

# Contraction Lipschitzienne des semigroupes Markoviens

Bertrand CLOEZ

Doctorant à l'université de Marne-la-Vallée sous la direction de Djaliil Chafaï

Journées des jeunes probabilistes de France, 18 avril 2012

# Semigroupe Markovien

Si  $(X_t)_{t \geq 0}$  est un processus Markovien, son semigroupe est défini par

$$P_t f(x) = \mathbb{E}[f(X_t) \mid X_0 = x].$$

# Semigroupe Markovien

Si  $(X_t)_{t \geq 0}$  est un processus Markovien, son semigroupe est défini par

$$P_t f(x) = \mathbb{E}[f(X_t) \mid X_0 = x].$$

Il vérifie naturellement

- $P_t 1 = 1$
- $f \geq 0 \Rightarrow P_t f \geq 0$
- $P_{t+s} = P_t \circ P_s$
- $\lim_{t \rightarrow 0} P_t f = f$

# Distance de Wasserstein

- Pour toutes lois  $\mu_1, \mu_2$  sur  $(E, d)$  :

$$\begin{aligned}\mathcal{W}(\mu_1, \mu_2) &= \inf_{\Pi} \int_{E \times E} d(x, y) \Pi(dx, dy) \\ &= \inf_{X_1 \sim \mu_1, X_2 \sim \mu_2} \mathbb{E}[d(X_1, X_2)].\end{aligned}$$

# Distance de Wasserstein

- Pour toutes lois  $\mu_1, \mu_2$  sur  $(E, d)$  :

$$\begin{aligned}\mathcal{W}(\mu_1, \mu_2) &= \inf_{\Pi} \int_{E \times E} d(x, y) \Pi(dx, dy) \\ &= \inf_{X_1 \sim \mu_1, X_2 \sim \mu_2} \mathbb{E}[d(X_1, X_2)].\end{aligned}$$

- Convergence en  $\mathcal{W}$   $\Leftrightarrow$  Convergence en loi + 1<sup>er</sup> moment

# Distance de Wasserstein

- Pour toutes lois  $\mu_1, \mu_2$  sur  $(E, d)$  :

$$\begin{aligned}\mathcal{W}(\mu_1, \mu_2) &= \inf_{\Pi} \int_{E \times E} d(x, y) \Pi(dx, dy) \\ &= \inf_{X_1 \sim \mu_1, X_2 \sim \mu_2} \mathbb{E}[d(X_1, X_2)].\end{aligned}$$

- Convergence en  $\mathcal{W} \Leftrightarrow$  Convergence en loi + 1<sup>er</sup> moment
- distance de Kantorovich, de Mallows, de Monge, de Fréchet, de transport, de couplage,  $L^1$  généralisé...

# Distance de Wasserstein

Formulation duale de Kantorovich-Rubinstein :

$$\mathcal{W}(\mu_1, \mu_2) = \sup_{\text{Lip}(f) \leq 1} \int f d\mu_1 - \int f d\mu_2$$

# Distance de Wasserstein

Formulation duale de Kantorovich-Rubinstein :

$$\mathcal{W}(\mu_1, \mu_2) = \sup_{\text{Lip}(f) \leq 1} \int f d\mu_1 - \int f d\mu_2$$

Sur  $\mathbb{R}$  :

$$\mathcal{W}(\mu_1, \mu_2) = \int_0^1 |F_1^{-1}(u) - F_2^{-1}(u)| du$$

et

$$\mathcal{W}(\mu_1, \mu_2) = \int_{\mathbb{R}} |F_1(x) - F_2(x)| dx.$$

Où  $F_1, F_2$  sont les fonctions de répartition et  $F_1^{-1}, F_2^{-1}$  leurs inverses généralisés.

# Courbure de Wasserstein

## Definition

La courbure de Wasserstein d'un semigroupe de Markov  $(P_t)_{t \geq 0}$  est la plus grande constante  $\rho$  tel que

$$\mathcal{W}(\mu P_t, \nu P_t) \leq e^{-\rho t} \mathcal{W}(\mu, \nu),$$

pour toutes mesures de probabilités  $\mu, \nu$  et pour tout  $t \geq 0$ .

# Courbure de Wasserstein

## Definition

La courbure de Wasserstein d'un semigroupe de Markov  $(P_t)_{t \geq 0}$  est la plus grande constante  $\rho$  tel que

$$\mathcal{W}(\mu P_t, \nu P_t) \leq e^{-\rho t} \mathcal{W}(\mu, \nu),$$

pour toutes mesures de probabilités  $\mu, \nu$  et pour tout  $t \geq 0$ .

- $\rho > 0 \Rightarrow$  existence et unicité d'une mesure invariante  $\pi$ . De plus

$$\mathcal{W}(\mu P_t, \pi) \leq e^{-\rho t} \mathcal{W}(\mu, \pi).$$

# Courbure de Wasserstein

## Definition

La courbure de Wasserstein d'un semigroupe de Markov  $(P_t)_{t \geq 0}$  est la plus grande constante  $\rho$  tel que

$$\mathcal{W}(\mu P_t, \nu P_t) \leq e^{-\rho t} \mathcal{W}(\mu, \nu),$$

pour toutes mesures de probabilités  $\mu, \nu$  et pour tout  $t \geq 0$ .

- $\rho > 0 \Rightarrow$  existence et unicité d'une mesure invariante  $\pi$ . De plus

$$\mathcal{W}(\mu P_t, \pi) \leq e^{-\rho t} \mathcal{W}(\mu, \pi).$$

- $\rho > 0 +$  réversibilité  $\Rightarrow$  un trou spectral :

$$\text{Var}_\pi(P_t f) \leq e^{-2\rho t} \text{Var}_\pi(f).$$

# Processus homogènes

Soit  $(L_t)_{t \geq 0}$  un processus de Lévy sur  $\mathbb{R}^d$ .

# Processus homogènes

Soit  $(L_t)_{t \geq 0}$  un processus de Lévy sur  $\mathbb{R}^d$ . C'est à dire

- $L_0 = 0$
- $L_t - L_s \stackrel{d}{=} L_{t-s}$
- $L_t - L_s$  est indépendant de  $L_s$

# Processus homogènes

Soit  $(L_t)_{t \geq 0}$  un processus de Lévy sur  $\mathbb{R}^d$ . C'est à dire

- $L_0 = 0$
- $L_t - L_s \stackrel{d}{=} L_{t-s}$
- $L_t - L_s$  est indépendant de  $L_s$

On définit un semigroupe de la manière suivante :

$$P_t f(x) = \mathbb{E}[f(x + L_t)]$$

# Processus homogènes

Soit  $(L_t)_{t \geq 0}$  un processus de Lévy sur  $\mathbb{R}^d$ . C'est à dire

- $L_0 = 0$
- $L_t - L_s \stackrel{d}{=} L_{t-s}$
- $L_t - L_s$  est indépendant de  $L_s$

On définit un semigroupe de la manière suivante :

$$P_t f(x) = \mathbb{E}[f(x + L_t)]$$

**Theorem (Courbure d'un processus de Lévy)**

*La courbure de Wasserstein de ce semigroupe est nulle. De plus,*

$$\mathcal{W}(\delta_x P_t, \delta_y P_t) = |x - y| = \mathcal{W}(\delta_x, \delta_y).$$

# Mouvement Brownien sur une variété

Soit  $(M, g)$  une surface Riemannienne (régulière) et  $(P_t)_{t \geq 0}$  le semigroupe solution de l'équation de la chaleur :

$$\forall t \geq 0, \partial_t P_t f = \Delta P_t f,$$

pour toute fonction  $f$  suffisamment régulière.  $\Delta$  représente l'opérateur de Laplace-Beltrami.

## Theorem (Sturm et Von Renesse, 2004)

On a

$$\forall x \in M, \forall v \in \mathbb{R}^n, \text{Ricci}_x(v, v) \geq k \|v\|^2 \Leftrightarrow \rho \geq k.$$

De plus, pour tout  $x, y \in M$ , il existe un couple de Mouvement Brownien  $(B_t^x, B_t^y)_{t \geq 0}$  tel que

$$\forall t \geq 0, d(B_t^x, B_t^y) \leq e^{-\rho t/2} d(x, y),$$

et  $B_0^x = x, B_0^y = y$ .

# Processus de naissance et de mort

On considère le semigroupe généré par

$$Lf(x) = b(x)(f(x+1) - f(x)) + d(x)(f(x-1) - f(x)).$$

# Processus de naissance et de mort

On considère le semigroupe généré par

$$Lf(x) = b(x)(f(x+1) - f(x)) + d(x)(f(x-1) - f(x)).$$

## Theorem (Chafaï et Joulin, 2012)

*On suppose que le processus est irréductible et non explosif. La courbure de Wasserstein est alors donnée par*

$$\rho = \inf_{n \geq 0} (d(n+1) - d(n) + b(n) - b(n+1)).$$

# Processus de naissance et de mort

On considère le semigroupe généré par

$$Lf(x) = b(x)(f(x+1) - f(x)) + d(x)(f(x-1) - f(x)).$$

## Theorem (Chafaï et Joulin, 2012)

*On suppose que le processus est irréductible et non explosif. La courbure de Wasserstein est alors donnée par*

$$\rho = \inf_{n \geq 0} (d(n+1) - d(n) + b(n) - b(n+1)).$$

→ la loi invariante est donnée  $\pi(\{i\}) = \text{Cst} \prod_{n=1}^i b(n-1)/d(n)$ .

# Processus TCP window size, $r(x) = ax + b$

On considère le semigroupe généré par

$$Lf(x) = f'(x) + r(x) \left( f\left(\frac{x}{2}\right) - f(x) \right).$$

# Processus TCP window size, $r(x) = ax + b$

On considère le semigroupe généré par

$$Lf(x) = f'(x) + r(x) \left( f\left(\frac{x}{2}\right) - f(x) \right).$$

Theorem ( Chafaï, Malrieu, Paroux, 2010)

Si  $r(x) = ax + b$  alors  $\rho = b$

# Processus TCP window size, $r(x) = ax + b$

On considère le semigroupe généré par

$$Lf(x) = f'(x) + r(x) \left( f\left(\frac{x}{2}\right) - f(x) \right).$$

Theorem ( Chafaï, Malrieu, Paroux, 2010)

Si  $r(x) = ax + b$  alors  $\rho = b$

→ Méthode : couplage au mieux des temps.

# Preuve quand $r(x) = r$

Quand  $r$  est constant, on peut construire un couplage  $(X_t, Y_t)_{t \geq 0}$  avec les même temps de sauts associés à un processus de Poisson  $(N_t)_{t \geq 0}$  :

$$\begin{aligned}\mathbb{E}(|X_t - Y_t|) &= \sum_{k=0}^{\infty} \mathbb{E}(|X_t - Y_t| \mathbb{1}_{\{N_t=k\}}) \\ &= |x - y| \sum_{k=0}^{\infty} 2^{-k} \mathbb{P}(N_t = k) \\ &= |x - y| e^{-rt/2}.\end{aligned}$$

Ainsi, si  $\Pi$  est un couplage des lois initiales :

$$\mathcal{W}(\mathcal{L}(X_t), \mathcal{L}(Y_t)) \leq e^{-rt/2} \int_{\mathbb{R}_+^2} |x - y| \Pi(dx, dy).$$

# Preuve quand $r(x) = r$

Quand  $r$  est constant, on peut construire un couplage  $(X_t, Y_t)_{t \geq 0}$  avec les même temps de sauts associés à un processus de Poisson  $(N_t)_{t \geq 0}$  :

$$\begin{aligned}\mathbb{E}(|X_t - Y_t|) &= \sum_{k=0}^{\infty} \mathbb{E}(|X_t - Y_t| \mathbb{1}_{\{N_t=k\}}) \\ &= |x - y| \sum_{k=0}^{\infty} 2^{-k} \mathbb{P}(N_t = k) \\ &= |x - y| e^{-rt/2}.\end{aligned}$$

Ainsi, si  $\Pi$  est un couplage des lois initiales :

$$\mathcal{W}(\mathcal{L}(X_t), \mathcal{L}(Y_t)) \leq e^{-rt/2} \int_{\mathbb{R}_+^2} |x - y| \Pi(dx, dy).$$

→ Ne marche pas si  $r$  n'est pas constant.

# Vitesse exponentielle quand $r$ décroît

## Theorem (C.12)

*Si  $r$  décroît alors la courbure vaut*

$$\rho = \frac{1}{2} \inf_{x \geq 0} (r(x) - xr'(x)).$$

# Vitesse exponentielle quand $r$ décroît

## Theorem (C.12)

*Si  $r$  décroît alors la courbure vaut*

$$\rho = \frac{1}{2} \inf_{x \geq 0} (r(x) - xr'(x)).$$

→ Même borne que la précédente.

# Vitesse exponentielle quand $r$ décroît

## Theorem (C.12)

*Si  $r$  décroît alors la courbure vaut*

$$\rho = \frac{1}{2} \inf_{x \geq 0} (r(x) - xr'(x)).$$

→ Même borne que la précédente.

→ Méthode : Relation de "commutation" ou d'"entrelacement" entre  $\nabla$  et  $L$ .

# Preuve

On montre d'abord que

$$\nabla \mathbb{E}_x[f(X_t)] = \mathbb{E}_x \left[ \nabla f(Y_t) e^{-\int_0^t V(Y_s) ds} \right]$$

avec

$$V(x) = \frac{1}{2}(r(x) - xr'(x)).$$

# Preuve

On montre d'abord que

$$\nabla \mathbb{E}_x[f(X_t)] = \mathbb{E}_x \left[ \nabla f(Y_t) e^{-\int_0^t V(Y_s) ds} \right]$$

avec

$$V(x) = \frac{1}{2}(r(x) - xr'(x)).$$

Puis par le théorème des accroissements finis :

$$\begin{aligned} |\mathbb{E}_x[f(X_t)] - \mathbb{E}_y[f(X_t)]| &\leq \sup_{\zeta \in [x,y]} \mathbb{E}_\zeta \left[ f'(Y_t) e^{-\int_0^t V(Y_s) ds} \right] |x - y| \\ &\leq e^{-t \inf_{z \geq 0} V(z)} |x - y|. \end{aligned}$$

# Merci de votre attention !