Langevin dynamics
at equilibrium and out of equilibrium
from hypocoercivity to efficient sampling

Gabriel STOLTZ

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

In collaboration with A. Iacobucci, S. Olla, S. Redon, J. Roussel, E. Vanden-Eijnden; and currently B. Leimkuhler and M. Sachs

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

Outline

- **A quick introduction to computational statistical physics**
- **Equilibrium Langevin dynamics**
  - Various convergence results
  - A focus on the approach by Dolbeault, Mouhot and Schmeiser
- **Various extensions/modifications**
  - Modified kinetic energies
  - Rates of convergence for nonequilibrium Langevin dynamics
  - Constructing control variates
  - Proving the convergence of Temperature Accelerated Dynamics

A quick introduction to computational statistical physics

Computational statistical physics

- *Predict macroscopic properties of matter from its microscopic description*

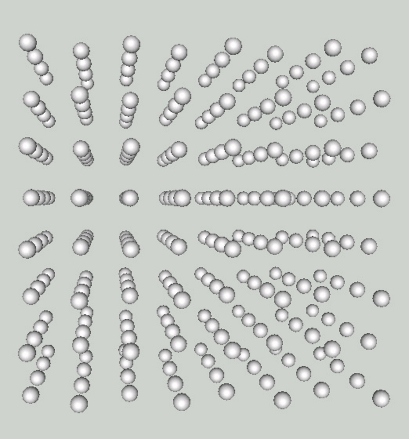
- **Microstate**
  - positions $q = (q_1, \dots, q_N)$ and momenta $p = (p_1, \dots, p_N)$
  - energy $V(q) + \displaystyle\sum_{i=1}^{N} \frac{p_i^2}{2m_i}$

- **Macrostate**
  - described by a probability measure μ
  - constraints fixed exactly or in average (number of particles, volume, energy)

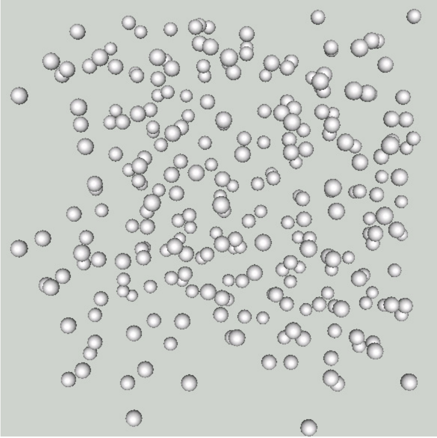
- **Properties :**
  - static $\langle A \rangle = \displaystyle\int_{\mathcal{E}} A(q,p)\, \mu(dq\, dp)$ (equation of state, heat capacity,...)
  - dynamic (transport coefficient, transition pathway, etc)

Examples of molecular systems (1)

What is the melting temperature of Argon?



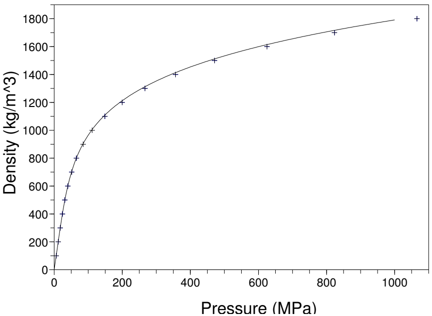
(a) Solid Argon (low temperature)

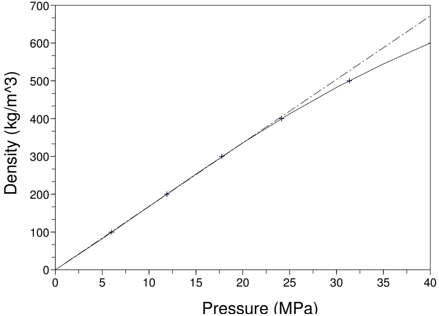


(b) Liquid Argon (high temperature)

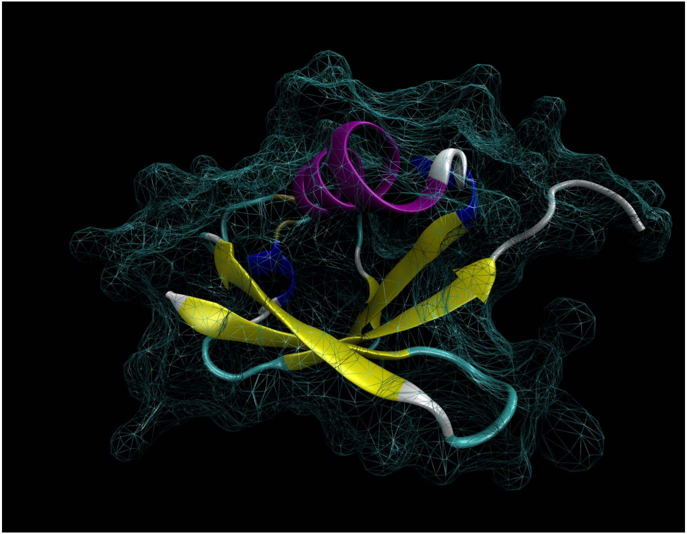
Examples of molecular systems (2)

Equation of state of Argon: density as a function of pressure at fixed temperature $T = 300$ K





Examples of molecular systems (3)



Ubiquitin: what is its structure? What are its conformational changes?

Some orders of magnitude...

- **Physical orders of magnitude**
  - distances $\sim 1\ \mathring{A} = 10^{-10}$ m
  - energy per particle $\sim k_{\mathrm{B}} T \sim 4 \times 10^{-21}$ J at 300 K
  - atomic masses $\sim 10^{-26}$ kg
  - typical times $\sim 10^{-15}$ s
  - number of particles $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

- **"Standards" simulations**
  - 10^6 particles ["heroic": from 10^9 particles on]
  - total time: (fraction of) ns ["heroic": (fraction of) μs]

- Computation of high dimensional integrals...

$\to$ Ergodic methods $\displaystyle \frac{1}{t}\int_0^t A(q_s, p_s)\, ds \xrightarrow[t \to +\infty]{} \langle A \rangle$

Equilibrium Langevin dynamics

Langevin dynamics (1)

- Positions $q \in \mathcal{D} = (L\mathbb{T})^d$ or $\mathbb{R}^d$ and momenta $p \in \mathbb{R}^d$

$\rightarrow$ phase-space $\mathcal{E} = \mathcal{D} \times \mathbb{R}^d$

- Hamiltonian $H(q,p) = V(q) + \frac{1}{2} p^T M^{-1} p$ (more general kinetic energies $U(p)$ can be considered[1])

Stochastic perturbation of the Hamiltonian dynamics

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

- Friction $\gamma > 0$ (could be a position-dependent matrix)

[1] Redon, Stoltz, Trstanova, *J. Stat. Phys.* (2016)

Langevin dynamics (2)

- Evolution semigroup $\left(\mathrm{e}^{t\mathcal{L}}\varphi\right)(q,p) = \mathbb{E}\left[\varphi(q_t,p_t)\,\middle|(q_0,p_0)=(q,p)\right]$

- Generator of the dynamics $\mathcal{L}$

$$\frac{d}{dt}\left(\mathbb{E}\left[\varphi(q_t,p_t)\,\middle|(q_0,p_0)=(q,p)\right]\right) = \mathbb{E}\left[(\mathcal{L}\varphi)(q_t,p_t)\,\middle|(q_0,p_0)=(q,p)\right]$$

Generator of the Langevin dynamics $\mathcal{L} = \mathcal{L}_{\mathrm{ham}} + \gamma\mathcal{L}_{\mathrm{FD}}$

$$\mathcal{L}_{\mathrm{ham}} = p^T M^{-1}\nabla_q - \nabla V^T \nabla_p, \qquad \mathcal{L}_{\mathrm{FD}} = -p^T M^{-1}\nabla_p + \frac{1}{\beta}\Delta_p$$

- Existence and uniqueness of the invariant measure characterized by

$$\forall \varphi \in C_0^\infty(\mathcal{E}), \qquad \int_{\mathcal{E}} \mathcal{L}\varphi \, d\mu = 0$$

- Here, canonical measure

$$\mu(dq\,dp) = Z^{-1}\mathrm{e}^{-\beta H(q,p)}\,dq\,dp = \nu(dq)\,\kappa(dp)$$

Fokker–Planck equations

- Evolution of the law $\psi(t,q,p)$ of the process at time $t \geqslant 0$

$$\frac{d}{dt}\left(\int_{\mathcal{E}} \varphi\,\psi(t)\right) = \int_{\mathcal{E}} (\mathcal{L}\varphi)\,\psi(t)$$

- Fokker–Planck equation (with $\mathcal{L}^\dagger$ adjoint of $\mathcal{L}$ on $L^2(\mathcal{E})$)

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

- It is convenient to work in $L^2(\mu)$ with $f(t) = \psi(t)/\mu$
  - denote the adjoint of $\mathcal{L}$ on $L^2(\mu)$ by $\mathcal{L}^*$

$$\mathcal{L}^* = -\mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{FD}}$$

  - Fokker–Planck equation $\partial_t f = \mathcal{L}^* f$
- Convergence results for $\mathrm{e}^{t\mathcal{L}}$ on $L^2(\mu)$ are very similar to the ones for $\mathrm{e}^{t\mathcal{L}^*}$

Hamiltonian and overdamped limits

- As $\gamma \to 0$, the Hamiltonian dynamics is recovered
- Overdamped limit $\gamma \to +\infty$ or $m \to 0$

$$q_{\gamma t} - q_0 = -\frac{1}{\gamma} \int_0^{\gamma t} \nabla V(q_s)\, ds + \sqrt{\frac{2}{\gamma\beta}} W_{\gamma t} - \frac{1}{\gamma}(p_{\gamma t} - p_0)$$

which converges to the solution of $dQ_t = -\nabla V(Q_t)\, dt + \sqrt{\frac{2}{\beta}}\, dW_t$

- In both cases, slow convergence to equilibrium
  - it takes time to change energy levels in the Hamiltonian limit[2]
  - for m fixed, time has to be rescaled by a factor γ

[2] Hairer and Pavliotis, *J. Stat. Phys.*, **131**(1), 175-202 (2008)

Ergodicity results (1)

- Almost-sure convergence[3] of ergodic averages $\widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(q_s, p_s)\, ds$

- Asymptotic variance of ergodic averages

$$\sigma_\varphi^2 = \lim_{t\to+\infty} t\mathbb{E}\left[\widehat{\varphi}_t^2\right] = 2\int_{\mathcal{E}} \left(-\mathcal{L}^{-1}\Pi_0\varphi\right)\Pi_0\varphi\, d\mu$$

where $\Pi_0\varphi = \varphi - \mathbb{E}_\mu(\varphi)$

- A central limit theorem holds[4] when the equation has a solution in $L^2(\mu)$

Poisson equation in $L^2(\mu)$

$$-\mathcal{L}\Phi = \Pi_0\varphi$$

- Well-posedness of such equations? Hypoelliptic operator

---

[3] Kliemann, *Ann. Probab.* **15**(2), 690-707 (1987)

[4] Bhattacharya, *Z. Wahrsch. Verw. Gebiete* **60**, 185–201 (1982)

Ergodicity results (2)

- Invertibility of $\mathcal{L}$ on subsets of $L_0^2(\mu) = \left\{ \varphi \in L^2(\mu) \,\middle|\, \int_{\mathcal{E}} \varphi \, d\mu = 0 \right\}$?

$$-\mathcal{L}^{-1} = \int_0^{+\infty} \mathrm{e}^{t\mathcal{L}} \, dt$$

- Prove exponential convergence of the semigroup $\mathrm{e}^{t\mathcal{L}}$
  - various Banach spaces $E \cap L_0^2(\mu)$
  - Lyapunov techniques[5,6,7] $L_W^\infty(\mathcal{E}) = \left\{ \varphi \text{ measurable}, \left\| \frac{\varphi}{W} \right\|_{L^\infty} < +\infty \right\}$
  - standard hypocoercive[8] setup $H^1(\mu)$
  - $E = L^2(\mu)$ after hypoelliptic regularization[9] from $H^1(\mu)$
  - coupling arguments[10]

---

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

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

[7] Mattingly, Stuart and Higham, *Stoch. Proc. Appl.* (2002)

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

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

[10] A. Eberle, A. Guillin and R. Zimmer, *arXiv preprint* **1703.01617** (2017)

Direct $L^2(\mu)$ approach

- Assume that the potential V is smooth and[11,12]
  - the marginal measure ν satisfies a Poincaré inequality
$$\|\Pi_0\varphi\|^2_{L^2(\nu)} \leqslant \frac{1}{C_\nu}\|\nabla_q\varphi\|^2_{L^2(\nu)}.$$
  - there exist $c_1 > 0$, $c_2 \in [0,1)$ and $c_3 > 0$ such that V satisfies
$$\Delta V \leqslant c_1 + \frac{c_2}{2}|\nabla V|^2, \quad |\nabla^2 V| \leqslant c_3\left(1 + |\nabla V|\right).$$

There exist $C > 0$ and $\lambda_\gamma > 0$ such that, for any $\varphi \in L^2_0(\mu)$,
$$\forall t \geqslant 0, \qquad \left\|\mathrm{e}^{t\mathcal{L}}\varphi\right\|_{L^2(\mu)} \leqslant C\mathrm{e}^{-\lambda_\gamma t}\|\varphi\|_{L^2(\mu)}.$$

with convergence rate of order $\min(\gamma,\gamma^{-1})$: there exists $\overline{\lambda} > 0$ such that
$$\lambda_\gamma \geqslant \overline{\lambda}\min(\gamma,\gamma^{-1}).$$

[11] Dolbeault, Mouhot and Schmeiser, *C. R. Math. Acad. Sci. Paris* (2009)

[12] Dolbeault, Mouhot and Schmeiser, *Trans. AMS*, **367**, 3807–3828 (2015)

Sketch of proof

- Modified square norm $\mathcal{H}[\varphi] = \frac{1}{2}\|\varphi\|^2 - \varepsilon \langle A\varphi, \varphi \rangle$ for $\varepsilon \in (-1,1)$ and

$$A = \Big(1 + (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)\Big)^{-1}(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*, \qquad \Pi_p\varphi = \int_{\mathbb{R}^D} \varphi \, d\kappa$$

- $A = \Pi_p A(1-\Pi_p)$ and $\mathcal{L}_{\mathrm{ham}}A$ are bounded so that $\mathcal{H} \sim \|\cdot\|^2_{L^2(\mu)}$

Coercivity in the scalar product $\langle\langle \cdot, \cdot \rangle\rangle$ induced by $\mathcal{H}$

$$\mathscr{D}[\varphi] := \langle\langle -\mathcal{L}\varphi, \varphi \rangle\rangle \geqslant \widetilde{\lambda}_\gamma \|\varphi\|^2,$$

- Idea: control of $\|(1-\Pi_p)\varphi\|^2$ by $\langle -\mathcal{L}_{\mathrm{FD}}\varphi, \varphi\rangle$ (Poincaré); for $\|\Pi_p\varphi\|^2$,

$$\|\mathcal{L}_{\mathrm{ham}}\Pi_p\varphi\|^2 \geqslant \frac{DC_\nu}{\beta m}\|\Pi_p\varphi\|^2, \qquad \text{hence } A\mathcal{L}_{\mathrm{ham}}\Pi_p \geqslant \lambda_{\mathrm{ham}}\Pi_p$$

- Gronwall inequality $\frac{d}{dt}\left(\mathcal{H}\left[\mathrm{e}^{t\mathcal{L}}\varphi\right]\right) = -\mathscr{D}\left[\mathrm{e}^{t\mathcal{L}}\varphi\right] \leqslant -\frac{2\widetilde{\lambda}_\gamma}{1+\varepsilon}\mathcal{H}\left[\mathrm{e}^{t\mathcal{L}}\varphi\right]$

Extensions/modifications

Using modified kinetic energies

- General kinetic energy function $U(p)$ in the Langevin dynamics
  - heavy/light tails
  - ∇U vanishes on open sets (ARPS)

- Can still prove convergence results (Lyapunov[13] or hypocoercivity[14]) although the dynamics is not hypoelliptic

- Dedicated numerical schemes to integrate non-globally Lipschitz ∇U
  - Strang splitting between Hamiltonian part and fluctuation/dissipation
  - Metropolization of FD using a HMC-like scheme (weak order 3/2)

- Possible reduction of metastability by good choices of U (e.g. $U = V$ in low dimensions)

---

[13] S. Redon, G. Stoltz and Z. Trstanova, *J. Stat. Phys.* (2016)

[14] G. Stoltz and Z. Trstanova, accepted in *Multiscale Model. Sim.* (2018)

Rates of convergence for nonequilibrium Langevin dynamics

- Compact position space $\mathcal{D} = (2\pi\mathbb{T})^d$, constant force $|F| = 1$

Langevin dynamics perturbed by a constant force term

$$\begin{cases} dq_t = \dfrac{p_t}{m} dt, \\ dp_t = (-\nabla V(q_t) + \tau F) dt - \gamma \dfrac{p_t}{m} dt + \sqrt{\dfrac{2\gamma}{\beta}} dW_t, \end{cases}$$

- Non-zero velocity in the direction F is expected in the steady-state
- F **does not derive from the gradient of a periodic function**
    - of course, $F = -\nabla W_F(q)$ with $W_F(q) = -F^T q$
    - ...but W_F is not periodic!

Rates of convergence for nonequilibrium Langevin dynamics

- Lyapunov approaches are non-perturbative but also non-quantitative
- Suboptimal results by the standard hypocoercive approach in $H^1(\mu)$
$\rightarrow$ nonequilibrium perturbation[15] of direct $L^2(\mu)$ strategy
- Invariant measure $\psi_\eta = h_\tau \mu$ with $h_\tau \in L^2(\mu)$ for $|\tau|$ small

Uniform rates for nonequilibrium perturbations

There exist $C, \delta_* > 0$ such that, for any $\delta \in [0, \delta^*]$, there is $\overline{\lambda}_\delta > 0$ for which, for all $\gamma \in (0, +\infty)$ and all $\tau \in [-\delta \min(\gamma, 1), \delta \min(\gamma, 1)]$,

$$\left\| \mathrm{e}^{t\mathcal{L}_{\gamma,\tau}^*} f - h_\tau \right\|_{L^2(\mu)} \leqslant C \mathrm{e}^{-\overline{\lambda}_\delta \min(\gamma, \gamma^{-1}) t} \| f - h_\tau \|_{L^2(\mu)}.$$

- As a corollary: lower bounds on the spectral gap of order $\min(\gamma, \gamma^{-1})$
$\rightarrow$ can be checked numerically [16]

---

[15] E. Bouin, F. Hoffmann, and C. Mouhot, *arXiv preprint* **1605.04121**

[16] A. Iacobucci, S. Olla and G. Stoltz, to appear in *Ann. Math. Quebec* (2017)

Constructing control variates

• The computation of transport coefficients by nonequilibrium steady-state techniques involves the computations of quantities of the form

$$\frac{\mathbb{E}_\eta(R)}{\eta}, \qquad |\eta| \ll 1$$

$\rightarrow$ **Magnification of the statistical error**

• Typical cases: $\mathcal{L}_\eta = \mathcal{L}_0 + \eta\widetilde{\mathcal{L}}$

- nonequilibrium perturbation
- coupling parameter between otherwise independent systems
- anharmonic part in an otherwise linear dynamics

• Control variate idea

- note that $\mathbb{E}_\eta(R - \mathcal{L}_\eta\Phi) = \mathbb{E}_\eta(R)$ for all Φ
- ...but it may happen that $\mathrm{Var}_\eta(R - \mathcal{L}_\eta\Phi) \ll \mathrm{Var}_\eta(R)$
- Optimal choice $\Phi = \mathcal{L}_\eta^{-1}(R - \mathbb{E}_\eta(R))$ unknown
- approximate it by $\mathcal{L}_0^{-1}(R - \mathbb{E}_0(R))$

Constructing control variates

- **Error estimates for Galerkin discretization**[17]
  - spectral basis (weighted Fourier modes + Hermites polynomials)
  - (non-)conformal formulation (basis functions of mean 0 or not)
  - estimation of consistency and approximation errors
- **Error estimates on the variance for approximate control variate**[18]
  - modified estimator based on $R + \mathcal{L}_\eta \Phi_0$
  - the asymptotic variance is of order η^2 when Φ_0 is exactly computed
  - applications to (i) 1D nonequilibrium Langevin, (2) thermal transport in 1D chains, (3) dimer in WCA solvent

[17] J. Roussel and G. Stoltz, to appear in *M2AN* (*arXiv preprint* **1702.04718**)
[18] J. Roussel and G. Stoltz, *arXiv preprint* **1712.08022**

Proving the convergence of TAMD

- Additional variable z (e.g. free energy computation), higher temperature $\bar{\beta}^{-1}$, acceleration factor $\delta^{-1} \to$ **Nonequilibrium** dynamics

$$\left\{ \begin{aligned} dq_t &= \delta^{-1} M^{-1} p_t \, dt, \\ dp_t &= -\delta^{-1} \nabla_q U_\kappa(q_t, z_t) \, dt - \delta^{-1} \gamma \, M^{-1} p_t \, dt + \sqrt{2\gamma(\beta\delta)^{-1}} \, dW_t^p, \\ dz_t &= -\nabla_z U_\kappa(q_t, z_t) \, dt + \sqrt{2\bar{\beta}^{-1}} \, dW_t^z. \end{aligned} \right.$$

- Effective dynamics as $\delta \to 0$

$$d\bar{z}_t = -\nabla_z A(\bar{z}_t) \, dt + \sqrt{2\bar{\beta}^{-1}} \, dW_t^z, \qquad \mathrm{e}^{-\beta A(z)} = \int_{\mathcal{D}_q} \mathrm{e}^{-\beta U(q,z)} dq$$

- **Convergence results**[19]
  - Exponential decay of semigroup, rate close to effective dynamics
  - Expansion of the invariant measure in powers of δ
  - Asymptotic variance agrees at first order in δ with effective dynamics

[19] G. Stoltz and E. Vanden-Eijnden, *arXiv preprint* **1708.08800**