SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

MICHEL ÉMERY

Information associée à un semigroupe

Séminaire de probabilités (Strasbourg), tome 11 (1977), p. 566-573
<http://www.numdam.org/item?id=SPS_1977__11__566_0>

Université de Strasbourg

Séminaire de Probabilités

INFORMATION ASSOCIEE A UN SEMIGROUPE

par

M. EMERY

L'objet de cet exposé est un théorème de Donsker et Varadhan (M. DONSKER et S.R.S. VARADHAN, Asymptotic evaluation of certain Markov process expectations for large time, Comm. Pure and Applied Math., 1975).

On considère un espace mesurable $(E,\mathcal{E})$ sur lequel opère un semigroupe $(P_t\,,\,t\geq 0)$ de noyaux markoviens. On désigne par B l'espace des fonctions mesurables bornées sur $(E,\mathcal{E})$, par B_o le sous-espace de B formé des fonctions f telles que, lorsque t tend vers zéro, P_tf tend vers f uniformément sur E , par A le générateur infinitésimal de P_t , dont le domaine D_A est dense dans B_o. On note B^+ (respectivement B_o^+ , D_A^+) le sous-ensemble de B (respectivement B_o , D_A) formé des fonctions dont la borne inférieure est strictement positive.

Pour toute probabilité μ sur $(E,\mathcal{E})$, on pose

$$I_t(\mu) = -\inf_{f\in B^+}\int_E \mathrm{Log}\,\frac{P_tf}{f}\,d\mu\ ;$$

$$I'(\mu) = -\inf_{f\in D_A^+}\int_E \frac{Af}{f}\,d\mu\ .$$

Ces deux quantités sont à valeurs dans $[0,\infty]$; pour tous t et s , $I_{t+s}(\mu)\leq I_t(\mu)+I_s(\mu)$.

PROPOSITION 1. Etant donnés t et μ, une fonction f de B^+ vérifie $I_t(\mu) = -\int_E \text{Log} \frac{P_t f}{f} d\mu$ si et seulement si $\frac{1}{f}\cdot\mu = (\frac{1}{P_t f}\cdot\mu)P_t$.

Démonstration : (L'indice t est omis.)

Si $I(\mu) = \int_E \text{Log} \frac{f}{Pf} d\mu$, alors, pour toute fonction g de B^+,

$$\int_E \text{Log} \frac{f}{Pf} d\mu \geq \int_E \text{Log} \frac{g}{Pg} d\mu .$$

Pour $h \in B$ telle que $\sup_E |h| \leq \frac{1}{2} \inf_E f$, $f + xh$ est dans B^+ pour tout x de $]-1,1[$, donc

$$\int_E \text{Log} \frac{Pf + xPh}{f + xh} d\mu \geq \int_E \text{Log} \frac{Pf}{f} d\mu .$$

Mais, pour $|x| < 1$, $|\frac{h}{f + xh}| \leq 1$ et $|\frac{Ph}{Pf + xPh}| \leq 1$, d'où, par dérivation sous le signe somme,

$$\frac{d}{dx} \int_E \text{Log} \frac{Pf + xPh}{f + xh} d\mu = \int_E (\frac{Ph}{Pf + xPh} - \frac{h}{f + xh})d\mu .$$

La dérivée doit s'annuler pour $x = 0$, et $\int_E \frac{Ph}{Pf} d\mu = \int_E \frac{h}{f} d\mu$. Ceci, ayant lieu pour toute fonction borélienne bornée h, entraîne l'égalité des mesures $\frac{1}{g}\cdot\mu$ et $(\frac{1}{Pf}\cdot\mu)P$. Réciproquement, si ces deux mesures sont égales, soit $g \in B^+$. L'application $(x,y) \to x \text{ Log} \frac{y}{x}$ étant concave sur $\mathbb{R}_+^* \times \mathbb{R}_+^*$, on peut écrire, pour toute probabilité ν sur $(E,\mathcal{E})$

$$\int_E f \text{ Log} \frac{g}{f} d\nu \leq \int_E f\, d\nu \text{ Log} \frac{\int_E g d\nu}{\int_E f d\nu} .$$

En prenant pour ν la probabilité $P(x,.)$, on obtient

$$P\ (f \text{ Log} \frac{g}{f}) \leq P\ f \text{ Log} \frac{P\ g}{P\ f} ,$$

d'où

$$\int_E \frac{P\ (f \text{ Log} \frac{g}{f})}{P\ f} d\mu \leq \int_E \text{Log} \frac{P\ g}{P\ f} d\mu .$$

Mais, par hypothèse, le premier membre vaut

$$\int_E f \text{ Log } \frac{g}{f} \, d[(\frac{1}{P f}.\mu)P\,] = \int_E f \text{ Log } \frac{g}{f} \, d(\frac{1}{f}.\mu) = \int_E \text{Log } \frac{g}{f} \, d\mu \ ,$$

et donc, pour tout $g \in B^+$,

$$\int_E \text{Log } \frac{P\,g}{g} \, d\mu \geq \int_E \text{Log } \frac{P\,f}{f} \, d\mu \ ,$$

ce qui achève la démonstration.

PROPOSITION 2. Soit, pour $x \in [0,1]$, $\varphi(x) = \text{Log } \frac{1}{1-x} - x$. Alors, pour tout $F \in \mathcal{E}$, $I_t(\mu) \geq \varphi(|\mu P_t(F) - \mu(F)|)$. En particulier, $I_t(\mu) = 0$ si et seulement si $\mu P_t = \mu$.

Démonstration : (On ommet encore l'indice t .)

Posant $x = \mu(F) - \mu P(F)$, on peut supposer $x > 0$ (le cas $x = 0$ est trivial ; si $x < 0$, remplacer F par son complémentaire). Soit $f = 1 + \alpha I_F \in B^+$, où la constante $\alpha > 0$ sera fixée plus tard.

$$I(\mu) = \sup_{g \in B^+} \int_E \text{Log } \frac{g}{Pg} \, d\mu \geq \int_E \text{Log } \frac{f}{Pf} \, d\mu$$

$$\geq \text{Log}(1+\alpha) \ \mu(F) - \int_E \text{Log}(1+\alpha P(.,F))d\mu$$

$$\geq \text{Log}(1+\alpha) \ \mu(F) - \alpha \ \mu P(F) \ .$$

Si $\mu P(F) = 0$, $I(\mu) \geq \text{Log}(1+\alpha) \ \mu(F)$ pour tout $\alpha > 0$, donc $I(\mu) = +\infty$ et il n'y a rien à démontrer.

Si $\mu P(F) > 0$, prenons $\alpha = \frac{x}{\mu P(F)}$.

$$I(\mu) \geq \text{Log}(1+\frac{x}{\mu P(F)}) \ \mu(F) - x = \text{Log}(1+\frac{x}{\mu P(F)})(\mu P(F) + x) - x$$

$$\geq \inf_{0 \leq y \leq 1-x} [\text{Log}(1+\frac{x}{y})(x+y) - x] \ .$$

C'est une fonction décroissante de y (dérivée : $\text{Log}(1+\frac{x}{y}) - \frac{x}{y} \leq 0$). L' inf est obtenu pour $y = 1-x$:

$$I(\mu) \geq \text{Log}(1+\frac{x}{1-x}) - x = \varphi(x) \ .$$

La deuxième partie de la proposition en découle, car si $\mu P_t = \mu$, la proposition 1 entraîne que $I_t(\mu) = -\int_E \text{Log}\,\frac{P_t 1}{1}\,d\mu = 0$. Cette dernière proposition éclaire le rôle de I_t : intuitivement, $I_t(\mu)$ contrôle le gain d'information entre μ et μP_t .

On suppose maintenant vérifiée la condition suivante :

(a) Pour toute probabilité ν sur $(E,\mathcal{E})$ et toute fonction $f \in B$, il existe une suite $(f_n)_{n \in \mathbb{N}}$ d'éléments de B_o telle que $\|f_n\|_B \leq \|f\|_B$ et que

$$\lim_n f_n = f \quad \nu\text{-presque partout.}$$

THEOREME 1. <u>Lorsque la condition</u> (a) <u>est satisfaite, alors, pour tout</u> t , $I_t(\mu) \leq tI'(\mu)$; <u>plus précisément,</u> $I'(\mu)$ <u>est la dérivée à droite en zéro de</u> $I_t(\mu)$.

<u>Démonstration</u> : Soit $f \in D_A^+$; si l'on pose $\psi(t) = \int_E \text{Log}\,\frac{P_t f}{f}\,d\mu$, par dérivation sous le signe somme, $\psi'(t) = \int_E \frac{AP_t f}{P_t f}\,d\mu$, et $\psi'(t) \geq -I'(\mu)$.

$$\int_E \text{Log}\,\frac{P_t f}{f}\,d\mu = \psi(t) = \int_0^t \psi'(s)ds \geq \int_0^t -I'(\mu)ds = -tI'(\mu) .$$

Ceci reste vrai pour f dans B_o^+ puisque D_A est dense dans B_o . Pour $f \in B^+$, il existe une suite (f_n) de fonction de B_o^+ , ayant au plus la norme de f , qui tendent vers f μ-presque partout et μP_t-presque partout. Par convergence dominée, on a encore

$$\int_E \text{Log}\,\frac{P_t f}{f}\,d\mu \geq -tI'(\mu) ,$$

d'où $I_t(\mu) \leq tI'(\mu)$.

Ensuite, pour $f \in D_A^+$, $\frac{1}{t}\,I_t(\mu) \geq -\frac{1}{t}\int_E \text{Log}\,\frac{P_t f}{f}\,d\mu$. Mais $P_t f = f + tAf + o(t)$, où $\frac{o(t)}{t}$ tend vers zéro dans B lorsque t tend vers zéro. Donc $\text{Log}\,\frac{P_t f}{f} = t\,\frac{Af}{f} + o(t)$, et

$$\frac{1}{t}\,I_t(\mu) \geq \int_E -\left(\frac{Af}{f} + \frac{o(t)}{t}\right)d\mu$$

$$\liminf_{t \to 0^+} \frac{1}{t}\,I_t(\mu) \geq -\int_E \frac{Af}{f}\,d\mu$$

$$\liminf \frac{1}{t} I_t(\mu) \geq I'(\mu) .$$

Ceci établit le théorème.

COROLLAIRE. Sous les hypothèses du thoérème, si μ est une probabilité telle que $I'(\mu) < \infty$, alors, uniformément en $F \in \mathcal{E}$, $\mu P_t(F)$ tend vers $\mu(F)$ lorsque t tend vers zéro.

Démonstration :

$$\tfrac{1}{2}[\mu P_t(F) - \mu(F)]^2 \leq \varphi(|\mu P_t(F) - \mu(F)|) \leq I_t(\mu) \leq tI'(\mu) .$$

Dans le cas auto-adjoint, il est possible de préciser la valeur de $I(\mu)$: nous supposons maintenant, outre la condition (a), que

(b) $P_t(x,dy) = p_t(x,y)\ \lambda(dy)$, où λ est une mesure positive σ-finie sur E, , et où les fonctions p_t sont mesurables et symétriques sur $E \times E$.

(c) Le sous-espace de $L^2(\lambda)$ formé des éléments dont un représentant est dans B_o est dense dans $L^2(\lambda)$.

Sous ces conditions, le semigroupe P_t opère dans L^2 . C'est un semi-groupe fortement continu, à contraction, auto-adjoint, admettant un générateur infinitésimal $\widetilde{A}$ auto-adjoint négatif. Soient $\sqrt{-\widetilde{A}}$ la racine carrée positive de $-\widetilde{A}$ et $D_{\sqrt{-\widetilde{A}}}$ son domaine.

THEOREME 2. Sous les condition (a), (b), (c) , soit μ une probabilité sur $(E,\mathcal{E})$. Alors $I'(\mu)$ est fini si et seulement si μ est absolument continue par rapport à λ , avec

$$g = \sqrt{\frac{d\mu}{d\lambda}} \in D_{\sqrt{-\widetilde{A}}} .$$

Lorsque c'est le cas, $I'(\mu) = \|\sqrt{-\widetilde{A}}\, g\|^2$.

Démonstration :

$$\text{i)}\quad I'(\mu) < \infty \Rightarrow \begin{cases} \mu = g^2.\lambda \\ g \in D_{\sqrt{-\widetilde{A}}} \\ \|\sqrt{-\widetilde{A}}\, g\|^2 \leq I'(\mu) \end{cases}$$

$$\text{ii)}\quad \left.\begin{matrix} \mu = g^2.\lambda \\ g \in D_{\sqrt{-\widetilde{A}}} \end{matrix}\right\} \Rightarrow \|\sqrt{-\widetilde{A}}\, g\|^2 \geq I'(\mu) \ .$$

i) On suppose $I'(\mu)$ fini. Soit F tel que $\mu(F) > 0$. Pour t assez petit, $\mu P_t(F) > 0$, d'où $\mu \ll \lambda$. Soit $g = \sqrt{\frac{d\mu}{d\lambda}} \in L^2(\lambda)$. Pour toute fonction $f \in B^+$, on peut écrire

$$\int_E \frac{f - P_t f}{f}\, d\mu \leq \int_E -\text{Log}\, \frac{P_t f}{f}\, d\mu \leq I_t(\mu) \leq tI'(\mu) \ .$$

En particulier, pour $g_n = \inf(g,n)$,

$$\int_E \frac{(I - P_t)g_n}{g_n + \varepsilon}\, g^2\, d\lambda \leq tI'(\mu) \ ,$$

et, à la limite,

$$\int_E \frac{(I - P_t)g}{g + \varepsilon}\, g^2\, d\lambda \leq tI'(\varepsilon)$$

$$\int_E (I - P_t)g\; g\; d\lambda \leq tI'(\mu) \ .$$

Si H_x est le sous-espace où $(-\widetilde{A} - xI)^+$ est nul, et E_x l'opérateur de projection sur H_x, on a les décompositions spectrales

$$-\widetilde{A} = \int_0^\infty x\, dE_x \ ; \ P_t = \int_0^\infty e^{-tx}\, dE_x \ .$$

En notant sur la mesure sur $\mathbb{R}_+$ telle que

$$m([0,x[) = \langle E_x g, g \rangle \ ,$$

il vient $\int_0^\infty (1 - e^{-tx})dm(x) \leq tI'(\mu)$.

Mais, lorsque t décroît vers zéro, $\frac{1}{t}(1-e^{-tx})$ croît vers x, et $\int_0^\infty x\,dm(x) \le I'(\mu)$, c'est-à-dire $g \in D_{\sqrt{-\tilde{A}}}$ avec $\|\sqrt{-\tilde{A}}\,g\|^2 \le I(\mu)$.

ii) On suppose $\mu = g^2.\lambda$, où $g \in D_{\sqrt{-\tilde{A}}}$. Il s'agit de démontrer que, pour $u \in D_A^+$,

$$\int_E -\frac{Au}{u}\,g^2\,d\lambda \le \langle -\tilde{A}\,g,g\rangle ,$$

ou encore, en notant $\bar{A}$ l'opérateur de L^2 défini par

$$\bar{A}h(x) = \tilde{A}h(x) - \frac{Au(x)}{u(x)}\,h(x) ,$$

que $\bar{A}$ est de type négatif.

Les notations E^x et X_t se référant au processus de Markov associé au semigroupe P_t, les processus

$$M_t = \int_0^t Au(X_s)ds \quad \text{et} \quad N_t = \int_0^t \frac{1}{u(X_s)}\,dM_s$$

sont des martingales pour les lois P^x.

La formule d'Ito

$$\operatorname{Log} Y_t = \operatorname{Log} Y_0 + \int_0^t \frac{dY_s}{Y_s} - \frac{1}{2}\int_0^t \frac{1}{Y_s^2}\,d\langle Y^c,Y^c\rangle_s$$

appliquée à $Y_t = u(X_t)$ fournit, P^x-presque sûrement,

$$\exp\left(-\int_0^t \frac{Au}{u}(X_s)ds\right) = \frac{u(X_0)}{u(X_t)}\,e^{-N_t - \frac{1}{2}\langle N^c,N^c\rangle_t}$$

$$\le \frac{\sup u}{\inf u}\,e^{-N_t - \frac{1}{2}\langle N^c,N^c\rangle_t}$$

d'où $\sup\limits_t \sup\limits_x E^x\left[\exp\left(-\int_0^t \frac{Au}{u}(X_s)ds\right)\right] \le \frac{\sup u}{\inf u} = C < \infty$.

Ceci étant, les opérateurs P_t^u définis par

$$P_t^u f(x) = E^x\left[f(X_t)\exp\left(-\int_0^t \frac{Au}{u}(X_s)ds\right)\right]$$

forment un semigroupe auto-adjoint qui opère dans $L^2(\lambda)$ et dont le générateur infinitésimal n'est autre que $\bar{A}$. L'inégalité de Schwarz entraîne

$$(P_t^u f)^2(x) \le E^x[\exp(-\int_0^t \frac{Au}{u}(X_s)ds)]\ E^x[f^2(X_t)\ \exp(-\int_0^t \frac{Au}{u}(X_s)ds)]$$

$$\le C\ P_t^u\ f^2(x)$$

$$\|P_t^u f\|^2 \le C < P_t^u f^2, 1 > = C < f^2, P_t^u 1 > \le C^2\|f\|^2$$

d'où $K_t = \|P_t^u\| \le C$.

Mais comme en outre

$$\|P_t^u f\|^2 = < P_t^u f, P_t^u f > = < P_{2t}^u f, f > \le K_{2t}\|f\|^2 ,$$

$$K_t \le \sqrt{K_{2t}} \text{ , et, en fin de compte, } \|P_t^u\| \le 1 .$$

Le générateur infinitésimal $\bar{A}$ est de type négatif, ce qui termine la démonstration.