Hypocoercivity of Langevin-like dynamics with Schur complements

Gabriel STOLTZ

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

Work in collaboration with E. Bernard, A. Levitt and M. Fathi
arXiv preprint 2003.00726

Bernoulli-IMS One World Symposium, August 2020

Langevin dynamics

- Positions and momenta $(q,p) \in \mathcal{E} = \mathcal{D} \times \mathbb{R}^d$ with $\mathcal{D} = (L\mathbb{T})^d$ or $\mathbb{R}^d$
- Hamiltonian $H(q,p) = V(q) + p^T M^{-1} p/2$

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

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

- Existence and uniqueness of the invariant probability measure

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

- Generator $\mathcal{L} = \mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{FD}}$: (anti)symmetric parts on $L^2(\mu)$

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

- Hypoelliptic operator (degeneracy of the diffusion)

Hamiltonian and overdamped limits

• As $\gamma \to 0$, the Hamiltonian dynamics is recovered
$\to$ time $\sim \dfrac{1}{\gamma}$ to change energy levels in the Hamiltonian limit[1]

• Overdamped limit $\gamma \to +\infty$ (or masses going to 0)

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

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)

Ergodicity results

- Almost-sure convergence[2] 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[3] when the equation has a solution in $L^2(\mu)$

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

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

- **Central question here: Well-posedness of such equations?**

---

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

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

Exponential convergence: an (incomplete) survey

- Invertibility of $\mathcal{L}$ on subsets of $L^2_0(\mu) = \left\{ \varphi \in L^2(\mu) \,\middle|\, \int_{\mathcal{E}} \varphi \, d\mu = 0 \right\}$?

$$-\mathcal{L}^{-1} = \int_0^{+\infty} \mathrm{e}^{t\mathcal{L}} \, dt, \qquad \left(\mathrm{e}^{t\mathcal{L}}\varphi\right)(q_0,p_0) = \mathbb{E}^{(q_0,p_0)}\left[\varphi(q_t,p_t)\right]$$

- Prove exponential convergence of the semigroup $\mathrm{e}^{t\mathcal{L}}$ on subsets of $L^2_0(\mu)$
  - Lyapunov techniques[4] $L^\infty_W(\mathcal{E}) = \left\{ \varphi \text{ measurable}, \left\| \frac{\varphi}{W} \right\|_{L^\infty} < +\infty \right\}$
  - standard hypocoercive[5] setup $H^1(\mu)$
  - $E = L^2(\mu)$ after hypoelliptic regularization[6] from $H^1(\mu)$
  - Directly[7] $E = L^2(\mu)$ (recently[8] Poincaré using $\partial_t - \mathcal{L}_{\mathrm{ham}}$)
  - coupling arguments[9] and other probabilistic proofs

---

[4] Wu ('01); Mattingly/Stuart/Higham ('02); Rey-Bellet ('06); Hairer/Mattingly ('11)
[5] Villani (2009), after Talay (2002), Eckmann/Hairer (2003), Hérau/Nier (2004), ...
[6] F. Hérau, *J. Funct. Anal.* **244**(1), 95-118 (2007)
[7] Hérau (2006), Dolbeault/Mouhot/Schmeiser (2009, 2015), ...
[8] Armstrong/Mourrat (2019), Cao/Lu/Wang (2019)
[9] A. Eberle, A. Guillin and R. Zimmer, *Ann. Probab.* **47**(4), 1982-2010 (2019)

Direct $L^2(\mu)$ approach: lack of coercivity and a way around

- Standard strategy for coercive generators: $\varphi \in L_0^2(\mu)$,

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

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

- Change of scalar product to use the antisymmetric part $\mathcal{L}_{\mathrm{ham}}$

- Example: modified square norm $\mathcal{H}[\varphi] = \dfrac{1}{2}\|\varphi\|^2_{L^2(\mu)} - \varepsilon\langle A\varphi, \varphi\rangle$ for $\varepsilon \in (-1, 1)$ and

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

- Approach less quantitative (optimize scalar product)

Obtaining directly bounds on the resolvent

- Decompose $L_0^2(\mu) = \mathcal{H} = \mathcal{H}_- \oplus \mathcal{H}_+$ with $\mathcal{H}_- = \Pi\mathcal{H}$

$$-\mathcal{L} = \begin{pmatrix} 0 & B^* \\ -B & A \end{pmatrix}, \qquad \begin{aligned} A &:= -(1-\Pi)\mathcal{L}(1-\Pi) : \mathcal{H}_+ \to \mathcal{H}_+, \\ B &:= (1-\Pi)\mathcal{L}_{\mathrm{ham}}\Pi : \mathcal{H}_- \to \mathcal{H}_+ \end{aligned}$$

- A invertible on $\mathcal{H}_+$ since $\langle \varphi, A\varphi \rangle_{L^2(\mu)} \geqslant \dfrac{\gamma}{m} \|(1-\Pi)\varphi\|^2_{L^2(\mu)}$

Formal action of the inverse: with $S = \left(B^* A^{-1} B\right)^{-1}$,

$$(-\mathcal{L})^{-1} = \begin{pmatrix} S & -SB^*A^{-1}, \\ A^{-1}BS & A^{-1} - A^{-1}BSB^*A^{-1} \end{pmatrix}$$

- Invertibility of S is the crucial element: assume "macroscopic coercivity"

$$\|B\varphi\|_{L^2(\mu)} \geqslant \rho \|\Pi\varphi\|_{L^2(\mu)}, \qquad B^*B \geqslant \rho^2 \Pi$$

Guaranteed here by a Poincaré inequality for $\nu(dq)$, with $\rho^2 = K_\nu^2/(\beta m)$, since $B^*B = \nabla_q^* \nabla_q \Pi$

Elements of the proof (general case)

- Orthogonal projector $P = B\left(B^*B\right)^{-1}B^*$ on $\mathrm{Ran}(B)$, and $\mathcal{H}_{+,1} = P\mathcal{H}_+$

$$-\mathcal{L} = \begin{pmatrix} 0 & B_1^* & 0 \\ -B_1 & A_{11} & A_{12} \\ 0 & A_{21} & A_{22} \end{pmatrix},$$

with $B_1 : \mathcal{H}_- \to \mathcal{H}_{+,1}$ the restriction of B to $\mathcal{H}_{1,+}$

Key result (again a Schur complement)

$$\left(B^*A^{-1}B\right)^{-1} = \left(B^*B\right)^{-1}B^*\left(A_{11} - A_{12}A_{22}^{-1}A_{21}\right)B\left(B^*B\right)^{-1}.$$

- Estimates on other terms in $\mathcal{L}^{-1}$ relying on some reversibility property:

$$\mathcal{R}^2 = \mathrm{Id}, \qquad \mathcal{R}\Pi = \Pi\mathcal{R} = \Pi, \qquad \mathcal{R}\mathcal{L}_{\mathrm{s}}\mathcal{R} = \mathcal{L}_{\mathrm{s}}, \qquad \mathcal{R}\mathcal{L}_{\mathrm{a}}\mathcal{R} = -\mathcal{L}_{\mathrm{a}}$$

For Langevin dynamics, $(\mathcal{R}\varphi)(q,p) = \varphi(q,-p)$

Scaling with the friction and the dimension

- Final estimate for Langevin dynamics: scaling $\max(\gamma, \gamma^{-1})$

$$\|\mathcal{L}^{-1}\|_{\mathcal{B}(L^2_0(\mu))} \leqslant \frac{2\beta\gamma}{K_\nu^2} + \frac{4m}{\gamma}\left(\frac{1}{2} + \left\|(1-\Pi)\mathcal{L}_{\mathrm{a}}^2\Pi\left(B^*B\right)^{-1}\right\|^2_{\mathcal{B}(\mathcal{H}_-,\mathcal{H}_+)}\right),$$

- Estimate $2\left(C + C' K_\nu^{-2}\right)$ for 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 $\Delta V \leqslant c_1 d + \dfrac{c_2\beta}{2}|\nabla V|^2$ (with $c_2 \leqslant 1$) and $\left|\nabla^2 V\right|^2 \leqslant c_3\left(d + |\nabla V|^2\right)$

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

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

Generalizations and perspectives

- **Approach works for other hypocoercive dynamics**[10]
  - linear Boltzmann/randomized HMC (replace $\mathcal{L}_{\mathrm{FD}}$ with $\Pi - 1$)
  - adaptive Langevin dynamics (additional Nosé–Hoover part)

- **Some work needed to extend it to more degenerate dynamics**
  - generalized Langevin dynamics
  - chains of oscillators

- **Current work also on obtaining...**
  - resolvent estimates $(\mathrm{i}\omega - \mathcal{L})^{-1}$ uniform in ω
  - space-time Poincaré inequalities with our algebraic framework

$$\left\| f - \langle f, \mathbf{1} \rangle_{L^2(\widetilde{\mu}_T)} \right\|_{L^2(\widetilde{\mu}_T)} \leqslant C_{1,T} \|(1-\Pi) f\|_{L^2(\widetilde{\mu}_T)} + C_{2,T} \|(1-\mathcal{L}_{\mathrm{s}})^{-1/2} \left(-\partial_t + \mathcal{L}_{\mathrm{a}}\right) f\|_{L^2(\widetilde{\mu}_T)}$$

[10] E. Bernard, M. Fathi, A. Levitt, G. Stoltz, *arXiv preprint* **2003.00726**