Computational statistical physics and hypocoercivity

Gabriel STOLTZ

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

Research funded by ANR SINEQ and ERC Synergy EMC2

ICMS, Edinburgh, September 2024

Outline of the talk

- **Computational statistical physics**
  - A general perspective
  - Langevin dynamics and its overdamped limit
  - Error estimates to compute average properties
- **Longtime convergence of overdamped Langevin dynamics**
  - Poincaré inequalities
  - Estimates on the asymptotic variance
- **Longtime convergence of "hypocoercive" ODEs**
- **Longtime convergence of Langevin dynamics**
  - The need for a modified scalar product
  - One L^2-hypocoercive approach for Langevin dynamics
  - Direct estimates on the variance
  - Space-time approaches

General references (1)

- Computational Statistical Physics
  - D. Frenkel and B. Smit, *Understanding Molecular Simulation, From Algorithms to Applications* (2002)
  - M. Tuckerman, *Statistical Mechanics: Theory and Molecular Simulation* (2010)
  - M. P. Allen and D. J. Tildesley, *Computer simulation of liquids* (2017)
  - D. C. Rapaport, *The Art of Molecular Dynamics Simulations* (1995)
  - T. Schlick, *Molecular Modeling and Simulation* (2002)

- Computational Statistics [my personal references... many more out there!]
  - J. Liu, *Monte Carlo Strategies in Scientific Computing*, Springer, 2008
  - W. R. Gilks, S. Richardson and D. J. Spiegelhalter (eds), *Markov Chain Monte Carlo in Practice* (Chapman & Hall, 1996)

- Machine learning and sampling
  - C. Bishop, *Pattern Recognition and Machine Learning* (Springer, 2006)
  - K.P. Murphy, *Probabilistic Machine Learning: An Introduction* (MIT Press, 2022)

General references (2)

- 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)

- Convergence of Markov chains
  - S. Meyn and R. Tweedie, *Markov Chains and Stochastic Stability* (Cambridge University Press, 2009)
  - R. Douc, E. Moulines, P. Priouret and P. Soulier, *Markov Chains* (Springer, 2018)

Computational statistical physics

Statistical physics (1)

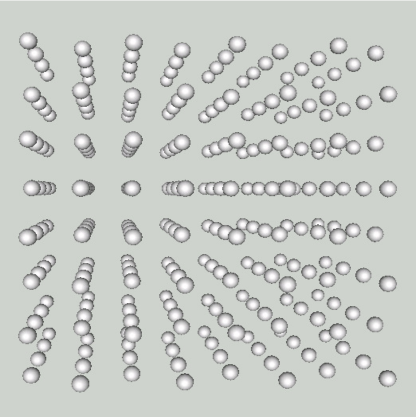
- **Aims of computational statistical physics**
  - numerical microscope
  - computation of average properties, static or dynamic

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

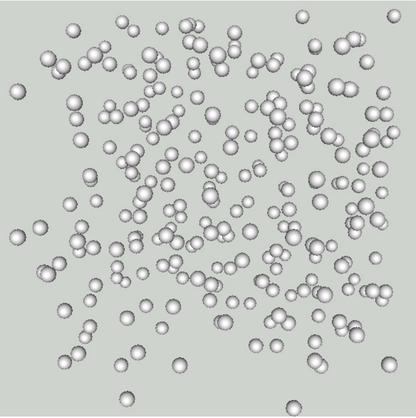
- **"Standard" simulations**
  - 10^6 particles ["world records": around 10^9 particles]
  - integration time: (fraction of) ns ["world records": (fraction of) μs]

Statistical physics (2)

What is the melting temperature of argon?



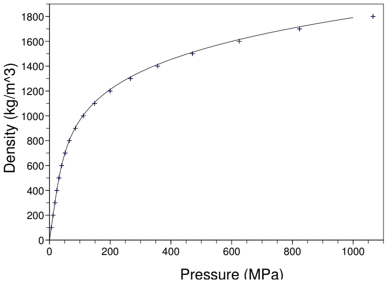
(a) Solid argon (low temperature)

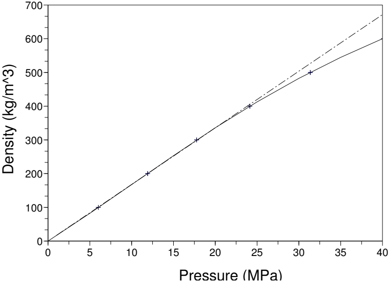


(b) Liquid argon (high temperature)

Statistical physics (3)

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

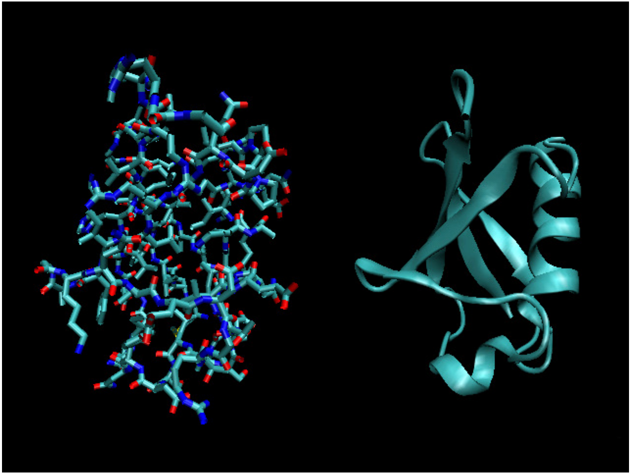




Equation of state (pressure/density diagram) for argon at $T = 300$ K

Statistical physics (4)

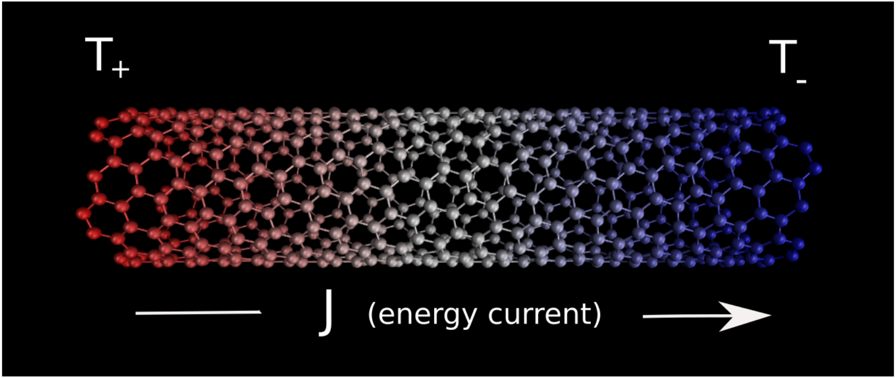
What is the structure of the protein? What are its typical conformations, and what are the transition pathways from one conformation to another?



Statistical physics (5)

Computation of transport coefficient, *e.g.* thermal conductivity

$$J = -\kappa \nabla T$$



Statistical physics (6)

• Microstate of a classical system of N particles:

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

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

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

All the physics is contained in V

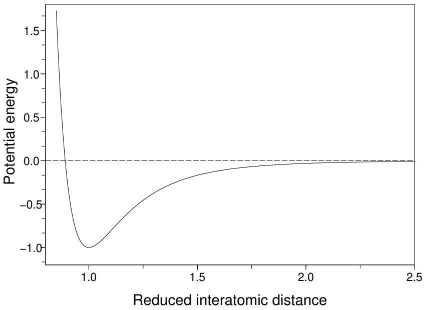
- ideally derived from quantum mechanical computations
- in practice, empirical potentials for large scale calculations

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

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

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

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



Statistical physics (8)

- Macrostate of the system described by a probability measure

Equilibrium thermodynamic properties (pressure,...)

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

- 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 $N \to +\infty$)

- Canonical ensemble = measure on (q,p), average energy fixed 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 $\displaystyle\int_{\mathcal{E}} H\, \rho\, dq\, dp = E_0$

Langevin dynamics (1)

Computation of high-dimensional integrals... Ergodic averages

$$\int_{\mathcal{E}} \varphi \, d\mu = \lim_{t \to +\infty} \widehat{\varphi}_t, \qquad \widehat{\varphi}_t = \frac{1}{t} \int_0^t \varphi(q_s, p_s) \, ds$$

Almost-sure convergence (Kliemann, *Ann. Probab.* 1987)

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

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

- Hamiltonian $H(q,p) = V(q) + \frac{1}{2} p^\top M^{-1} p$

Stochastic perturbation of the Hamiltonian dynamics (friction $\gamma > 0$)

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

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

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

- Existence and uniqueness of the invariant measure characterized by

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

- 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 $\mathcal{L}^*$

$$\mathcal{L}^* = -\mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{FD}}, \quad \mathcal{L}_{\mathrm{FD}} = -\frac{1}{\beta}\sum_{i=1}^d \partial_{p_i}^* \partial_{p_i}, \quad \mathcal{L}_{\mathrm{ham}} = \frac{1}{\beta}\sum_{i=1}^d \partial_{p_i}^* \partial_{q_i} - \partial_{q_i}^* \partial_{p_i}$$

  - Fokker–Planck equation $\partial_t f = \mathcal{L}^* f$

Convergence results for $\mathrm{e}^{t\mathcal{L}}$ on $L^2(\mu)$ very similar to the ones for $\mathrm{e}^{t\mathcal{L}^*}$

Hamiltonian and overdamped limits

- As $\gamma \to 0$, the Hamiltonian dynamics is recovered

$$\frac{d}{dt}\mathbb{E}\left[H(q_t,p_t)\right] = -\gamma\left(\mathbb{E}\left[p_t^\top M^{-2} p_t\right] - \frac{1}{\beta}\mathrm{Tr}(M^{-1})\right) dt$$

Time $\sim \gamma^{-1}$ to change energy levels in this limit[1]

- Overdamped limit $\gamma \to +\infty$ with $M = \mathrm{Id}$: rescaling of time γt

$$\begin{aligned} q_{\gamma t} - q_0 &= -\frac{1}{\gamma}\int_0^{\gamma t} \nabla V(q_s)\, ds + \sqrt{\frac{2}{\gamma\beta}} W_{\gamma t} - \frac{1}{\gamma}\left(p_{\gamma t} - p_0\right) \\ &= -\int_0^t \nabla V(q_{\gamma s})\, ds + \sqrt{2\beta^{-1}} B_t - \frac{1}{\gamma}\left(p_{\gamma t} - p_0\right) \end{aligned}$$

which converges to the solution of $dQ_t = -\nabla V(Q_t)\, dt + \sqrt{2\beta^{-1}}\, dB_t$

- In both cases, slow convergence, with rate scaling as $\min\left(\gamma, \gamma^{-1}\right)$

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

Estimating average properties: Types of errors

Estimators of $\mathbb{E}_\mu(\varphi)$

$$\widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(q_s,p_s)\,ds, \qquad \widehat{\varphi}_{\Delta t}^N = \frac{1}{N}\sum_{n=1}^N \varphi(q^n,p^n)$$

Statistical error (variance of the estimator)

- dictated by the central limit theorem for continuous dynamics
- discrete dynamics: asymptotic variance coincides[2] at order Δt^α

Bias (expectation of the estimator)

- finite time integration time $\rightarrow$ bias $\mathrm{O}\left(\frac{1}{t}\right)$
- discretization of the dynamics $\rightarrow$ bias $\mathrm{O}(\Delta t^\alpha)$

[2] B. Leimkuhler, C. Matthews and G. Stoltz, *IMA J. Numer. Anal.* (2016)

Finite time integration bias

Bias $\mathrm{O}(1/t)$, typically smaller than statistical error $\mathrm{O}(1/\sqrt{t})$

$$\left|\mathbb{E}\left(\widehat{\varphi}_t\right) - \mathbb{E}_\mu(\varphi)\right| \leqslant \frac{K}{t}$$

Key equality for the proofs: introduce $-\mathcal{L}\Phi = \Pi\varphi := \varphi - \mathbb{E}_\mu(\varphi)$, write

$$\begin{aligned} \widehat{\varphi}_t - \mathbb{E}_\mu(\varphi) &= \frac{1}{t}\int_0^t \Pi\varphi(q_s,p_s)\,ds \\ &= \frac{\Phi(q_0,p_0)-\Phi(q_t,p_t)}{t} + \sqrt{\frac{2\gamma}{\beta}}\,\frac{1}{t}\int_0^t \nabla_p\Phi(q_s,p_s)^\top dW_s \end{aligned}$$

with Ito calculus $d\Phi(q_s,p_s) = \mathcal{L}\Phi(q_s,p_s) + \sqrt{2\gamma\beta^{-1}}\nabla_p\Phi(q_s,p_s)^\top dW_s$

Also allows to prove CLT: martingale part dominant, with variance

$$\frac{2\gamma}{\beta t^2}\int_0^t \mathbb{E}\left[\left|\nabla_p\Phi(q_s,p_s)\right|^2\right] ds \sim \frac{2\gamma}{\beta t}\int_{\mathcal{E}} \left|\nabla_p\Phi\right|^2\,d\mu = \frac{2\gamma}{\beta t}\int_{\mathcal{E}} \Phi(-\mathcal{L}\Phi)\,d\mu$$

Statistical error (1)

- **Asymptotic variance** $\sigma_\varphi^2 = \lim_{t\to+\infty} t \, \mathrm{Var}_\mu(\widehat{\varphi}_t)$: with $\Pi\varphi = \varphi - \int_{\mathcal{E}} \varphi \, d\mu$,

$$\sigma_\varphi^2 = \lim_{t\to+\infty} \int_0^t \left(1 - \frac{s}{t}\right) \mathbb{E}_\mu \left[\Pi\varphi(q_t,p_t)\Pi\varphi(q_0,p_0)\right] ds$$

$$= 2\int_0^{+\infty} \int_{\mathcal{E}} (\mathrm{e}^{s\mathcal{L}}\Pi\varphi)\Pi\varphi \, d\mu \, ds = 2\int_{\mathcal{E}} (-\mathcal{L}^{-1}\Pi\varphi)\Pi\varphi \, d\mu$$

Well-defined provided $-\mathcal{L}\Phi = \Pi\varphi$ has a solution in $L_0^2(\mu) = \Pi L^2(\mu)$

A Central Limit Theorem holds[3] in this case: $\boxed{\widehat{\varphi}_t - \mathbb{E}_\mu(\varphi) \simeq \frac{\sigma_\varphi}{\sqrt{t}}\mathcal{G}}$

- Sufficient condition: integrability of the semigroup, *e.g.*

$$\left\|\mathrm{e}^{t\mathcal{L}}\right\|_{\mathcal{B}(L_0^2(\mu))} \leqslant C\mathrm{e}^{-\lambda t}, \qquad -\mathcal{L}^{-1} = \int_0^{+\infty} \mathrm{e}^{s\mathcal{L}} \, ds$$

Question: dependence of σ_φ^2 on friction γ, potential V, ...

[3] R. N. Bhattacharya, *Z. Wahrsch. Verw. Gebiete* (1982)

Statistical error (2)

Prove exponential convergence of the semigroup $\mathrm{e}^{t\mathcal{L}}$ on $E \subset L^2_0(\mu)$

- Lyapunov techniques[4] $L^\infty_{\mathscr{K}}(\mathcal{E}) = \left\{ \varphi \text{ measurable, } \sup \left| \frac{\varphi}{\mathscr{K}} \right| < +\infty \right\}$
- "historic" hypocoercive[5] setup $H^1(\mu)$
- $L^2(\mu)$ after hypoelliptic regularization[6] from $H^1(\mu)$
- direct transfer from $H^1(\mu)$ to $L^2(\mu)$ by spectral argument[7]
- directly[8] $L^2(\mu)$ (recently[9] Poincaré using $\partial_t - \mathcal{L}_{\mathrm{ham}}$)
- coupling arguments[10]
- direct estimates on the resolvent using Schur complements[11]

Rate of convergence $\min\left(\gamma, \gamma^{-1}\right)$ so variance $\sim \max\left(\gamma, \gamma^{-1}\right)$

---

[4] Wu ('01); Mattingly/Stuart/Higham ('02); Rey-Bellet ('06); Hairer/Mattingly ('11)

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

[6] Hérau, *J. Funct. Anal.* (2007)

[7] Deligiannidis/Paulin/Doucet, *Ann. Appl. Probab.* (2020)

[8] Hérau (2006), Dolbeault/Mouhot/Schmeiser (2009, 2015)

[9] Albritton/Armstrong/Mourrat/Novack (2019), Cao/Lu/Wang (2019), Brigatti (2021), Dietert/Hérau/Hutridurga/Mouhot (2022), Brigati/Stoltz (2023)

[10] Eberle/Guillin/Zimmer, *Ann. Probab.* (2019)

[11] Bernard/Fathi/Levitt/Stoltz, *Annales Henri Lebesgue* (2022)

Convergence of overdamped Langevin dynamics

Overdamped Langevin dynamics and its generator

- Generator of overdamped Langevin dynamics (advection/diffusion)

$$\mathcal{L}_{\mathrm{ovd}} = -\nabla V(q) \cdot \nabla_q + \frac{1}{\beta}\Delta_q = -\frac{1}{\beta}\sum_{i=1}^{d} \partial_{q_i}^* \partial_{q_i}$$

hence self-adjoint on $L^2(\nu)$ with $\nu(dq) = Z_\nu^{-1} \mathrm{e}^{-\beta V(q)}\, dq$. Indeed,

$$\int_{\mathcal{D}} (\partial_{q_i}\varphi)\, \phi\, d\nu = -\int_{\mathcal{D}} \varphi\, (\partial_{q_i}\phi)\, d\nu - \int_{\mathcal{D}} \varphi\phi\, \partial_{q_i}\nu$$

so that $\partial_{q_i}^* = -\partial_{q_i} + \beta \partial_{q_i} V$

- Generator unitarily equivalent to a Schrödinger operator on $L^2(\mathbb{R}^d)$

$$-\widetilde{\mathcal{L}}_{\mathrm{ovd}} = \frac{1}{\beta}\Delta + \mathcal{V}, \qquad \mathcal{V} = \frac{1}{2}\left(\frac{\beta}{2}|\nabla V|^2 - \Delta V\right)$$

by considering $\widetilde{\mathcal{L}}_{\mathrm{ovd}} g = \nu^{1/2} \mathcal{L}_{\mathrm{ovd}}(\nu^{-1/2} g)$

Time evolution and decay estimates

- Solution $\varphi(t) = \mathrm{e}^{t\mathcal{L}_{\mathrm{ovd}}}\varphi_0$ to $\partial_t \varphi(t) = \mathcal{L}_{\mathrm{ovd}}\varphi(t)$: mass preservation

$$\frac{d}{dt}\left(\int_{\mathcal{D}} \varphi(t)\,\nu\right) = \int_{\mathcal{D}} \mathcal{L}_{\mathrm{ovd}}\varphi(t)\,\nu = \int_{\mathcal{D}} \varphi(t)\,(\mathcal{L}_{\mathrm{ovd}}\mathbf{1})\,\nu = 0$$

- Suggests the longtime limit $\varphi(t) \xrightarrow[t\to+\infty]{} \int_{\mathcal{D}} \varphi_0\,d\nu$

- Can assume w.l.o.g. that $\int_{\mathcal{D}} \varphi_0\,d\nu = 0$ (subspace $L^2_0(\nu)$ of $L^2(\nu)$)

- Decay estimate

$$\frac{d}{dt}\left(\frac{1}{2}\left\|\varphi(t)\right\|^2_{L^2(\nu)}\right) = \left\langle \mathcal{L}_{\mathrm{ovd}}\varphi(t), \varphi(t)\right\rangle_{L^2(\nu)} = -\frac{1}{\beta}\left\|\nabla_q \varphi(t)\right\|^2_{L^2(\nu)}$$

Poincaré inequality and convergence of the semigroup

- Assume that a Poincaré inequality holds:

$$\forall \phi \in H^1(\nu) \cap L_0^2(\nu), \qquad \|\phi\|_{L^2(\nu)} \leqslant \frac{1}{K_\nu} \|\nabla_q \phi\|_{L^2(\nu)}$$

Various sufficient conditions (V uniformly convex, confining, etc)

Exponential decay of the semigroup

ν satisfies a Poincaré inequality with constant $K_\nu > 0$ if and only if

$$\left\| \mathrm{e}^{t\mathcal{L}} \right\|_{\mathcal{B}(L_0^2(\nu))} \leqslant \mathrm{e}^{-K_\nu^2 t/\beta}.$$

Proof: Gronwall inequality $\dfrac{d}{dt}\left(\dfrac{1}{2}\|\varphi(t)\|_{L^2(\nu)}^2\right) \leqslant -\dfrac{K_\nu^2}{\beta}\|\varphi(t)\|_{L^2(\nu)}^2$

Several remarks:

- The prefactor for the exponential convergence is 1
- The convergence rate is *not degraded* (but is it improved?) when one adds an antisymmetric part $\mathcal{A} = F \cdot \nabla$ to $\mathcal{L}$ (with $\mathrm{div}(F\mathrm{e}^{-\beta V}) = 0$)

Longtime convergence of hypocoercive ODEs

A paradigmatic example of hypocoercive ODE

- ODE $\dot{X} = LX \in \mathbb{R}^2$ with (for $\gamma > 0$)

$$-L = A + \gamma S, \qquad A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

- **Structure of $-L$:**
  - Degenerate symmetric part $S \geqslant 0$
  - Antisymmetric part A coupling the kernel and the image of S
  - Smallest real part of eigenvalues (spectral gap) of order $\min(\gamma, \gamma^{-1})$

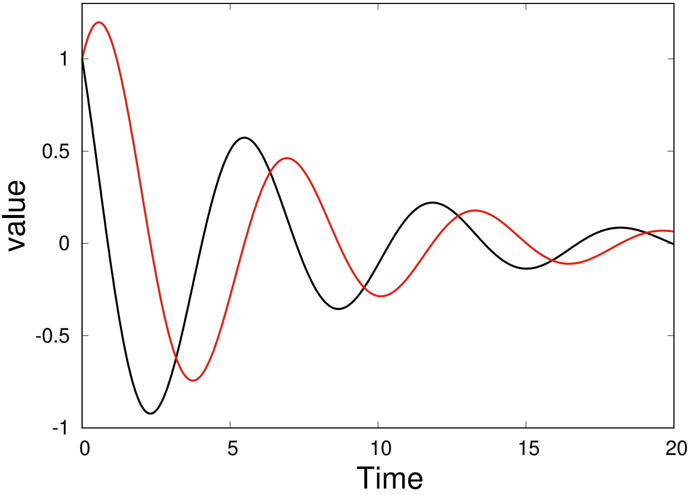
    determinant 1, trace γ, so eigenvalues $\lambda_\pm = \frac{\gamma}{2} \pm \left(\frac{\gamma^2}{4} - 1\right)^{1/2}$

- **Longtime convergence of e^{tL}?** Use $\mathrm{e}^{tL} = U^{-1} \begin{pmatrix} \mathrm{e}^{-t\lambda_+} & 0 \\ 0 & \mathrm{e}^{-t\lambda_-} \end{pmatrix} U$

Decay rate provided by the spectral gap $\lambda = \min\{\mathrm{Re}(\lambda_-), \mathrm{Re}(\lambda_+)\}$

$$X(t) = \mathrm{e}^{tL} X(0), \qquad |X(t)| \leqslant C\mathrm{e}^{-\lambda t}|X(0)|$$

Longtime convergence of hypocoercive ODE: illustration



Values $X_1(t), X_2(t)$ for $X(0) = (1, 1)$ and $\gamma = 0.5$

Longtime convergence of this hypocoercive ODE (1)

- **"Elliptic PDE way":** $\dfrac{d}{dt}\left(\dfrac{1}{2}|X(t)|^2\right) = -\gamma X(t)^\top S X(t) = -\gamma X_2(t)^2$

No dissipation in X_1... cannot conclude that $|X(t)|$ converges to 0...

- Change the scalar product with P positive definite:

$$|X|_P^2 = X^\top P X, \qquad \frac{d}{dt}\left(|X(t)|_P^2\right) = X(t)^\top \left(PL + L^\top P\right) X(t)$$

- **Fundamental idea: couple X_1 and X_2.** Start perturbatively:

$$P = \mathrm{Id} - \varepsilon \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

so that $-(PL + L^\top P) = 2\gamma PS + 2\varepsilon \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \sim 2 \begin{pmatrix} \varepsilon & 0 \\ 0 & \gamma \end{pmatrix}$

This provides some (small...) dissipation in X_1!

Longtime convergence of this hypocoercive ODE (2)

• Optimal choice[12] for P? Think of "$L^\top P \geqslant \lambda P$" and diagonalize $L^\top$

$$P = a_- X_- \overline{X}_-^\top + a_+ X_+ \overline{X}_+^\top, \qquad a_\pm > 0, \qquad L^\top X_\pm = \lambda_\pm X_\pm$$

Then $-(PL + L^\top P) \geqslant 2\lambda P$

• Therefore, $|X(t)|_P^2 \leqslant \mathrm{e}^{-2\lambda t}|X_0|_P^2$ (no prefactor here), and so, by equivalence of scalar products,

$$|X(t)| \leqslant \min\left(1, C\mathrm{e}^{-\lambda t}\right)|X_0|$$

Decay rate given by spectral gap

• Prefactor $C \geqslant 1$ really needed!
Exponential convergence with $C = 1$ if and only if $-L$ is coercive (*i.e.* $-X^\top L X \geqslant \alpha |X|^2$ with $\alpha > 0$)

[12] F. Achleitner, A. Arnold, and D. Stürzer, *Riv. Math. Univ. Parma*, 6(1):1–68, 2015.

Convergence of Langevin dynamics

Direct $L^2(\mu)$ approach: lack of coercivity

• The generator, considered on $L^2(\mu)$, is the sum of...

- a degenerate symmetric part $\mathcal{L}_{\mathrm{FD}} = -p^\top M^{-1}\nabla_p + \frac{1}{\beta}\Delta_p$
- an antisymmetric part $\mathcal{L}_{\mathrm{ham}} = p^\top M^{-1}\nabla_q - \nabla V^\top \nabla_p$

• Standard strategy for coercive generators: consider φ with average 0 with respect to μ and compute

$$\begin{aligned}\frac{d}{dt}\left(\left\|\mathrm{e}^{t\mathcal{L}}\varphi\right\|^2_{L^2(\mu)}\right) &= \left\langle \mathrm{e}^{t\mathcal{L}}\varphi, \mathcal{L}\mathrm{e}^{t\mathcal{L}}\right\rangle_{L^2(\mu)} = \left\langle \mathrm{e}^{t\mathcal{L}}\varphi, \mathcal{L}_{\mathrm{FD}}\mathrm{e}^{t\mathcal{L}}\right\rangle_{L^2(\mu)} \\ &= -\frac{1}{\beta}\left\|\nabla_p \mathrm{e}^{t\mathcal{L}}\varphi\right\|^2_{L^2(\mu)} \leqslant 0,\end{aligned}$$

but no control of $\|\phi\|_{L^2(\mu)}$ by $\|\nabla_p\phi\|_{L^2(\mu)}$ for a Gronwall estimate...

Two options:

- change of scalar product (use antisymmetric part)
- average in time (dissipation vanishes only exceptionally)

Almost direct $L^2(\mu)$ approach: convergence result

- Assume that the potential V is smooth and[13,14]
  - the marginal measure ν satisfies a Poincaré inequality

$$\|\varphi - \nu(\varphi)\|_{L^2(\nu)} \leqslant \frac{1}{K_\nu}\|\nabla_q \varphi\|_{L^2(\nu)}$$

  - there exist $c_1 > 0$, $c_2 \in [0,1)$ and $c_3 > 0$ such that V satisfies

$$\Delta V \leqslant c_1 + \frac{c_2}{2}|\nabla V|^2, \qquad \left|\nabla^2 V\right| \leqslant c_3\left(1 + |\nabla V|\right)$$

There exist $C > 0$ and $\lambda_\gamma > 0$ such that, for any $\varphi \in L_0^2(\mu)$,

$$\forall t \geqslant 0, \qquad \left\|\mathrm{e}^{t\mathcal{L}}\varphi\right\|_{L^2(\mu)} \leqslant C\mathrm{e}^{-\lambda_\gamma t}\|\varphi\|_{L^2(\mu)}$$

with convergence rate of order $\min(\gamma, \gamma^{-1})$: there is $\overline{\lambda} > 0$ for which

$$\lambda_\gamma \geqslant \overline{\lambda}\min(\gamma, \gamma^{-1})$$

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

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

Sketch of proof (1)

- Change of scalar product to use the antisymmetric part $\mathcal{L}_{\mathrm{ham}}$:
  - bilinear form $\mathcal{H}[\varphi] = \frac{1}{2}\|\varphi\|^2_{L^2(\mu)} - \varepsilon \langle R\varphi, \varphi\rangle$ with[15]

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

  - $R = \Pi_p R(1-\Pi_p)$ and $\mathcal{L}_{\mathrm{ham}} R$ are bounded
  - modified square norm $\mathcal{H} \sim \|\cdot\|^2_{L^2(\mu)}$ for $\varepsilon \in (-1,1)$
  - Approach not fully quantitative (optimize scalar product, here ε)

- **Interest:** $(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p) = \beta^{-1}\nabla_q^*\nabla_q$ coercive in q, and

$$R\mathcal{L}_{\mathrm{ham}}\Pi_p = \frac{(\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)}{1 + (\mathcal{L}_{\mathrm{ham}}\Pi_p)^*(\mathcal{L}_{\mathrm{ham}}\Pi_p)}$$

---
[15] Hérau (2006), Dolbeault/Mouhot/Schmeiser (2009, 2015), ...

Sketch of proof (2)

- Recall Poincaré inequalities: $\nabla_p^* \nabla_p \geqslant K_\kappa^2 (1 - \Pi_p)$ and $\nabla_q^* \nabla_q \geqslant K_\nu^2 \Pi_p$

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

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

- Upon controlling the remainder terms (some elliptic estimates)

$$\begin{aligned}
\mathscr{D}[\varphi] &= \gamma \left\langle -\mathcal{L}_{\mathrm{FD}}\varphi, \varphi \right\rangle + \varepsilon \left\langle R\mathcal{L}_{\mathrm{ham}}\Pi_p \varphi, \varphi \right\rangle + \mathrm{O}(\gamma\varepsilon) \\
&= \frac{\gamma}{\beta} \|\nabla_p \varphi\|_{L^2(\mu)}^2 + \varepsilon \left\langle \frac{\nabla_q^* \nabla_q}{\beta + \nabla_q^* \nabla_q} \Pi_p \varphi, \Pi_p \varphi \right\rangle + \mathrm{O}(\gamma\varepsilon) \\
&\geqslant \frac{\gamma K_\kappa^2}{\beta} \|(1 - \Pi_p)\varphi\|_{L^2(\mu)}^2 + \frac{\varepsilon K_\nu^2}{\beta + K_\nu^2} \|\Pi_p \varphi\|_{L^2(\mu)}^2 + \mathrm{O}(\gamma\varepsilon)
\end{aligned}$$

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

Obtaining directly bounds on the resolvent (1)

"Saddle-point like" structure[16] for typical hypocoercive operators on $L_0^2(\mu)$

$$\mathcal{L} = \begin{pmatrix} 0 & \mathcal{A}_{0+} \\ \mathcal{A}_{+0} & \mathcal{L}_{++} \end{pmatrix}, \qquad \mathcal{H} = \mathcal{H}_0 \oplus \mathcal{H}_+, \qquad \mathcal{H}_0 = \Pi_p \mathcal{H}, \qquad \mathcal{A} = \mathcal{L}_{\mathrm{ham}}$$

Formal inverse with Schur complement $\mathfrak{S}_0 = \mathcal{A}_{+0}^* \mathcal{L}_{++}^{-1} \mathcal{A}_{+0}$

$$\mathcal{L}^{-1} = \begin{pmatrix} \mathfrak{S}_0^{-1} & -\mathfrak{S}_0^{-1} \mathcal{A}_{0+} \mathcal{L}_{++}^{-1} \\ -\mathcal{L}_{++}^{-1} \mathcal{A}_{+0} \mathfrak{S}_0^{-1} & \mathcal{L}_{++}^{-1} + \mathcal{L}_{++}^{-1} \mathcal{A}_{+0} \mathfrak{S}_0^{-1} \mathcal{A}_{0+} \mathcal{L}_{++}^{-1} \end{pmatrix}$$

Invertibility of $\mathfrak{S}_0$ is the crucial element: two ingredients

- $-\dfrac{1}{2}(\mathcal{L} + \mathcal{L}^*) \geqslant s\Pi_+ = s(1 - \Pi_p)$ (Poincaré on $\kappa(dp)$ for Langevin)
- "macroscopic coercivity" $\|\mathcal{A}_{+0}\varphi\|_{L^2(\mu)} \geqslant a \|\Pi_p \varphi\|_{L^2(\mu)}$
  Amounts to $\mathcal{A}_{+0}^* \mathcal{A}_{+0} \geqslant a^2 \Pi_p$
  Guaranteed here by a Poincaré inequality for $\nu(dq)$, with $a^2 = K_\nu^2/\beta$

[16] E. Bernard, M. Fathi, A. Levitt and G. Stoltz, *Annales Henri Lebesgue* (2022)

Obtaining directly bounds on the resolvent (2)

- Further decompose $\mathcal{L}$ using $\Pi_1 = \mathcal{A}_{+0}\left(\mathcal{A}_{+0}^*\mathcal{A}_{+0}\right)^{-1}\mathcal{A}_{+0}^*$

$$\mathcal{L} = \begin{pmatrix} 0 & \mathcal{A}_{01} & 0 \\ \mathcal{A}_{10} & \mathcal{L}_{11} & \mathcal{L}_{12} \\ 0 & \mathcal{L}_{21} & \mathcal{L}_{22} \end{pmatrix}, \qquad \mathcal{A}_{01} = -\mathcal{A}_{10}^*.$$

- Additional technical assumptions ($\mathcal{S} = \gamma\mathcal{L}_{\mathrm{FD}}$ symmetric part):
  - There exists an involution $\mathcal{R}$ (e.g. momentum flip) on $\mathcal{H}$ such that

$$\mathcal{R}\Pi_0 = \Pi_0\mathcal{R} = \Pi_0, \qquad \mathcal{R}\mathcal{S}\mathcal{R} = \mathcal{S}, \qquad \mathcal{R}\mathcal{A}\mathcal{R} = -\mathcal{A}$$

  - The operators $\mathcal{S}_{11}$ and $\mathcal{L}_{21}\mathcal{A}_{10}\left(\mathcal{A}_{+0}^*\mathcal{A}_{+0}\right)^{-1}$ are bounded

Abstract resolvent estimates

$$\|\mathcal{L}^{-1}\| \leqslant 2\left(\frac{\|\mathcal{S}_{11}\|}{a^2} + \frac{\|\mathcal{R}_{22}\|\|\mathcal{L}_{21}\mathcal{A}_{10}(\mathcal{A}_{+0}^*\mathcal{A}_{+0})^{-1}\|^2}{s}\right) + \frac{3}{s}$$

Scaling with the friction and the dimension

- Final estimate for Fokker–Planck operators: scaling $\max(\gamma, \gamma^{-1})$

$$\left\|\mathcal{L}^{-1}\right\|_{\mathcal{B}(L_0^2(\mu))} \leqslant \frac{2\beta\gamma}{K_\nu^2} + \frac{4}{\gamma}\left(\frac{3}{4} + \left\|\Pi_+ \mathcal{L}_{\mathrm{ham}}^2 \Pi_p \left(\mathcal{A}_{+0}^* \mathcal{A}_{+0}\right)^{-1}\right\|^2\right)$$

- Estimate $2\left(C + C' K_\nu^{-2}\right)$ for squared operator norm on r.h.s.
  - $C = 1$ and $C' = 0$ when V is convex;
  - $C = 1$ and $C' = K$ when $\nabla_q^2 V \geqslant -K\mathrm{Id}$ for some $K \geqslant 0$;
  - $C = 2$ and $C' = \mathrm{O}(\sqrt{d})$ when $\Delta V \leqslant c_1 d + \frac{c_2\beta}{2}|\nabla V|^2$ (with $c_2 \leqslant 1$) and $\left|\nabla^2 V\right|^2 \leqslant c_3^2\left(d + |\nabla V|^2\right)$

- Better scaling $C' = \mathrm{O}(\log d)$ when logarithmic Sobolev inequality and

$$\forall x \in \mathbb{R}^d, \qquad \left\|\nabla^2 V(q)\right\|_{\mathcal{B}(\ell^2)} \leqslant c_3\left(1 + |\nabla V(q)|_\infty\right)$$

Space-time approaches

Average decay[17] over $[t, t+\tau]$ for $\tau > 0$: with $\mathrm{U}_\tau(dt) = \mathbf{1}_{[0,\tau]}(s)\dfrac{dt}{\tau}$,

$$\frac{d}{dt}\left(\int_0^\tau \|f(t+s,\cdot,\cdot)\|^2_{L^2(\mu)}\,\mathrm{U}_\tau(ds)\right) \leqslant -2\gamma \int_0^\tau \|\nabla_p f(t+s,\cdot,\cdot)\|^2_{L^2(\mu)}\,\mathrm{U}_\tau(ds)$$

- For $h(t) = \mathrm{e}^{t\mathcal{L}} h_0$, control dissipation with full space-time antisymmetric part

$$\|(\partial_t - \mathcal{L}_{\mathrm{ham}})h\|_{L^2(\mathrm{U}_\tau\otimes\nu;H^{-1}(\kappa))} \leqslant \gamma \|\nabla_p h\|_{L^2(\mathrm{U}_\tau\otimes\mu)}$$

- Space-time-velocity Poincaré inequality ($\mu(h) = 0$)

$$\begin{aligned}\bar{\lambda}\|h\|^2_{L^2(\mathrm{U}_\tau\otimes\mu)} &\leqslant \|(\partial_t - \mathcal{L}_{\mathrm{ham}})h\|^2_{L^2(\mathrm{U}_\tau\otimes\nu;H^{-1}(\kappa))} + \|(\mathrm{Id}-\Pi_p)h\|^2_{L^2(\mathrm{U}_\tau\otimes\mu)} \\ &\leqslant \|(\partial_t - \mathcal{L}_{\mathrm{ham}})h\|^2_{L^2(\mathrm{U}_\tau\otimes\nu;H^{-1}(\kappa))} + \frac{1}{K_\kappa}\|\nabla_p h\|^2_{L^2(\mathrm{U}_\tau\otimes\mu)}\end{aligned}$$

Combination leads to exponential convergence through Gronwall estimate (**explicit constants:** scaling in γ, τ, dimension, Poincaré constants, etc.)

[17] G. Brigati and G. Stoltz, *arXiv preprint* **2302.14506**

Space-time-velocity Poincaré inequality

Aim: sufficient to control $\Pi_p h \to$ space-time functions (no velocity)

Two key ingredients: a Poincaré–Lions inequality

$$\left\| g - \iint_{[0,\tau]\times\mathcal{D}} g(t,q)\, \mathrm{U}_\tau(dt)\, \nu(dq) \right\|^2_{L^2(\mathrm{U}_\tau\otimes\nu)} \leqslant C_\tau^{\mathrm{Lions}} \|\nabla_{t,q} g\|^2_{H^{-1}(\mathrm{U}_\tau\otimes\nu)}$$

and an averaging result

$$\|\nabla_{t,q}\Pi_p h\|^2_{H^{-1}} \leqslant K_{\mathrm{avg}}\Big(\|(\mathrm{Id}-\Pi_p)h\|^2_{L^2} + \|(\partial_t - \mathcal{L}_{\mathrm{ham}})h\|^2_{L^2(\mathrm{U}_\tau\otimes\nu, H^{-1}(\kappa))}\Big)$$

Directly leads to $\bar{\lambda} = \dfrac{1}{1 + C_\tau^{\mathrm{Lions}} K_{\mathrm{avg}}}$

Same conditions on V as DMS approach

Averaging lemma

Based on identities such as $(z = z(t,q))$

$$\begin{aligned}\int_0^\tau \int_{\mathcal{D}} (\partial_t \Pi_p h)\, z \, d\mathrm{U}_\tau \, d\mu &= \int_0^\tau \int_{\mathcal{D}} \int_{\mathbb{R}^3} [(\partial_t - \mathcal{L}_{\mathrm{ham}}) \Pi_p h]\, z \, d\mathrm{U}_\tau \, d\nu \, d\kappa \\ &= \int_0^\tau \int_{\mathcal{D}} \int_{\mathbb{R}^3} [(\partial_t - \mathcal{L}_{\mathrm{ham}}) h]\, z \, d\mathrm{U}_\tau \, d\mu \\ &\quad + \int_0^\tau \int_{\mathcal{D}} \int_{\mathbb{R}^3} [(\mathrm{Id} - \Pi_p) h]\, (\partial_t - \mathcal{L}_{\mathrm{ham}}) z \, d\mathrm{U}_\tau \, d\mu \\ &\leqslant \|(\partial_t - \mathcal{L}_{\mathrm{ham}}) h\|_{L^2(\mathrm{U}_\tau \otimes \nu, H^{-1}(\kappa))} \|z\|_{L^2(\mathrm{U}_\tau \otimes \nu)} \|1\|_{H^1(\kappa)} \\ &\quad + \|(\mathrm{Id} - \Pi_p) h\|_{L^2(\mathrm{U}_\tau \otimes \mu)} \|(\partial_t - \mathcal{L}_{\mathrm{ham}}) z\|_{L^2(\mathrm{U}_\tau \otimes \mu)}\end{aligned}$$

for $z \in \mathrm{H}^1_{\mathrm{DC}}(\mathrm{U}_\tau \otimes \nu)$ with $\|z\|^2_{H^1(\mathrm{U}_\tau \otimes \nu)} \leqslant 1$

(Dirichlet boundary conditions $z(0) = z(\tau) = 0$ used for integration by parts in time)

Explicit expression for K_{avg} in terms of kinetic energy $E_{\mathrm{kin}}(p)$

Poincaré–Lions inequality (1/2)

Reduction to divergence equation: for $f \in L^2_0(\mathrm{U}_\tau \otimes \nu)$, find a solution $Z = (Z_0, Z_1, \dots, Z_d) \in H^1_{\mathrm{DC}}(\mathrm{U}_\tau \otimes \nu)^{d+1}$ satisfying

$$-\partial_t Z_0 + \sum_{i=1}^{d} \partial_{q_i}^\star Z_i = f$$

with estimates $\|Z\|_{H^1(\mathrm{U}_\tau \otimes \nu)} \leqslant C_\tau^{\mathrm{div}} \|f\|_{L^2(\mathrm{U}_\tau \otimes \nu)}$

Beware the boundary conditions in time for Z!

Proceed by duality (f with average 0 w.r.t. $\mathrm{U}_\tau \otimes \nu$)

$$\begin{aligned} \|f\|^2_{\mathrm{L}^2(\mathrm{U}_\tau \otimes \nu)} &= \int_0^\tau \int_{\mathcal{D}} \left(-\partial_t Z_0 + \sum_{i=1}^{d} \partial_{q_i}^\star Z_i \right) f \, d\mathrm{U}_\tau \, d\nu \\ &= \langle \nabla_{t,q} f, Z \rangle_{H^{-1}(\mathrm{U}_\tau \otimes \nu), H^1_{\mathrm{DC}}(\mathrm{U}_\tau \otimes \nu)} \\ &\leqslant C_\tau^{\mathrm{div}} \|f\|_{L^2(\mathrm{U}_\tau \otimes \nu)} \|\nabla_{t,q} f\|_{H^{-1}(\mathrm{U}_\tau \otimes \nu)} \end{aligned}$$

Poincaré–Lions inequality (2/2)

Decompose f in $\mathcal{N}$ and its orthogonal, with $L = \left(\nabla_q^\star \nabla_q\right)^{1/2}$ and

$$\mathcal{N} = \left\{ \mathrm{e}^{-tL} g_+ + \mathrm{e}^{-(\tau - t)L} g_-, \quad g_+, g_- \in L_0^2(\nu) \right\}$$

By construction, $(-\partial_t^2 + \nabla_q^\star \nabla_q) g = 0$ for $g \in \mathcal{N}$

Explicit solution to divergence equation (non unique)

$$Z = \nabla_{t,q} \mathscr{W}^{-1} \mathscr{P}_{\mathcal{N}^\perp} f + \begin{pmatrix} F_0(t,L) \\ \partial_{q_1} F_1(t,L) \\ \vdots \\ \partial_{q_d} F_1(t,L) \end{pmatrix} \mathscr{P}_{\mathcal{N},+} f + \begin{pmatrix} F_0(\tau - t,L) \\ \partial_{q_1} F_1(\tau - t,L) \\ \vdots \\ \partial_{q_d} F_1(\tau - t,L) \end{pmatrix} \mathscr{P}_{\mathcal{N},-} f$$

where $\mathcal{W} = -\partial_t^2 + \nabla_q^\star \nabla_q$ with Neumann BC in time, and

$$-\partial_t \Big[\underbrace{F_0(t,L) \mathrm{e}^{-tL}}_{P_0(\mathrm{e}^{-tL})} \Big] + L^2 \underbrace{F_1(t,L) \mathrm{e}^{-tL}}_{P_1(\mathrm{e}^{-tL})} = \mathrm{e}^{-tL}$$

Generalizations/perspectives

These approaches works for other hypocoercive dynamics

- non-quadratic kinetic energies (but still Poincaré inequality)[18]
- weak confinements and/or heavy tail distributions of velocities[19]
- adaptive Langevin dynamics (additional Nosé–Hoover part)[20]
- linear Boltzmann (HMC)/piecewise deterministic Markov processes

Possibly stretched exponential or algebraic convergence rates

Some work needed to extend the approaches to...

- more degeneracy: generalized Langevin,[21] chains of oscillators[22]
- non-gradient forcings (steady-state nonequilibrium dynamics)[23]

---

[18] G. Stoltz and Z. Trstanova (2018)

[19] M. Grothaus and F.-Y. Wang (2019); E. Bouin, J. Dolbeault and L. Ziviani (2024); G. Brigati, G. Stoltz, A. Wang and L. Wang (2024)

[20] B. Leimkuhler, M. Sachs and G. Stoltz (2020)

[21] M. Ottobre and G. Pavliotis (2011), G. Pavliotis, G. Stoltz and U. Vaes (2021)

[22] A. Menegaki (2020)

[23] H. Dietert (2023)