(I(x))\Big]$ and $h(\theta) = \left(\sum_{i=1}^{d} \frac{\theta_\star(i)}{\theta(i)}\right)^{-1} (\theta_\star - \theta)$

- Issue: make sure that $\theta_n(i)$ remains positive
- Idea of proofs:
  - η_n is a "small, random" perturbation
  - the mean-field function h ensures the convergence to $\theta_\star$ in the absence of noise: there is a Lyapunov function W such that $\langle \nabla W, h\rangle < 0$ when $\theta \neq \theta_\star$
  - conditions on the step-sizes

The Wang-Landau algorithm (4)

• The density ν is such that $\sup_{\mathcal{D}} \nu < \infty$ and $\inf_{\mathcal{D}} \nu > 0$. In addition, $\theta_\star(i) > 0$.

• For any $\theta \in \Theta$, P_θ is a Metropolis-Hastings dynamics with invariant distribution ν_θ and symmetric proposal distribution with density $T(x,y)$ satisfying $\inf_{\mathcal{D}^2} T > 0$.

• The sequence $(\gamma_n)_{n \geqslant 1}$ is a non-negative determinstic sequence such that

(a) $(\gamma_n)_n$ is a non-increasing sequence converging to 0;

(b) $\sup_n \gamma_n \leqslant 1$;

(c) $\sum_n \gamma_n = \infty$;

(d) $\sum_n \gamma_n^2 < \infty$;

(e) $\sum_n |\gamma_n - \gamma_{n-1}| < \infty$.

Examples of acceptable step-sizes: $\gamma_n = \dfrac{\gamma_*}{n^\alpha}$ with $\alpha \in (1/2, 1]$

The Wang-Landau algorithm (5)

Under the previous assumptions, the convergence follows from general results of SA[44]

Weak stability result

The weight sequence almost surely comes back to a compact subset of Θ

$$\limsup_{n\to\infty} \left(\min_{1 \leqslant j \leqslant d} \theta_n(j) \right) > 0 \quad \text{a.s.}$$

Convergence result

The sequence $\{\theta_n\}$ almost surely converges to $\theta_\star$, and

$$\frac{1}{n} \sum_{k=1}^{n} f\left(q^k\right) \xrightarrow{\text{a.s.}} \int f(q)\, \nu_{\theta_\star}(q)\, dx$$

Various ways to recover averages with respect to ν (instead of $\nu_{\theta_\star}$).

[44] C. Andrieu, E. Moulines and P. Priouret, *SIAM J. Control Opt.* (2005)

Adaptive dynamics: extensions and open issues

- Obtain convergence rates for Wang-Landau? (Efficiency)
  - Only (very) partial results, such as the precise study of exit times out of metastable states[45]
  - adaptive dynamics allow to go from exponential scalings of the exit times to power-law scalings

- Convergence of other adaptive methods using trajectory averages?
  - Study discrete-in-time versions of SHUS and ABF
  - stochastic approximation with random time steps

- ABF for Langevin?

[45] G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre and G. Stoltz, *arXiv* **1207.6880**

References

- **Constrained stochastic dynamics** (equilibrium and nonequilibrium)

  - G. Ciccotti, T. Lelièvre and E. Vanden-Eijnden, Projection of diffusions on submanifolds: Application to mean force computation, *Commun. Pure Appl. Math.*, **61**(3) (2008) 371–408
  - T. Lelièvre, M. Rousset and G. Stoltz, Computation of free energy differences through nonequilibrium stochastic dynamics: the reaction coordinate case, *J. Comput. Phys.*, **222**(2) (2007) 624–643
  - T. Lelièvre, M. Rousset and G. Stoltz, Langevin dynamics with constraints and computation of free energy differences, *Math. Comput.*, **81** (2012) 2071–2125

- **Selection mechanims**

  - M. Rousset and G. Stoltz, An interacting particle system approach for molecular dynamics, *J. Stat. Phys.*, **123**(6) (2006) 1251-1272
  - T. Lelièvre, M. Rousset and G. Stoltz, Computation of free energy profiles with parallel adaptive dynamics, *J. Chem. Phys.*, **126** (2007) 134111
  - C. Chipot, T. Lelièvre and K. Minoukadeh, Potential of mean force calculations: a multiple-walker adaptive biasing force approach, *J. Chem. Theor. Comput.*, **6**(4) (2010) 1008-1017

References

- **Adaptive dynamics**

- T. Lelièvre, M. Rousset and G. Stoltz, Long-time convergence of an Adaptive Biasing Force method, *Nonlinearity*, **21** (2008) 1155–1181
- B. Dickson, F. Legoll, T. Lelièvre, G. Stoltz and P. Fleurat-Lessard, Free energy calculations: An efficient adaptive biasing potential method, *J. Phys. Chem. B*, **114**(17) (2010) 5823–5830
- B. Jourdain, T. Lelièvre and R. Roux, Existence, uniqueness and convergence of a particle approximation for the Adaptive Biasing Force process, *Math. Model. Numer. Anal.*, **44** (2010) 831-865
- T. Lelièvre and K. Minoukadeh, Long-time convergence of an Adaptive Biasing Force method : the bi-channel case, *Arch. Ration. Mech. Anal.*, **202**(1) (2011) 1–34
- G. Fort, B. Jourdain, E. Kuhn, T. Lelièvre and G. Stoltz, Convergence and efficiency of the Wang-Landau algorithm, *arXiv preprint* **1207.6880** (2012)

- T. Lelièvre, M. Rousset and G. Stoltz, *Free Energy Computations: A Mathematical Perspective* (Imperial College Press, 2010)

Computation of transport coefficients

Computation of transport properties

- There are three main types of techniques
  - Equilibrium techniques: Green-Kubo formula (autocorrelation)
  - Transient methods
  - Steady-state nonequilibrium techniques
    - boundary driven
    - bulk driven

- Definitions use analogy with macroscopic evolution equations

- Example of mathematical questions:
  - (equilibrium) integrability of correlation functions
  - (steady-state nonequilibrium): existence and uniqueness of an invariant probability measure

Steady-state nonequilibrium dynamics: some examples

- Perturbations of equilibrium dynamics by

Non-gradient forces (periodic potential V, $q \in \mathbb{T}$)

$$
(1) \qquad \left\{ \begin{aligned} dq_t &= M^{-1} p_t \, dt \\ dp_t &= \Big(-\nabla V(q_t) + \xi F \Big) dt - \gamma M^{-1} p_t \, dt + \sqrt{\frac{2\gamma}{\beta}} \, dW_t \end{aligned} \right.
$$

Fluctuation terms with different temperatures

$$
\left\{ \begin{aligned} dq_i &= p_i \, dt \\ dp_i &= \Big(v'(q_{i+1} - q_i) - v'(q_i - q_{i-1}) \Big) dt, \qquad i \neq 1, N \\ dp_1 &= \Big(v'(q_2 - q_1) - v'(q_1) \Big) dt - \gamma p_1 \, dt + \sqrt{2\gamma(T + \Delta T)} \, dW_t^1 \\ dp_N &= -v'(q_N - q_{N-1}) \, dt - \gamma p_N \, dt + \sqrt{2\gamma(T - \Delta T)} \, dW_t^N \end{aligned} \right.
$$

- Definition of nonequilibrium systems in physics: existence of currents (energy, particles, ...)

Invariant measure for nonequilibrium steady-states

- Mathematical definition of nonequilibrium systems?

  The generator of the dynamics is not self-adjoint with respect to $L^2(\mu)$, where μ is the invariant measure.

Often, μ replaced by invariant measure of related reference dynamics

- Quantification of the reversibility defaults by entropy production

$$\mathcal{R}\mathcal{A}^*\mathcal{R} = \mathcal{A} - \sigma, \qquad \sigma(q,p) = \xi\beta p^T M^{-1} F \text{ for } (1)$$

- Prove existence/uniqueness of μ: find a Lyapunov function

- May be difficult, *e.g.* 1D atom chains[46,47,48]

- Hypocoercivity? (works on $L^2(\psi_0)$...)

[46] L Rey-Bellet and L. Thomas, *Commun. Math. Phys.* (2002)
[47] P. Carmona, *Stoch. Proc. Appl.* (2007)
[48] J.-P. Eckmann and M. Hairer, *Commun. Math. Phys.* (2000)

Invariant measure for nonequilibrium steady-states

• For equilibrium systems, local perturbations in the dynamics induce local perturbations in the invariant measure

$$dx_t = \Big(-\nabla V(x_t) + \nabla \widetilde{V}(x_t) \Big) dt + \sqrt{\frac{2}{\beta}}\, dW_t$$

so that $\mu(dx) = Z^{-1} \mathrm{e}^{-\beta(V(x) - \widetilde{V}(x))}\, dx$

• For nonequilibrium systems, the invariant measure depends non-trivially on the details of the dynamics and perturbations are non-local!

• For the dynamics $dx_t = \Big(-\widetilde{V}'(x_t) + F \Big) dt + \sqrt{2}\, dW_t$ on $\mathbb{T}$,

$$\mu(dx) = Z^{-1} \mathrm{e}^{-\widetilde{V}(x) + Fx} \left(\int_x^{x+1} \mathrm{e}^{\widetilde{V}(y) - Fy}\, dy \right) dx$$

Variance reduction techniques?

- Importance sampling? Invariant probability measures ψ_∞, ψ_∞^A for

$$dq_t = b(q_t)\,dt + \sigma dW_t, \qquad dq_t = \Big(b(q_t) + \nabla A(q_t)\Big)dt + \sigma dW_t$$

In general $\psi_\infty^A \neq Z^{-1}\psi_\infty \mathrm{e}^{A}$ (consider $b(q) = F$ and $A = \widetilde{V}$)

- Stratification? (as in TI...) Consider $x \in \mathbb{T}^2$, $\psi_\infty = \mathbf{1}_{\mathbb{T}^2}$

$$\left\{\begin{aligned} dx_t^1 &= \partial_{x_2} H(x_t^1, x_t^2) + \sqrt{2}\,dW_t^1 \\ dx_t^2 &= -\partial_{x_1} H(x_t^1, x_t^2) + \sqrt{2}\,dW_t^2 \end{aligned}\right.$$

Constraint $\xi(x) = x_2$, constrained dynamics

$$dx_t^1 = f(x_t^1)\,dt + \sqrt{2}\,dW_t^1, \qquad f(x^1) = \partial_{x_2} H(x^1, 0).$$

Then $\psi_\infty(x^1) = Z^{-1} \displaystyle\int_0^1 \mathrm{e}^{V(x^1+y) - V(x^1) - Fy}\,dy \neq \mathbf{1}_{\mathbb{T}}(x^1)$

where $F = \displaystyle\int_0^1 f$ and $V(x^1) = \displaystyle\int_0^{x^1} (f(s) - F)\,ds$

Linear response (1)

- Generator of the perturbed dynamics $\mathcal{A}_0 + \xi \mathcal{A}_1$, on $L^2(\psi_0)$ (where ψ_0 is the unique invariant measure of the dynamics generated by $\mathcal{A}_0$)
- Fokker-Planck equation: $(\mathcal{A}_0^* + \xi \mathcal{A}_1^*) f_\xi = 0$ with $\displaystyle \int f_\xi \psi_0 = 1$

Series expansion of the invariant measure $\psi_\xi = f_\xi \psi_0$

$$f_\xi = (\mathcal{A}_0^* + \xi \mathcal{A}_1^*)^{-1} \mathcal{A}_0^* \mathbf{1} = \left(1 + \sum_{n=1}^{+\infty} \xi^n \left[- \left(\mathcal{A}_0^*\right)^{-1} \mathcal{A}_1^* \right]^n \right) \mathbf{1}$$

- These computations can be made rigorous for ξ sufficiently small when...
  - (equilibrium) $\mathrm{Ker}(\mathcal{A}_0^*) = \mathbf{1}$ and $\mathcal{A}_0^*$ invertible on

$$\mathcal{H} = \left\{ f \in L^2(\psi_0) \,\middle|\, \int f \psi_0 = 0 \right\} = L^2(\psi_0) \cap \{\mathbf{1}\}^\perp$$

  - (perturbation) $\mathrm{Ran}(\mathcal{A}_1^*) \subset \mathcal{H}$ and $\left(\mathcal{A}_0^*\right)^{-1} \mathcal{A}_1^*$ bounded on $\mathcal{H}$, *e.g.* when $\|\mathcal{A}_1 \varphi\|_{L^2(\psi_0)} \leqslant a \|\mathcal{A}_0 \varphi\|_{L^2(\psi_0)} + b \|\varphi\|_{L^2(\psi_0)}$

Linear response (2)

- Response property $R \in \mathcal{H}$, conjugated response $S = \mathcal{A}_1^* \mathbf{1}$

Linear response from Green-Kubo type formulas

$$\alpha = \lim_{\xi \to 0} \frac{\langle R \rangle_\xi}{\xi} = -\int_{\mathcal{E}} \left[\mathcal{A}_0^{-1} R\right] \left[\mathcal{A}_1^* \mathbf{1}\right] \psi_0 = \int_0^{+\infty} \mathbb{E}\Big(R(x_t) S(x_0)\Big) dt$$

using the formal equality $-\mathcal{A}_0^{-1} = \int_0^{+\infty} \mathrm{e}^{t\mathcal{A}_0}\, dt$ (as operators on $\mathcal{H}$)

- Autocorrelation of R recovered for perturbations such that $\mathcal{A}_1^* \mathbf{1} \propto R$
- For general property: consider $\displaystyle \lim_{\xi \to 0} \frac{\langle R \rangle_\xi - \langle R \rangle_0}{\xi}$
- In practice:
  - Identify the response function
  - Construct a physically meaningful perturbation
  - Equivalent non physical perturbations ("Synthetic NEMD")

Example 1: Autodiffusion (1)

- Periodic potential V, constant external force F

$$\begin{cases} dq_t = M^{-1} p_t \, dt \\ dp_t = \Big(-\nabla V(q_t) + \xi F \Big) dt - \gamma M^{-1} p_t \, dt + \sqrt{\dfrac{2\gamma}{\beta}} \, dW_t \end{cases}$$

- In this case, $\mathcal{A}_1 = F \cdot \partial_p$ and so $\mathcal{A}_1^* \mathbf{1} = -\beta F \cdot M^{-1} p$

- Response: $R(q,p) = F \cdot M^{-1} p =$ average velocity in the direction F

- Linear response result:

Definition of the mobility

$$\alpha = \lim_{\xi \to 0} \frac{\left\langle F \cdot M^{-1} p \right\rangle_\xi}{\xi} = \beta \int_0^{+\infty} \mathbb{E}_{\mathrm{eq}} \Big((F \cdot M^{-1} p_t)(F \cdot M^{-1} p_0) \Big) \, dt$$

(Expectation over canonical initial conditions and realizations of the dynamics)

Example 1: Autodiffusion (2)

- Einstein formulation: diffusive time-scale for the equilibrium dynamics

Definition of the diffusion

$$D = \lim_{T \to +\infty} \frac{\Big(F \cdot \mathbb{E}_{\mathrm{eq}}(q_T - q_0)\Big)^2}{2T}$$

- Relation between mobility and diffusion

$$\alpha = \beta D$$

since $\dfrac{\Big(F \cdot \mathbb{E}(q_T - q_0)\Big)^2}{2T} = \displaystyle\int_0^T \mathbb{E}\Big((F \cdot M^{-1} p_t)(F \cdot M^{-1} p_0)\Big) \left(1 - \frac{t}{T}\right) dt$

- Various extensions:
  - Time-dependent forcings $F(t)$ (stochastic resonance)
  - Random forcings
  - Space-time dependent[49] forcings $F(t,q)$

[49] R. Joubaud, G. Pavliotis and G. Stoltz, in preparation

Example 2: Thermal transport in atom chains (1)

- Hamiltonian $H(q,p) = \sum_{i=1}^{N} \frac{p_i^2}{2} + \sum_{i=1}^{N-1} v(q_{i+1} - q_i) + v(q_1)$

- Hamiltonian dynamics with Langevin at the boundaries

- Perturbation $\mathcal{A}_1 = \gamma(\partial_{p_1}^2 - \partial_{p_N}^2)$

- Response function: Total energy current

$$J = \sum_{i=1}^{N-1} j_{i+1,i}, \qquad j_{i+1,i} = -v'(q_{i+1} - q_i)\frac{p_i + p_{i+1}}{2}$$

- Motivation: Local conservation of the energy (in the bulk)

$$\frac{d\varepsilon_i}{dt} = j_{i-1,i} - j_{i,i+1}, \qquad \varepsilon_i = \frac{p_i^2}{2} + \frac{1}{2}\Big(v(q_{i+1} - q_i) + v(q_i - q_{i-1})\Big)$$

Example 2: Thermal transport in atom chains (2)

• Definition of the thermal conductivity: linear response

$$\kappa_N = \lim_{\Delta T \to 0} \frac{\langle J \rangle_{\Delta T}}{\Delta T} = \frac{2\beta^2}{N-1} \int_0^{+\infty} \mathbb{E}\Big(J(q_t,p_t) J(q_0,p_0) \Big)\, dt$$

• Synthetic dynamics: fixed temperatures of the thermostats but external forcings $\to$ bulk driven dynamics (convergence may be faster?)

- Non-gradient perturbation $-\xi\Big(v'(q_{i+1}-q_i) + v'(q_i - q_{i-1})\Big)$
- Hamiltonian perturbation $H_0 + \xi H_1$ with $H_1(q,p) = \displaystyle\sum_{i=1}^N i\varepsilon_i$

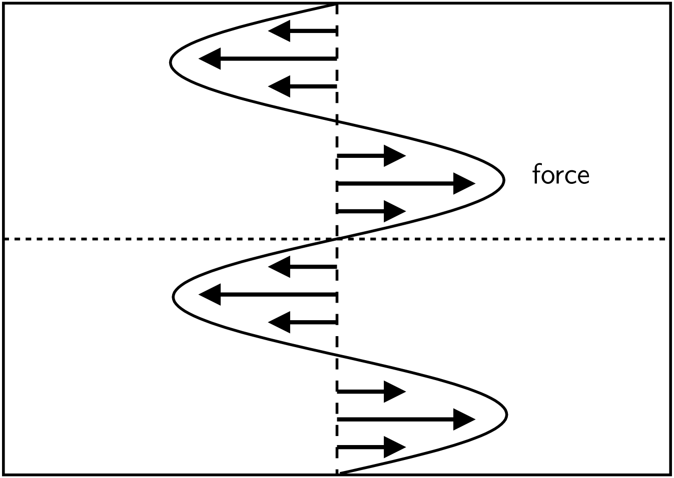
In both cases, $\mathcal{A}_1^* = -\mathcal{A}_1 + cJ$

• Necessary and sufficient conditions for κ_N to have a limit as $N \to +\infty$? (use of stochastic perturbations[50], numerical studies, ...)

[50] S. Olla, C. Bernardin, ...

Shear viscosity in fluids (1)

2D system to simplify notation: $\mathcal{D} = (L_x\mathbb{T} \times L_y\mathbb{T})^N$



Shear viscosity in fluids (2)

- Add a smooth nongradient force in the x direction, depending on y

Langevin dynamics under flow

$$\begin{cases} dq_{i,t} = \dfrac{p_{i,t}}{m}\,dt, \\[2ex] dp_{xi,t} = -\nabla_{q_{xi}} V(q_t)\,dt + \xi F(q_{yi,t})\,dt - \gamma_x \dfrac{p_{xi,t}}{m}\,dt + \sqrt{\dfrac{2\gamma_x}{\beta}}\,dW_t^{xi}, \\[2ex] dp_{yi,t} = -\nabla_{q_{yi}} V(q_t)\,dt - \gamma_y \dfrac{p_{yi,t}}{m}\,dt + \sqrt{\dfrac{2\gamma_y}{\beta}}\,dW_t^{yi}, \end{cases}$$

- Existence/uniqueness of a smooth invariant measure provided $\gamma_x, \gamma_y > 0$
- Perturbation $\mathcal{A}_1 = \sum_{i=1}^{N} F(q_{y,i}) \partial_{p_{x,i}}$ $\mathcal{A}_0$-bounded since

$\|\mathcal{A}_1 \varphi\|^2 \leqslant |\langle \varphi, \mathcal{A}_0 \varphi \rangle|$

- Linear response: $\lim_{\xi \to 0} \dfrac{\langle \mathcal{A}_0 h \rangle_\xi}{\xi} = -\dfrac{\beta}{m} \left\langle h, \sum^{N} p_{xi} F(q_{yi}) \right\rangle$

Shear viscosity in fluids (3)

- Average longitudinal velocity $u_x(Y) = \lim_{\varepsilon \to 0} \lim_{\xi \to 0} \frac{\langle U_x^\varepsilon(Y,\cdot)\rangle_\xi}{\xi}$ where

$$U_x^\varepsilon(Y,q,p) = \frac{L_y}{Nm} \sum_{i=1}^N p_{xi} \chi_\varepsilon\left(q_{yi} - Y\right)$$

- Average off-diagonal stress $\sigma_{xy}(Y) = \lim_{\varepsilon \to 0} \lim_{\xi \to 0} \frac{\langle \dots \rangle_\xi}{\xi}$, where ... =

$$\frac{1}{L_x}\left(\sum_{i=1}^N \frac{p_{xi}p_{yi}}{m} \chi_\varepsilon\left(q_{yi} - Y\right) - \sum_{1 \leqslant i < j \leqslant N} \mathcal{V}'(|q_i - q_j|) \frac{q_{xi} - q_{xj}}{|q_i - q_j|} \int_{q_{yj}}^{q_{yi}} \chi_\varepsilon(s - Y)\, ds\right)$$

- Local conservation of momentum[51]: replace h by U_x^ε (with $\overline{\rho} = N/|\mathcal{D}|$)

$$\frac{d\sigma_{xy}(Y)}{dY} + \gamma_x \overline{\rho} u_x(Y) = \overline{\rho} F(Y)$$

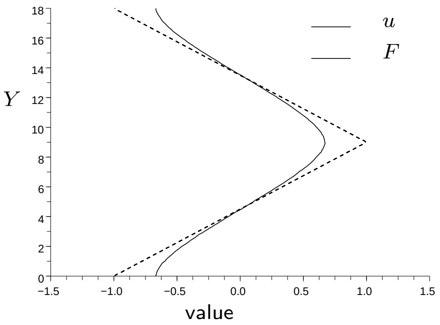
[51] Irving and Kirkwood, *J. Chem. Phys.* **18** (1950)

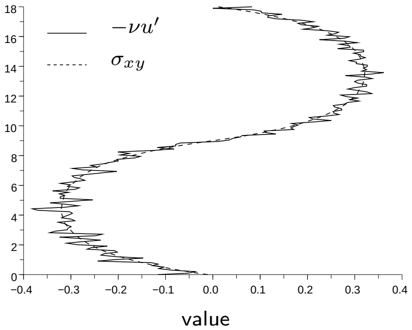
Shear viscosity in fluids (4)

- Definition $\sigma_{xy}(Y) := -\eta(Y)\dfrac{du_x(Y)}{dY}$, closure assumption $\eta(Y) = \eta > 0$

Velocity profile in Langevin dynamics under flow

$$-\eta u_x''(Y) + \gamma_x \overline{\rho} u_x(Y) = \overline{\rho} F(Y)$$





Transient techniques

- Onsager: *The return to equilibrium of a macroscopic perturbation is governed by the same laws as the equilibrium fluctuations*

- Perturbed initial condition of Gibbs type (with $A \in \mathcal{H}$ *i.e.* $\langle A \rangle_0 = 0$)

$$\psi_\eta = Z_\eta \, \mathrm{e}^{-\beta\eta A} \, \psi_0 = \Big(1 - \beta\eta A\Big)\psi_0 + \mathrm{O}(\eta^2)$$

- Evolution of some observable B under the equilibrium dynamics $\mathcal{A}_0$:

$$\langle B \rangle_\eta(t) = \int_{\mathcal{X}} \mathrm{e}^{t\mathcal{A}_0} B \, \psi_\eta = \langle B \rangle_0 - \beta\eta\mathbb{E}\Big(B(x_t)A(x_0)\Big) + \mathrm{O}(\eta^2)$$

- A Green-Kubo type formula is recovered upon integration (for $B \in \mathcal{H}$)

$$\lim_{\eta \to 0} \int_0^{+\infty} \frac{\langle B \rangle_\eta(t)}{\eta} \, dt = -\beta \int_0^{+\infty} \mathbb{E}\Big(B(x_t)A(x_0)\Big) dt$$

- Autodiffusion: Start from the canonical distribution associated with

$$H_\eta(q,p) = \frac{1}{2}\Big(p - \eta F\Big)^T M^{-1}\Big(p - \eta F\Big) + V(q)$$

Elements of numerical analysis (in preparation...)

- Autodiffusion case: same splitting scheme as equilibrium dynamics with decentered Ornstein-Uhlenbeck process (generator C_ξ)

$$dp_t = \xi F\,dt - \gamma M^{-1} p_t\,dt + \sqrt{\frac{2\gamma}{\beta}}\,dW_t$$

- Existence and uniqueness of an invariant measure $\mu_{\Delta t,\xi}$

Talay-Tubaro like estimates

For a splitting scheme of order p when $\xi = 0$,

$$\int_{\mathcal{E}} \psi\,d\mu_{\Delta t,\xi} = \int_{\mathcal{E}} \psi\Big(1 + \xi f_{0,1} + \Delta t^p f_{1,0} + \xi \Delta t^p f_{1,1}\Big) d\mu + a^\psi_{\Delta t,\xi}$$

with $|a^\psi_{\Delta t,\xi}| \leqslant K(\xi^2 + \Delta t^{p+1})$ and $|a^\psi_{\Delta t,\xi} - a^\psi_{\Delta t,0}| \leqslant K\xi(\xi + \Delta t^{p+1})$

- Allows to control errors on the transport coefficients (only $f_{1,1}$ remains)
- Error estimates on the Green-Kubo formula (recover the precision of the scheme)

References

- **Some introductory references...**
  - L. Rey-Bellet, Open classical systems, *Lecture Notes in Mathematics*, **1881** (2006) 41–78
  - D. J. Evans and G. P. Morriss, *Statistical Mechanics of Nonequilibrium Liquids* (Cambridge University Press, 2008)
  - M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (Oxford, 2010)
  - G. Stoltz, *Molecular Simulation: Nonequilibrium and Dynamical Problems*, Habilitation Thesis (2012) [Chapter 3]

- And many reviews on **specific topics**! For instance, thermal transport in one dimensional systems