Numerical methods for computational statistical physics

Gabriel STOLTZ

gabriel.stoltz@enpc.fr

(CERMICS, Ecole des Ponts & MATHERIALS team, INRIA Paris)

Work also supported by ANR Funding ANR-14-CE23-0012 ("COSMOS")

CIMPA school, IIT Kanpur, July 2017

Outline

- **Some elements of statistical physics** [Lecture 1]
- **Sampling the microcanonical ensemble** [Lectures 1-3]
  - Hamiltonian dynamics and ergodic assumption
  - Longtime numerical integration of the Hamiltonian dynamics
- **Sampling the canonical ensemble** [Lectures 2-4-5]
  - Stochastic differential equations (Langevin dynamics)
  - Markov chain approaches (Metropolis-Hastings)
- **Lab sessions**
  - integration of Hamiltonian dynamics
  - Metropolis algorithm

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)
  - B. Leimkuhler and C. Matthews, *Molecular Dynamics: With Deterministic and Stochastic Numerical Methods* (Springer, 2015).
  - T. Lelièvre and G. Stoltz, Partial differential equations and stochastic methods in molecular dynamics, *Acta Numerica* **25**, 681-880 (2016)

Some elements of statistical physics

General perspective (1)

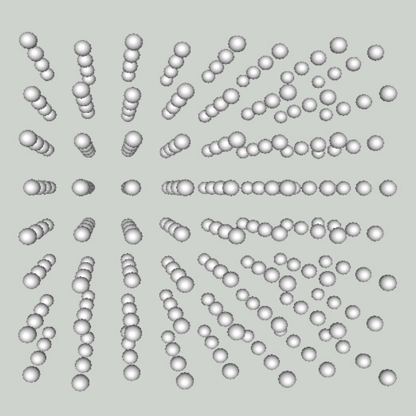
- **Aims of computational statistical physics**
  - numerical microscope
  - computation of average properties, static or dynamic

- **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 $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

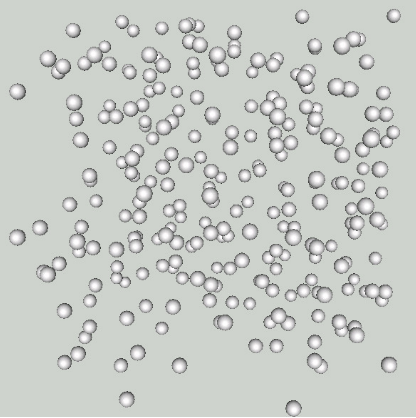
- **"Standard" simulations**
  - 10^6 particles ["world records": around 10^9 particles]
  - integration time: (fraction of) ns ["world records": (fraction of) μs]

General perspective (2)

What is the melting temperature of argon?



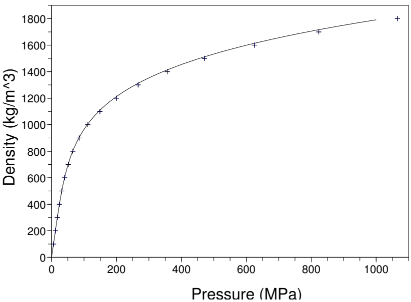
(a) Solid argon (low temperature)

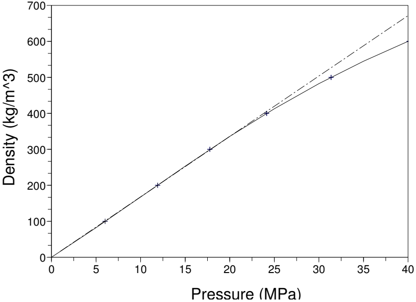


(b) Liquid argon (high temperature)

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?"

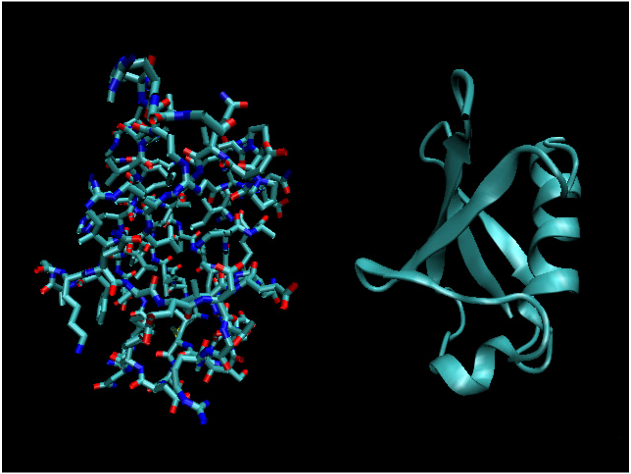




Equation of state (pressure/density diagram) for argon at $T = 300$ K

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

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

Microscopic description: interaction laws

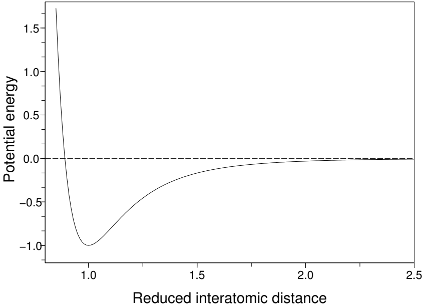
- All the physics is contained in V
  - ideally derived from quantum mechanical computations
  - in practice, empirical potentials for large scale calculations

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

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

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

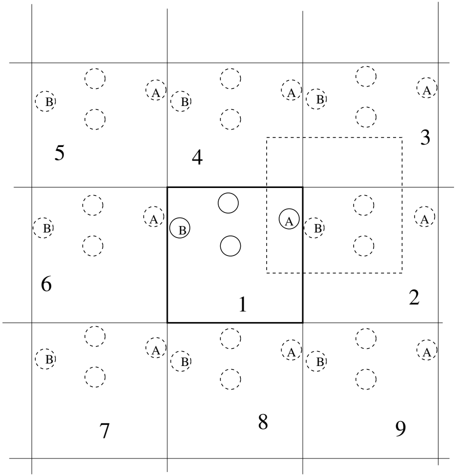
$$\text{Argon:}\quad \begin{cases} \sigma = 3.405\times 10^{-10}\ \text{m} \\ \varepsilon/k_{\mathrm{B}} = 119.8\ \ \text{K} \end{cases}$$



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

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

Main issue

Computation of high-dimensional integrals... Ergodic averages

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

Sampling the microcanonical ensemble

Outline

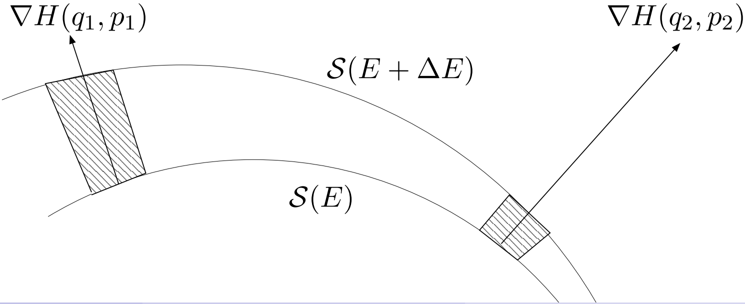
- **Sampling the microcanonical measure**
  - Definition of the microcanonical measure
  - The Hamiltonian dynamics and its properties
  - The ergodic assumption
- **Standard numerical analysis of ordinary differential equations**
  - Consistency, stability, convergence
  - Standard examples
- **Longtime numerical integration of the Hamiltonian dynamics**
  - Failure of standard schemes
  - Symplecticity and construction of symplectic schemes
  - Elements of backward error analysis

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 (1)

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)

- Flow: $\phi_t(q_0,p_0)$ solution at time t starting from initial condition (q_0,p_0)
- Why Hamiltonian formalism? (instead of working with velocities?)
  - Note that the vector field is divergence-free

$$\mathrm{div}_q\Big(\nabla_p H(q(t),p(t))\Big) + \mathrm{div}_p\Big(-\nabla_q H(q(t),p(t))\Big) = 0$$

  - Volume preservation $\displaystyle\int_{\phi_t(B)} dq\,dp = \int_B dq\,dp$

The Hamiltonian dynamics (2)

- Other properties
  - Preservation of energy $H \circ \phi_t = H$

$$\frac{d}{dt}\Big[H\big(q(t),p(t)\big)\Big] = \nabla_q H(q(t),p(t))\cdot\frac{dq(t)}{dt} + \nabla_p H(q(t),p(t))\cdot\frac{dp(t)}{dt} = 0$$

  - Time-reversibility $\phi_{-t} = S \circ \phi_t \circ S$ where $S(q,p) = (q,-p)$

    Proof: use $S^2 = \mathrm{Id}$ and note that

$$S \circ \phi_{-t}(q_0,p_0) = \big(q(-t), -p(-t)\big)$$

    is a solution of the Hamiltonian dynamics starting from $(q_0,-p_0)$, as is $\phi_t \circ S(q_0,p_0)$. Conclude by uniqueness of solution.

  - Symmetry $\phi_{-t} = \phi_t^{-1}$ (in general, $\phi_{t+s} = \phi_t \circ \phi_s$)

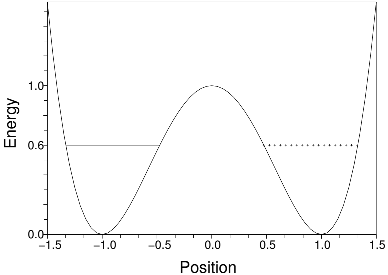
Ergodicity of the Hamiltonian dynamics

- Invariance of the microcanical measure by 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 conserved, 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

Outline

- **Sampling the microcanonical measure**
  - Definition of the microcanonical measure
  - The Hamiltonian dynamics and its properties
  - The ergodic assumption
- **Standard numerical analysis of ordinary differential equations**
  - consistency, stability, convergence
  - standard examples
- **Longtime numerical integration of the Hamiltonian dynamics**
  - Failure of standard schemes
  - Symplecticity and construction of symplectic schemes
  - Elements of backward error analysis

Some fundaments of numerical integration of ODEs

- Consider an ordinary differential equation $\frac{dy(t)}{dt} = f\Big(y(t)\Big)$
- Assume that it is well posed (unique solution for all initial conditions)

$$y(t) = \phi_t\Big(y(0)\Big) = y(0) + \int_0^t f\Big(y(s)\Big)\,ds$$

- Introduce y^n, approximation of $y(t_n)$ with $t_n = n\Delta t$ (fixed time step)

One step method

$$y^{n+1} = \Phi_{\Delta t}\left(y^n\right)$$

- Simplest example: Explicit Euler

$$y^{n+1} = y^n + \Delta t\, f(y^n)$$

in which case $\Phi_{\Delta t}(y) = y + \Delta t\, f(y)$

Further examples

- Explicit methods
  - Heun: $y^{n+1} = y^n + \frac{\Delta t}{2}\Big(f(y^n) + f\big(y^n + \Delta t f(y^n)\big)\Big)$
  - Fourth order Runge-Kutta scheme

$$y^{n+1} = y^n + \Delta t \frac{f(y^n) + 2f(Y^{n+1}) + 2f(Y^{n+2}) + f(Y^{n+3})}{6}$$

with $Y^{n+1} = y^n + f(y^n)\frac{\Delta t}{2}$, $Y^{n+2} = y^n + f(Y^{n+1})\frac{\Delta t}{2}$, and $Y^{n+3} = y^n + f(Y^{n+2})\Delta t$

- Implicit methods [solve using a fixed-point iteration for instance]
  - Implicit Euler: $y^{n+1} = y^n + \Delta t\, f\left(y^{n+1}\right)$
  - Trapezoidal rule: $y^{n+1} = y^n + \frac{\Delta t}{2}\left(f(y^n) + f(y^{n+1})\right)$
  - Midpoint: $y^{n+1} = y^n + \Delta t\, f\left(\frac{y^n + y^{n+1}}{2}\right)$

Standard error analysis

- Error on the trajectory over finite times
  - local error at each time step (consistency + rounding off error)
  - accumulation of the errors (stability)
- A numerical method is convergent when the global error satisfies

$$\lim_{\Delta t \to 0} \left(\max_{0 \leqslant n \leqslant N} \|y^n - y(n\Delta t)\| \right) = 0$$

- Order p consistency: quantification of the error over one time step

$$e(y_0) = y(\Delta t) - \Phi_{\Delta t}(y_0) = \mathrm{O}(\Delta t^{p+1})$$

- Example: explicit Euler is of order 1 → Taylor expansion

$$y(\Delta t) - \Big(y_0 + \Delta t\, f(y_0) \Big) = \frac{\Delta t^2}{2} y''(\theta \Delta t), \qquad y''(\tau) = \partial_y f\big(y(\tau)\big) \cdot f\big(y(\tau)\big)$$

Standard error analysis

• Stability: for all sequences $y^{n+1} = \Phi_{\Delta t}(y^n)$ and $z^{n+1} = \Phi_{\Delta t}(z^n) + \delta^n$, it holds ($S$ independent of Δt)

$$\max_{0 \leqslant n \leqslant N} \|y^n - z^n\| \leqslant S \left(\left|y^0 - z^0\right| + \sum_{n=0}^{N} \|\delta^n\| \right)$$

True when $\|\Phi_{\Delta t}(y_1) - \Phi_{\Delta t}(y_2)\| \leqslant \Lambda \|y_1 - y_2\|$

• A method which is stable and consistent is convergent
(take $z^n = y(n\Delta t)$ exact solution, so that δ_n is the local truncation error)

• For a method of order p, there are $N = [T/\Delta t]$ integration steps

$$\max_{0 \leqslant n \leqslant N} \|y^n - y(t_n)\| \leqslant C(T)\Delta t^p$$

with a prefator which typically grows exponentially with T...

Outline

- **Sampling the microcanonical measure**
  - Definition of the microcanonical measure
  - The Hamiltonian dynamics and its properties
  - The ergodic assumption

- **Standard numerical analysis of ordinary differential equations**
  - consistency, stability, convergence
  - standard examples

- **Longtime numerical integration of the Hamiltonian dynamics**
  - Failure of standard schemes
  - Symplecticity and construction of symplectic schemes
  - Elements of backward error analysis

Longtime integration: failure of default schemes

- Appropriate notion of stability: longtime energy preservation

Hamiltonian dynamics as a first-order differential equation

$$y = \begin{pmatrix} q \\ p \end{pmatrix}, \qquad \dot{y} = J\nabla H(y), \qquad J = \begin{pmatrix} 0 & I_{3N} \\ -I_{3N} & 0 \end{pmatrix}$$

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

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

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

Which qualitative properties are important?

- Time reversibility $\Phi_{\Delta t} \circ S = S \circ \Phi_{-\Delta t}$ usually verified

Check it for Explicit Euler $\Phi_{\Delta t}^{\mathrm{Euler}}(q,p) = \left(q + \Delta t M^{-1} p, p - \Delta t \nabla V(q)\right)$

$$\Phi_{\Delta t}^{\mathrm{Euler}}(q,-p) = \begin{pmatrix} q - \Delta t M^{-1} p \\ -p - \Delta t \nabla V(q) \end{pmatrix} = S \begin{pmatrix} q - \Delta t M^{-1} p \\ p + \Delta t \nabla V(q) \end{pmatrix} = S\left(\Phi_{-\Delta t}^{\mathrm{Euler}}(q,p)\right)$$

- Symmetry $\Phi_{\Delta t}^{-1} = \Phi_{-\Delta t}$ is not trivial at all

- Oriented volume preservation: linear case in 2D
  - two independent vectors $q = (x,y)$ and $q' = (x',y')$, oriented volume

$$q \wedge q' = xy' - xy = q^T J q', \qquad J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$$

  - linear transformation A, so that $q \to Aq$ and $q' \to Aq'$

$$q^T J q' \to q^T A^T J A q'$$

  - unchanged provided $A^T J A = J$

Longtime integration: symplecticity (1)

- Generalization to higher dimensions and nonlinear transformations
  - mapping $g(q,p) = (g_1(q,p), \dots, g_{6N}(q,p))^T$
  - Jacobian matrix $g'(q,p)$

$$g'(q,p) = \begin{pmatrix} \dfrac{\partial g_1}{\partial q_1} & \dots & \dfrac{\partial g_1}{\partial q_{3N}} & \dfrac{\partial g_1}{\partial p_1} & \dots & \dfrac{\partial g_1}{\partial p_{3N}} \\ & \ddots & & & \ddots & \\ \dfrac{\partial g_{6N}}{\partial q_1} & \dots & \dfrac{\partial g_{6N}}{\partial q_{3N}} & \dfrac{\partial g_{6N}}{\partial p_1} & \dots & \dfrac{\partial g_{6N}}{\partial q_{2dN}} \end{pmatrix}.$$

Symplectic mapping

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

- A mapping is symplectic if and only if it is (locally) the flow of a Hamiltonian system

- A composition of symplectic mappings is symplectic

Longtime integration: symplecticity (2)

- Proof: A Hamiltonian mapping is symplectic

Derive the Jacobian matrix $\psi(t,y) = \dfrac{\partial \phi_t(y)}{\partial y}$

$$\frac{d\psi}{dt} = \frac{\partial}{\partial y}\left(\frac{d\phi_t(y)}{dt}\right) = \frac{\partial}{\partial y}\big(J\nabla H(\phi_t(y))\big) = J\left(\nabla^2 H(\phi_t(y))\right)\frac{\partial \phi_t(y)}{\partial y}$$

so that, using $J^T = -J$

$$\frac{d}{dt}\Big(\psi(t)^T J\psi(t)\Big) = \psi(t)^T\left(\nabla^2 H(\phi_t(y))\right)J^T J\psi(t) + \psi(t)^T\left(\nabla^2 H(\phi_t(y))\right)J^2\psi(t) = 0$$

The conclusion follows since $\psi(0)^T J\psi(0) = J$. Converse statement: "integrability Lemma" (see Hairer/Lubich/Wanner, Theorem VI.2.6 and Lemma VI.2.7)

- Composition of symplectic mappings g, h: use $(g\circ h)' = (g'\circ h)h'$ and

$$h'(q,p)^T\Big(g'(h(q,p))\Big)^T J\Big(g'(h(q,p)\Big)h'(q,p) = [h'(q,p)]^T Jh'(q,p) = J$$

Longtime integration: symplecticity (3)

- Stability result

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]

- Does not say anything on the statistical behavior! (except for integrable systems)

Near energy preservation is a necessary condition

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

Longtime integration: constructing symplectic schemes (1)

- Splitting strategy for a general ODE $\dot{y}(t) = f(y)$, flow ϕ_t
  - Decompose the vector field as $f(y) = f_1(y) + f_2(y)$
  - Define the flows ϕ_t^i associated with each elementary ODE $\dot{z}(t) = f_i(z)$
  - Motivation: (almost) analytical integration of elementary ODEs
  - Generalization to a decomposition into $m \geqslant 2$ parts

- Trotter splitting (first order accurate)

$$\phi_{\Delta t} = \phi^1_{\Delta t} \circ \phi^2_{\Delta t} + \mathrm{O}(\Delta t^2) = \phi^2_{\Delta t} \circ \phi^1_{\Delta t} + \mathrm{O}(\Delta t^2)$$

- Strang splitting (second order)

$$\phi_{\Delta t} = \phi^1_{\Delta t/2} \circ \phi^2_{\Delta t} \circ \phi^1_{\Delta t/2} + \mathrm{O}(\Delta t^3) = \phi^2_{\Delta t/2} \circ \phi^1_{\Delta t} \circ \phi^2_{\Delta t/2} + \mathrm{O}(\Delta t^3)$$

- Extension to higher order schemes (Suzuki, Yoshida)

Longtime integration: constructing symplectic schemes (2)

- Splitting Hamiltonian systems: $\left\{ \begin{array}{l} \dot{q} = M^{-1}\,p \\ \dot{p} = \quad 0 \end{array} \right.$ and $\left\{ \begin{array}{l} \dot{q} = \quad 0 \\ \dot{p} = -\nabla V(q) \end{array} \right.$

- 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

$$\left\{ \begin{aligned} p^{n+1/2} &= p^n - \frac{\Delta t}{2} \nabla V(q^n) \\ q^{n+1} &= q^n + \Delta t \; M^{-1} p^{n+1/2} \\ p^{n+1} &= p^{n+1/2} - \frac{\Delta t}{2} \nabla V(q^{n+1}) \end{aligned} \right.$$

- 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

Molecular constraints

- In some cases, mechanical systems are constrained
- Numerical motivation: highly oscillatory systems
  - Fast oscillations of the system, *e.g.* vibrations of bonds and bond angles
  - Severe limitations on admissible time steps since $\omega \Delta t < 2$
  - Remove the limitation by constraining these degrees of freedom
  - Introduces some sampling errors, which can be corrected
- Other motivation: computation of free energy difference with thermodynamic integration
- The Hamiltonian dynamics has to be modified consistently, and appropriate numerical schemes have to be devised (RATTLE)

Outline

- **Sampling the microcanonical measure**
  - Definition of the microcanonical measure
  - The Hamiltonian dynamics and its properties
  - The ergodic assumption

- **Standard numerical analysis of ordinary differential equations**
  - consistency, stability, convergence
  - standard examples

- **Longtime numerical integration of the Hamiltonian dynamics**
  - Failure of standard schemes
  - Symplecticity and construction of symplectic schemes
  - Elements of backward error analysis

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

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

- 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

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

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) $\rightarrow$ impossible...
  - Stochastic differential equations
  - Markov chain methods
  - 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
- **Stochastic differential equations**
  - General perspective (convergence results, ...)
  - Overdamped Langevin dynamics (Einstein-Schmolukowski)
  - Langevin dynamics
  - Extensions: DPD, Generalized Langevin

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 $\min\left(1, r(q^n, \widetilde{q}^{n+1}\right)$ where

$$r(q, q') = \frac{T(q', q)\,\nu(q')}{T(q, q')\,\nu(q)}, \qquad \nu(dq) \propto \mathrm{e}^{-\beta V(q)}.$$

  If acception, set $q^{n+1} = \tilde{q}^{n+1}$; 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)

- The normalization constant in the canonical measure needs not be known
- Transition kernel: accepted moves + rejection

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

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

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

- The canonical measure is reversible with respect to ν

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

This implies invariance: $\displaystyle\int_{\mathcal{D}} \psi(q') P(q, dq')\, \nu(dq) = \int_{\mathcal{D}} \psi(q)\, \nu(dq)$

Metropolis-Hastings algorithm (3)

- Proof: Detailed balance on the absolutely continuous parts

$$\begin{aligned}\min\left(1, r(q,q')\right) T(q,dq')\nu(dq) &= \min\left(1, r(q',q)\right) r(q,q') T(q,dq')\nu(dq) \\ &= \min\left(1, r(q',q)\right) T(q',dq)\nu(dq')\end{aligned}$$

using successively $\min(1,r) = r\min\left(1,\dfrac{1}{r}\right)$ and $r(q,q') = \dfrac{1}{r(q',q)}$

- Equality on the singular parts $(1-\alpha(q))\,\delta_q(dq')\nu(dq) = (1-\alpha(q'))\delta_{q'}(dq)\nu(dq')$

$$\begin{aligned}\int_{\mathcal{D}}\int_{\mathcal{D}} \phi(q,q')\,(1-\alpha(q))\,\delta_q(dq')\nu(dq) &= \int_{\mathcal{D}} \phi(q,q)(1-\alpha(q))\nu(dq) \\ &= \int_{\mathcal{D}}\int_{\mathcal{D}} \phi(q,q')(1-\alpha(q'))\delta_{q'}(dq)\nu(dq')\end{aligned}$$

- Note: other acceptance ratios $R(r)$ possible as long as $R(r) = rR(1/r)$, but the Metropolis ratio $R(r) = \min(1,r)$ is optimal in terms of asymptotic variance[10]

---

[10] P. Peskun, *Biometrika* (1973)

Metropolis-Hastings algorithm (4)

- Irreducibility: for almost all q_0 and any set $\mathcal{S}$ of positive measure, there exists n such that

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

- Assume also aperiodicity (comes from rejections)

- Pathwise ergodicity[11] $\displaystyle \lim_{N_{\rm iter}\to+\infty} \frac{1}{N_{\rm iter}} \sum_{n=1}^{N_{\rm iter}} A(q^n) = \int_{\mathcal{D}} A(q)\,\nu(dq)$

- Central limit theorem for Markov chains under additional assumptions:

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

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

Metropolis-Hastings algorithm (5)

- 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 0.5)[12]

- Refined Monte Carlo moves such as
  - “non physical” moves
  - parallel tempering
  - replica exchanges
  - Hybrid Monte-Carlo

- A way to stabilize discretization schemes for SDEs

[12] Roberts/Gelman/Gilks (1997), ..., Jourdain/Lelièvre/Miasojedow (2012)

Outline

- **Markov chain methods**
  - Metropolis-Hastings algorithm
- **Stochastic differential equations**
  - General perspective (convergence results, ...)
  - Overdamped Langevin dynamics (Einstein-Schmolukowski)
  - Langevin dynamics
  - Extensions: DPD, Generalized Langevin

Langevin dynamics

- Stochastic perturbation of the Hamiltonian dynamics : friction $\gamma > 0$

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

- Motivations
  - Ergodicity can be proved and is indeed observed in practice
  - Many useful extensions (dissipative particle dynamics, rigorous NPT and μVT samplings, etc)
- Aims
  - Understand the meaning of this equation
  - Understand why it samples the canonical ensemble
  - Implement appropriate discretization schemes
  - Estimate the errors (systematic biases vs. statistical uncertainty)

An intuitive view of the Brownian motion (1)

- Independant Gaussian increments whose variance is proportional to time

$$\forall\, 0 < t_0 \leqslant t_1 \leqslant \cdots \leqslant t_n, \qquad W_{t_{i+1}} - W_{t_i} \sim \mathcal{N}(0, t_{i+1} - t_i)$$

where the increments $W_{t_{i+1}} - W_{t_i}$ are independent

- $G \sim \mathcal{N}(m, \sigma^2)$ distributed according to the probability density

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-m)^2}{2\sigma^2}\right)$$

- The solution of $dq_t = \sigma dW_t$ can be thought of as the limit $\Delta t \to 0$

$$q^{n+1} = q^n + \sigma\sqrt{\Delta t}\, G^n, \qquad G^n \sim \mathcal{N}(0,1) \text{ i.i.d.}$$

where q^n is an approximation of $q_{n\Delta t}$

- Note that $q^n \sim \mathcal{N}(q^0, \sigma n \Delta t)$

- Multidimensional case: $W_t = (W_{1,t}, \dots, W_{d,t})$ where W_i are independent

An intuitive view of the Brownian motion (2)

- Analytical study of the process: law $\psi(t, q)$ of the process at time t

$\rightarrow$ distribution of all possible realizations of q_t for

  - a given initial distribution $\psi(0, q)$, *e.g.* δ_{q^0}
  - and all realizations of the Brownian motion

Averages at time t

$$\mathbb{E}\Big(A(q_t)\Big) = \int_{\mathcal{D}} A(q)\, \psi(t, q)\, dq$$

- Partial differential equation governing the evolution of the law

Fokker-Planck equation

$$\partial_t \psi = \frac{\sigma^2}{2} \Delta \psi$$

Here, simple heat equation $\rightarrow$ "diffusive behavior"

An intuitive view of the Brownian motion (3)

- Proof: Taylor expansion, beware random terms of order $\sqrt{\Delta t}$

$$A\left(q^{n+1}\right) = A\left(q^n + \sigma\sqrt{\Delta t}\,G^n\right)$$
$$= A\left(q^n\right) + \sigma\sqrt{\Delta t}G^n \cdot \nabla A\left(q^n\right) + \frac{\sigma^2 \Delta t}{2}\left(G^n\right)^T\left(\nabla^2 A\left(q^n\right)\right)G^n + \mathrm{O}\Big(\Delta t^{3/2}\Big)$$

Taking expectations (Gaussian increments G^n independent from the current position q^n)

$$\mathbb{E}\left[A\left(q^{n+1}\right)\right] = \mathbb{E}\left[A\left(q^n\right) + \frac{\sigma^2\Delta t}{2}\Delta A\left(q^n\right)\right] + \mathrm{O}\Big(\Delta t^{3/2}\Big)$$

Therefore, $\mathbb{E}\left[\frac{A\left(q^{n+1}\right) - A\left(q^n\right)}{\Delta t} - \frac{\sigma^2}{2}\Delta A\left(q^n\right)\right] \to 0$. On the other hand,

$$\mathbb{E}\left[\frac{A\left(q^{n+1}\right) - A\left(q^n\right)}{\Delta t}\right] \to \partial_t\Big(\mathbb{E}\left[A(q_t)\right]\Big) = \int_{\mathcal{D}} A(q)\partial_t\psi(t,q)\,dq.$$

This leads to

$$0 = \int_{\mathcal{D}} A(q)\partial_t\psi(t,q)\,dq - \frac{\sigma^2}{2}\int_{\mathcal{D}} \Delta A(q)\,\psi(t,q)\,dq = \int_{\mathcal{D}} A(q)\left(\partial_t\psi(t,q) - \frac{\sigma^2}{2}\Delta\psi(t,q)\right)dq$$

This equality holds for all observables A.

General SDEs (1)

• State of the system $X \in \mathbb{R}^d$, m-dimensional Brownian motion, diffusion matrix $\sigma \in \mathbb{R}^{d\times m}$

$$dX_t = b(X_t)\,dt + \sigma(X_t)\,dW_t$$

to be thought of as the limit as $\Delta t \to 0$ of (X^n approximation of $X_{n\Delta t}$)

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

• Generator

$$\mathcal{L} = b(x)\cdot\nabla + \frac{1}{2}\sigma\sigma^T(x) : \nabla^2 = \sum_{i=1}^d b_i(x)\partial_{x_i} + \frac{1}{2}\sum_{i,j=1}^d \left[\sigma\sigma^T(x)\right]_{i,j}\partial_{x_i}\partial_{x_j}$$

• Proceeding as before, it can be shown that

$$\partial_t\Big(\mathbb{E}\left[A(q_t)\right]\Big) = \int_{\mathcal{X}} A\,\partial_t\psi = \mathbb{E}\Big[\left(\mathcal{L}A\right)\left(X_t\right)\Big] = \int_{\mathcal{X}} \left(\mathcal{L}A\right)\psi$$

General SDEs (2)

Fokker-Planck equation

$$\partial_t \psi = \mathcal{L}^* \psi$$

where $\mathcal{L}^*$ is the adjoint of $\mathcal{L}$

$$\int_{\mathcal{X}} (\mathcal{L}A)(x)\, B(x)\, dx = \int_{\mathcal{X}} A(x)\; (\mathcal{L}^* B)(x)\, dx$$

• Invariant measures are stationary solutions of the Fokker-Planck equation

Invariant probability measure $\psi_\infty(x)\,dx$

$$\mathcal{L}^* \psi_\infty = 0, \qquad \int_{\mathcal{X}} \psi_\infty(x)\, dx = 1, \qquad \psi_\infty \geqslant 0$$

• When $\mathcal{L}$ is elliptic (*i.e.* $\sigma\sigma^T$ has full rank: the noise is sufficiently rich), the process can be shown to be irreducible = accessibility property

$$P_t(x, \mathcal{S}) = \mathbb{P}(X_t \in \mathcal{S} \,|\, X_0 = x) > 0$$

General SDEs (3)

- Sufficient conditions for ergodicity
  - irreducibility
  - existence of an invariant probability measure $\psi_\infty(x)\,dx$

Then the invariant measure is unique and

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

- 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 $\sigma_\varphi^2 = 2\,\mathbb{E}\left[\int_0^{+\infty} \widetilde{\varphi}(X_t)\widetilde{\varphi}(X_0)dt\right]$ (proof: later, discrete time setting)

SDEs: numerics (1)

- Numerical discretization: various schemes (Markov chains in all cases)

- Example: Euler-Maruyama

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

- Standard notions of error: fixed integration time $T < +\infty$
  - Strong error $\sup_{0 \leqslant n \leqslant T/\Delta t} \mathbb{E}|X^n - X_{n\Delta t}| \leqslant C \Delta t^p$
  - Weak error: $\sup_{0 \leqslant n \leqslant T/\Delta t} \left| \mathbb{E}\left[\varphi\left(X^n\right)\right] - \mathbb{E}\left[\varphi\left(X_{n\Delta t}\right)\right] \right| \leqslant C \Delta t^p$ (for any φ)
  - "mean error" *vs.* "error of the mean"

- Example: for Euler-Maruyama, weak order 1, strong order $1/2$ (1 when σ constant)

Generating (pseudo) random numbers (1)

- The basis is the generation of numbers uniformly distributed in $[0,1]$

- Deterministic sequences which look like they are random...
  - Early methods: linear congruential generators ("chaotic" sequences)

$$x_{n+1} = ax_n + b \mod c, \qquad u_n = \frac{x_n}{c-1}$$

  - Known defects: short periods, point alignments, etc, which can be (partially) patched by cleverly combining several generators

- More recent algorithms: shift registers, such as Mersenne-Twister

$\rightarrow$ defaut choice in *e.g.* Scilab, available in the GNU Scientific Library

- Randomness tests: various flavors

Generating (pseudo) random numbers (2)

- Standard distributions are obtained from the uniform distribution by...
  - inversion of the cumulative function $F(x) = \int_{-\infty}^{x} f(y)\,dy$ (which is an increasing function from $\mathbb{R}$ to $[0,1]$)

$$X = F^{-1}(U) \sim f(x)\,dx$$

Proof: $\mathbb{P}\{a < X \leqslant b\} = \mathbb{P}\{a < F^{-1}(X) \leqslant b\} = \mathbb{P}\{F(a) < U \leqslant F(b)\} = F(b) - F(a) = \int_a^b f(x)\,dx$

Example: exponential law of density $\lambda \mathrm{e}^{-\lambda x} \mathbf{1}_{\{x \geqslant 0\}}$, $F(x) = \mathbf{1}_{\{x \geqslant 0\}}(1 - \mathrm{e}^{-\lambda x})$, so that $X = -\frac{1}{\lambda} \ln U$

  - change of variables: standard Gaussian $G = \sqrt{-2 \ln U_1} \cos(2\pi U_2)$

Proof: $\mathbb{E}(f(X,Y)) = \frac{1}{2\pi} \int_{\mathbb{R}^2} f(x,y)\, \mathrm{e}^{-(x^2+y^2)/2}\,dx\,dy = \int_0^{+\infty} f\left(\sqrt{r}\cos\theta, \sqrt{r}\sin\theta\right) \frac{1}{2}\mathrm{e}^{-r/2}\,dr\,\frac{d\theta}{2\pi}$

  - using the rejection method

Find a probability density g and a constant $c \geqslant 1$ such that $0 \leqslant f(x) \leqslant cg(x)$. Generate i.i.d. variables $(X^n, U^n) \sim g(x)\,dx \otimes \mathcal{U}[0,1]$, compute $r^n = \frac{f(X^n)}{cg(X^n)}$, and accept X^n if $r^n \geqslant U^n$

SDEs: numerics (2)

- Trajectorial averages: estimator $\Phi_{N_{\rm iter}} = \dfrac{1}{N_{\rm iter}} \displaystyle\sum_{n=1}^{N_{\rm iter}} \varphi(X^n)$

- Numerical scheme ergodic for the probability measure $\psi_{\infty,\Delta t}$

- Two types of errors to compute averages w.r.t. invariant measure
  - Statistical error, quantified using a Central Limit Theorem

$$\Phi_{N_{\rm iter}} = \int \varphi\, \psi_{\infty,\Delta t} + \frac{\sigma_{\Delta t,\varphi}}{\sqrt{N_{\rm iter}}} \mathscr{G}_{N_{\rm iter}}, \qquad \mathscr{G}_{N_{\rm iter}} \sim \mathcal{N}(0,1)$$

  - Systematic errors
    - perfect sampling bias, related to the finiteness of Δt

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

    - finite sampling bias, related to the finiteness of $N_{\rm iter}$

SDEs: numerics (3)

Expression of the asymptotic variance: correlations matter!

$$\sigma^2_{\Delta t,\varphi} = \mathrm{Var}(\varphi) + 2\sum_{n=1}^{+\infty} \mathbb{E}\Big(\widetilde{\varphi}(X^n)\widetilde{\varphi}(X^0)\Big), \qquad \widetilde{\varphi} = \varphi - \int \varphi\, \psi_{\infty,\Delta t}$$

where $\mathrm{Var}(\varphi) = \displaystyle\int_{\mathcal{X}} \widetilde{\varphi}^2 \psi_{\infty,\Delta t} = \int_{\mathcal{X}} \varphi^2 \psi_{\infty,\Delta t} - \left(\int_{\mathcal{X}} \varphi \psi_{\infty,\Delta t}\right)^2$

Proof: compute $N_{\mathrm{iter}}\mathbb{E}\left(\widetilde{\Phi}^2_{N_{\mathrm{iter}}}\right) = \dfrac{1}{N_{\mathrm{iter}}} \displaystyle\sum_{n,m=0}^{N_{\mathrm{iter}}} \mathbb{E}\Big(\widetilde{\varphi}(X^n)\widetilde{\varphi}(X^m)\Big)$

Stationarity $\mathbb{E}\Big(\widetilde{\varphi}(X^n)\widetilde{\varphi}(X^m)\Big) = \mathbb{E}\Big(\widetilde{\varphi}(X^{n-m})\widetilde{\varphi}(X^0)\Big)$ implies

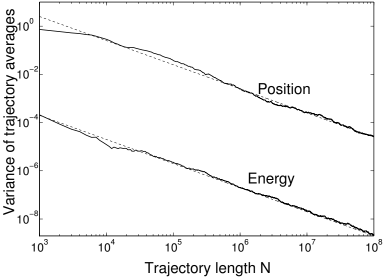
$$N_{\mathrm{iter}}\mathbb{E}\left(\widetilde{\Phi}^2_{N_{\mathrm{iter}}}\right) = \mathbb{E}\Big(\widetilde{\varphi}\left(X^0\right)^2\Big) + 2\sum_{n=1}^{+\infty}\left(1 - \frac{n}{N_{\mathrm{iter}}}\right)\mathbb{E}\Big(\widetilde{\varphi}(X^n)\widetilde{\varphi}(X^0)\Big)$$

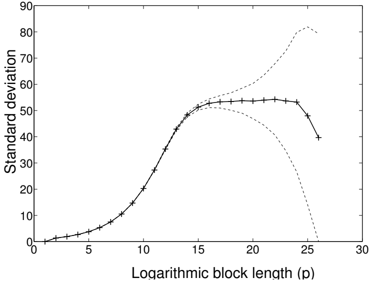
- Useful rewriting: number of correlated steps $\sigma^2_{\Delta t,\varphi} = N_{\mathrm{corr}}\mathrm{Var}(\varphi)$
- Note also that $\sigma^2_{\Delta t,\varphi} \sim \dfrac{2}{\Delta t}\mathbb{E}\left[\displaystyle\int_0^{+\infty} \widetilde{\varphi}(X_t)\widetilde{\varphi}(X_0)\, dt\right]$

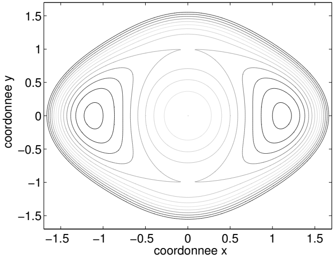
- Estimation of $\sigma_{\Delta t,\varphi}$ by block averaging (batch means)

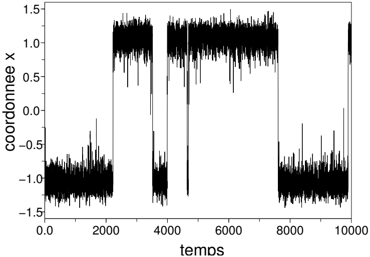
$$\sigma^2_{\Delta t,\varphi} = \lim_{N,M\to+\infty} \frac{N}{M}\sum_{k=1}^{M}\left(\Phi_N^k - \Phi_{NM}^1\right)^2, \quad \Phi_N^k = \frac{1}{N}\sum_{i=(k-1)N+1}^{kN} \varphi(q^i,p^i)$$

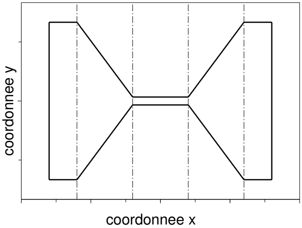
Expected $\Phi_N^k \sim \displaystyle\int_{\mathcal{X}} \varphi\,\psi_{\infty,\Delta t} + \frac{\sigma_{\Delta t,\varphi}}{\sqrt{N}}\mathscr{G}^k$, with $\mathscr{G}^k$ i.i.d.

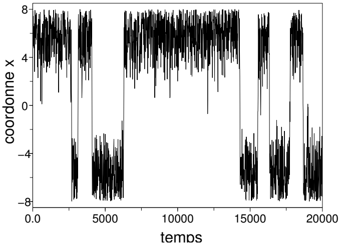












Need for variance reduction techniques! (more on Friday)

Outline

- **Markov chain methods**
  - Metropolis-Hastings algorithm
- **Stochastic differential equations**
  - General perspective (convergence results, ...)
  - Overdamped Langevin dynamics (Einstein-Schmolukowski)
  - Langevin dynamics
  - Extensions: DPD, Generalized Langevin

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 $\mathcal{L} = -\nabla V(q) \cdot \nabla_q + \frac{1}{\beta}\Delta_q$
    - invariance of ψ_0: adjoint $\mathcal{L}^*\varphi = \mathrm{div}_q \left((\nabla V)\varphi + \frac{1}{\beta}\nabla_q \varphi \right)$
    - 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)

Langevin dynamics (1)

- Stochastic perturbation of the Hamiltonian dynamics

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

- γ,σ may be matrices, and may depend on q

- Generator $\mathcal{L} = \mathcal{L}_{\rm ham} + \mathcal{L}_{\rm thm}$

$$\mathcal{L}_{\rm ham} = p^T M^{-1} \nabla_q - \nabla V(q)^T \nabla_p = \sum_{i=1}^{dN} \frac{p_i}{m_i} \partial_{q_i} - \partial_{q_i} V(q) \partial_{p_i}$$

$$\mathcal{L}_{\rm thm} = -p^T M^{-1} \gamma^T \nabla_p + \frac{1}{2} \left(\sigma \sigma^T \right) : \nabla_p^2 \qquad \left(= \frac{\sigma^2}{2} \Delta_p \text{ for scalar } \sigma \right)$$

- Irreducibility can be proved (control argument)

Langevin dynamics (2)

- Invariance of the canonical measure to conclude to ergodicity?

Fluctuation/dissipation relation

$$\sigma\sigma^T = \frac{2}{\beta}\gamma \qquad \text{implies} \qquad \mathcal{L}^*\left(\mathrm{e}^{-\beta H}\right) = 0$$

- Proof for scalar γ, σ: a simple computation shows that

$$\mathcal{L}_{\mathrm{ham}}^* = -\mathcal{L}_{\mathrm{ham}}, \qquad \mathcal{L}_{\mathrm{ham}} H = 0$$

- Overdamped Langevin analogy $\mathcal{L}_{\mathrm{thm}} = \gamma\left(-p^T M^{-1}\nabla_p + \frac{1}{\beta}\Delta_p\right)$

$\rightarrow$ Replace q by p and $\nabla V(q)$ by $M^{-1}p$

$$\mathcal{L}_{\mathrm{thm}}^*\left[\exp\left(-\beta\frac{p^T M^{-1} p}{2}\right)\right] = 0$$

- Conclusion: $\mathcal{L}_{\mathrm{ham}}^*$ and $\mathcal{L}_{\mathrm{thm}}^*$ both preserve $\mathrm{e}^{-\beta H(q,p)}\,dq\,dp$

Langevin dynamics (3)

- Prove exponential convergence of the semigroup $\mathrm{e}^{t\mathcal{L}}$
  - various Banach spaces $E \cap L_0^2(\mu)$
  - Lyapunov techniques[13,14]

$$L_W^\infty(\mathcal{E}) = \left\{ \varphi \,\text{measurable}, \ \left\| \frac{\varphi}{W} \right\|_{L^\infty} < +\infty \right\}$$

  - standard hypocoercive[15] setup $H^1(\mu)$
  - $E = L^2(\mu)$ after hypoelliptic regularization[16] from $H^1(\mu)$
  - Direct $L^2(\mu)$ approach[17]

---

[13] L. Rey-Bellet, *Lecture Notes in Mathematics* (2006)

[14] Hairer and Mattingly, *Progr. Probab.* **63** (2011)

[15] Villani (2009) and before Talay (2002), Eckmann/Hairer (2003), Hérau/Nier (2004)

[16] F. Hérau, *J. Funct. Anal.* **244**(1), 95-118 (2007)

[17] Dolbeault, Mouhot and Schmeiser (2009, 2015)

Numerical integration of the Langevin dynamics (1)

- Splitting strategy: Hamiltonian part + fluctuation/dissipation

$$\left\{ \begin{aligned} dq_t &= \ M^{-1}\, p_t\, dt \\ dp_t &= -\nabla V(q_t)\, dt \end{aligned} \right. \qquad \oplus \qquad \left\{ \begin{aligned} dq_t &= 0 \\ dp_t &= -\gamma M^{-1} p_t\, dt + \sqrt{\frac{2\gamma}{\beta}}\, dW_t \end{aligned} \right.$$

- Hamiltonian part integrated using a Verlet scheme

- Analytical integration of the fluctuation/dissipation part

$$d\left(\mathrm{e}^{\gamma M^{-1} t} p_t\right) = \mathrm{e}^{\gamma M^{-1} t}\left(dp_t + \gamma M^{-1} p_t\, dt\right) = \sqrt{\frac{2\gamma}{\beta}} \mathrm{e}^{\gamma M^{-1} t}\, dW_t$$

so that

$$p_t = \mathrm{e}^{-\gamma M^{-1} t}\, p_0 + \sqrt{\frac{2\gamma}{\beta}} \int_0^t \mathrm{e}^{-\gamma M^{-1}(t-s)}\, dW_s$$

It can be shown that $\displaystyle\int_0^t f(s)\, dW_s \sim \mathcal{N}\left(0, \int_0^t f(s)^2 ds\right)$

Numerical integration of the Langevin dynamics (2)

- Trotter splitting (define $\alpha_{\Delta t} = \mathrm{e}^{-\gamma M^{-1}\Delta t}$, choose $\gamma M^{-1}\Delta t \sim 0.01 - 1$)

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

Error estimate on the invariant measure $\mu_{\Delta t}$ of the numerical scheme

There exist a function f such that, for any smooth observable ψ,

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

- Strang splitting more expensive and not more accurate

Some extensions (1)

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

Dissipative Particle Dynamics

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

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

- Invariance of the canonical measure, preservation of $\displaystyle\sum_{i=1}^{N} p_i$
- Ergodicity is an issue[18]
- Numerical scheme: splitting strategy[19]

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

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

Some extensions (2)

• Mori-Zwanzig derivation[20] from a generalized Hamiltonian system: particle coupled to harmonic oscillators with a distribution of frequencies

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

$$\left\{\begin{aligned} dq &= p_t\,dt \\ dp_t &= -\nabla V(q_t)\,dt + R_t\,dt \\ \varepsilon\,dR_t &= -R_t\,dt - \gamma p_t\,dt + \sqrt{\frac{2\gamma}{\beta}}\,dW_t \end{aligned}\right.$$

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

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

• Ergodicity proofs (hypocoercivity): as for the Langevin equation[21]

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

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