Convergence to equilibrium of Markov processes (eventually piecewise deterministic)

A. Guillin

Université Blaise Pascal and IUF

Rennes

joint works with D. Bakry, F. Barthe, F. Bolley, P. Cattiaux, R. Douc, G. Fort, I. Gentil, F. Malrieu, P-A. Zitt

Our main goal:

$\hookrightarrow$ study the speed of convergence to equilibrium of Markov Process.

- Generator L
- Semigroup P_t ergodic
- Invariant measure μ

Carré du champ $\Gamma(f,g) = \frac{1}{2}(L(fg) - fLg - gLf)$.
Energy $\mathcal{E}(f) = \int -fLf d\mu = \int \Gamma(f) d\mu$

Examples :

• Continuous time Markov chain: $L = P - I$ where P is the transition matrix.

• Reversible diffusion:

$$L = \Delta - \nabla V.\nabla, \qquad \Gamma(f) = |\nabla f|^2$$

• Kinetic Fokker Planck:

$$Lf(x,v) = v\nabla_x f + \Delta_v - \nabla V(x).\nabla_v f - v.\nabla_v f, \qquad \Gamma(f) = |\nabla_v f|^2$$

• Hypocoercive walk:

$$Lf(x,y) = y\partial_x f(x,y) + a(f(x,-y) - f(x,y))$$

Examples (continued) :

- TCP process: with $0 < r < 1$, $l(x) = c$ or x

$$Lf(x) = f'(x) + l(x)(f(rx) - f(x))$$

- Ergodic Telegraph process:

$$Lf(y,w) = w\partial_y f(y,w) + (a + (b-a)1_{\{yw>0\}})(f(y,-w - -f(y,w))$$

- Random switching:

$$Lf(x,i) = L^i f(x,i) + \sum_j a(i,j)(f(x,j) - f(x,i))$$

What is the speed of convergence of P_t to μ? in which distance?

Various techniques:

- Coupling

  ↪ TV distance, weighted TV distance, Wasserstein distance

  ↪Lyapunov's conditions, control hitting times.

  ↪Ad-Hoc Coupling

- Functional inequalities

  ↪ L^2 distance, entropy, Wasserstein distance

  ↪Poincaré, log-Sobolev inequalities, WJ, curvature...

Meyn-Tweedie's approach : Lyapunov condition and coupling

Lyapunov condition: find $W \geq 1$, a nice set C, $b > 0$ and a positive φ such that

$$LW \leq -\varphi \times W + b\,1_C.$$

This condition expresses "the strength" which pushes the process to a nice region of the space.

Remark: one may consider this condition as the supermartingale statement

$$\mathbb{E}_x(W(X_t)) + \mathbb{E}_x\left[\int_0^t \varphi(X_s)W(X_s)ds\right] \leq W(x) + b\mathbb{E}_x\left[\int_0^t 1_C(X_s)ds\right]$$

Examples

– Ornstein-Uhlenbeck : $L = \Delta - x.\nabla$.

$$W(x) = 1 + |x|^2, \qquad LW = 2n - 2|x|^2 \le -W(x) + 2(n-1)1_{\{|x|^2 \le 2n\}}$$

but with another choice

$$W(x) = e^{a|x|^2}, \qquad LW = \left(2an + 4a\left(a - \frac{1}{2}\right)|x|^2\right)W(x) \le -\lambda|x|^2 W(x) + b1_{\{|x| \le R\}}$$

Examples

– Ornstein-Uhlenbeck : $L = \Delta - x.\nabla$.

$$W(x) = 1 + |x|^2, \qquad LW = 2n - 2|x|^2 \le -W(x) + 2(n-1)1_{\{|x|^2 \le 2n\}}$$

but with another choice

$$W(x) = e^{a|x|^2}, \qquad LW = \left(2an + 4a\left(a - \frac{1}{2}\right)|x|^2\right)W(x) \le -\lambda|x|^2 W(x) + b1_{\{|x| \le R\}}$$

Examples

– Ornstein-Uhlenbeck : $L = \Delta - x.\nabla$.

$$W(x) = 1 + |x|^2, \qquad LW = 2n - 2|x|^2 \leq -W(x) + 2(n-1)1_{\{|x|^2 \leq 2n\}}$$

but with another choice

$$W(x) = e^{a|x|^2}, \qquad LW = \left(2an + 4a\left(a - \frac{1}{2}\right)|x|^2\right)W(x) \leq -\lambda|x|^2 W(x) + b1_{\{|x| \leq R\}}$$

Examples

– Exponential type process: $L = \Delta - \frac{x}{|x|}.\nabla$.
choose $a < 1$

$$W(w) = e^{a|x|}, \qquad LW \leq -c\, W(x) + b1_{\{|x|\leq R\}}$$

– Cauchy type process: $L = \Delta - (n+\alpha)\frac{\nabla V}{V}.\nabla$ et V convexe.

choose $2 < k < \alpha(1-\varepsilon) + n\varepsilon + 2$ and small enough ε

$$W(x) = 1 + |x|^k, \qquad LW \leq -c\,(W(x))^{\frac{k-2}{k}} + b1_{\{|x|\leq R\}}$$

Coupling by small set

It is a probabilist approach: Lyapunov condition expresses that the process is pushed to nice regions

- how to express this strength? $\rightarrow$ control of hitting times!
- Nice regions? $\rightarrow$ small set

Coupling by small set

It is a probabilist approach: Lyapunov condition expresses that the process is pushed to nice regions

- how to express this strength? $\rightarrow$ control of hitting times!
- Nice regions? $\rightarrow$ small set

Coupling by small set

It is a probabilist approach: Lyapunov condition expresses that the process is pushed to nice regions

- how to express this strength? $\rightarrow$ control of hitting times!
- Nice regions? $\rightarrow$ small set

Mathematically:

We have to build two processes X_t and Y_t such that:

- X_t and Y_t has the same law but starting from x and y
- after some random time T, the two processes sticks together

Indeed,

$$\begin{aligned} \|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} &\leq \inf_{couplage} \mathbb{E}(1_{X_t \neq Y_t}) \\ &\leq \mathbb{P}(T > t) \end{aligned}$$

For the coupling construction we will use a minorization condition and to control T the Lyapunov condition.

Coupling construction

Minorization condition:
Suppose that for some set C and $t^* > 0$, there exists $\varepsilon > 0$ such thate

$$(Minor) \qquad \forall x \in C, \qquad P_{t^*}(x,\cdot) \geq \varepsilon\nu(\cdot)$$

It enables to write a mixture of probability for P_t: if $x \in C$

$$P_{t^*}(x,\cdot) = \varepsilon\,\nu(\cdot) + (1-\varepsilon)\,\frac{P_{t^*}(x,\cdot) - \varepsilon\,\nu(\cdot)}{1-\varepsilon}$$

Intuition : if my process is in C there is a probability ε to be distributed at time t^* as ν independently of my position in C!

Definition of (X_t, Y_t)

1. $X_0 = x,\ Y_0 = y$.
2. If $t_0 = \inf\{t;\ (X_t, Y_t) \in C \times C\}$ and $t_n = \inf\{t \geq t_{n-1} + t^*;\ (X_t, Y_t) \in C \times C\}$. If we have not coupled at t_i then:
   - with probability ε, $X_{t_i+t^*} = Y_{t_i+t^*} = Z$ with $Z \sim \nu$ and we have coupled, then set $T = t_i + t^*$!.
   - with probability $1 - \varepsilon$, generate $X_{t_i+t^*}$ et $Y_{t_i+t^*}$ independently with the residual kernel

We need to control the return times in $C \times C$.

Remark that before coupling, the two processes are generated by $L \otimes L$.

Denote $\tau_C(t^*) = \inf\{t \geq t^*;\, X_t \in C\}$. Suppose the Lyapunov conditions

$$LW \leq -\delta \times W + b1_C$$

then

$$\forall x \notin C,\, \mathbb{E}_x(e^{\delta \tau_C(0)}) \leq W(x)$$

or the weaker: with strictly concave φ, $b > 0$ such that

$$LW \leq -\varphi(W) + b1_C$$

$$\forall x \notin C,\, \mathbb{E}_x\left(H_\varphi^{-1}(\tau_C(t^*)) \leq W(x) + c_{b,\varphi,\theta^*}\right.$$

with $H_\varphi(u) = \int_1^u \frac{1}{\varphi} ds$.

Denote $\tau_C(t^*) = \inf\{t \geq t^*;\ X_t \in C\}$. Suppose the Lyapunov conditions

$$LW \leq -\delta \times W + b1_C$$

then

$$\forall x \notin C,\ \mathbb{E}_x(e^{\delta \tau_C(0)}) \leq W(x)$$

or the weaker: with strictly concave φ, $b > 0$ such that

$$LW \leq -\varphi(W) + b1_C$$

$$\forall x \notin C,\ \mathbb{E}_x\left(H_\varphi^{-1}(\tau_C(t^*)\right) \leq W(x) + c_{b,\varphi,\theta^*}$$

with $H_\varphi(u) = \int_1^u \frac{1}{\varphi} ds$.

Denote $\tau_C(t^*) = \inf\{t \geq t^*;\, X_t \in C\}$. Suppose the Lyapunov conditions

$$LW \leq -\delta \times W + b1_C$$

then

$$\forall x \notin C,\, \mathbb{E}_x(e^{\delta \tau_C(0)}) \leq W(x)$$

or the weaker: with strictly concave φ, $b > 0$ such that

$$LW \leq -\varphi(W) + b1_C$$

$$\forall x \notin C,\, \mathbb{E}_x\left(H_\varphi^{-1}(\tau_C(t^*)\right) \leq W(x) + c_{b,\varphi,\theta^*}$$

with $H_\varphi(u) = \int_1^u \frac{1}{\varphi} ds$.

Denote $\tau_C(t^*) = \inf\{t \geq t^*;\, X_t \in C\}$. Suppose the Lyapunov conditions

$$LW \leq -\delta \times W + b1_C$$

then

$$\forall x \notin C,\, \mathbb{E}_x(e^{\delta \tau_C(0)}) \leq W(x)$$

or the weaker: with strictly concave φ, $b > 0$ such that

$$LW \leq -\varphi(W) + b1_C$$

$$\forall x \notin C,\, \mathbb{E}_x\left(H_\varphi^{-1}(\tau_C(t^*)\right) \leq W(x) + c_{b,\varphi,\theta^*}$$

with $H_\varphi(u) = \int_1^u \frac{1}{\varphi} ds$.

Proof

Let apply Itô's formula $e^t W(x)$:

$$\begin{aligned} \mathbb{E}_x\left(e^{at\wedge\tau_C}\right) &\leq \mathbb{E}_x\left(e^{at\wedge\tau_C} W(X_{t\wedge\tau_C}))\right) \\ &\leq W(x) + \mathbb{E}_x\left(\int_0^{t\wedge\tau_C} (aW + LW)(X_s) e^{as} ds\right) \\ &\leq W(x) + \mathbb{E}_x\left(\int_0^{t\wedge\tau_C} (a-\delta) W(X_s) e^{as} ds\right) \\ &\leq W(x) \end{aligned}$$

if $a \leq \delta$.

The subexponential case is a little bit more complicated.

With the Lyapunov condition

$$LW \leq -\delta \times W + b1_C$$

and the minorization condition with $C = \{W \leq K\}$ then for $\rho = \rho(\varepsilon, C, t^*, W) < 1\ K > 1$,

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} \leq K\rho^t\,(W(x) + W(y))$$

and if for φ strictly concave, $b > 0$such that

$$LW \leq -\varphi(W) + b1_C$$

then for $K = K(\varepsilon, C, \varphi, W)$

$$\|P_t(x,\cdot) - P_t(y,\cdot)\|_{TV} \leq \frac{K}{H_\varphi^{-1}(t - t^*)}\,(W(x) + W(y))$$

Advantages

- generic method (non reversible, discrete, discrete/continuous)
- sharp qualitative results

Disadvantages

- poor quantitative results because of the minorization condition
- can be difficult to find the Lyapunov function W (oscillators,...).

We thus often have to consider ad-hoc coupling deduced from the dynamics or other distances: for example

$$W_p^p(\nu,\mu) := \inf\left\{\mathbb{E}(d(X,Y)^p\,;\, X\sim\nu,\ Y\sim\mu\right\}$$

- 2-Wasserstein distance and synchronous coupling (convexity)
- 1-Wasserstein distance and reflection coupling (convexity at infinity)

Advantages: work for hypoelliptic diffusions, non linear process such as McKean-Vlasov models.

For more examples see the talks of Bardet, Cloez, Guerin...

Functional inequalities

Let us recall some well known facts:

- L^2 exponential decay $\iff$ Poincaré inequality

$$C\, Var_\mu(f) \le \int \Gamma(f) d\mu$$

Indeed

$$\begin{aligned}\frac{d}{dt} Var_\mu(P_t f) &= 2\int P_t f L P_t f d\mu \\ &= -2\mathcal{E}(P_t f) \\ &\le -2C\, Var_\mu(P_t f)\end{aligned}$$

and Gronwall's lemma gives $Var_\mu(P_t f) \le e^{-2Ct}\, Var_\mu(P_t f)$.

- Entropic decay $\iff$ log-Sobolev inequality (Diffusion)

$$C\, Ent_\mu(f^2) \le 2\int \Gamma(f) d\mu$$

How to get a Poincaré or log-Sobolev inequality ?

One well known criterion: Γ_2 criterion.

$$BE(K,\infty): \qquad \Gamma_2(f) \geq K\,\Gamma(f)$$

where $\Gamma_2(f,g) := \frac{1}{2}(L\Gamma(f,g) - \gamma(f,Lg) - \Gamma(g,Lf))$.

Theorem : $BE(K,\infty) \implies$ log-Sobolev and Poincaré inequality with constant K.

Example: $L = \Delta - \nabla V.\nabla$, $\Gamma_2(f) = \|\mathrm{Hess} f\|_2^2 + \nabla f^t\, \mathrm{Hess} V\, \nabla f$

$$BE(K,\infty) \iff \mathrm{Hess} V \geq K\, Id.$$

Indeed: $BE(K,\infty) \iff \Gamma(P_t f) \leq e^{-2Kt} P_t\Gamma(f)$

(Trick show $\psi'(s) \geq 0$, with $\psi(s) = P_s(\Gamma(P_{t-s}f))$.

After that:

$$\begin{aligned} Var_\mu(f) &= -\int\int_0^\infty \partial_s(P_s f)^2 ds d\mu \\ &= 2\int_0^\infty \int \Gamma(P_s f) d\mu ds \\ &\overset{c}{\leq} 2\int_0^\infty e^{-2Ks} ds \int P_s\Gamma(f) d\mu \\ &= \frac{1}{K}\int \Gamma(f) d\mu \end{aligned}$$

Commutation is crucial (in this proof...), but one can give integrated criterion. See the talk of Joulin for refinements around this commutation property.

Indeed: $BE(K,\infty) \iff \Gamma(P_t f) \le e^{-2Kt} P_t\Gamma(f)$

(Trick show $\psi'(s) \ge 0$, with $\psi(s) = P_s(\Gamma(P_{t-s}f))$.

After that:

$$\begin{aligned} Var_\mu(f) &= -\int\int_0^\infty \partial_s(P_s f)^2 ds d\mu \\ &= 2\int_0^\infty \int \Gamma(P_s f) d\mu ds \\ &\overset{c}{\le} 2\int_0^\infty e^{-2Ks} ds \int P_s\Gamma(f) d\mu \\ &= \frac{1}{K}\int \Gamma(f) d\mu \end{aligned}$$

Commutation is crucial (in this proof...), but one can give integrated criterion. See the talk of Joulin for refinements around this commutation property.

Links between functional inequalities and Poincaré inequalities?

In fact, in the reversible case $\mu(fLg) = \mu(gLf)$, one has in "regular situation"

$$\text{Lyapunov Condition} \iff \text{Poincaré inequalities}$$

Indeed let us suppose that $LW \leq -\lambda W + b 1_C$ and a local Poincaré inequality

$$\int_C (f - \mu(f 1_C))^2 d\mu \leq \kappa_C \int_C \Gamma(f) d\mu$$

then Poincaré inequality holds with constant $\frac{1}{\lambda}(1 + b\kappa_C)$

Proof:

$$\begin{aligned} Var_\mu(f) &\leq \int (f - \mu(f1_C))^2 d\mu \\ &\leq \frac{1}{\lambda} \int (f - \mu(f1_C))^2 \frac{-LW}{W} d\mu + \frac{1}{\lambda} \int_C (f - \mu(f1_C))^2 d\mu \end{aligned}$$

Use then local Poincaré inequality and the large deviations bound (or direct calculus for diffusion) in the reversible case

$$\int f^2 \frac{-LW}{W} d\mu \leq \int \Gamma(f) d\mu$$

There are also variants for

- subexponential decay via weak Poincaré
- decay in entropy via logarithmic Sobolev inequality
- decay in Wasserstein distance via WJ inequality
- in L^2 weighted space in non reversible case via Lyapunov-Poincaré inequalities

Advantages

- generic method with other consequences (concentration,...)
- variety of criterion

Disadvantages

• non reversible case is difficult, a refined spectral analysis may be used (see Monmarché)

- sharp constants can be hard to get.

Conclusion

Two different methods which can be generically tested, at least for qualitative results....

In general, one has to refine these methods to get a good answer to a particular problem.....

it will be the subject of quite a few talks here...

Facultative slide : let consider P_t the Heat semigroup on $\mathbb{R}^n$.

By synchronous coupling it is not difficult to get that for densities (wrt Lebesgue) f, g

$$W_2(P_t f, P_t g) \leq W_2(f, g)$$

but we recently prove by analytical method (with Bolley and Gentil) that

$$W_2^2(P_t f, P_t g) \leq W_2^2(f, g) - \frac{2}{n} \int_0^t (Ent_\mu(P_s f) - Ent_\mu(P_s g))^2 ds$$

Can you find a coupling proof????