Sharp convergence in the total variation distance for Langevin dynamics

Gerardo Barrera
(University of Helsinki, Finland)

Joint work with
Milton Jara (IMPA, Rio de Janeiro, Brazil),
Conrado da Costa (Durham University, Durham, U.K.).

04th/April/2023

Plan

Mixing-time

Mixing-time

Well-posed problem

1. Let $(X_t(x_0))_{t\geq 0}$ be a dynamical system (model) such that the current state at time t, $X_t(x_0)$, converges to its limiting state μ as $t \to \infty$ for some suitable (distance/discrepancy) dist.
2. Assume that the map $t \mapsto \mathrm{dist}(X_t(x_0), \mu)$ is non-increasing.
3. Well-posed problem: given a prescribed error $\eta > 0$, the η-mixing-time is defined as

$$\tau_\eta(x_0) := \inf\{t \geq 0 : \mathrm{dist}(X_t(x_0), \mu) \leq \eta\}.$$

It is a hard problem in general! For instance, in ergodic Markov chains theory we have the existence of constants $C(x_0, N)$ and $\delta(N)$ satisfying

$$\left\| X_t^N(x_0) - \mu^N \right\|_{\mathbb{TV}} \leq C(x_0, N) e^{-\delta(N) t} \quad \text{for all} \quad t \geq 0.$$

Here N denotes the cardinality of the space state. Good knowledge of $C(x_0, N)$ and $\delta(N)$ is needed. What about lower bound estimates?

A warm-up example: the Ornstein–Uhlenbeck process

Let $(X_t(x_0))_{t\geq 0}$ be the solution of the SDE

$$\mathrm{d}X_t = -\lambda X_t \mathrm{d}t + \sqrt{\epsilon}\mathrm{d}W_t, \qquad X_0 = x_0,$$

with λ and ϵ being positive numbers. Integration by parts formula yields

$$X_t(x_0) = e^{-\lambda t}x_0 + \sqrt{\epsilon}e^{-\lambda t}\int_0^t e^{\lambda s}\mathrm{d}W_s.$$

Then Itô's isometry implies

$$X_t(x_0) \stackrel{\mathcal{D}}{=} \mathcal{N}\left(e^{-\lambda t}x_0, \frac{\epsilon}{2\lambda}(1-e^{-2\lambda t})\right),$$

and the limiting law is given by

$$\mu \stackrel{\mathcal{D}}{=} \mathcal{N}\left(0, \frac{\epsilon}{2\lambda}\right).$$

OUP total variation mixing time

For computing the η-mixing time, one needs to find the smallest $t > 0$ for which

$$\|X_t(x_0) - \mu\|_{\mathbb{TV}} = \left\| \mathcal{N}\left(e^{-\lambda t}x_0, \frac{\epsilon}{2\lambda}(1-e^{-2\lambda t})\right) - \mathcal{N}\left(0, \frac{\epsilon}{2\lambda}\right) \right\|_{\mathbb{TV}}$$
$$= \left\| \mathcal{N}\left(\sqrt{\frac{2\lambda}{\epsilon}}e^{-\lambda t}x_0, 1-e^{-2\lambda t}\right) - \mathcal{N}(0,1) \right\|_{\mathbb{TV}} \leq \eta.$$

Since

$$|\sqrt{1-e^{-2\lambda t}} - 1| = \frac{e^{-2\lambda t}}{\sqrt{1-e^{-2\lambda t}}+1}, \quad \text{morally, one needs to solve}$$

$$\left\| \mathcal{N}\left(\sqrt{\frac{2\lambda}{\epsilon}}e^{-\lambda t}x_0, 1\right) - \mathcal{N}(0,1) \right\|_{\mathbb{TV}} = \frac{2}{\sqrt{2\pi}} \int_0^{\sqrt{\frac{\lambda}{2\epsilon}}e^{-\lambda t}|x_0|} e^{-z^2/2}\mathrm{d}z \leq \eta,$$

which for t large yields

$$\frac{2}{\sqrt{2\pi}}\sqrt{\frac{\lambda}{2\epsilon}}e^{-\lambda t}|x_0| = \sqrt{\frac{\lambda}{\pi\epsilon}}e^{-\lambda t}|x_0| \leq \eta.$$

OUP Wasserstein mixing time

To compute the η-mixing time, one needs to find the smallest $t > 0$ satisfying

$$\mathcal{W}_2(X_t(x_0), \mu) = \mathcal{W}_2\left(\mathcal{N}\left(e^{-\lambda t}x_0, \frac{\epsilon}{2\lambda}(1-e^{-2\lambda t})\right), \mathcal{N}\left(0, \frac{\epsilon}{2\lambda}\right)\right)$$
$$= \sqrt{e^{-2\lambda t}x_0^2 + \frac{\epsilon}{2\lambda}\left(\sqrt{1-e^{-2\lambda t}}-1\right)^2} \leq \eta.$$

Since

$$\left(\sqrt{1-e^{-2\lambda t}}-1\right)^2 = \frac{e^{-4\lambda t}}{(\sqrt{1-e^{-2\lambda t}}+1)^2},$$

morally, we have

$$e^{-\lambda t}|x_0| \leq \eta.$$

Cutoff phenomenon (in total variation distance)

In the eighties, P. Diaconis and M. Shahshahani asked:

- How many shufflings do we need for the deck of cards to be well-mixed?
- When is it random enough?

Behavior

Table 1. Distance to stationarity for repeated shuffles of 52 cards

	k									
	1	2	3	4	5	6	7	8	9	10
$\|P^k - \pi\|$	1.000	1.000	1.000	1.000	0.924	0.624	0.312	0.161	0.083	0.041

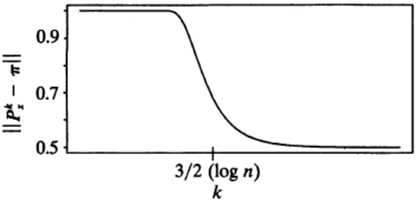


FIG. 1. The cutoff phenomenon for repeated riffle shuffles of n = 52 cards.

Figure: P. Diaconis. *The cut-off phenomenon in finite Markov chains*, PNAS, USA, 93, 1659-1664, 1996.

Random transpositions

- **Space of state**:
  S_n = set of biyections (permutations) on $\{1, 2, \ldots, n\}$. Note that $|S_n| = n!$ is too big.

$$|S_{52}| \gg \text{Avogadro's number} = 6.022 \times 10^{23}.$$

- **Dynamics**: Pairs of cards are exchanged randomly.
- **Equilibrium**: Uniform probability on S_n, that is,

$$\mathbb{P}(\sigma) = \frac{1}{n!} \quad \text{for each permutation} \quad \sigma \in S_n.$$

- **Well-mixed**: $\frac{1}{2} n \log(n)$ steps. The proof can be done by Fourier analysis and representation theory.

The model: unidimensional Langevin dynamics

$$\mathrm{d}X_t^\epsilon = -V'(X_t^\epsilon)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}W_t, \qquad X_0^\epsilon = x_0,$$

where

i) $V : \mathbb{R} \to [0, \infty)$ is a smooth convex degenerate potential,

ii) $\epsilon \in (0, 1]$ controls the amplitude of the noise,

iii) $W = (W_t)_{t \geq 0}$ is a unidimensional standard Brownian motion,

iv) x_0 is an initial datum.

Limiting distribution

- Under a suitable growth condition on V, the Langevin system possesses a unique invariant probability measure μ^ϵ.
- The density $\mu^\epsilon(\mathrm{d}z)$ is given by

$$\mu^\epsilon(\mathrm{d}z) = \frac{1}{\mathcal{Z}^\epsilon} \exp\left(-\frac{2V(z)}{\epsilon}\right) \mathrm{d}z,$$

  where $\mathcal{Z}^\epsilon$ is the partition function.
- **Weak ergodicity:** For any initial condition x_0, the law of the marginal $X_t^\epsilon(x_0)$ converges in distribution to μ^ϵ as $t \to \infty$, i.e.,

$$\lim_{t\to\infty} \mathbb{E}\left[f(X_t^\epsilon(x_0))\right] = \int_{\mathbb{R}} f(z)\mu^\epsilon(\mathrm{d}z) \quad \text{for all} \quad f \in \mathcal{C}_b(\mathbb{R},\mathbb{R}).$$

Strong ergodicity

- **Strong ergodicity:** For any initial condition x_0, the law of the marginal $X_t^\epsilon(x_0)$ converges in the total variation distance to μ^ϵ as $t \to \infty$, i.e.,

$$\lim_{t\to\infty} \sup_{|f|_\infty \le 1} \left| \mathbb{E}\left[f(X_t^\epsilon(x_0))\right] - \int_{\mathbb{R}} f(z)\mu^\epsilon(\mathrm{d}z) \right| = 0,$$

  where the supremum is taken over all measurable functions $f : \mathbb{R} \to [-1, 1]$.
- For short, we write $\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}}$ instead of

$$\sup_{|f|_\infty \le 1} \left| \mathbb{E}\left[f(X_t^\epsilon(x_0))\right] - \int_{\mathbb{R}} f(z)\mu^\epsilon(\mathrm{d}z) \right|.$$

Rate of convergence

- **Exponential ergodicity:** For any $\epsilon > 0$ there exists a positive constant δ_ϵ such that for any $x_0 \in \mathbb{R}$, there is a positive constant $C_\epsilon(x_0)$ satisfying

$$\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} \leq C_\epsilon(x_0)e^{-\delta_\epsilon t} \quad \text{for all} \quad t \geq 0.$$

- $C_\epsilon(x_0)$ is hard to compute and/or estimate.
- δ_ϵ is related with the spectral gap of the generator $\mathcal{L}$ of the Markov process $(X_t^\epsilon(x_0))_{t\geq 0}$.
- Lower bounds for $\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}}$ are also very difficult to obtain.

Properties of the variation distance

i) $[0,\infty) \ni t \mapsto \|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} \in [0,1]$ is non-increasing.

ii) Variational representation:

$$\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} = \inf_{\pi \text{ Coupling}} \pi(\text{ outside of the diagonal }).$$

iii) Dual representation:

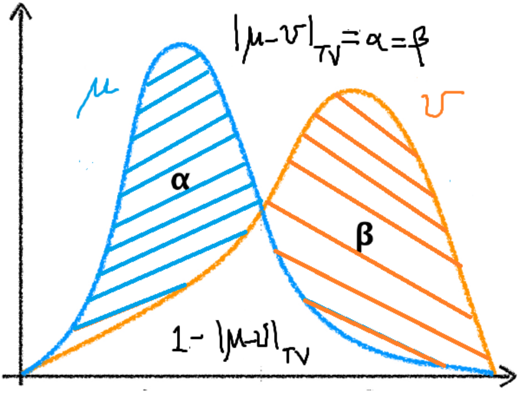
$$\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} = \sup_{|f|_\infty \leq 1} |\mathbb{E}[f(X_t^\epsilon(x_0))] - \mathbb{E}_{\mu^\epsilon}[f]| \,.$$

iv) Smooth representation:

$$\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} = \frac{1}{2}\int_{\mathbb{R}} |f^{\epsilon,x_0}(z,t) - f^\epsilon(z)|\mathrm{d}z,$$

where $\mathbb{P}(X_t^\epsilon(x_0) \in \mathrm{d}z) = f^{\epsilon,x_0}(z,t)\mathrm{d}z$ and $\mu^\epsilon(\mathrm{d}z) = f^\epsilon(z)\mathrm{d}z$.

Description of the variation distance



Goal

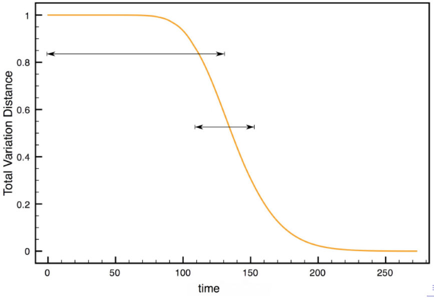
Our aim is to analyze the shape of the convergence to zero of

$$d_t^\epsilon(x_0) = \|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}}$$

for each datum $x_0 \in \mathbb{R}$ and degenerate smooth convex potentials V.
In particular, there is cutoff at $(a_\epsilon)_{\epsilon>0}$ if

$$\lim_{\epsilon\to 0^+} d_{a_\epsilon t}^\epsilon(x_0) = \begin{cases} 1 & \text{for} \quad t \in (0,1) \\ 0 & \text{for} \quad t \in (1,+\infty). \end{cases}$$

The cutoff phenomenon is associated to a switching phenomenon.



Stability: degenerate vs hyperbolic fixed points

Let V such that $V'(0) = 0$ and $V''(x) \geq 0$ and consider the IVP

$$\dot{\varphi}_t^{x_0} = -V'(\varphi_t^{x_0}), \qquad \varphi_0^{x_0} = x_0.$$

Degenerate

$$\begin{aligned} V''(0) &= 0, \\ V(x) &= x^4, \\ V(x) &= x^4 + x^6. \end{aligned}$$

Polynomial rate
$|\varphi_t^{x_0}| \sim C_D(x_0) t^{-1/2}$

Hyperbolic

$$\begin{aligned} V''(0) &> 0, \\ V(x) &= x^2, \\ V(x) &= x^2 + x^4. \end{aligned}$$

Exponential rate
$|\varphi_t^{x_0}| \sim C_H(x_0) e^{-2t}$

Theorem (Hyperbolic case, B. & Jara, 2016, JSP)

Let $V:\mathbb{R}\to[0,\infty)$ *be a* $\mathcal{C}^2$*-function with* $V(0)=V'(0)=0$ *satisfying* **the coercivity condition:** *there is a positive* δ *such that* $V''(x)\geq\delta$ *for all* $x\in\mathbb{R}$. *Consider the unique strong solution of the SDE*

$$\begin{cases} \mathrm{d}X_t^\epsilon = -V'(X_t^\epsilon)\mathrm{d}t+\sqrt{\epsilon}\mathrm{d}B_t \quad \text{for} \quad t>0,\\ X_0^\epsilon = x_0, \end{cases}$$

where x_0 *is an initial condition,* $(B_t)_{t\geq 0}$ *is a standard Brownian motion. Then for any* $x_0\neq 0$ *and* $\rho\in\mathbb{R}$ *it follows that*

$$\lim_{\epsilon\to 0^+}\left\|X^\epsilon_{t_\epsilon+\rho\cdot w_\epsilon}(x_0)-\mu^\epsilon\right\|_{\mathbb{TV}}=\frac{2}{\sqrt{2\pi}}\int_0^{\lambda(x_0)e^{-\rho}}\exp\left(-\frac{z^2}{2}\right)\mathrm{d}z,$$

where $t_\epsilon:=\frac{1}{2V''(0)}\ln(1/\epsilon)$, $w_\epsilon:=\frac{1}{V''(0)}+\mathrm{o}_{\epsilon\to 0^+}(1)$,

$$\lambda(x_0):=\frac{\sqrt{2V''(0)}|c(x_0)|}{2}\quad\text{with}\quad c(x_0):=\lim_{t\to\infty}e^{V''(0)t}X_t^0(x_0)>0.$$

Idea for the proof

Linear nonhomogeneous approximation:
Freidlin–Wentzell first order

↓

Cutoff for the nonhomogeneous linear approximation
Hartman–Grobman Theorem

↓

Coupling the nonlinear and linear approximation
Short time coupling

↓

Coupling the limiting measures
Bootstrap argument/Direct computation dimension one

Main result

Theorem (Degenerate case, B., da Costa & Jara, 2022)

Let $V : \mathbb{R} \to [0, \infty)$ *be a* $\mathcal{C}^2$*, convex, even function with* $V(0) = 0$ *that satisfies the following conditions:*

(1) Local behavior at zero: *there exist constants* $C_0 > 0$ *and* $\alpha > 0$ *such that for any* $K > 0$ *we have*

$$\lim_{\lambda \to 0^+} \sup_{|x| \leq K} \left| \frac{V'(\lambda x)}{\lambda^{1+\alpha}} - C_0 |x|^{1+\alpha} \mathrm{sgn}(x) \right| = 0,$$

where $\mathrm{sgn}(x) := {}^{x}/_{|x|}$ *for* $x \neq 0$ *and* $\mathrm{sgn}(0) := 0$.

(2) Growth condition at infinity: *there exist* $c_0, R_0 \in (0, \infty)$*, and* $\beta \in (-1, \infty)$ *such that*

$$V'(x) \geq c_0 x^{1+\beta} \quad \textit{for all} \quad x \geq R_0. \qquad \textbf{(G)}$$

Main result

Theorem (Degenerate case)

Consider the unique strong solution of the SDE

$$\begin{cases} \mathrm{d}X_t^\epsilon = -V'(X_t^\epsilon)\mathrm{d}t + \sqrt{\epsilon}\mathrm{d}B_t \quad \text{for} \quad t > 0, \\ X_0^\epsilon = x_0, \end{cases}$$

where x_0 is an initial condition, $(B_t)_{t\geq 0}$ is a standard Brownian motion. Then for $t > 0$ it follows that

$$\lim_{\epsilon \to 0^+} \left\| X_{a_\epsilon t}^\epsilon(x_0) - \mu^\epsilon \right\|_{\mathbb{TV}} = \| Y_t(x_0) - \nu \|_{\mathbb{TV}} \in (0,1),$$

where $a_\epsilon = \epsilon^{-\frac{\alpha}{2+\alpha}}$, $(Y_t(x_0))_{t\geq 0}$ is the unique solution of the SDE

$$\begin{cases} \mathrm{d}Y_t = -C_0|Y_t|^{1+\alpha}\mathrm{sgn}(Y_t)\mathrm{d}t + \mathrm{d}W_t \quad \text{for} \quad t > 0, \\ Y_0 = \mathrm{sgn}(x_0)\infty, \end{cases} \tag{1}$$

and ν is the unique invariant probability measure for (1).

No cutoff phenomenon

Corollary (No cutoff phenomenon)

With the assumptions and notations of the previous theorem, for any $x_0 \in \mathbb{R}$, the family of processes $(X^\epsilon(x_0))_{\epsilon\in(0,1]}$ does not exhibit cutoff in the total variation distance as $\epsilon \to 0$. In other words, there is no time scale $(t_\epsilon)_{\epsilon\in(0,1]}$ with $t_\epsilon \to \infty$ as $\epsilon \to 0$ and

$$\lim_{\epsilon\to 0^+} \left\| X^\epsilon_{\delta t_\epsilon}(x) - \mu^\epsilon \right\|_{\mathbb{TV}} = \mathbb{1}_{(0,1)}(\delta) \quad \textit{for any} \quad \delta > 0,\ \delta \neq 1.$$

Proof of the previous corollary

Let $A_\epsilon \to \infty$ as $\epsilon \to 0^+$.
i) Assume that $A_\epsilon \le Ca_\epsilon$ for some positive constant C and for all $\epsilon \ll 1$. Then for all $t \ge 0$ we have

$$\left\|X^{\epsilon}_{Ca_\epsilon t}(x_0) - \mu^{\epsilon}\right\|_{\mathbb{TV}} \le \left\|X^{\epsilon}_{A_\epsilon t}(x_0) - \mu^{\epsilon}\right\|_{\mathbb{TV}}$$

and consequently the main theorem for all $t > 0$ yields

$$\begin{aligned} 0 < \|Y_{Ct}(x_0) - \nu\|_{\mathbb{TV}} &= \liminf_{\epsilon \to 0^+} \left\|X^{\epsilon}_{Ca_\epsilon t}(x_0) - \mu^{\epsilon}\right\|_{\mathbb{TV}} \\ &\le \liminf_{\epsilon \to 0^+} \left\|X^{\epsilon}_{A_\epsilon t}(x_0) - \mu^{\epsilon}\right\|_{\mathbb{TV}}. \end{aligned}$$

In particular, for $t > 1$ we have

$$\liminf_{\epsilon \to 0^+} \left\|X^{\epsilon}_{A_\epsilon t}(x_0) - \mu^{\epsilon}\right\|_{\mathbb{TV}} > 0.$$

ii) The case $a_\epsilon \le CA_\epsilon$ for all $\epsilon \ll 1$ is completely analogous using that the limiting profile $\|Y_t(x_0) - \nu\|_{\mathbb{TV}} < 1$ for all $t > 0$.

No cutoff phenomenon via scaling procedure

By the Chapman–Kolmogorov equation, for any $x_0 \in \mathbb{R}$ and $\epsilon \in (0,1]$ it follows that the map

$$t \mapsto d_t^\epsilon(x_0) := \|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} \qquad \text{is non-increasing.}$$

Lemma (Scaling argument)

Let $x_0 \in \mathbb{R}$ and assume that there is a sequence $(a_\epsilon(x_0))_{\epsilon \in (0,1]}$ for which the following conditions hold true:

i) $\lim_{\epsilon \to 0} a_\epsilon(x_0) = \infty$.

ii) *For any $t > 0$*

$$0 < \liminf_{\epsilon \to 0^+} d^\epsilon_{a_\epsilon(x_0)t}(x_0) \leq \limsup_{\epsilon \to 0^+} d^\epsilon_{a_\epsilon(x_0)t}(x_0) < 1.$$

Then there is no cutoff for the family $(X^\epsilon(x_0))_{\epsilon \in (0,1]}$ in total variation distance as ϵ tends to zero.

Mixing times asymptotics

Lemma (Scaling argument)

If in addition, the function

$$t \mapsto d_t^\epsilon(x_0) := \|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} \quad \textit{is continuous and striclty decreasing,}$$

and for all $t > 0$ the limit

$$\lim_{\epsilon \to 0} d^\epsilon_{a_\epsilon(x_0)t}(x_0) = G_{x_0}(t) \in (0,1)$$

then for any $\eta \in (0,1)$ it follows that

$$\lim_{\epsilon \to 0} \frac{\tau_{mix}^{\epsilon, x_0}(\eta)}{a_\epsilon(x_0)} = \inf\{t \geq 0 : G_{x_0}(t) \leq \eta\}.$$

Asymptotics

Corollary (Mixing time asymptotics)

Suppose that all assumptions and notation made in previous theorem hold true. For any $x_0 \in \mathbb{R}$ and $\eta \in (0,1)$, the η-mixing time

$$\tau_{mix}^{\epsilon,x_0}(\eta) := \inf\{t \geq 0 : \|X_t^{\epsilon}(x_0) - \mu^{\epsilon}\|_{\mathbb{TV}} \leq \eta\}$$

satisfies the limiting behavior

$$\lim_{\epsilon \to 0} \frac{\tau_{mix}^{\epsilon,x_0}(\eta)}{a_\epsilon} = \inf\{t \geq 0 : \|Y_t(\mathrm{sgn}(x_0)\infty) - \nu\|_{\mathbb{TV}} \leq \eta\}.$$

That is,

$$\tau_{mix}^{\epsilon,x_0}(\eta) \approx a_\epsilon \cdot \inf\{t \geq 0 : \|Y_t(\mathrm{sgn}(x_0)\infty) - \nu\|_{\mathbb{TV}} \leq \eta\} \quad \textit{for} \quad \epsilon \approx 0,$$

where $a_\epsilon = \epsilon^{-\frac{\alpha}{2+\alpha}}$.

Regularization: exponential convergence

By the growth condition we have that there exist $c > 0$, $R > 1$ and $\beta > -1$ such that

$$xV'(x) \geq c|x|^{2+\beta} \quad \text{for} \quad |x| \geq R.$$

Hence, one can prove that for any $\epsilon > 0$ and $\mathfrak{a} > 0$ there exists positive constants $C_1(\epsilon, \mathfrak{a})$ and $C_2(\epsilon, \mathfrak{a})$ such that for all $x_0 \in \mathbb{R}$ it follows

$$\|X_t^\epsilon(x_0) - \mu^\epsilon\|_{\mathbb{TV}} \leq C_1(\epsilon, \mathfrak{a}) \left(e^{\mathfrak{a}|x_0|} + \int_{\mathbb{R}} e^{\mathfrak{a}|z|} \mu^\epsilon(\mathrm{d}z) \right) e^{-C_2(\epsilon,\mathfrak{a})t}$$

for all $t \geq 0$. Note that $C_1(\epsilon, \mathfrak{a})$ and $C_2(\epsilon, \mathfrak{a})$ also depend on β.

Coming down from infinity: ODE

For simplicity, let us assume that $V(x) = x^4/4$ for all $x \in \mathbb{R}$.
In the sequel, we define the solution of

$$\dot{\varphi}_t = -(\varphi_t)^3 \quad \text{with} \quad \varphi_0 = \infty. \tag{2}$$

Consider the Cauchy problem

$$\dot{\varphi}_t^\ell = -(\varphi_t^\ell)^3 \quad \text{with} \quad \varphi_0^\ell = \ell > 0.$$

- Its explicit solution is given by $\varphi_t^\ell = \sqrt{\frac{1}{2t+\ell^{-2}}}$ for all $t \geq 0$.
- Hence, $\varphi_t^\infty = \sqrt{\frac{1}{2t}}$, $t \geq 0$ solves the differential equation (2).
- $\varphi_t^\infty \in \mathbb{R}$ for any $t > 0$.

Coming down from infinity: SDE

In the sequel, we present the main steps for proving that the following SDE is well-posed:

$$\mathrm{d}Y_t = -(Y_t)^3 \mathrm{d}t + \mathrm{d}W_t \quad \text{for} \quad Y_0 = \infty.$$

Let $(Y_t(\ell), t \geq 0)$ is the unique strong solution of the SDE

$$\mathrm{d}Y_t(\ell) = -(Y_t(\ell))^3 \mathrm{d}t + \mathrm{d}W_t \quad \text{with} \quad Y_0(\ell) = \ell.$$

- **Monotonicity:** Set $Y_t^* = \lim_{\ell \uparrow \infty} Y_t(\ell)$ for $t \geq 0$. Note that $Y_0^* = \infty$.
- **Tightness w.r.t. the i. c.:** $\mathbb{E}[(Y_t(\ell))^2] \leq \psi_t < \infty$ for $\ell, t > 0$ $\underbrace{\Longrightarrow}_{\text{Fatou's Lemma}}$ $Y_t^* \in \mathbb{R}$ for $t > 0 \Longrightarrow$ **Continuous extension**.
- **Markovianity:** $(\mathcal{P}_\ell(\cdot))_{\ell \in [-\infty, \infty]}$.

Scale analysis and limiting shape for a toy model

$V(x) = x^4/4, \quad a_\epsilon = \epsilon^{-1/2}, \quad b_\epsilon = \epsilon^{1/4}, \quad Y_t^\epsilon(x_0) := (1/b_\epsilon) X_{a_\epsilon t}^\epsilon(x_0).$

Goal: analyze

$$d_{a_\epsilon t}^\epsilon(x_0) := \left\| X_{a_\epsilon t}^\epsilon(x_0) - \mu^\epsilon \right\|_{\mathbb{TV}} = \left\| Y_t^\epsilon(x_0) - \mu^1 \right\|_{\mathbb{TV}}.$$

$$\begin{cases} \mathrm{d}Y_t^\epsilon = -(Y_t^\epsilon)^3 \mathrm{d}t + \mathrm{d}W_t, \\ Y_0^\epsilon = x_0/b_\epsilon, \end{cases} \qquad \begin{cases} \mathrm{d}Y_t = -Y_t^3 \mathrm{d}t + \mathrm{d}W_t, \\ Y_0 = \mathrm{sgn}(x_0)\infty. \end{cases}$$

$(Y_t^\epsilon(x_0))_{t\geq 0}$ converges, as $\epsilon \to 0^+$, to $(Y_t(x_0))_{t\geq 0}$.

Hence, for $t > 0$ we deduce

$$\lim_{\epsilon \to 0^+} d_{a_\epsilon t}^\epsilon(x_0) = \left\| Y_t(x_0) - \mu^1 \right\|_{\mathbb{TV}} \in (0, 1).$$

Scheme of the proof

$$\left\| X^{\epsilon}_{a_\epsilon t}(x_0) - \mu^{\epsilon} \right\|_{\mathbb{TV}}$$

$$\Big\downarrow \quad \epsilon \to 0^+$$

$\| Y_t(x_0) - \mu^1 \|_{\mathbb{TV}} \in (0,1) \longrightarrow$ No cutoff at scale $a_\epsilon \to \infty$

$$\Big\downarrow$$

Asymptotics for mixing and no cutoff for any scale A_ϵ

General degenerate potential: main ideas

Let $a_\epsilon > 0$ and $b_\epsilon > 0$ be scaling parameters to be properly fixed. Define

$$Y_t^\epsilon(x_0) := (1/b_\epsilon) X^\epsilon_{a_\epsilon t}(x_0) \quad \text{for} \quad t \geq 0.$$

By the Itô formula we see that

$$\begin{cases} \mathrm{d}Y_t^\epsilon(x_0) = -\frac{a_\epsilon}{b_\epsilon} V'(b_\epsilon Y_t^\epsilon(x_0))\mathrm{d}t + \frac{\sqrt{\epsilon a_\epsilon}}{b_\epsilon}\mathrm{d}W_t \quad \text{for} \quad t \geq 0, \\ Y_0^\epsilon(x_0) = x_0/b_\epsilon, \end{cases}$$

Define a_ϵ and b_ϵ as the unique solution to the system

$$\begin{cases} \frac{\sqrt{\epsilon a_\epsilon}}{b_\epsilon} = 1, \\ a_\epsilon b_\epsilon^\alpha = 1. \end{cases} \quad \text{i.e.} \quad a_\epsilon = \epsilon^{-\frac{\alpha}{2+\alpha}} \quad \text{and} \quad b_\epsilon = \epsilon^{\frac{1}{2+\alpha}}.$$

Hypothesis **(1) Local behavior at zero:** there exist constants $C_0 > 0$ and $\alpha > 0$ such that for any $K > 0$ we have

$$\lim_{\lambda \to 0^+} \sup_{|x| \leq K} \left| \frac{V'(\lambda x)}{\lambda^{1+\alpha}} - C_0 |x|^{1+\alpha} \mathrm{sgn}(x) \right| = 0,$$

where $\mathrm{sgn}(x) := x/|x|$ for $x \neq 0$ and $\mathrm{sgn}(0) := 0$.

General degenerate potential: main ideas

Since the total variation distance is invariant by scaling, we deduce

$$d^{\epsilon}_{a_\epsilon t}(x_0) = \left\|X^{\epsilon}_{a_\epsilon t}(x_0) - \mu^{\epsilon}\right\|_{\mathbb{TV}} = \|Y^{\epsilon}_t(x_0) - b_\epsilon \mu^{\epsilon}(b_\epsilon \mathrm{d}z)\|_{\mathbb{TV}}.$$

By the triangle inequality we have

$$\Big|\underbrace{d^{\epsilon}_{a_\epsilon t}(x_0)}_{\text{Objective}} - \underbrace{\|Y_t(x_0) - \nu\|_{\mathbb{TV}}}_{\text{Limiting profile}}\Big| \leq \underbrace{\|Y^{\epsilon}_t(x_0) - Y_t(x_0)\|_{\mathbb{TV}}}_{\text{Coupling SDEs}} + \underbrace{\|\nu - b_\epsilon \mu^{\epsilon}(b_\epsilon \mathrm{d}z)\|_{\mathbb{TV}}}_{\text{Laplace method}},$$

where $(Y_t(x_0))_{t\geq 0}$ is the solution of the following SDE

$$\begin{cases} \mathrm{d}Y_t(x_0) = -C_0|Y_t(x_0)|^{1+\alpha}\mathrm{sgn}(Y_t)\mathrm{d}t + \mathrm{d}W_t \quad \text{for} \quad t \geq 0, \\ Y_0(x_0) = \mathrm{sgn}(x_0)\infty, \end{cases}$$

where C_0 and α are the positive constants that appears in local assumption, and ν represents the unique invariant probability measure for the preceding random dynamics.

Coupling around the origin

Hypothesis **(2) Growth condition at infinity:** there exist $c_0, R_0 \in (0,\infty)$, and $\beta \in (-1,\infty)$ such that

$$V'(x) \geq c_0 x^{1+\beta} \quad \text{for all} \quad x \geq R_0. \qquad \textbf{(G)}$$

Observe that $\| Y_t^{\epsilon}(x_0) - Y_t(x_0)\|_{\mathbb{TV}}$ is a complicated term. However,

$$\| Y_t^{\epsilon}(x_0) - Y_t(x_0)\|_{\mathbb{TV}} \leq \underbrace{\left\| Y_t^{\epsilon}(x_0) - \widetilde{Y}_t^{\epsilon}(x_0)\right\|_{\mathbb{TV}}}_{\text{Synchronous coupling and maximal inequalities}} + \underbrace{\left\| \widetilde{Y}_t^{\epsilon}(x_0) - Y_t(x_0)\right\|_{\mathbb{TV}}}_{\text{Girsanov coupling or Kabanov's coupling}}.$$

Gaussian-setting: Cameron–Martin–Girsanov's Theorem, Fokker–Planck estimates, Kabanov, Y. et alt. estimates. For instance, it is known that

$$\|X_t(z) - Y_t(z)\|^2_{\mathbb{TV}} \leq 16 \int_0^t \mathbb{E}[|F(X_s(z)) - G(Y_s(z))|^2] \mathrm{d}s,$$

where F and G are the fields for $(X_t(z))_{t\geq 0}$ and $(Y_t(z))_{t\geq 0}$, respectively.

Main references

- G. Barrera, C. da Costa & M. Jara. *Sharp convergence for degenerate Langevin dynamics*. ArXiv:2209.11026
- G. Barrera & M. Jara: **Abrupt convergence of stochastic small perturbations of one dimensional dynamical systems**. Journal of Statistical Physics 2016.
- G. Barrera & M. Jara: **Thermalisation for small random perturbations of dynamical systems**. Annals of Applied Probability 2020.
- G. Barrera & J. C. Pardo. **Cutoff phenomenon for the Ornstein-Uhlenbeck processes driven by Lévy processes**. Electronic Journal of Probability 2020.
- G. Barrera, J. C. Pardo & M. Högele: **The cutoff phenomenon in total variation for nonlinear Langevin systems with small layered stable noise**. Electronic Journal of Probability 2021.
- G. Barrera, J. C. Pardo & M. Högele: **Cutoff thermalization for Ornstein-Uhlenbeck systems with small Lévy noise in the Wasserstein distance**. Journal of Statistical Physics 2021.
- G. Barrera, J. C. Pardo & M. Högele: **The cutoff phenomenon in Wasserstein distance for nonlinear stable Langevin systems with small Lévy noise**. Journal of Dynamics and Differential Equations 2022.