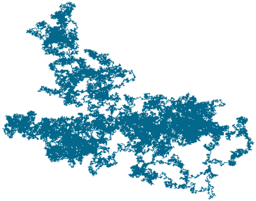
Ergodicity and convergence to equilibrium for Langevin dynamics with general potentials

David P. Herzog
Iowa State University



September 9th, 2024
BIRS Workshop on Stochastics and Geometry
Banff, Alberta

Talk Overview

- *Goal*: Introduction to methods used to establish convergence to equilibrium for Langevin dynamics.

Talk Overview

- *Goal*: Introduction to methods used to establish convergence to equilibrium for Langevin dynamics.
- *Two different paths*: Probabilistic versus Poincaré:

$$\|\cdot\|_{TV} \qquad \text{versus} \qquad \|\cdot\|_{H(\nu)} .$$

Talk Overview

- *Goal*: Introduction to methods used to establish convergence to equilibrium for Langevin dynamics.
- *Two different paths*: Probabilistic versus Poincaré:

$$\| \cdot \|_{TV} \qquad \text{versus} \qquad \| \cdot \|_{H(\nu)} .$$

- *Paths cross*: Overlaps between the methods.

Talk Overview

- *Goal*: Introduction to methods used to establish convergence to equilibrium for Langevin dynamics.
- *Two different paths*: Probabilistic versus Poincaré:

$$\| \cdot \|_{TV} \qquad \text{versus} \qquad \| \cdot \|_{H(\nu)}.$$

- *Paths cross*: Overlaps between the methods.
- *Main technical issues*: System is *degenerately damped*; randomness is also *degenerate*. Types of potential functions can make arguments harder (nonsingular vs singular).

Talk Overview

- *Goal*: Introduction to methods used to establish convergence to equilibrium for Langevin dynamics.
- *Two different paths*: Probabilistic versus Poincaré:

$$\|\cdot\|_{TV} \qquad \text{versus} \qquad \|\cdot\|_{H(\nu)}.$$

- *Paths cross*: Overlaps between the methods.
- *Main technical issues*: System is *degenerately damped*; randomness is also *degenerate*. Types of potential functions can make arguments harder (nonsingular vs singular).
- Interfaces statistical mechanics, MCMC, geometry and Boltzmann.

Doeblin's condition

Markov chain $\{X_n\}$ on a finite state space $\{1, 2, \ldots, d\}$ with transition P. Suppose there exists $\epsilon \in (0, 1)$ and a probability η on $\{1, 2, \ldots, d\}$ such that

$$P(x, A) \geq \epsilon\eta(A)$$

for all $x \in \{1, 2, \ldots, d\}$, $A \subset \{1, 2, \ldots, d\}$. Then for all $k \geq 1$

$$\|P^k(x, \cdot) - P^k(y, \cdot)\|_{TV} \leq (1 - \epsilon)^k \|\delta_x - \delta_y\|_{TV}.$$

Doeblin's condition

Markov chain $\{X_n\}$ on a finite state space $\{1, 2, \ldots, d\}$ with transition P. Suppose there exists $\epsilon \in (0, 1)$ and a probability η on $\{1, 2, \ldots, d\}$ such that

$$P(x, A) \geq \epsilon\eta(A)$$

for all $x \in \{1, 2, \ldots, d\}$, $A \subset \{1, 2, \ldots, d\}$. Then for all $k \geq 1$

$$\|P^k(x, \cdot) - P^k(y, \cdot)\|_{TV} \leq (1 - \epsilon)^k \|\delta_x - \delta_y\|_{TV}.$$

Proof. Let μ_1, μ_2 be probability measures on $\{1, 2, \ldots, d\}$. Define $Q(x, \cdot) = \frac{1}{1-\epsilon}P(x, \cdot) - \frac{\epsilon}{1-\epsilon}\eta(\cdot)$. Then

$$\mu_1 P - \mu_2 P = (1 - \epsilon)(\mu_1 Q - \mu_2 Q)$$

Doeblin's condition

Markov chain $\{X_n\}$ on a finite state space $\{1, 2, \ldots, d\}$ with transition P. Suppose there exists $\epsilon \in (0, 1)$ and a probability η on $\{1, 2, \ldots, d\}$ such that

$$P(x, A) \geq \epsilon\eta(A)$$

for all $x \in \{1, 2, \ldots, d\}$, $A \subset \{1, 2, \ldots, d\}$. Then for all $k \geq 1$

$$\|P^k(x, \cdot) - P^k(y, \cdot)\|_{TV} \leq (1 - \epsilon)^k \|\delta_x - \delta_y\|_{TV}.$$

Proof. Let μ_1, μ_2 be probability measures on $\{1, 2, \ldots, d\}$. Define $Q(x, \cdot) = \frac{1}{1-\epsilon} P(x, \cdot) - \frac{\epsilon}{1-\epsilon}\eta(\cdot)$. Then

$$\mu_1 P - \mu_2 P = (1 - \epsilon)(\mu_1 Q - \mu_2 Q)$$

$$\implies \|\mu_1 P - \mu_2 P\|_{TV} \leq (1 - \epsilon)\|\mu_1 - \mu_2\|_{TV}.$$

DOEBLIN'S CONDITION

Let $\{X_n\}$ be a Markov chain on state space $\mathcal{X}$.

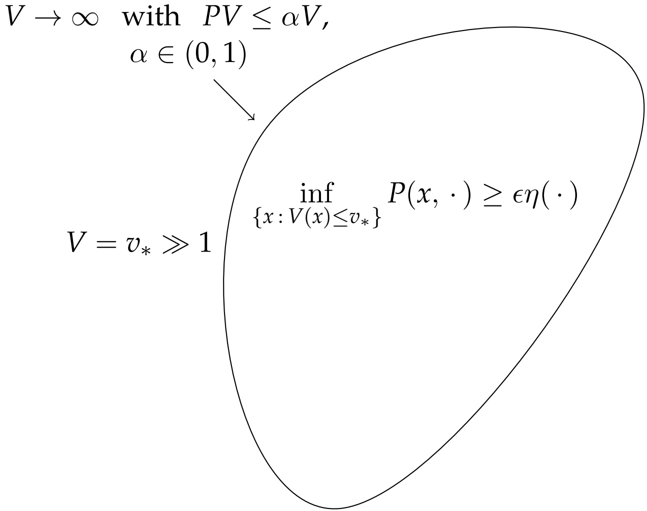
(DC) There exists $\epsilon \in (0, 1)$ and a probability measure η on $\mathcal{X}$ such that

$$P(x, A) \geq \epsilon\eta(A)$$

for all $x \in \mathcal{X}$, $A \subset \mathcal{X}$ measurable.

- If $\mathcal{X} = \{1, 2, \ldots, d\}$ and $\{X_n\}$ irreducible and aperiodic, then (DC) follows for $\mathcal{P} := P^k$ (e.g. take $\nu(A) = \delta_1(A)$).
- If (DC) is not satisfied globally, need return times to a "small" set where (DC) is true to have exponential moments (i.e. it takes log time on average to return to small set). Use of Lyapunov structure.

Convergence Picture [1]



[1] Harris ’54; Hasminskii ’80; Meyn, Tweedie ’92/’93; Hairer-Mattingly ’08

Example

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t.$$

Example

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t.$$

- B_t is a standard, d-dimensional Brownian motion;
- $U \in C^\infty(\mathbf{R}^d; [0, \infty))$ satisfies:
  - $U(x) \to \infty$ as $|x| \to \infty$;
  - $\Delta U - |\nabla U|^2 \leq -cU + d$ for some constants $c, d > 0$;
  - $\int e^{-U(x)}\,dx < \infty$.

EXAMPLE

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\, dt + \sqrt{2}\, dB_t.$$

- B_t is a standard, d-dimensional Brownian motion;
- $U \in C^\infty(\mathbf{R}^d; [0, \infty))$ satisfies:
  - $U(x) \to \infty$ as $|x| \to \infty$;
  - $\Delta U - |\nabla U|^2 \leq -cU + d$ for some constants $c, d > 0$;
  - $\int e^{-U(x)}\, dx < \infty$.

Lyapunov structure: We have

$$LU = -|\nabla U|^2 + \Delta U \leq -cU + d \implies P^t U \leq e^{-ct} U + \frac{d}{c}.$$

EXAMPLE CON'T

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\, dt + \sqrt{2}\, dB_t.$$

[2] J. Evans '18.

Example con't

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t.$$

Doeblin Condition:

- Fundamental solutions of the Kolmogorov equations $(\partial_t \pm L)p = 0$, $(\partial_t \pm L^*)p = 0$ are smooth and strictly positive on $(0,\infty) \times \mathbf{R}^d$;
- Transition density $p_t(q, q')$ is smooth and strictly positive on $(0,\infty) \times \mathbf{R}^d \times \mathbf{R}^d$.
- (DC) follows using Lebesgue measure on a bounded set.
- The ϵ in (DC) is typically existential $\implies$ quantitative minorization?[2]

[2] J. Evans '18.

Poincaré

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t.$$

The process q_t has a unique stationary distribution μ given by

$$\mu(dq) \propto e^{-U(q)}\,dq.$$

POINCARÉ

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t.$$

The process q_t has a unique stationary distribution μ given by

$$\mu(dq) \propto e^{-U(q)}\,dq.$$

Let $\varphi \in L^2(\mu)$ have $\mu(\varphi) = 0$. WTS that $\|P^t\varphi\|_{L^2(\mu)} \to 0$ as $t \to \infty$ exponentially fast.

Poincaré

Stochastic gradient dynamics on $\mathbf{R}^d$:

$$dq_t = -\nabla U(q_t)\,dt + \sqrt{2}\,dB_t.$$

The process q_t has a unique stationary distribution μ given by

$$\mu(dq) \propto e^{-U(q)}\,dq.$$

Let $\varphi \in L^2(\mu)$ have $\mu(\varphi) = 0$. WTS that $\|P^t\varphi\|_{L^2(\mu)} \to 0$ as $t \to \infty$ exponentially fast. Then

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\|P^t\varphi\|^2_{L^2(\mu)} = \langle \tfrac{d}{dt}P^t\varphi, P^t\varphi\rangle &= \langle LP^t\varphi, P^t\varphi\rangle \\ &= \tfrac{1}{2}\mu(L(P^t\varphi)^2) - \|\nabla P^t\varphi\|^2_{L^2(\mu)} \\ &= -\|\nabla P^t\varphi\|^2_{L^2(\mu)}.\end{aligned}$$

POINCARÉ[3]

Note that μ satisfies a *Poincaré inequality*. That is, there exists $\rho > 0$ such that for all $\varphi \in H^1(\mu)$ with $\mu(\varphi) = 0$ we have

$$\|\nabla\varphi\|^2_{L^2(\mu)} \geq \rho\|\varphi\|^2_{L^2(\mu)}.$$

[3] Talay '00; Eckmann, Hairer '03; Hérau, Nier '04; Helffer, Nier '05

POINCARÉ[3]

Note that μ satisfies a *Poincaré inequality*. That is, there exists $\rho > 0$ such that for all $\varphi \in H^1(\mu)$ with $\mu(\varphi) = 0$ we have

$$\|\nabla \varphi\|_{L^2(\mu)}^2 \geq \rho \|\varphi\|_{L^2(\mu)}^2.$$

Combining the above with the formal calculation gives:

$$\frac{1}{2}\frac{d}{dt}\|P^t\varphi\|_{L^2(\mu)}^2 = -\|\nabla P^t \varphi\|_{L^2(\mu)}^2 \leq -\rho\|P^t\varphi\|_{L^2(\mu)}^2.$$

[3] Talay '00; Eckmann, Hairer '03; Hérau, Nier '04; Helffer, Nier '05

Poincaré[3]

Note that μ satisfies a *Poincaré inequality*. That is, there exists $\rho > 0$ such that for all $\varphi \in H^1(\mu)$ with $\mu(\varphi) = 0$ we have

$$\|\nabla\varphi\|^2_{L^2(\mu)} \geq \rho\|\varphi\|^2_{L^2(\mu)}.$$

Combining the above with the formal calculation gives:

$$\frac{1}{2}\frac{d}{dt}\|P^t\varphi\|^2_{L^2(\mu)} = -\|\nabla P^t\varphi\|^2_{L^2(\mu)} \leq -\rho\|P^t\varphi\|^2_{L^2(\mu)}.$$

Hence

$$\|P^t\varphi\|_{L^2(\mu)} \leq e^{-\rho t}\|\varphi\|_{L^2(\mu)}.$$

[3] Talay ’00; Eckmann, Hairer ’03; Hérau, Nier ’04; Helffer, Nier ’05

POINCARÉ[3]

Note that μ satisfies a *Poincaré inequality*. That is, there exists $\rho > 0$ such that for all $\varphi \in H^1(\mu)$ with $\mu(\varphi) = 0$ we have

$$\|\nabla\varphi\|^2_{L^2(\mu)} \geq \rho\|\varphi\|^2_{L^2(\mu)}.$$

Combining the above with the formal calculation gives:

$$\frac{1}{2}\frac{d}{dt}\|P^t\varphi\|^2_{L^2(\mu)} = -\|\nabla P^t\varphi\|^2_{L^2(\mu)} \leq -\rho\|P^t\varphi\|^2_{L^2(\mu)}.$$

Hence

$$\|P^t\varphi\|_{L^2(\mu)} \leq e^{-\rho t}\|\varphi\|_{L^2(\mu)}.$$

- Proofs of Poincaré inequality often use bounds like $\Delta U - |\nabla U|^2 \leq -cU + d$.

[3]Talay '00; Eckmann, Hairer '03; Hérau, Nier '04; Helffer, Nier '05

Langevin dynamics

Consider the following SDE for $x_t = (q_t, p_t)$ on $\mathcal{X} \subset \mathbf{R}^d \times \mathbf{R}^d$:

$$
\begin{aligned}
dq_t &= p_t\, dt \\
dp_t &= -\gamma p_t\, dt - \nabla U(q_t)\, dt + \sqrt{2\gamma}\, dB_t.
\end{aligned}
$$

- B_t is a standard d-dimensional Brownian motion, $\mathcal{X} = \{U(q) < \infty\} \times \mathbf{R}^d$, $\gamma > 0$ is the friction coefficient;
- $U \in C^\infty(\mathcal{X}; [0, \infty))$ satisfies
  - $|\nabla U| \to \infty$ as $U \to \infty$;
  - $|\nabla^2 U| \leq \epsilon |\nabla U|^2 + C_\epsilon$.
- Hamiltonian $H(q, p) = |p|^2/2 + U(q)$ with stationary distribution ν on $\mathcal{X}$

$$
\nu(dqdp) \propto e^{-H(q,p)}\, dq\, dp = e^{-\frac{|p|^2}{2}} e^{-U(q)}\, dq\, dp.
$$

LANGEVIN DYNAMICS

Consider the following SDE for $x_t = (q_t, p_t) \in \mathcal{X} \subset \mathbf{R}^d \times \mathbf{R}^d$:

$$\begin{aligned} dq_t &= p_t\, dt \\ dp_t &= -\gamma p_t\, dt - \nabla U(q_t)\, dt + \sqrt{2\gamma}\, dB_t. \end{aligned}$$

Lyapunov?

- Generator: $\mathcal{L} = p \cdot \nabla_q - \gamma p \cdot \nabla_p - \nabla U(q) \cdot \nabla_p + \gamma \Delta_p$.
- $\mathcal{L}H(q,p) = -\gamma |p|^2 + \gamma d$.

LANGEVIN DYNAMICS

Consider the following SDE for $x_t = (q_t, p_t) \in \mathcal{X} \subset \mathbf{R}^d \times \mathbf{R}^d$:

$$\begin{aligned} dq_t &= p_t \, dt \\ dp_t &= -\gamma p_t \, dt - \nabla U(q_t) \, dt + \sqrt{2\gamma} \, dB_t. \end{aligned}$$

Lyapunov?

- Generator: $\mathcal{L} = p \cdot \nabla_q - \gamma p \cdot \nabla_p - \nabla U(q) \cdot \nabla_p + \gamma \Delta_p$.
- $\mathcal{L} H(q,p) = -\gamma |p|^2 + \gamma d$.

Poincaré?

- If $\varphi \in L^2(\nu)$ with $\nu(\varphi) = 0$, then

$$\frac{1}{2} \frac{d}{dt} \|\mathcal{P}^t \varphi\|^2_{L^2(\nu)} = \langle \mathcal{L} \mathcal{P}^t \varphi, \mathcal{P}^t \varphi \rangle = -\gamma \|\nabla_p \mathcal{P}^t \varphi\|^2_{L^2(\nu)}.$$

Langevin dynamics

Consider the following SDE for $x_t = (q_t, p_t) \in \mathcal{X} \subset \mathbf{R}^d \times \mathbf{R}^d$:

$$
\begin{aligned}
dq_t &= p_t \, dt \\
dp_t &= -\gamma p_t \, dt - \nabla U(q_t) \, dt + \sqrt{2\gamma} \, dB_t.
\end{aligned}
$$

Lyapunov?

- Generator: $\mathcal{L} = p \cdot \nabla_q - \gamma p \cdot \nabla_p - \nabla U(q) \cdot \nabla_p + \gamma \Delta_p$.
- $\mathcal{L} H(q,p) = -\gamma |p|^2 + \gamma d$.

Poincaré?

- If $\varphi \in L^2(\nu)$ with $\nu(\varphi) = 0$, then

$$
\frac{1}{2} \frac{d}{dt} \|\mathcal{P}^t \varphi\|^2_{L^2(\nu)} = \langle \mathcal{L} \mathcal{P}^t \varphi, \mathcal{P}^t \varphi \rangle = -\gamma \|\nabla_p \mathcal{P}^t \varphi\|^2_{L^2(\nu)}.
$$

Conclusion: Langevin dynamics is not pointwise contractive.

Averaging

Example: $d = 1$, $\gamma = 1$, $U(q) = \frac{|q|^4}{4} + \frac{1}{2|q|^2}$

More particles

Example: $d = 3$, $\gamma = 1$, $U(q) = \sum_i |q_i|^2 + \sum_{i \neq j} |q_i - q_j|^{-1.3}$

Averaging

Example: $d = 1,\ \gamma = 1, U(q) = \frac{|q|^4}{4} + \frac{1}{2|q|^2}$

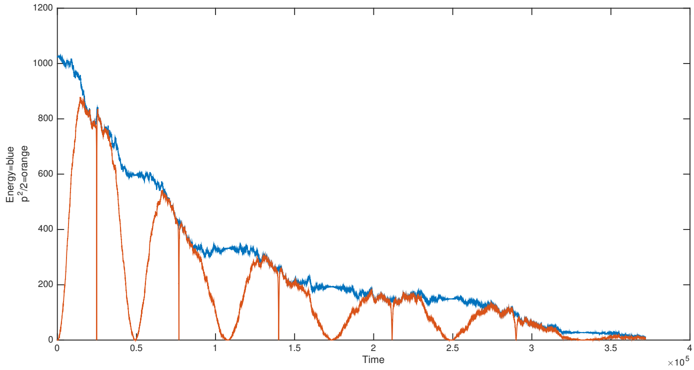


Figure: $p(t)^2/2$ and $H(q(t), p(t))$ plotted over $t \in [0, 4]$.

AVERAGING: HOW TO LEVERAGE?

Lyapunov: Let $\mathrm{Av}(f)(q,p)$ be the average value of f along Hamiltonian orbit containing (q,p), and

$$\mathcal{H} = p \cdot \nabla_q - \nabla U \cdot \nabla_p.$$

Then

$$\int_0^t |p_s|^2 \, ds = t\mathrm{Av}(|P|^2)(q,p) + \int_0^t |p_s|^2 - \mathrm{Av}(|P|^2)(q,p) \, ds.$$

Averaging: How to leverage?

Lyapunov: Let $\mathrm{Av}(f)(q,p)$ be the average value of f along Hamiltonian orbit containing (q,p), and

$$\mathcal{H} = p \cdot \nabla_q - \nabla U \cdot \nabla_p.$$

Then

$$\int_0^t |p_s|^2\, ds = t\mathrm{Av}(|P|^2)(q,p) + \int_0^t |p_s|^2 - \mathrm{Av}(|P|^2)(q,p)\, ds.$$

Use $V = H + \psi$ where ψ is lower-order and "satisfies"

$$\mathcal{H}\psi = |p|^2 - \mathrm{Av}(|P|^2)(q,p).$$

THE "pq TRICK"

For $d = 1$ and $U(q) = q^{2n}/2n$,

$$\mathcal{H}(pq) = (1+n)p^2 - 2nH(q,p) = (1+n)p^2 - \frac{1}{n+1}\mathrm{Av}(P^2)(q,p).$$

The "pq trick"

For $d = 1$ and $U(q) = q^{2n}/2n$,

$$\mathcal{H}(pq) = (1+n)p^2 - 2nH(q,p) = (1+n)p^2 - \frac{1}{n+1}\mathrm{Av}(P^2)(q,p).$$

Polynomial-like potentials ($|\nabla^2 U| \leq C|\nabla U|^1 + D$):

- D.Talay '00;
- L. Wu '01;
- Mattingly/Stuart/Higham '02;
- Rey-Bellet '06;
- *Zimmer '17 and Eberle, Guillin, Zimmer '19*.

Different choices of ψ and general potentials ($1 \mapsto 2$) :

- Cooke, H, Mattingly, McKinley, Schmidler '17;
- H, Mattingly '19;
- Lu, Mattingly '20.

Poincaré?

Recall for $\varphi \in L^2(\nu)$ with $\nu(\varphi) = 0$:

$$\|\mathcal{P}^t\varphi\|^2_{L^2(\nu)} - \|\varphi\|^2_{L^2(\nu)} = -2\gamma \int_0^t \|\nabla_p \mathcal{P}^s\varphi\|^2_{L^2(\nu)}\, ds,$$

so we hope that

$$\int_0^t \|\mathcal{P}^s\varphi\|^2_{L^2(\nu)}\, ds \lesssim \int_0^t \|\nabla_p \mathcal{P}^s\varphi\|^2_{L^2(\nu)}\, ds.$$

Poincaré?

Recall for $\varphi \in L^2(\nu)$ with $\nu(\varphi) = 0$:

$$\|\mathcal{P}^t \varphi\|_{L^2(\nu)}^2 - \|\varphi\|_{L^2(\nu)}^2 = -2\gamma \int_0^t \|\nabla_p \mathcal{P}^s \varphi\|_{L^2(\nu)}^2 \, ds,$$

so we hope that

$$\int_0^t \|\mathcal{P}^s \varphi\|_{L^2(\nu)}^2 \, ds \lesssim \int_0^t \|\nabla_p \mathcal{P}^s \varphi\|_{L^2(\nu)}^2 \, ds.$$

Idea 1: Follow the flow: If $\varphi_t = \mathcal{P}^t \varphi$ and $\|\cdot\| = \|\cdot\|_{L^2(\nu)}$, then

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|\nabla_p \varphi_t\|^2 &= \langle \nabla_p \mathcal{L} \varphi_t, \nabla_p \varphi_t \rangle \\ &= \langle [\nabla_p, \mathcal{L}] \varphi_t, \nabla_p \varphi_t \rangle + \langle \mathcal{L} \nabla_p \varphi_t, \nabla_p \varphi_t \rangle \\ &= \langle (\nabla_q - \gamma \nabla_p) \varphi_t, \nabla_p \varphi_t \rangle - \gamma \|\nabla_p^2 \varphi_t\|^2. \end{aligned}$$

Averaging: How to leverage this?

Idea 1: Follow the flow: If $\varphi_t = \mathcal{P}^t \varphi$ and $\tilde{\nabla} := \nabla_q - \gamma \nabla_p$, then

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dt}\|\tilde{\nabla}\varphi_t\|^2_{L^2(\nu)} &= \langle [\tilde{\nabla}, \mathcal{L}]\varphi_t, \tilde{\nabla}\varphi_t\rangle + \langle \mathcal{L}\tilde{\nabla}\varphi_t, \tilde{\nabla}\varphi_t\rangle \\
&= -\gamma\|\tilde{\nabla}\varphi_t\|^2 - \gamma\|\nabla_p\tilde{\nabla}\varphi_t\|^2 - \langle \nabla^2 U \nabla_p \varphi_t, \tilde{\nabla}\varphi_t\rangle
\end{aligned}
$$

Averaging: How to leverage this?

Idea 1: Follow the flow: If $\varphi_t = \mathcal{P}^t \varphi$ and $\tilde{\nabla} := \nabla_q - \gamma \nabla_p$, then

$$\frac{1}{2}\frac{d}{dt}\|\tilde{\nabla}\varphi_t\|^2_{L^2(\nu)} = \langle [\tilde{\nabla}, \mathcal{L}]\varphi_t, \tilde{\nabla}\varphi_t\rangle + \langle \mathcal{L}\tilde{\nabla}\varphi_t, \tilde{\nabla}\varphi_t\rangle$$
$$= -\gamma\|\tilde{\nabla}\varphi_t\|^2 - \gamma\|\nabla_p\tilde{\nabla}\varphi_t\|^2 - \langle \nabla^2 U \nabla_p \varphi_t, \tilde{\nabla}\varphi_t\rangle$$

Hence use the modified $H^1(\nu)$ norm

$$|||\varphi|||^2 := c_1\|\varphi\|^2_{L^2(\nu)} + c_2\|\nabla_p\varphi\|^2 + c_3\|\tilde{\nabla}\varphi\|^2.$$

- Desvillettes, Villani ’01; Hérau, Nier ’04; Hellfer, Nier ’05; Mouhot, Neumann ’06, Hérau ’07;
- Villani ’09;
- Conrad and Grothaus ’10; Grothaus and Stilgenbauer ’15;
- Baudoin ’17, Monmarché ’19;
- Cattiaux, Guillin, Monmarché, Zhang ’17 and Baudoin, Gordina, H ’21.

DMS: The direct $L^2(\nu)$ approach

Idea 2: Construct a norm equivalent to $L^2(\nu)$ instead:

$$\|\varphi\|_{1+\delta A}^2 := \|\varphi\|_{L^2(\nu)}^2 + \delta\langle A\varphi, \varphi\rangle.$$

If $\varphi_t = \mathcal{P}^t\varphi$ and $\nu(\varphi) = 0$, then

$$\begin{aligned} \frac{d}{dt}\langle A\varphi_t, \varphi_t\rangle &= \langle(\mathcal{L}^\dagger A + A\mathcal{L})\varphi_t, \varphi_t\rangle \\ &= \langle A\mathcal{H}\Pi\varphi_t, \varphi_t\rangle + R(\varphi_t) \end{aligned}$$

where

$$\Pi\varphi(q) = \frac{1}{\sqrt{(2\pi)^d}} \int_{\mathbf{R}^d} \varphi(q,p) e^{-\frac{|p|^2}{2}}\, dp.$$

DMS: The direct $L^2(\nu)$ approach

Idea 2: Construct a norm equivalent to $L^2(\nu)$ instead:

$$\frac{d}{dt}\langle A\varphi_t, \varphi_t\rangle = \langle A\mathcal{H}\Pi\varphi_t, \varphi_t\rangle + R(\varphi_t),$$

so pick $A = -(\mathcal{H}\Pi)^\dagger$ so that

$$\langle A\mathcal{H}\Pi\varphi_t, \varphi_t\rangle = -\|\mathcal{H}\Pi\varphi_t\|^2 = -\|p \cdot \nabla_q \Pi\varphi_t\|^2 = -c\|\nabla_q \Pi\varphi_t\|^2.$$

DMS: The direct $L^2(\nu)$ approach

Idea 2: Construct a norm equivalent to $L^2(\nu)$ instead:

$$\frac{d}{dt}\langle A\varphi_t, \varphi_t\rangle = \langle A\mathcal{H}\Pi\varphi_t, \varphi_t\rangle + R(\varphi_t),$$

so pick $A = -(\mathcal{H}\Pi)^\dagger$ so that

$$\langle A\mathcal{H}\Pi\varphi_t, \varphi_t\rangle = -\|\mathcal{H}\Pi\varphi_t\|^2 = -\|p \cdot \nabla_q \Pi\varphi_t\|^2 = -c\|\nabla_q \Pi\varphi_t\|^2.$$

Note: A above is not bounded on $L^2(\nu)$ so need to renormalize:

$$A\varphi = -(1 + (\mathcal{H}\Pi)^\dagger(\mathcal{H}\Pi))^{-1}(\mathcal{H}\Pi)^\dagger\varphi = \mathbf{E}_q \int_0^\infty e^{-s}\mathcal{H}\Pi\varphi(q_s)\, ds$$

DMS: THE DIRECT $L^2(\nu)$ APPROACH

Idea 2: Construct a norm equivalent to $L^2(\nu)$ instead:

$$\|\varphi\|^2_{1+\delta A} := \|\varphi\|^2_{L^2(\nu)} + \delta\langle A\varphi, \varphi\rangle.$$

- Hérau '06;
- Dolbeault, Mouhot, Schmeiser '09, '15;
- Grothaus and F-Y Wang '19;
- Leimkuhler, Sachs, Stoltz '20;
- Camrud, Gordina, H, Stoltz '21

Don't change the norm

Idea 3: Don't change the norm!
In other words, show that for $\varphi \in L^2(\nu)$ with $\nu(\varphi) = 0$:

$$\frac{1}{\tau}\int_t^{t+\tau} \|\mathcal{P}^s\varphi\|_{L^2(\nu)}^2 \, ds \leq \|\varphi\|_{L^2(\nu)}^2 e^{-\lambda(\tau)t}.$$

DON'T CHANGE THE NORM

Idea 3: Don't change the norm!
In other words, show that for $\varphi \in L^2(\nu)$ with $\nu(\varphi) = 0$:

$$\frac{1}{\tau}\int_t^{t+\tau} \|\mathcal{P}^s\varphi\|_{L^2(\nu)}^2\,ds \le \|\varphi\|_{L^2(\nu)}^2 e^{-\lambda(\tau)t}.$$

Time-averaged Poincaré?

$$c\int_t^{t+\tau} \|\mathcal{P}^s\varphi\|_{L^2(\nu)}^2\,ds \le \int_t^{t+\tau} \|\nabla_p\mathcal{P}^s\varphi\|_{L^2(\nu)}^2\,ds$$

DON'T CHANGE THE NORM

Hörmander's condition:
Let $U \subset \mathbf{R}^d$ be open, bounded and $X_0, X_1, \ldots, X_r$ be $C^\infty(U)$ vector fields. We say that $X_0, X_1, \ldots, X_r$ satisfies *Hörmander's condition on U* if for every $x \in U$, the list

$$\begin{array}{ll} X_{j_1}(x), & j_1 = 0, 1, \ldots, r \\ [X_{j_1}, X_{j_2}](x), & j_1, j_2 = 0, 1, \ldots, r \\ [X_{j_1}[X_{j_2}, X_{j_3}]](x), & j_1, j_2, j_3 = 0, 1, \ldots, r \\ \vdots & \vdots \end{array}$$

contains a basis of $\mathbf{R}^d$.

DON'T CHANGE THE NORM

Hörmander's condition:
Let $U \subset \mathbf{R}^d$ be open, bounded and $X_0, X_1, \ldots, X_r$ be $C^\infty(U)$ vector fields. We say that $X_0, X_1, \ldots, X_r$ satisfies *Hörmander's condition on U* if for every $x \in U$, the list

$$\begin{array}{ll} X_{j_1}(x), & j_1 = 0, 1, \ldots, r \\ [X_{j_1}, X_{j_2}](x), & j_1, j_2 = 0, 1, \ldots, r \\ [X_{j_1}[X_{j_2}, X_{j_3}]](x), & j_1, j_2, j_3 = 0, 1, \ldots, r \\ \vdots & \vdots \end{array}$$

contains a basis of $\mathbf{R}^d$.

Example. $X_0 = p\partial_q - U'(q)\partial_p - p\partial_p$ and $X_1 = \partial_p$. Note that $[X_1, X_0] = \partial_q - \partial_p$.

Theorem (Hörmander 1967)

Let $K \Subset U$ and suppose $\mathcal{M} = X_0 + \sum_{j=1}^{r} X_j^2$ and $X_0, X_1, \ldots, X_r$ satisfies Hörmander's condition on U. Then there exists $s, C > 0$ such that

$$\|u\|_{H^s} \leq C(\|\mathcal{M}u\|_{L^2} + \|u\|_{L^2})$$

for all $u \in C_0^\infty(K)$.

Theorem (Hörmander 1967)

Let $K \Subset U$ *and suppose* $\mathcal{M} = X_0 + \sum_{j=1}^r X_j^2$ *and* $X_0, X_1, \ldots, X_r$ *satisfies Hörmander's condition on* U*. Then there exists* $s, C > 0$ *such that*

$$\|u\|_{H^s} \leq C(\|\mathcal{M}u\|_{L^2} + \|u\|_{L^2})$$

for all $u \in C_0^\infty(K)$.

Actually:

$$\|u\|_{H^s} \leq C\Big(|||u||| + |||X_0 u|||' \Big),$$

$$|||u||| := \|u\|_{L^2} + \sum_{j=1}^r \|X_j u\|_{L^2}, \quad |||u|||' := \sup_{|||\varphi||| \leq 1} \int u\varphi \, dx$$

EXAMPLE IN $d = 1$

Question: How does this help?

Example. For Langevin in $d = 1$ with $\gamma = 1$, $\varphi_t = \mathcal{P}^t \varphi$. Then

$$\partial_t \varphi_t = \mathcal{L} \varphi_t = (\mathcal{H} + \partial_p^2) \varphi_t.$$

EXAMPLE IN $d = 1$

Question: How does this help?

Example. For Langevin in $d = 1$ with $\gamma = 1$, $\varphi_t = \mathcal{P}^t \varphi$. Then

$$\partial_t \varphi_t = \mathcal{L} \varphi_t = (\mathcal{H} + \partial_p^2) \varphi_t.$$

Formally setting $u = \varphi_t$ in estimate on $K \Subset U \subset (0, \infty) \times \mathbf{R}$:

$$\begin{aligned} \|\varphi_t\|_{H^s} &\leq C(\|\varphi_t\|_{L^2} + \|\partial_p \varphi_t\|_{L^2} + |||(\partial_t - \mathcal{H})\varphi_t|||') \\ &\leq C'(\|\varphi_t\|_{L^2} + \|\partial_p \varphi_t\|_{L^2}). \end{aligned}$$

EXAMPLE IN $d = 1$

Question: How does this help?

Example. For Langevin in $d = 1$ with $\gamma = 1$, $\varphi_t = \mathcal{P}^t\varphi$. Then

$$\partial_t\varphi_t = \mathcal{L}\varphi_t = (\mathcal{H} + \partial_p^2)\varphi_t.$$

Formally setting $u = \varphi_t$ in estimate on $K \Subset U \subset (0, \infty) \times \mathbf{R}$:

$$\begin{aligned}\|\varphi_t\|_{H^s} &\leq C(\|\varphi_t\|_{L^2} + \|\partial_p\varphi_t\|_{L^2} + |||(\partial_t - \mathcal{H})\varphi_t|||') \\ &\leq C'(\|\varphi_t\|_{L^2} + \|\partial_p\varphi_t\|_{L^2}).\end{aligned}$$

Conclusion: Try to obtain Poincaré inequality of the form

$$c\int_0^T \|\varphi_s\|_{L^2(\nu)}\, ds \leq \int_0^T \|\nabla_p\varphi_s\|_{L^2(\nu)}\, ds + |||(\partial_t - \mathcal{H})\varphi_t|||'_\nu$$

Time-dependent Poincaré inequalities

Try to obtain time-dependent Poincaré of the form:

$$c \int_0^\tau \|\varphi_s\|_{L^2(\nu)}\, ds \leq \int_0^\tau \|\nabla_p \varphi_s\|_{L^2(\nu)}\, ds + |||(\partial_t - \mathcal{H})\varphi_t|||'_\nu$$

- Y. Guo '02;
- Strain and Guo '04;
- Albritton, Armstrong, Mourrat and Novack '21;
- Cao, Lu and Wang '19;
- *Bedrossian and Liss '21: 2D Galerkin Navier-Stokes* .
- Brigatti '22;
- Brigatti and Stoltz '23.

Thank you!