Intertwinings for Markov processes

Aldéric Joulin - University of Toulouse

Joint work with :
Michel Bonnefont - Univ. Bordeaux

Workshop 2
"Piecewise Deterministic Markov Processes"
Rennes - May 15-17, 2013

Real-valued Ornstein-Uhlenbeck process, solution to SDE

$$dX_t = \sqrt{2} dB_t - X_t dt.$$

Instantaneous distribution is known: semigroup $(P_t)_{t \geq 0}$ given by

$$\begin{aligned} P_t f(x) &= \int_{\mathbb{R}} f(e^{-t} x + \sqrt{1 - e^{-2t}} y) \, \mu(dy) \\ &\underset{t \to \infty}{\longrightarrow} \int_{\mathbb{R}} f \, d\mu =: \mu(f), \end{aligned}$$

where invariant (and reversible) measure μ is $\mathcal{N}(0,1)$.

Generator:

$$\mathcal{L} f(x) = f''(x) - x \, f'(x).$$

Commutation relation between semigroup and gradient:

$$(P_t f)' = e^{-t} P_t(f'),$$

related to long-time behaviour:

- Convergence in Wasserstein distance.
- Convergence in $L^2(\mu)$:

$$\|P_t f - \mu(f)\|_{L^2(\mu)} \leq e^{-\lambda_1 t} \|f - \mu(f)\|_{L^2(\mu)},$$

via Poincaré (or spectral gap) inequality: $\lambda_1 = 1$, where λ_1 best constant in inequality:

$$\lambda_1 \operatorname{Var}_\mu(f) := \lambda_1 \left(\mu(f^2) - \mu(f)^2\right) \leq -\int_{\mathbb{R}} f \mathcal{L} f \, d\mu = \int_{\mathbb{R}} |f'|^2 \, d\mu,$$

i.e. given by variational formula

$$\lambda_1 = \inf_f \frac{-\int_{\mathbb{R}} f \mathcal{L} f \, d\mu}{\operatorname{Var}_\mu(f)}.$$

- Exponential decay in entropy:

$$\mathrm{Ent}_\mu(P_t f) \quad \leq \quad e^{-2t}\mathrm{Ent}_\mu(f),$$

via log-Sobolev inequality:

$$\mathrm{Ent}_\mu(f) := \mu(f \log f) - \mu(f) \log \mu(f) \quad \leq \quad -2 \int_{\mathbb{R}} \sqrt{f} \mathcal{L} \sqrt{f} \, d\mu.$$

Other consequences:

- Hypercontractivity.
- Measure concentration.
- Etc...

Generalization to process solution to SDE

$$dX_t = \sqrt{2} dB_t - U'(X_t) dt,$$

where U smooth potential.

Invariant (and reversible) measure:

$$\mu(dx) \propto e^{-U(x)} \, dx.$$

Generator:

$$\mathcal{L}f = f'' - U'f'.$$

Bakry-Émery's criterion through Γ_2-calculus: if

$$U'' \geq \rho,$$

then (sub-) commutation relation between gradient and semigroup

$$|(P_t f)'| \leq e^{-\rho t} P_t \left(|f'| \right).$$

In particular if $\rho > 0$, then good ergodicity properties:

$$\lambda_1 \geq \rho,$$

i.e. Poincaré inequality

$$\rho \operatorname{Var}_\mu(f) \leq -\int_{\mathbb{R}} f \mathcal{L} f d\mu = \int_{\mathbb{R}} |f'|^2 \, d\mu,$$

meaning long-time convergence in $L^2(\mu)$:

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

Also log-Sobolev, hypercontractivity, measure concentration, etc...

How to obtain Poincaré from Bakry-Émery ?

$$
\begin{aligned}
\mathrm{Var}_\mu(f) &= -2\int_0^{+\infty}\int_{\mathbb{R}} P_t f\,\mathcal{L}P_t f\,d\mu\,dt \\
&\overset{IBP}{=} 2\int_0^{+\infty}\int_{\mathbb{R}} |(P_t f)'|^2\,d\mu\,dt \\
&\overset{Bakry-Emery}{\leq} 2\int_0^{+\infty} e^{-2\rho t}\int_{\mathbb{R}} P_t(|f'|^2)\,d\mu\,dt \\
&\overset{Invariance}{=} 2\int_0^{+\infty} e^{-2\rho t}\,dt\int_{\mathbb{R}} |f'|^2\,d\mu \\
&\overset{\rho>0}{=} \frac{1}{\rho}\int_{\mathbb{R}} |f'|^2\,d\mu.
\end{aligned}
$$

Hence $\mathrm{Sign}(\rho)$ is important to obtain long-time convergence.

Intertwining between gradient and generators:

$$(\mathcal{L}f)' = \mathcal{L}^V(f'),$$

with $\mathcal{L}^V$ Schrödinger operator $\mathcal{L} - V$ with potential $V := U''$.

At level of semigroups:

$$(P_t f)'(x) = P_t^V(f')(x) := \mathbb{E}_x \left[f'(X_t) \exp\left(- \int_0^t U''(X_s)\, ds \right) \right],$$

where $(P_t^V)_{t\geq 0}$ associated Feynman-Kac semigroup.

Hence Jensen's inequality entails Bakry-Émery's criterion.

General diffusion satisfying the SDE

$$dX_t = \sqrt{2}\sigma(X_t)dB_t + b(X_t)dt.$$

Invariant (and reversible) measure:

$$\mu(dx) \propto \frac{1}{\sigma(x)^2} \exp\left(\int_{x_0}^{x} \frac{b(u)}{\sigma(u)^2}\, du\right) dx.$$

Generator of Sturm-Liouville type:

$$\mathcal{L}f = \sigma^2 f'' + bf'.$$

Bakry-Émery's criterion through Γ_2-calculus: if

$$V_\sigma := \frac{\mathcal{L}\sigma}{\sigma} - b' \geq \rho_\sigma,$$

then (sub-) commutation relation holds:

$$|\sigma(P_t f)'| \leq e^{-\rho_\sigma t} P_t\left(|\sigma f'|\right).$$

Probabilistic interpretation of V_σ ? Using method of tangent process, i.e. differentiate w.r.t. initial point x:

$$\partial_x X_t^x = 1 + \sqrt{2} \int_0^t \sigma'(X_s^x)\, \partial_x X_s^x \, dB_s + \int_0^t b'(X_s^x)\, \partial_x X_s^x \, ds,$$

hence by Itô's formula, process $\partial_x X^x$ given by

$$\begin{aligned} \partial_x X_t^x &= \frac{\sigma(X_t^x)}{\sigma(x)} \exp\left(- \int_0^t \left(\frac{\mathcal{L}\sigma}{\sigma} - b' \right) (X_s^x)\, ds \right) \\ &= \frac{\sigma(X_t^x)}{\sigma(x)} \exp\left(- \int_0^t V_\sigma(X_s^x)\, ds \right). \end{aligned}$$

Hence

$$\begin{aligned}(P_t f)'(x) &= \mathbb{E}\left[f'(X_t^x)\,\partial_x X_t^x\right] \\ &= \frac{1}{\sigma(x)}\,\mathbb{E}\left[\sigma(X_t^x)\,f'(X_t^x)\ \exp\left(-\int_0^t V_\sigma(X_s^x)\,ds\right)\right].\end{aligned}$$

Denoting the weighted gradient $\nabla_\sigma f := \sigma f'$, it rewrites as the intertwining

$$\begin{aligned}\nabla_\sigma P_t f(x) &= \mathbb{E}\left[\nabla_\sigma f(X_t^x)\,\exp\left(-\int_0^t V_\sigma(X_s^x)\,ds\right)\right] \\ &=: P_t^{V_\sigma}(\nabla_\sigma f)(x).\end{aligned}$$

Can be recovered through straightforward infinitesimal version:

$$\nabla_\sigma \mathcal{L} = \mathcal{L}^{V_\sigma} \nabla_\sigma .$$

If

$$\rho_\sigma := \inf V_\sigma > 0,$$

then Poincaré inequality:

$$\rho_\sigma \, \mathrm{Var}_\mu(f) \leq - \int_{\mathbb{R}} f \mathcal{L} f \, d\mu = \int_{\mathbb{R}} |\nabla_\sigma f|^2 \, d\mu .$$

But if $\rho_\sigma \leq 0$, how to use intertwining to obtain Poincaré, thus L^2-convergence ?

Idea: change the gradient.

Denote the weighted gradient $\nabla_a f := af'$ for positive function a. We have the intertwining

$$\nabla_a \mathcal{L} f = \mathcal{L}_a(\nabla_a f) - V_a \nabla_a f =: \mathcal{L}_a^{V_a}(\nabla_a f),$$

where new Sturm-Liouville generator is

$$\mathcal{L}_a f \; = \; \sigma^2 f'' + b_a f',$$

where

$$b_a \; := \; b + 2\sigma\sigma' - 2\sigma^2 \, \frac{a'}{a} \; = \; b + 2\sigma^2 \, \frac{h'}{h}, \quad \text{with} \quad h := \frac{\sigma}{a},$$

and potential

$$V_a := \frac{\mathcal{L}_a(a)}{a} - b'.$$

At level of semigroups,

$$\nabla_a P_t f(x) = P_{a,t}^{V_a}(\nabla_a f)(x) := \mathbb{E}_x \left[\nabla_a f(X_t^a) \exp\left(-\int_0^t V_a(X_s^a)\, ds \right) \right],$$

with $(P_{a,t})_{t\geq 0}$ semigroup of process X^a associated to $\mathcal{L}_a$.

If $a = \sigma$ then $h' = 0$ thus $\mathcal{L}_a = \mathcal{L}$ and

$$V_a \;=\; \frac{\mathcal{L}_a(a)}{a} - b' \;=\; \frac{\mathcal{L}\sigma}{\sigma} - b' \;=\; V_\sigma.$$

In particular if

$$\rho_a := \inf V_a > 0,$$

then convergence in (distorting) Wasserstein distance (exponentially fast, depending on function a) and also L^2-convergence via Poincaré:

$$
\begin{aligned}
\mathrm{Var}_\mu(f) &= 2\int_0^{+\infty}\int_{\mathbb{R}} |\nabla_\sigma P_t f|^2 \, d\mu \, dt \\
&= 2\int_0^{+\infty}\int_{\mathbb{R}} |\nabla_a P_t f|^2 \left(\frac{\sigma}{a}\right)^2 d\mu \, dt \\
&\overset{\textit{Intertwining}}{=} 2\int_0^{+\infty}\int_{\mathbb{R}} |P_{a,t}^{V_a}(\nabla_a f)|^2 \left(\frac{\sigma}{a}\right)^2 d\mu \, dt \\
&\overset{\textit{Jensen}}{\leq} 2\int_0^{+\infty} e^{-2\rho_a t}\int_{\mathbb{R}} P_{a,t}(|\nabla_a f|^2) \left(\frac{\sigma}{a}\right)^2 d\mu \, dt \\
&\overset{\textit{Invariance}}{=} 2\int_0^{+\infty} e^{-2\rho_a t}\, dt \int_{\mathbb{R}} |\nabla_a f|^2 \left(\frac{\sigma}{a}\right)^2 d\mu \\
&\overset{\rho_a > 0}{=} \frac{1}{\rho_a}\int_{\mathbb{R}} |\nabla_\sigma f|^2 \, d\mu,
\end{aligned}
$$

since $(\sigma/a)^2 d\mu$ is invariant for $(P_{a,t})_{t\geq 0}$.

Hence

$$\lambda_1 \geq \sup_a \rho_a.$$

Recovers Chen-Wang's Theorem on spectral gap, obtained originally via coupling.

Is our criterion optimal ? Yes.

Take $a = 1/g'$ where $g' > 0$. Then

$$V_a = -\frac{(\mathcal{L}g)'}{g'}.$$

Hence if g eigenvector associated to λ_1,

$$V_a \equiv \lambda_1.$$

Another consequence of intertwining: if

$$V_\sigma := \frac{\mathcal{L}\sigma}{\sigma} - b' > 0,$$

then

$$\lambda_1 \geq \frac{1}{\int_{\mathbb{R}} \frac{1}{V_\sigma} \, d\mu}.$$

Convenient when Bakry-Émery criterion does not apply (for instance $V_\sigma > 0$ but tends to 0 at infinity).

Some classical examples:

Let $U(x) = \frac{|x|^\alpha}{\alpha}$. Recall

$$dX_t = \sqrt{2} dB_t - U'(X_t)\, dt, \qquad \mu(dx) \propto e^{-U(x)}\, dx.$$

- $\alpha = 2$ (O.U.):

  with $\quad a = \sigma \equiv 1, \qquad \lambda_1 = 1 \quad$ (Bakry-Émery).

- $\alpha = 1$:

  with $\quad a(x) = e^{-|x|/2}, \qquad \lambda_1 = 1/4 \quad$ (Bobkov-Ledoux).

- $\alpha = 4$:

$$\text{with} \qquad a(x) = e^{-\frac{U(x)}{2} + \frac{\varepsilon x^2}{2}}, \qquad \lambda_1 \geq \varepsilon_{\max} = \sqrt{\frac{3}{2}}.$$

- $1 < \alpha < 2$:

$$V_\sigma(x) \;=\; (\alpha - 1)\,|x|^{\alpha-2} \underset{x\to\infty}{\longrightarrow} 0.$$

Hence

$$\begin{aligned} \lambda_1 \;&\geq\; \frac{1}{\int_{\mathbb{R}} \frac{1}{V_\sigma}\, d\mu} \\ &=\; (\alpha - 1)\, \alpha^{1-2/\alpha} \frac{\Gamma(1/\alpha)}{\Gamma((3-\alpha)/\alpha)}. \end{aligned}$$

Consider new process

$$dX_t = \sqrt{2}\,\sigma(X_t)\,dB_t + \left(2\sigma\sigma' - \sigma^2 U'\right)(X_t)\,dt,$$

also reversible w.r.t.

$$\mu(dx) \propto e^{-U(x)}\,dx.$$

Generator:

$$\mathcal{L}f = \sigma^2 f'' + \left(2\sigma\sigma' - \sigma^2 U'\right)\,f'.$$

Choice of diffusion constant σ ?

Heavy-tailed case: weighted Poincaré.

Take

$$U(x) := \beta \log(1 + x^2), \quad \text{with} \quad \beta > 1/2,$$

leading to Cauchy-type distribution

$$\mu(dx) \propto \frac{dx}{(1 + x^2)^\beta}.$$

Then with choice $\sigma(x) := \sqrt{1 + x^2}$,

$$V_\sigma(x) = \frac{2\beta - 1}{1 + x^2} \underset{x \to \infty}{\longrightarrow} 0.$$

Hence for $\beta > 3/2$,

$$\begin{aligned}\lambda_1 &\geq \frac{1}{\int_{\mathbb{R}} \frac{1}{V_\sigma}\, d\mu} \\ &= \frac{(2\beta-1)(\beta-3/2)}{\beta-1} \\ &=: C_\beta,\end{aligned}$$

meaning that weighted Poincaré holds:

$$\begin{aligned}C_\beta \operatorname{Var}_\mu(f) &\leq \int_{\mathbb{R}} |\nabla_\sigma f|^2 \, d\mu \\ &= \int_{\mathbb{R}} (1+x^2)\, |f'(x)|^2 \, \mu(dx).\end{aligned}$$

Coming back to general case, why process X^a and potential

$$V_a := \frac{\mathcal{L}_a(a)}{a} - b',$$

appear in the intertwining relation ?

Answer: (Doob's) h-transform.

$(P_t^V)_{t\geq 0}$ original Feynman-Kac semigroup with some potential V.
New dynamics: "multiply inside and divide outside by some function h".

$$P_t^{V^{(h)}} f := \frac{P_t^V(hf)}{h}.$$

At level of generator (of Schrödinger type):

$$\begin{aligned}
\mathcal{L}^{V^{(h)}} f &= \frac{\mathcal{L}^V (hf)}{h} \\
&= \mathcal{L} f + 2\,\frac{\Gamma(f,h)}{h}\, f' + \left(\frac{\mathcal{L} h}{h} - V \right) f \\
&= \underbrace{\sigma^2 f'' + \left(b + 2\sigma^2 \frac{h'}{h} \right) f'}_{generator\ of\ h-transform} + \underbrace{\left(\frac{\mathcal{L} h}{h} - V \right)}_{h-potential} f .
\end{aligned}$$

In case $V = 0$ then Doob's h-transform is Markov if and only if h is $\mathcal{L}$-harmonic, i.e. $\mathcal{L}h = 0$.

Group structure:

- hk-transform is the h-transform of the k-transform.

- h-transform and original dynamics have the same distribution if and only if h is constant.

If

$$V = V_\sigma := \frac{\mathcal{L}\sigma}{\sigma} - b' \quad \text{and} \quad h := \frac{\sigma}{a},$$

then X^a is Doob's h-transform of X, and

$$\begin{aligned}
\mathcal{L}^{V(\sigma/a)} f &= \mathcal{L}_a f + \left(\frac{\mathcal{L}h}{h} - V_\sigma \right) f \\
&= \mathcal{L}_a f + \left(\underbrace{\frac{\mathcal{L}(\sigma/a)}{\sigma/a}}_{\text{by } h-\text{transform}} - \underbrace{\frac{\mathcal{L}\sigma}{\sigma} - b'}_{\text{by classical intertwining}} \right) f \\
&= \mathcal{L}_a f - \left(\frac{\mathcal{L}_a(a)}{a} - b' \right) f \\
&= \mathcal{L}_a^{V_a} f.
\end{aligned}$$

Probabilistic interpretation of intertwining:

• first perform the classical intertwining by using method of tangent process:

$$\nabla_\sigma P_t f = P_t^{V_\sigma}(\nabla_\sigma f), \quad \text{with potential} \quad V_\sigma := \frac{\mathcal{L}\sigma}{\sigma} - b';$$

• then use h-transform with $h := \sigma/a$ to obtain the final intertwining

$$\nabla_a P_t f = P_{a,t}^{V_a}(\nabla_a f), \quad \text{with potential} \quad V_a := \frac{\mathcal{L}_a(a)}{a} - b'.$$