

# 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](http://www.numdam.org/item?id=SPS_1977__11__566_0)

© Springer-Verlag, Berlin Heidelberg New York, 1977, tous droits réservés.

L'accès aux archives du séminaire de probabilités (Strasbourg) (<http://portail.mathdoc.fr/SemProba/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme  
Numérisation de documents anciens mathématiques

<http://www.numdam.org/>

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  $B_0$  le sous-espace de  $B$  formé des fonctions  $f$  telles que, lorsque  $t$  tend vers zéro,  $P_t f$  tend vers  $f$  uniformément sur  $E$ , par  $A$  le générateur infinitésimal de  $P_t$ , dont le domaine  $D_A$  est dense dans  $B_0$ . On note  $B^+$  (respectivement  $B_0^+, D_A^+$ ) le sous-ensemble de  $B$  (respectivement  $B_0, D_A$ ) formé des fonctions dont la borne inférieure est strictement positive.

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

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

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

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  $\mu$ , une fonction  $f$  de  $B^+$  vérifie

$$I_t(\mu) = - \int_E \text{Log} \frac{P_t f}{f} d\mu \quad \text{si et seulement si} \quad \frac{1}{f} \cdot \mu = \left( \frac{1}{P_t f} \cdot \mu \right) P_t .$$

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

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

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  $\left( \frac{1}{Pf} \cdot \mu \right) P$ . Réciproquement, si ces deux mesures sont égales, soit  $g \in B^+$ . L'application  $(x, y) \rightarrow x \text{Log} \frac{y}{x}$  étant concave sur  $\mathbb{R}_+^* \times \mathbb{R}_+^*$ , on peut écrire, pour toute probabilité  $\nu$  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  $\nu$  la probabilité  $P(x, \cdot)$ , on obtient

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

d'où

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

Mais, par hypothèse, le premier membre vaut

$$\int_E f \operatorname{Log} \frac{g}{f} d\left[\left(\frac{1}{P} \cdot \mu\right)^P\right] = \int_E f \operatorname{Log} \frac{g}{f} d\left(\frac{1}{f} \cdot \mu\right) = \int_E \operatorname{Log} \frac{g}{f} d\mu,$$

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

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

ce qui achève la démonstration.

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

Démonstration : (On omet 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.

$$\begin{aligned} I(\mu) &= \sup_{g \in B^+} \int_E \operatorname{Log} \frac{g}{Pg} d\mu \geq \int_E \operatorname{Log} \frac{f}{Pf} d\mu \\ &\geq \operatorname{Log}(1 + \alpha) \mu(F) - \int_E \operatorname{Log}(1 + \alpha P(\cdot, F)) d\mu \\ &\geq \operatorname{Log}(1 + \alpha) \mu(F) - \alpha \mu^P(F). \end{aligned}$$

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

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

$$\begin{aligned} I(\mu) &\geq \operatorname{Log}\left(1 + \frac{x}{\mu^P(F)}\right) \mu(F) - x = \operatorname{Log}\left(1 + \frac{x}{\mu^P(F)}\right) (\mu^P(F) + x) - x \\ &\geq \inf_{0 \leq y \leq 1-x} \left[ \operatorname{Log}\left(1 + \frac{x}{y}\right) (x + y) - x \right]. \end{aligned}$$

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

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

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  $\mu$  et  $\mu P_t$ .

On suppose maintenant vérifiée la condition suivante :

(a) Pour toute probabilité  $\nu$  sur  $(E, \mathcal{L})$  et toute fonction  $f \in B$ , il existe une suite  $(f_n)_{n \in \mathbb{N}}$  d'éléments de  $B_0$  telle que  $\|f_n\|_B \leq \|f\|_B$  et que

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

**THEOREME 1.** Lorsque la condition (a) est satisfaite, alors, pour tout  $t$ ,

$I_t(\mu) \leq tI'(\mu)$ ; plus précisément,  $I'(\mu)$  est la dérivée à droite en zéro de  $I_t(\mu)$ .

Démonstration : 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  $B_0^+$  puisque  $D_A$  est dense dans  $B_0$ . Pour  $f \in B^+$ , il existe une suite  $(f_n)$  de fonction de  $B_0^+$ , ayant au plus la norme de  $f$ , qui tendent vers  $f$   $\mu$ -presque partout et  $\mu P_t$ -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

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

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

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

Ceci établit le théorème.

COROLLAIRE. Sous les hypothèses du théorème, si  $\mu$  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 :

$$\frac{1}{2} [\mu P_t(F) - \mu(F)]^2 \leq \varphi(|\mu P_t(F) - \mu(F)|) \leq I_t(\mu) \leq t I'(\mu) .$$

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  $\lambda$  est une mesure positive  $\sigma$ -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  $B_0$  est dense dans  $L^2(\lambda)$ .

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

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

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

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

Démonstration :

$$i) \quad I'(\mu) < \infty \Rightarrow \begin{cases} \mu = g^2 \cdot \lambda \\ g \in D \\ \|\sqrt{-\tilde{A}} g\|^2 \leq I'(\mu) \end{cases}$$

$$ii) \quad \begin{cases} \mu = g^2 \cdot \lambda \\ g \in D \\ \sqrt{-\tilde{A}} \end{cases} \Rightarrow \|\sqrt{-\tilde{A}} g\|^2 \geq I'(\mu) .$$

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

$$\int_E \frac{f - P_t f}{f} d\mu \leq \int_E -\text{Log} \frac{P_t f}{f} d\mu \leq I_t(\mu) \leq t I'(\mu) .$$

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

$$\int_E \frac{(I - P_t)g_n}{g_n + \varepsilon} g^2 d\lambda \leq t I'(\mu) ,$$

et, à la limite,

$$\int_E \frac{(I - P_t)g}{g + \varepsilon} g^2 d\lambda \leq t I'(\mu)$$

$$\int_E (I - P_t)g g d\lambda \leq t I'(\mu) .$$

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

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

En notant sur la mesure sur  $R_+$  telle que

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

il vient  $\int_0^\infty (1 - e^{-tx}) dm(x) \leq t I'(\mu)$ .

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

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

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

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

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

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

$$\begin{aligned} \exp\left(-\int_0^t \frac{Au}{u}(X_s) \, ds\right) &= \frac{u(X_0)}{u(X_t)} e^{-N_t - \frac{1}{2} \langle N^C, N^C \rangle_t} \\ &\leq \frac{\sup u}{\inf u} e^{-N_t - \frac{1}{2} \langle N^C, N^C \rangle_t} \end{aligned}$$

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

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

$$\begin{aligned} (P_t^u f)^2(x) &\leq E^x[\exp(-\int_0^t \frac{Au}{u}(X_s)ds)] E^x[f^2(X_t) \exp(-\int_0^t \frac{Au}{u}(X_s)ds)] \\ &\leq C P_t^u f^2(x) \end{aligned}$$

$$\|P_t^u f\|^2 \leq C \langle P_t^u f^2, 1 \rangle = C \langle f^2, P_t^u 1 \rangle \leq C^2 \|f\|^2$$

d'où  $K_t = \|P_t^u\| \leq C$ .

Mais comme en outre

$$\|P_t^u f\|^2 = \langle P_t^u f, P_t^u f \rangle = \langle P_{2t}^u f, f \rangle \leq K_{2t} \|f\|^2,$$

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

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

