



European Research Council
Established by the European Commission



Longtime convergence of evolution semigroups in molecular dynamics

Gabriel STOLTZ

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

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

- **A quick introduction to molecular dynamics**
- **(Non)equilibrium Langevin dynamics**
 - Various convergence results
 - The hypocoercive approach by Dolbeault, Mouhot and Schmeiser
 - Various extensions and modifications
- **Feynmann–Kac dynamics**
 - Reformulation in terms of evolution semigroups
 - Proof for compact position spaces
 - Statement of the result in the general case
 - Elements of proof

A quick introduction to molecular dynamics

Computational statistical physics (1)

- *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 $H(q, p) = V(q) + \sum_{i=1}^N \frac{p_i^2}{2m_i}$

- **Macrostate**

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

- **Properties :**

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

Computational statistical physics (2)

- Positions $q \in \mathcal{D} = (LT)^d$ or \mathbb{R}^d and momenta $p \in \mathbb{R}^d$
→ phase-space $\mathcal{E} = \mathcal{D} \times \mathbb{R}^d$
- The very **high dimensional** average $\langle \varphi \rangle$ is computed using **time averages** of dynamics ergodic for μ :

$$\langle \varphi \rangle = \lim_{T \rightarrow +\infty} \frac{1}{T} \int_0^T \varphi(q_t, p_t) dt$$

- **Examples of dynamics:**
 - Deterministic dynamics (Hamiltonian, Nosé–Hoover and its variations)
 - **Stochastic differential equations**
 - Markov chains (Metropolis schemes, discretizations of SDEs)
 - Piecewise deterministic Markov processes

Convergence results for evolution semigroups of Langevin dynamics

Langevin dynamics (1)

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

Stochastic perturbation of the Hamiltonian dynamics

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

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

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

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

- In both cases, **slow convergence to equilibrium**

Stochastic differential equations and their generators

- General SDE $dx_t = b(x_t) dt + \sigma(x_t) dW_t$ on \mathcal{X}
- **Generator** of the dynamics $\left. \frac{d}{dt} \left(\mathbb{E} \left[\varphi(x_t) \mid x_0 = x \right] \right) \right|_{t=0} = (\mathcal{L}\varphi)(x)$
- It holds $\mathcal{L} = b^T \nabla + \frac{1}{2} \sigma \sigma^T : \nabla^2 = \sum_{i=1}^d b_i \partial_{x_i} + \sum_{i,j=1}^d (\sigma \sigma^T)_{ij} \partial_{x_i, x_j}^2$
- Invariance of the probability measure $\pi(dx)$ characterized by

$$\forall \varphi \in C_0^\infty(\mathcal{X}), \quad \int_{\mathcal{X}} \mathcal{L}\varphi d\pi = 0$$

- **Evolution semigroup** $(e^{t\mathcal{L}}\varphi)(x) = \mathbb{E} \left[\varphi(x_t) \mid x_0 = x \right]$
- The latter quantity is expected to converge to $\int_{\mathcal{X}} \varphi d\pi$

Fokker–Planck equations

- **Dual viewpoint:** convergence of the distribution rather than convergence of observables (Schrödinger vs. Heisenberg)
- Evolution of the law $\psi(t, x)$ of the process at time $t \geq 0$

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

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

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

- It is expected that $\psi(t, x) dx$ converges to $\pi(dx)$

Langevin dynamics (2)

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

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

- Preserves the **canonical measure**

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

- 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}^*

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

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

Ergodicity results (1)

- Almost-sure convergence¹ of **ergodic averages** $\widehat{\varphi}_t = \frac{1}{t} \int_0^t \varphi(q_s, p_s) ds$
- **Asymptotic variance** of ergodic averages

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

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

- A central limit theorem holds² 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

¹Kliemann, *Ann. Probab.* **15**(2), 690-707 (1987)

²Bhattacharya, *Z. Wahrsch. Verw. Gebiete* **60**, 185-201 (1982)

Ergodicity results (2)

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

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

- Prove **exponential convergence** of the semigroup $e^{t\mathcal{L}}$
 - various Banach spaces $E \cap L_0^2(\mu)$
 - **Lyapunov** techniques^{3,4,5} $L_W^\infty(\mathcal{E}) = \left\{ \varphi \text{ measurable, } \left\| \frac{\varphi}{W} \right\|_{L^\infty} < +\infty \right\}$
 - standard **hypocoercive**⁶ setup $H^1(\mu)$
 - $E = L^2(\mu)$ after hypoelliptic regularization⁷ from $H^1(\mu)$
 - **coupling** arguments⁸

³L. Rey-Bellet, *Lecture Notes in Mathematics* (2006)

⁴Hairer and Mattingly, *Progr. Probab.* **63** (2011)

⁵Mattingly, Stuart and Higham, *Stoch. Proc. Appl.* (2002)

⁶Villani (2009) and before Talay (2002), Eckmann/Hairer (2003), Hérau/Nier (2004)

⁷F. Hérau, *J. Funct. Anal.* **244**(1), 95-118 (2007)

⁸A. Eberle, A. Guillin and R. Zimmer, *arXiv preprint* **1703.01617** (2017)

Direct $L^2(\mu)$ approach

- Assume that the potential V is **smooth** and^{9,10}
 - the marginal measure ν satisfies a **Poincaré** inequality

$$\|\Pi_0 \varphi\|_{L^2(\nu)}^2 \leq \frac{1}{C_\nu} \|\nabla_q \varphi\|_{L^2(\nu)}^2.$$

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

$$\Delta V \leq c_1 + \frac{c_2}{2} |\nabla V|^2, \quad |\nabla^2 V| \leq c_3 (1 + |\nabla V|).$$

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

$$\forall t \geq 0, \quad \|e^{t\mathcal{L}} \varphi\|_{L^2(\mu)} \leq C e^{-\lambda_\gamma t} \|\varphi\|_{L^2(\mu)}.$$

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

$$\lambda_\gamma \geq \bar{\lambda} \min(\gamma, \gamma^{-1}).$$

⁹Dolbeault, Mouhot and Schmeiser, *C. R. Math. Acad. Sci. Paris* (2009)

¹⁰Dolbeault, 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
$$A = \left(1 + (\mathcal{L}_{\text{ham}} \Pi_p)^*(\mathcal{L}_{\text{ham}} \Pi_p)\right)^{-1} (\mathcal{L}_{\text{ham}} \Pi_p)^*, \quad \Pi_p \varphi = \int_{\mathbb{R}^D} \varphi d\kappa$$
- $A = \Pi_p A (1 - \Pi_p)$ and $\mathcal{L}_{\text{ham}} A$ are bounded so that $\mathcal{H} \sim \|\cdot\|_{L^2(\mu)}^2$

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

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

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

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

- Gronwall inequality $\frac{d}{dt} (\mathcal{H} [e^{t\mathcal{L}}\varphi]) = -\mathcal{D} [e^{t\mathcal{L}}\varphi] \leq -\frac{2\tilde{\lambda}_\gamma}{1+\varepsilon} \mathcal{H} [e^{t\mathcal{L}}\varphi]$

Extensions/modifications/variations

- **General kinetic energy** function $U(p)$ in the Langevin dynamics¹¹
 - heavy/light tails
 - ∇U vanishes on open sets (generator not hypoelliptic)
- **Galerkin discretization** and variance reduction¹²
- Convergence of certain nonequilibrium methods for computing free energy differences¹³
- **One more precise result:** nonequilibrium Langevin dynamics with external forcing

¹¹G. Stoltz and Z. Trstanova, accepted in *Multiscale Model. Sim.* (2018)

¹²J. Roussel and G. Stoltz, *M2AN*, 2018

¹³G. Stoltz and E. Vanden-Eijnden, *Nonlinearity*, 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

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

- 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 $H^1(\mu)$
→ nonequilibrium perturbation¹⁴ 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 $\bar{\lambda}_\delta > 0$ for which, for all $\gamma \in (0, +\infty)$ and all $\tau \in [-\delta \min(\gamma, 1), \delta \min(\gamma, 1)]$,

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

- As a corollary: lower bounds on the **spectral gap** of order $\min(\gamma, \gamma^{-1})$
→ can be checked numerically¹⁵

¹⁴E. Bouin, F. Hoffmann, and C. Mouhot, *arXiv preprint* **1605.04121**

¹⁵A. Iacobucci, S. Olla and G. Stoltz, to appear in *Ann. Math. Quebec* (2017)

Convergence of Feynmann–Kac dynamics

Feynmann–Kac averages

- Diffusion process X_t , **weighted with an exponential factor** $\int_0^t f(X_s) ds$
- Evolution of probability measures as

$$\Theta_t(\mu)(\varphi) = \frac{\mathbb{E} \left[\varphi(X_t) e^{\int_0^t f(X_s) ds} \mid x_0 \sim \mu \right]}{\mathbb{E} \left[e^{\int_0^t f(X_s) ds} \mid x_0 \sim \mu \right]},$$

Convergence of $\Theta_t(\mu)$?

Show that there exists a unique probability measure μ_f^* such that $\Theta_t(\mu)(\varphi) \rightarrow \mu_f^*(\varphi)$ as $t \rightarrow +\infty$, and quantify the rate of convergence.

- Applications in **Diffusion Monte Carlo** and computation of **large deviations** estimates

Analytical reformulation

- Evolution semigroup $(P_t^f \varphi)(x) = \mathbb{E}^x \left(\varphi(X_t) e^{\int_0^t f(X_s) ds} \right)$
- In fact, $P_t^f = e^{t(\mathcal{L}+f)}$ where \mathcal{L} is the generator of X_t , so that

$$\Theta_t(\mu)(\varphi) = \frac{\int_{\mathcal{X}} e^{t(\mathcal{L}+f)} \varphi d\mu}{\int_{\mathcal{X}} e^{t(\mathcal{L}+f)} \mathbf{1} d\mu}.$$

- One expects that $\Theta_t(\mu)$ converges to some probability measure
- Convergence rate related to some **spectral gap**
- Simple analysis for **compact spaces** $\mathcal{X} = \mathbb{T}^d$ or for **self-adjoint generators**

A simple case: additive noise, compact space \mathcal{D} (1)

- Dynamics $dX_t = b(X_t) dt + \sqrt{2} dW_t$
 - Invariant probability measure $\nu(dx)$ (unknown expression)
 - Generator $\mathcal{L} = b^T \nabla + \Delta$, considered on $L^2(\nu)$, **discrete spectrum**
 - **First eigenvectors** of \mathcal{L} and \mathcal{L}^* : **positive**, unique up to normalization

$$(\mathcal{L} + f)\widehat{h}_f = \lambda_f \widehat{h}_f, \quad (\mathcal{L}^* + f)h_f = \lambda_f h_f, \quad \int_{\mathcal{D}} h_f d\nu = \int_{\mathcal{D}} \widehat{h}_f d\nu = 1$$

- Then $e^{t(\mathcal{L}+f-\lambda_f)}g$ converges **exponentially fast** to $\frac{\langle g, h_f \rangle_{L^2(\nu)}}{\langle h_f, \widehat{h}_f \rangle_{L^2(\nu)}} h_f$
- This allows to identify the limiting probability measure $\mu_f^* \propto h_f d\nu$

Convergence in the general case (1)

- Unstructured dynamics: **Lyapunov approach**

Assumption 1 (Lyapunov conditions)

There is a $C^2(\mathcal{X})$ function $W : \mathcal{X} \rightarrow [1, +\infty)$ going to infinity at infinity such that

$$W^{-1}(\mathcal{L} + f)W \xrightarrow{|x| \rightarrow +\infty} -\infty.$$

In addition, there exist a $C^2(\mathcal{X})$ function $\mathscr{W} : \mathcal{X} \rightarrow [1, +\infty)$ and a constant $c \geq 0$ such that

$$\varepsilon(x) := \frac{\mathscr{W}(x)}{W(x)} \xrightarrow{|x| \rightarrow +\infty} 0, \quad \mathscr{W}^{-1}(\mathcal{L} + f)\mathscr{W} \leq c.$$

- Typical choice: $W(x) = e^{\alpha V(x)}$ and $\mathscr{W}(x) = e^{\alpha' V(x)}$ with $\alpha' \leq \alpha$
- Example: $\sigma(x) = \sqrt{2}$, $b(x) = -\nabla V(x)$, with, for some $a \in (1/2, 1)$,

$$\lim_{|x| \rightarrow +\infty} \left(-\beta(1-a)|\nabla V|^2 + a\Delta V + f \right) = -\infty$$

Convergence in the general case (2)

Assumption 2 (Regularity and positivity of the transition kernel)

The functions f and σ are continuous and, for any $t > 0$, the transition kernel P_t^f has a continuous positive density with respect to the Lebesgue measure: $P_t^f(x, dy) = p_t^f(x, y) dy$ with $p_t^f(x, y) > 0$ for all $x, y \in \mathcal{X}$.

- Introduce $B_W^\infty(\mathcal{X}) = \left\{ \varphi \text{ measurable, } \sup_{x \in \mathcal{X}} \left| \frac{\varphi(x)}{W(x)} \right| < +\infty \right\}$

Theorem (Ferré/Rousset/Stoltz, 2018)

There exist a unique invariant measure μ_f^* and $\kappa > 0$ such that, for any initial measure $\mu \in \mathcal{P}(\mathcal{X})$ with $\mu(W) < +\infty$, there is $C_\mu > 0$ for which

$$\forall \varphi \in B_W^\infty(\mathcal{X}), \quad \forall t > 0, \quad \left| \Theta_t(\mu)(\varphi) - \mu_f^*(\varphi) \right| \leq C_\mu e^{-\kappa t} \|\varphi\|_{B_W^\infty}.$$

Moreover, the invariant measure satisfies $\mu_f^*(W) < +\infty$.

Sketch of proof (1)

- Reduction to time-discrete case: $Q^f = e^{t_0(\mathcal{L}+f)}$ for some fixed $t_0 > 0$

Key result

The operator Q^f considered on $B_W^\infty(\mathcal{X})$ has a zero essential spectral radius, admits its spectral radius $\Lambda > 0$ as a largest eigenvalue (in modulus), and has an associated eigenfunction $h \in B_W^\infty(\mathcal{X})$, normalized so that $\|h\|_{B_W^\infty} = 1$, and which satisfies $0 < h(x) < +\infty$ for all $x \in \mathcal{X}$.

- It is then possible to consider the **Markov kernel** $Q_h\phi = \Lambda^{-1}h^{-1}Q^f(h\phi)$
- It suffices to understand the convergence of Q_h since

$$\Theta_{kt_0}(\mu)(\varphi) = \frac{\mu(h(Q_h)^k(h^{-1}\varphi))}{\mu(h(Q_h)^k h^{-1})}$$

- Denoting by μ_h the **invariant measure for Q_h** ,

$$\mu_f^*(\varphi) = \frac{\mu_h(h^{-1}\varphi)}{\mu_h(h^{-1})}$$

Sketch of proof (2)

- Convergence of Q_h : standard convergence results for Markov operators¹⁶

Lyapunov condition

There exist a function $\mathcal{K} : \mathcal{X} \rightarrow [1, +\infty)$ and constants $C \geq 0$, $\gamma \in (0, 1)$ such that $Q\mathcal{K} \leq \gamma\mathcal{K} + C$.

The Lyapunov function for Q_h is $Wh^{-1} : \mathcal{X} \rightarrow [1, +\infty)$.

Minorization

There exist $\alpha \in (0, 1)$ and $\eta \in \mathcal{P}(\mathcal{X})$ such that $\inf_{x \in \mathcal{C}} Q(x, \cdot) \geq \alpha\eta(\cdot)$, where $\mathcal{C} = \{x \in \mathcal{X} \mid \mathcal{W}(x) \leq R + 1\}$ for some $R > 2C/(1 - \gamma)$.

Then, Q has a unique invariant measure μ_* , which is such that $\mu_*(\mathcal{W}) < +\infty$. Moreover, there exist $K > 0$ and $\bar{\alpha} \in (0, 1)$ such that,

$$\forall \varphi \in B_{\mathcal{W}}^{\infty}(\mathcal{X}), \quad \forall k \geq 0, \quad \|Q^k \varphi - \mu_*(\varphi)\|_{B_{\mathcal{W}}^{\infty}} \leq K \bar{\alpha}^k \|\varphi - \mu_*(\varphi)\|_{B_{\mathcal{W}}^{\infty}}.$$

¹⁶Hairer and Mattingly, *Progr. Probab.* **63** (2011)

Elements of proof of the key result

- The **essential spectral radius** θ of Q^f is **zero**: rely on the decomposition

$$(Q^f)^3 = (\mathbf{1}_K Q^f \mathbf{1}_K)^2 Q^f + \mathbf{1}_{K^c} Q^f (\mathbf{1}_K Q^f)^2 + Q^f \mathbf{1}_{K^c} (Q^f)^2 + Q^f \mathbf{1}_K Q^f \mathbf{1}_{K^c} Q^f$$

with $(\mathbf{1}_K Q^f \mathbf{1}_K)^2$ compact (using some continuity property and Ascoli) while $\mathbf{1}_{K^c} Q^f$ tends to 0 as K increases

- The **spectral radius** Λ of Q^f (considered as an operator on $B_W^\infty(\mathcal{X})$) is **positive** [rely on minorization conditions]
- **Krein–Rutman theorem** on the cone $\mathbb{K}_W = \{u \in B_W^\infty(\mathcal{X}) \mid u \geq 0\}$:
 - the cone is total (the norm closure of $\mathbb{K}_W - \mathbb{K}_W$ is $B_W^\infty(\mathcal{X})$)
 - The positiveness of $Q^f \in B_W^\infty(\mathcal{X})$ shows that $Q^f \mathbb{K}_W \subset \mathbb{K}_W$.
 - $\theta < \Lambda$

This shows that Λ is an eigenvalue of Q^f with an eigenvector in \mathbb{K}_W .