JEAN SAINT RAYMOND

Quelques remarques sur un article de Donsker et Varadhan

Séminaire de probabilités (Strasbourg), tome 12 (1978), p. 468-481
<http://www.numdam.org/item?id=SPS_1978__12__468_0>

Université de Strasbourg
Séminaire de Probabilités 1976/77

QUELQUES REMARQUES SUR UN ARTICLE DE M.D. DONSKER ET S.R.S. VARADHAN

par Jean Saint Raymond

Dans l'article [2], Donsker et Varadhan se placent alternativement dans les deux situations suivantes :

A. $\mathfrak{X}$ est un espace polonais, $\pi(x,dy)$ une probabilité de transition de Feller[1] sur $\mathfrak{X}$, π le noyau associé ($\pi u(x) = \int u(y)\pi(x,dy)$), $(\Omega,\underline{\underline{F}},(\underline{\underline{F}}_n),X_n)$ une chaîne de Markov dont $\mathfrak{X}$ est l'espace d'états et dont π engendre le semi-groupe de transition. On note $N_n(\omega)$ la mesure sur $\mathfrak{X}$ définie par

$$< N_n(\omega),f > = \frac{1}{n}\sum_{i=0}^{n-1} f\circ X_n(\omega)$$

que l'on considère comme un élément de l'espace $\underline{\underline{M}}(\mathfrak{X})$ des mesures de probabilité sur $\mathfrak{X}$, muni de la topologie étroite.

Pour tout x de $\mathfrak{X}$ et tout $n\geq 1$, la loi P^x induit par N_n une loi de probabilité $Q_{n,x}$ sur $\underline{\underline{M}}(\mathfrak{X})$, image de P^x par $\omega\mapsto N_n(\omega)$.

B. $\mathfrak{X}$ est un espace polonais, $(P_t)_{t\in\mathbb{R}_+}$ un semi-groupe de Feller sur $\mathfrak{X}$, L le générateur infinitésimal du semi-groupe (P_t), $\underline{\underline{D}}$ le domaine de L dans l'espace $\underline{\underline{C}}_b(\mathfrak{X})$ des fonctions continues bornées sur $\mathfrak{X}$, $(\Omega,\underline{\underline{F}},(\underline{\underline{F}}_t),X_t)$ un processus de Markov (mesurable[2]) dont $\mathfrak{X}$ est l'espace d'états et (P_t) le semi-groupe de transition. Pour $t>0$ et $\omega\in\Omega$, on note $N_t(\omega)$ la mesure définie par

$$< N_t(\omega),f > = \frac{1}{t}\int_0^t f\circ X_t(\omega)$$

qui appartient à $\underline{\underline{M}}(\mathfrak{X})$ comme ci-dessus.

Pour tout x de $\mathfrak{X}$ et tout $t>0$, la loi P^x sur Ω induit par N_t une loi de probabilité $Q_{t,x}$ sur $\underline{\underline{M}}(\mathfrak{X})$.

L'objet de l'article [2] est d'étudier le comportement asymptotique des lois $Q_{n,x}$ (resp. $Q_{t,x}$) quand n (resp. t) tend vers l'infini. Plus précisément, on veut montrer, sous certaines conditions restrictives, que si Φ est une fonction bornée étroitement continue sur $\underline{\underline{M}}(\mathfrak{X})$, on a

---

1. π (ou plus bas P_t) applique dans lui même l'espace $\underline{\underline{C}}_b(\mathfrak{X})$ des fonctions continues bornées sur $\mathfrak{X}$.
2. $\mathfrak{X}$ n'étant pas localement compact, le caractère fellérien de (P_t) n'entraîne pas l'existence de versions continues à droite.

dans le cas A

$$\lim_{n\to\infty} \frac{1}{n} \log \langle Q_{n,x}, e^{-n\Phi} \rangle = - \inf_{\mu \in \underline{\underline{M}}(\mathfrak{X})} (I(\mu) + \Phi(\mu))$$

avec $I(\mu) = \sup_{u \in \underline{\underline{U}}} \int \log \frac{u}{\pi u}(x)\, \mu(dx)$,

et $\underline{\underline{U}} = \{ u \in \underline{\underline{C}}_b(\mathfrak{X}) \mid u>0 \text{ et } \log u \text{ borné} \}$

dans le cas B

$$\lim_{t\to\infty} \frac{1}{t} \log \langle Q_{t,x}, e^{-t\Phi} \rangle = - \inf_{\mu \in \underline{\underline{M}}(\mathfrak{X})} (I(\mu) + \Phi(\mu))$$

avec $I(\mu) = \sup_{u \in \underline{\underline{D}}^+} \int \frac{-Lu}{u}(x)\mu(dx)$

Ces égalités sont obtenues par découpage de $\underline{\underline{M}}(\mathfrak{X})$ à partir des inégalités suivantes

dans le cas A

(1) Pour C fermé de $\underline{\underline{M}}(\mathfrak{X})$ $\quad \overline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(C) \leqq - \inf_{\mu\in C} I(\mu)$

(2) Pour G ouvert de $\underline{\underline{M}}(\mathfrak{X})$ $\quad \underline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(G) \geqq - \inf_{\mu\in G} I(\mu)$

dans le cas B

(3) Pour C fermé de $\underline{\underline{M}}(\mathfrak{X})$ $\quad \overline{\lim}_{t\to\infty} \frac{1}{t}\log Q_{t,x}(C) \leqq - \inf_{\mu\in C} I(\mu)$

(4) Pour G ouvert de $M(\mathfrak{X})$ $\quad \underline{\lim}_{t\to\infty} \frac{1}{t}\log Q_{t,x}(G) \geqq - \inf_{\mu\in G} I(\mu)$

Le but de ce travail est de simplifier quelques unes des démonstrations de Donsker et Varadhan, ce qui permet d'obtenir les formules (1) et (3) sous des hypothèses plus générales que dans [2], mais peut être moins aisées à utiliser, et une construction plus courte du processus ergodique qui est l'outil essentiel dans la démonstration des formules (2) et (4). Cette construction repose en fait sur les mêmes idées que dans [2].

On simplifie aussi le système d'hypothèses sous lequel est obtenue (4) en montrant que l'une des hypothèses (hypothèse (H_3)) est entraînée par le fait que le semi-groupe est fellérien. et qu'une autre (hypothèse (H_2)) est vérifiée dès que la topologie de $\mathfrak{X}$ est définie par le sous-espace B_{oo} de $\underline{\underline{C}}_b(\mathfrak{X})$. On a ajouté enfin un théorème de régularité sur les trajectoires des éléments de B_o .

I. Bornes supérieures dans les cas A et B

Soit Y un espace compact métrisable contenant $\mathfrak{X}$ comme sous-espace dense. Si f est une fonction sur $\mathfrak{X}$ à valeurs dans $\overline{\mathbb{R}}$, on désigne par $f^{\vee}$ la fonction sur Y définie par

$$f^{\vee}(y) = \liminf_{x \to y, x \in \mathfrak{X}} f(x)$$

La fonction $f^{\vee}$ est s.c.i. de Y dans $\overline{\mathbb{R}}$ et prolonge f si f est continue (ou même seulement s.c.i.) sur $\mathfrak{X}$.

On définit $\underline{\underline{U}}$ comme l'ensemble des $f \in \underline{\underline{C}}_b(\mathfrak{X})$ strictement positives à logarithme borné, et $\underline{\underline{D}}^{+} = \underline{\underline{D}} \cap \underline{\underline{U}}$.

<u>Dans le cas A</u> , $\underline{\underline{V}}$ est l'ensemble des $v \in \underline{\underline{C}}(Y)$ tels que la fonction φ_v définie sur $\mathfrak{X}$ par

$$\varphi_v(x) = \sup_n E^x[\exp(\sum_{i=0}^{n-1} v \circ X_i)]$$

soit localement bornée (E^x désignant comme d'habitude l'espérance pour la mesure P^x).

<u>Dans le cas B</u>, $\underline{\underline{V}}$ est l'ensemble des $v \in \underline{\underline{C}}(Y)$ tels que la fonction φ_v définie sur $\mathfrak{X}$ par

$$\varphi_v(x) = \sup_t E^x[\exp(\int_0^t v \circ X_s ds)]$$

soit localement bornée.

Dans les deux cas on pose, pour $\mu \in \underline{\underline{M}}(Y)$

$$J(\mu) = \sup_{v \in \underline{\underline{V}}} < \mu, v > \qquad \text{(on a } 0 \in \underline{\underline{V}}\text{, donc } J(\mu) \geqq 0 \text{)}$$

<u>Lemme 1. $\underline{\underline{V}}$ est une partie convexe de $\underline{\underline{C}}(Y)$.</u>

En effet, il résulte de la convexité de la fonction exponentielle que l'on a, pour $v, w \in \underline{\underline{V}}$ et $t \in [0,1]$

$$\varphi_{tv+(1-t)w} \leqq t\varphi_v + (1-t)\varphi_w$$

d'où le résultat.

<u>Lemme 2. La fonction J est convexe et s.c.i. sur $\underline{\underline{M}}(Y)$.</u>

En effet, J est l'enveloppe supérieure des fonctions affines continues $\mu \mapsto <\mu, v>$ sur $\underline{\underline{M}}(Y)$, v parcourant $\underline{\underline{V}}$.

<u>Lemme 3. Si μ est portée par $\mathfrak{X}$, on a $I(\mu) \leqq J(\mu)$.</u>

<u>Dans le cas A</u>. Soit $u \in \underline{\underline{U}}$. On démontre par récurrence sur n que

$$u(x) = E^x[u \circ X_n \exp(\sum_{i=0}^{n-1} \log \frac{u}{\pi u} \circ X_i)]$$

Il en résulte que pour toute $v \in \underline{\underline{C}}(Y)$, inférieure à $(\log \frac{u}{\pi u})^{\vee}$, on a

$$\varphi_v(x) \leqq u(x)/\alpha \quad \text{où} \quad \alpha = \inf_{x \in \mathfrak{X}} u(x)$$

donc que $v \in \underline{\underline{V}}$. La fonction $\log \frac{u}{\pi u} = h$ étant continue, $h^{\vee}$ prolonge h et l'on a si μ est portée par $\mathfrak{X}$

$$\int_{\mathfrak{X}} h(x)\mu(dx) = \int_Y h^\vee(y)\mu(dy) = \sup_{v\in\underline{\underline{C}}(Y),\, v\leq h^\vee} \int v(y)\mu(dy)$$
$$\leq \sup_{v\in\underline{\underline{V}}} \int v(y)\mu(dy) = J(\mu)$$

Donc $I(\mu)\leq J(\mu)$.

Dans le cas B. Soit $u\in\underline{\underline{D}}^+=\underline{\underline{D}}\cap\underline{\underline{U}}$. On a l'égalité

$$u(x) = E^x[u\circ X_t \exp\left(\int_0^t - \frac{Lu}{u}\circ X_s ds \right)] \tag{5}$$

Cette formule est prouvée dans [2] à partir de la formule de Feynman-Kac et de l'unicité de la solution d'une équation fonctionnelle. On peut aussi la montrer directement comme plus loin au lemme 7.

Il résulte de la formule (5) que si l'on désigne à nouveau par α la borne inférieure de u, par h la fonction $-Lu/u$, on a pour toute $v\in\underline{\underline{C}}(Y)$ inférieure à $h^\vee$

$$\varphi_v(x) \leq u(x)/\alpha$$

et donc $v\in\underline{\underline{V}}$. La même chaîne d'inégalités que plus haut montre alors que $\int_{\mathfrak{X}} h(x)\mu(dx) \leq \sup_{v\in\underline{\underline{V}}} \int_Y v(y)\mu(dy)$ si μ est portée par $\mathfrak{X}$, autrement dit, que $I(\mu)\leq J(\mu)$.

Théorème 4. Pour tout fermé Γ de $\underline{\underline{M}}(Y)$ on a

(6) Dans le cas A . $\overline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(\Gamma) \leq - \inf_{\mu\in\Gamma} J(\mu)$

(7) Dans le cas B . $\overline{\lim}_{t\to\infty} \frac{1}{t}\log Q_{t,x}(\Gamma) \leq - \inf_{\mu\in\Gamma} J(\mu)$

Soit $a < \inf_{\mu\in\Gamma} J(\mu)$. La semi-continuité de J entraîne l'existence d'un voisinage U de Γ dans $\underline{\underline{M}}(Y)$ tel que

$$a \leq \inf_{\mu\in U} J(\mu)$$

La compacité de Γ et la convexité locale de $\underline{\underline{M}}(Y)$ fournissent alors une suite finie $(\Gamma_1,\Gamma_2,\ldots,\Gamma_m)$ de parties convexes compactes de $\underline{\underline{M}}(Y)$ telles que

$$\Gamma \subset \bigcup_{j=1}^m \Gamma_j \subset U$$

On a alors $Q_{n,x}(\Gamma) \leq m.\sup_j Q_{n,x}(\Gamma_j)$, donc

$$\overline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(\Gamma) \leq \sup_j[\, \overline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(\Gamma_j)]$$

et d'autre part $- \inf_{\mu\in\Gamma_j} J(\mu) \geq - \inf_{\mu\in U} J(\mu) \geq -a$. Il nous suffit donc de prouver (6) pour chacun des Γ_j , autrement dit on peut se ramener au cas où Γ est compact convexe. Soit $v\in\underline{\underline{V}}$; on a pour tout (n,x)

$$\varphi_v(x) \geq E^x[\, \exp(\sum_0^{n-1} v\circ X_i)] = < Q_{n,x},\ e^{n<\cdot,v>} > \geq$$

$$\geqq Q_{n,x}(\Gamma)\ \exp(n.\inf_{\mu\in\Gamma} \langle\mu,v\rangle)$$

donc

$$\frac{1}{n}\log Q_{n,x}(\Gamma) \leqq \frac{1}{n}\log \varphi_v(x) - \inf_{\mu\in\Gamma} \langle\mu,v\rangle$$

Nous avons donc, uniformément sur tout compact de $\mathfrak{X}$ (puisque φ_v est borné sur tout compact de $\mathfrak{X}$)

$$\overline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(\Gamma) \leqq - \inf_{\mu\in\Gamma} \langle\mu,v\rangle$$

donc (en perdant peut être l'uniformité)

$$\overline{\lim}_{n\to\infty} \frac{1}{n}\log Q_{n,x}(\Gamma) \leqq - \sup_{v\in\underline{\underline{V}}} \inf_{\mu\in\Gamma} \langle \mu,v \rangle$$

$\underline{\underline{V}}$ étant convexe, Γ convexe compact, on peut intervertir le sup et l'inf grâce à un théorème de minimax, et en déduire la formule (6).

La démonstration de (7) est exactement la même.

<u>Théorème 5</u>. <u>Sous l'hypothèse</u>

(H_*) $\quad J(\mu) = +\infty$ <u>pour toute</u> $\mu\in\underline{\underline{M}}(Y)$ <u>qui charge</u> $Y\backslash\mathfrak{X}$

<u>on a pour tout fermé</u> C <u>de</u> $\underline{\underline{M}}(\mathfrak{X})$

<u>Dans le cas A</u> . $\qquad \overline{\lim}_{n\to\infty} \frac{1}{n} \log Q_{n,x}(C) \leqq - \inf_{\mu\in C} I(\mu)$

<u>Dans le cas B</u> . $\qquad \overline{\lim}_{n\to\infty} \frac{1}{t} \log Q_{t,x}(C) \leqq - \inf_{\mu\in C} I(\mu)$

<u>Démonstration</u>. Si l'on prend $\Gamma=\overline{C}$, on a $C=\Gamma\cap\underline{\underline{M}}(\mathfrak{X})$, donc $Q_{n,x}(C)=Q_{n,x}(\Gamma)$ (resp. $Q_{t,x}(C)=Q_{t,x}(\Gamma)$). On a d'après le lemme 3 et l'hypothèse H_*

$$\inf_{\mu\in C} I(\mu) \leqq \inf_{\mu\in C} J(\mu) \leqq \inf_{\mu\in\Gamma} J(\mu)$$

Les inégalités ci-dessus (c'est à dire les inégalités (1) et (3)) résultent alors des inégalités (6) et (7).

Les hypothèses (H^*) de [2] sont rappelées au cours de la démonstration.

<u>Lemme 6</u>. <u>Les hypothèses</u> (H^*) <u>de</u> [2] <u>entraînent</u> (H_*), <u>donc les inégalités</u> (1) <u>et</u> (3).

Dans le cas A, l'hypothèse H^* signifie qu'il existe $u \geqq 1$ telle que πu soit localement borné et que, pour tout $\sigma>0$, $\{ x \mid \frac{u}{\pi u}(x) \leqq \sigma \}$ soit compact. On voit comme dans le lemme 3 que pour toute $v\in\underline{\underline{C}}(Y)$ inférieure à h^v , où $h = \frac{u}{\pi u}$, on a pour tout x et tout n

$$u(x) \geqq E^x[\ u\circ X_n \exp(\sum_0^{n-1} v\circ X_i)] \geqq E^x[\exp(\sum_0^{n-1} v\circ X_i)]$$

donc en appliquant π

$$e^{v(x)}\pi u(x) \geqq e^{v(x)}E^x[\exp(\sum_1^{n} v\circ X_i)] = E^x[\exp(\sum_0^{n} v\circ X_i)]$$

et finalement $\varphi_v(x) \leq \pi u(x)e^{v(x)}$, donc $v\in\underline{\underline{V}}$. D'où comme dans le lemme 3 $J(\mu) \geq \int_Y h^\vee(y)\mu(dy)$, ce qui entraîne $J(\mu)=+\infty$ si $\mu(Y\backslash\mathfrak{X})>0$ puisque $h^\vee=+\infty$ sur $Y\backslash\mathfrak{X}$.

De même, dans le cas B l'hypothèse (H^*) s'énonce ainsi : il existe une suite (u_n) dans $\underline{\underline{D}}$ telle que $u_n \geq 1$, que $\sup_n u_n(x)$ soit borné sur tout compact, que $\sup_{n,x} \frac{Lu_n}{u_n}(x) < \infty$ et que la fonction définie par

$$h = \lim_{n\to\infty} - \frac{Lu_n}{u_n}$$

existe et soit telle que $\{ x \mid h(x) \leq \sigma$ soit compact pour tout σ (en particulier, h est s.c.i. sur $\mathfrak{X}$). Il résulte de la formule (5) que si $v\in\underline{\underline{C}}(Y)$ est inférieure à $h^\vee$, on a

$$\varphi_v(x) \leq \sup_n u_n(x)$$

donc que $v\in\underline{\underline{V}}$ et par suite que

$$J(\mu) \geq \sup_{\substack{v\leq h^\vee \\ v\in\underline{\underline{C}}(Y)}} \int v d\mu = \int h^\vee d\mu$$

et ceci entraîne que $J(\mu)=+\infty$ si μ charge $Y\backslash\mathfrak{X}$, puisque $h^\vee = +\infty$ sur $Y\backslash\mathfrak{X}$.

<u>Lemme 7</u>. <u>Pour</u> $x\in\mathfrak{X}$, $u\in\underline{\underline{D}}^+$ <u>et</u> $t\geq 0$, <u>on a</u>

$$(5) \qquad u(x) = E^x[\, u\circ X_t \exp(- \int_0^t \frac{Lu}{u}\circ X_s \, ds) \,] .$$

<u>Démonstration</u>. Désignons par $\psi(x,t)$ le second membre de (5). On a

$$\psi(x,t+h) = E^x[\, u\circ X_{t+h} \exp(-\int_0^t \frac{Lu}{u}\circ X_s ds) \exp(-\int_t^{t+h} \frac{Lu}{u}\circ X_s ds) \,]$$

Rappelons que Lu, u et 1/u sont bornées. Si nous pouvons montrer que

$$(*) \qquad \exp(-\int_t^{t+h} \frac{Lu}{u}\circ X_s \, ds) = \frac{u}{P_h u}\circ X_t + h\varepsilon(h) \qquad (0<h<1)$$

où $\varepsilon(h)$ désigne une fonction sur Ω qui tend vers 0 lorsque $h\to 0$ en restant bornée, P^x-p.s. pour tout x, nous aurons par convergence dominée

$$\psi(x,t+h) = E^x[u\circ X_{t+h} \exp(-\int_0^t \ldots)\frac{u}{P_h u}\circ X_t \,] + o(h) = \psi(x,t)+o(h)$$

puisque (P_s) est le semi-groupe de transition du processus (X_t), donc $\frac{\partial\psi}{\partial t} = 0$ et $\psi(x,t) = \psi(x,0) = u(x)$, ce qui achèvera la démonstration.

Démontrons (*). D'une part, puisque $u\in\underline{\underline{D}}$, on a $u\in B_0$ et $Lu\in B_0$, où B_0 désigne l'ensemble des $f\in\underline{\underline{C}}_b(\mathfrak{X})$ tels que $\lim_{h\to 0} \|P_h f-f\|=0$. Il en résulte, d'après le théorème 13 ci-dessous, que les processus $(u\circ X_s)$ et $(Lu\circ X_s)$ admettent des modifications continues à droite (w_s) et (z_s), donc que presque sûrement pour toute loi initiale

$$\frac{1}{h}\int_t^{t+h} \frac{Lu}{u}\circ X_s\, ds = \frac{1}{h}\int_t^{t+h} \frac{z_s}{w_s}\, ds \rightarrow \frac{z_t}{w_t} = \frac{Lu}{u}\circ X_t$$

le côté gauche restant, par ailleurs, borné. On peut donc écrire

$$\begin{aligned}\exp(-\int_t^{t+h}\ldots) &= \exp(\,-h\frac{Lu}{u}\circ X_t + h\varepsilon(h))\\ &= 1-h\frac{Lu}{u}\circ X_t + h\varepsilon(h)\end{aligned}$$

où la fonction $\varepsilon(h)$ varie de place en place, mais tend vers 0 en restant bornée lorsque $h\to 0$. D'autre part, on a

$$\frac{1}{h}(P_h u-u) = \frac{1}{h}\int_0^h P_s Lu\, ds = Lu + \varepsilon(h)$$

et par conséquent

$$\frac{P_h u}{u} = 1+ h\frac{Lu}{u} + h\varepsilon(h)\ , \qquad \frac{u}{P_h u} = 1-h\frac{Lu}{u} + h\varepsilon(h)$$

et finalement $\exp(-\int_t^{t+h}\ldots) = \frac{u}{P_h u}\circ X_t + h\varepsilon(h)$, la formule cherchée.

II . Construction du processus ergodique $\bar{P}$.

La démonstration des inégalités (2) et (4) est beaucoup plus difficile que celle de (1) et (3). Au lieu d'avoir deux démonstrations parallèles pour les cas A et B, on procède par réduction du cas continu au cas discret. Dans ce dernier cas, l'étape essentielle est la construction d'un certain processus ergodique, que l'on va présenter ici avec une démonstration un peu moins générale, mais plus simple que dans [2].

Définissons, comme dans [2], $\underline{\underline{U}}_1$ (resp. $\underline{\underline{U}}_2$) comme l'ensemble des fonctions boréliennes strictement positives sur $\mathfrak{X}$ (resp. $\mathfrak{X}\times\mathfrak{X}$) dont le logarithme est borné, posons pour $\mu\in\underline{\underline{M}}(\mathfrak{X})$, qui reste fixée ci-dessous

$$I(\mu) = \sup_{u\in\underline{\underline{U}}_1} \int \log \frac{u}{\pi u}(x)\ \mu(dx)$$

Définissons la mesure α sur $\mathfrak{X}\times\mathfrak{X}$ par

$$\alpha(dx,dy) = \mu(dx)\pi(x,dy)$$

Et enfin, pour $\lambda\in\underline{\underline{M}}(\mathfrak{X}\times\mathfrak{X})$

$$(*)\qquad \bar{I}(\lambda) = \sup_{u\in\underline{\underline{U}}_2} \left[\int\log u(x,y)\ \lambda(dx,dy) - \log \int u(x,y)\alpha(dx,dy)\right]$$

On démontre dans [1] que $\bar{I}(\lambda)$ vaut $+\infty$ si λ n'est pas absolument continue par rapport à α , et vaut dans le cas contraire

$$\int(\log \frac{d\lambda}{d\alpha})d\lambda = \int \frac{d\lambda}{d\alpha}(\log \frac{d\lambda}{d\alpha}\ )d\alpha$$

Nous désignerons par λ^1 et λ^2 les projections de λ, et par $\underline{\underline{M}}_\mu$ l'ensemble des mesures λ sur $\mathfrak{X}\times\mathfrak{X}$ telles que $\lambda^1=\lambda^2=\mu$

<u>Lemme 8</u>. <u>On a</u> $I(\mu) = \inf_{\lambda\in\underline{\underline{M}}_\mu} \bar{I}(\lambda)$.

On voit d'abord, comme dans [2], que si $g \in \underline{\underline{U}}_1$ et dans (*)

$$u(x,y) = \frac{g(y)}{\pi g(x)} \in \underline{\underline{U}}_2$$

on a pour tout λ dans $\underline{\underline{M}}_\mu$

$$\bar{I}(\lambda) \geqq \int \log g(y).\lambda^2(dy) - \int \log \pi g(x).\lambda^1(dx) - \log\left[\int \frac{g(y)}{\pi g(x)}\pi(x,dy)\mu(dx)\right]$$

$$\geqq \int \log \frac{g(y)}{\pi g(x)}\,\mu(dx)$$

d'où $\bar{I}(\lambda) \geqq I(\mu)$ et enfin $I(\mu) \leqq \inf_{\lambda \in \underline{\underline{M}}_\mu} \bar{I}(\lambda)$.

Inversement, supposons $I(\mu) < \inf_{\lambda \in \underline{\underline{M}}_\mu} \bar{I}(\lambda)$, et montrons que cela conduit à une contradiction.

L'énoncé du lemme ne dépendant que de la structure borélienne de X, qui est standard, on peut supposer X compact métrisable. $\bar{I}(\lambda)$ peut alors être calculé en prenant la borne supérieure sur $\underline{\underline{U}}_2 \cap \underline{\underline{C}}(\mathfrak{X}\times\mathfrak{X})$, et $\bar{I}$ est donc convexe s.c.i. sur le convexe compact $\underline{\underline{M}}(\mathfrak{X}\times\mathfrak{X})$. Si j est tel que $I(\mu)<j<\inf_{\lambda \in \underline{\underline{M}}_\mu} \bar{I}(\lambda)$, l'ensemble

$$\Lambda = \{\ \lambda \in \underline{\underline{M}}(\mathfrak{X}\times\mathfrak{X}) \mid \bar{I}(\lambda) \leqq j\ \}$$

est un convexe compact disjoint de $\underline{\underline{M}}_\mu$, et si p désigne l'application affine continue de $\underline{\underline{M}}(\mathfrak{X}\times\mathfrak{X})$ dans $\underline{\underline{M}}(\mathfrak{X})\times\underline{\underline{M}}(\mathfrak{X})$ qui à une mesure associe ses deux projections, $p(\Lambda)$ est un convexe compact qui ne contient pas (μ,μ). Il existe donc, d'après le théorème de Hahn-Banach, deux fonctions u_1 et v continues sur $\mathfrak{X}$ telles que

$$\begin{cases} <\mu,u_1> - <\mu,v> \leqq 0 \\ <\lambda^1,u_1> - <\lambda^2,v> \geqq 1 \text{ si } \lambda\in\Lambda \end{cases}$$

La constante $<\mu,v-u_1>$ est positive. Posant $u=u_1+<\mu,v-u_1>$ on a

(8) $$<\mu,u> = <\mu,v>$$

(9) $$<\lambda^1,u> - <\lambda^2,v> \geqq 1 \quad \text{si } \lambda\in\Lambda\ .$$

Si l'on pose alors $M(\theta) = \int e^{\theta[u(x)-v(y)]}\alpha(dx,dy)$, $\log M$ est une fonction différentiable et convexe s'annulant en 0. Soit maintenant $\eta>0$. On va montrer l'existence d'un $\theta_0 \in \mathbb{R}$ tel que $\log M(\theta_0) \leqq -j+\eta$. Prenant alors $f(y) = e^{-\theta_0 v(y)}$ et $g(x)=e^{-\theta_0 u(x)}$ on aura, en utilisant la définition de $I(\mu)$, la formule (8) et l'inégalité de Jensen

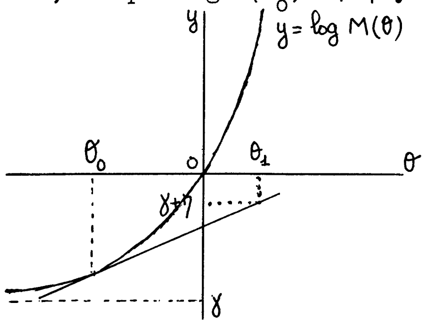
$$-I(\mu) \leqq \int \log \frac{\pi f(x)}{f(x)}\,\mu(dx) = \int \log \frac{\pi f(x)}{g(x)}\mu(dx) \leqq \log \int \frac{\pi f(x)}{g(x)}\,\mu(dx)$$

$$\leqq \log \int \frac{f(y)}{g(x)}\,\pi(x,dy)\mu(dx) = \log M(\theta_0) \leqq \eta - j$$

d'où la contradiction cherchée si $0<\eta<j-I(\mu)$.

L'existence de θ_0 est immédiate si $\gamma = \inf_\theta \log M(\theta) = -\infty$. Si $\gamma > -\infty$, nous allons construire θ_0 tel que $|\frac{M'}{M}(\theta_0)|<1$ et $\theta_0.|\frac{M'}{M}(\theta_0)| \leqq \eta$, après quoi nous montrerons qu'il répond à la question.

Si M' s'annule en x, il suffit de prendre $\theta_0 = x$. Si M' ne s'annule pas, il reste de signe constant, et on a $\log M(\theta) < 0$ si $\theta M'(0)<0$, d'où $\gamma < 0$. En choisissant $\eta < -\gamma$, il existe un θ_0 tel que la tangente en θ_0 à la courbe $y=\log M(\theta)$ passe par le point $(\theta_1, \gamma+\eta)$, où $|\theta_1|>\eta$ et $\theta_1 M'(0) >0$, et que $\log M(\theta_0)< \gamma+\eta$. On a alors $\theta_0 - \theta_1 < 0$ puisque $\log M(\theta_0)<0$,



et

$$\eta \geqq (\gamma-\eta) - \log M(\theta_0) = (\theta_1-\theta_0)\frac{M'}{M}(\theta_0) > 0$$

d'où

$$(|\theta_1|+|\theta_0|)|\frac{M'}{M}(\theta_0)| \leqq \eta$$

et

$$\begin{cases} |\theta_0||\frac{M'}{M}(\theta_0)| \leqq \eta \\ |\frac{M'}{M}(\theta_0)| \leqq \frac{\eta}{|\theta_1|} < 1 \end{cases}$$

les conditions cherchées. Il nous reste à vérifier que $\log M(\theta_0) \leqq -j+\eta$, ou encore à trouver une mesure $\lambda_0 \in \underline{\underline{M}}(\mathfrak{X}\times\mathfrak{X})$ telle que $\overline{I}(\lambda_0) \leqq \eta - \log M(\theta_0)$ et que $j \leqq \overline{I}(\lambda_0)$, autrement dit que $\lambda_0 \notin \Lambda$. Nous prendrons

$$\lambda_0(dx,dy) = \frac{1}{M(\theta_0)} e^{\theta_0[u(x)-v(y)]}\alpha(dx,dy)$$

de sorte que

$$\int(u(x)-v(y))\lambda_0(dx,dy) = \frac{M'}{M}(\theta_0) < 1$$

et que $\lambda_0 \notin \Lambda$ d'après (9). On a aussi

$$\overline{I}(\lambda_0) = \int \log (\frac{d\lambda_0}{d\alpha})\, d\lambda_0 = \theta_0\frac{M'}{M}(\theta_0) - \log M(\theta_0) \leqq \eta - \log M(\theta_0) \ .$$

Cela achève la démonstration du lemme 8. La démonstration a une autre conséquence : puisque $\overline{I}$ est s.c.i. et que $\underline{\underline{M}}_\mu$ est compact, il existe une mesure $\overline{\lambda}$ dans $\underline{\underline{M}}_\mu$ telle que $\overline{I}(\overline{\lambda}) = I(\mu)$. Puisque $\overline{\lambda}\in\underline{\underline{M}}_\mu$, on trouve par désintégration une probabilité de transition $\overline{\pi}(x,dy)$ telle que

$$\overline{\lambda}(dx,dy) = \mu(dx)\overline{\pi}(x,dy)$$

et l'on a , puisque $\overline{\lambda}^2 = \mu$, que μ est invariante par $\overline{\pi}$.

<u>Lemme</u> 9. <u>Il existe deux fonctions mesurables positives</u> a <u>et</u> b <u>telles que</u>

$$\overline{\lambda}(dx,dy) = a(x)b(y)\alpha(dx,dy) \ .$$

Il existe, pour tout $n\in\mathbf{N}$, une f_n positive dans $\underline{\underline{C}}(\mathfrak{X})$ telle que

$$\int \log \left(\frac{f_n(x)}{\pi f_n(x)} \right) \mu(dx) \geqq I(\mu)-2^{-n} = \overline{I}(\overline{\lambda})-2^{-n} \ .$$

Si l'on pose alors $\nu_n(dx,dy) = \frac{f_n(y)}{\pi f_n(x)}\,\alpha(dx,dy) \in \underline{\underline{M}}(\mathfrak{X}\times\mathfrak{X})$, on a $\overline{\lambda}\ll\alpha\approx\nu_n$, donc avec $\varphi_n = \frac{d\overline{\lambda}}{d\nu_n}$

$$\int\varphi_n\log(\varphi_n)d\nu_n = \int\log(\frac{d\overline{\lambda}}{d\nu_n})d\overline{\lambda} = \int\log(\frac{d\overline{\lambda}}{d\alpha})d\overline{\lambda} - \int\log(\frac{d\nu_n}{d\alpha})d\overline{\lambda}$$
$$= \overline{I}(\overline{\lambda}) - \int\log(\frac{f_n(x)}{\pi f_n(x)})\,\mu(dx) \leqq 2^{-n}$$

Il existe une constante $c>0$ telle que pour tout $t\geqq 0$ on ait

$$t\log t - (t-1) \geqq c\,\frac{(t-1)^2}{1+t}$$

On en tire

$$c\int\frac{(\varphi_n-1)^2}{1+\varphi_n}d\nu_n \leqq \int\varphi_n\log(\varphi_n)d\nu_n - \int\frac{d\overline{\lambda}}{d\nu_n}d\nu_n + 1 \leqq 2^{-n}-1+1 = 2^{-n}$$

et par l'inégalité de Schwarz appliquée à $|\varphi_n-1|/\sqrt{1+\varphi_n}$ et $\sqrt{1+\varphi_n}$

$$\|\overline{\lambda}-\nu_n\|^2 = [\int|\varphi_n-1|d\nu_n]^2 \leqq [\int\frac{(\varphi_n-1)^2}{1+\varphi_n}d\nu_n][\int(1+\varphi_n)d\nu_n] \leqq 2^{-n}\times(1+1) = 2^{1-n}$$

D'où $\|\frac{d\overline{\lambda}}{d\alpha} - \frac{d\nu_n}{d\alpha}\|_{L^1(\alpha)} \longrightarrow 0$. Il en résulte que $\frac{d\overline{\lambda}}{d\alpha}$ est adhérent dans $L^1(\alpha)$ à l'ensemble Δ des éléments $\theta(x,y)$ qui vérifient pour $\alpha\times\alpha$ presque tout (x,y,x',y')

$$\theta(x,y)\theta(x',y') = \theta(x,y')\theta(x',y)$$

Or cet ensemble est fermé dans $L^1(\alpha)$, et on conclut en remarquant que si $\theta\in\Delta$ n'est pas nul, il existe au moins un (x',y') tel que l'on ait pour α-presque tout (x,y)

$$\theta(x,y) = \frac{\theta(x,y')\theta(x',y)}{\theta(x',y')}$$

<u>Théorème 10.</u> <u>S'il existe une mesure</u> β <u>sur</u> $\mathfrak{X}$ <u>telle que, pour tout</u> x <u>de</u> $\mathfrak{X}$, $\pi(x,dy)$ <u>soit équivalente à</u> $\beta(dy)$, <u>et si</u> $I(\mu)<+\infty$, <u>il existe une fonction mesurable positive</u> h <u>sur</u> $\mathfrak{X}\times\mathfrak{X}$ <u>et une chaîne de Markov stationnaire et ergodique de loi initiale</u> μ <u>et de probabilité de transition</u> $\overline{\pi}(x,dy)$, <u>avec</u>

$$\overline{\pi}(x,dy) = h(x,y)\pi(x,dy)\ ;\ \int\log h(x,y)\mu(dx)\overline{\pi}(x,dy) = I(\mu).$$

Cet énoncé est un peu moins général que celui du théorème 2.6 de [2], mais suffit partout après dans la suite de l'article. La démonstration qui suit est essentiellement celle de [2].

Puisque $\overline{\lambda} \ll \alpha \approx \mu\times\beta$, on a $\mu=\overline{\lambda}^2 \ll \beta$. Si $\overline{\pi}(x,dy)$ est la probabilité de transition construite juste avant le lemme 9, nous savons que μ est invariante par $\overline{\pi}$. Nous savons aussi que $h(x,y) = \frac{d\overline{\lambda}}{d\alpha} = a(x)b(y)$, et que $\alpha(dx,dy) = \mu(dx)\pi(x,dy)$. Comme $\overline{\pi}$ n'est définie qu'à un ensemble μ-négligeable près, nous pouvons prendre pour tout x

$$\overline{\pi}(x,dy) = h(x,y)\pi(x,dy).$$

Soit (X_n) le processus de Markov admettant la probabilité de transition $\bar{\pi}$ et la loi initiale invariante μ. Si ce processus n'est pas ergodique, il existe une partie A de $\mathfrak{X}$ telle que $0<\mu(A)<1$ et

$$X_0 \in A \Rightarrow X_n \in A \text{ pour tout } n\ , \quad X_0 \in A^c \Rightarrow X_n \in A^c \text{ pour tout } n.$$

Comme $\bar{\lambda}$ est la loi du couple (X_0,X_1), on a alors

$$\bar{\lambda}(A\times A^c)=\bar{\lambda}(A^c\times A)=0$$

et nous allons montrer que cela entraîne $\mu(A)=0$ ou 1, d'où la conclusion. Posons $E=\{x \mid a(x)>0\}$ et $F=\{y \mid b(y)>0\}$; les relations précédentes entraînent

$$\alpha[(A\cap E)\times(A^c\cap F)] = \alpha[(A^c\cap E)\times(A\cap F)]=0$$

et, puisque $\alpha \approx \mu\times\beta$

$$\mu(A\cap E)\beta(A^c\cap F) = \mu(A^c\cap E)\beta(A\cap F) = 0$$

De plus, $\bar{\lambda}(E^c\times\mathfrak{X}) = 1-\mu(E) = \int_{E^c\times\mathfrak{X}} a(x)b(y)\alpha(dx,dy)=0$

et $\quad \bar{\lambda}(\mathfrak{X}\times F^c) = 1-\mu(F) = \int_{\mathfrak{X}\times F^c} a(x)b(y)\alpha(dx,dy)=0$

Donc $\mu(E)=\mu(F)=1$, et $\beta(F)\neq 0$ puisque $\mu\ll\beta$. Il en résulte que l'on a, soit $\beta(A^c\cap F)\neq 0$ soit $\beta(A\cap F)\neq 0$, donc soit $\mu(A)=\mu(A\cap E)=0$, soit $\mu(A^c)=\mu(A^c\cap E)=0$. On en conclut que $\mu(A)=0$ ou 1 et la démonstration est finie.

III. Remarque sur l'hypothèse (H_3).

On note B_o le centre fortement continu du semi-groupe, c'est à dire l'ensemble des $f\in \underline{\underline{C}}_b(\mathfrak{X})$ telles que $\lim_{h\to 0} \|P_hf-f\| = 0$. On rappelle que (P_t) est un semi-groupe de Feller, c'est à dire markovien et vérifiant

$$\forall t\geq 0 \quad \forall f\in\underline{\underline{C}}_b(\mathfrak{X}) \qquad P_tf \in \underline{\underline{C}}_b(\mathfrak{X})$$
$$\forall x\in\mathfrak{X} \quad \forall f\in\underline{\underline{C}}_b(\mathfrak{X}) \qquad f(x) = \lim_{t\to 0} P_tf(x)$$

Nous allons démontrer que ces propriétés entraînent l'hypothèse (H_3) de [2], p.445 , c'est à dire

Pour toute mesure $\mu\in\underline{\underline{M}}(\mathfrak{X})$ et toute f mesurable bornée, il existe des $f_n\in B_o$ telles que $\|f_n\|\leq\|f\|$ et $\lim_n f_n=f$ μ-p.p..

<u>Démonstration</u>. Soit U_p l'opérateur $\int_0^\infty e^{-pt}P_tdt$ de la résolvante $(p>0)$
Notons les faits suivants, bien connus :

- Si g est mesurable bornée, on a $U_pg\in B_o$.
- Si g est continue bornée, $pU_pg \to g$ lorsque $p\to\infty$ en restant majoré par $\|g\|$, donc $pU_pg\to g$ dans $L^1(\mu)$.

Soit alors f μ-mesurable bornée. Il existe pour tout n une $g_n\in\underline{\underline{C}}_b(\mathfrak{X})$ telle que

$$\|g_n\| \leq \|f\| \quad , \int |f-f_n| d\mu \leq 2^{-n}$$

Soit $f_n = p_n U_{p_n} g_n$, où p_n est choisi assez grand pour que $\int |g_n - f_n| d\mu \leq 2^{-n}$.
On a alors

$$f_n \in B_o \, , \quad \|f_n\| \leq \|f\| \quad , \quad \Sigma_n \int |f-f_n| d\mu \leq 2\Sigma_n \, 2^{-n} < +\infty$$

et cette dernière propriété entraîne que $f_n \to f$ μ-presque partout.

IV. Remarque sur B_{oo} et sur l'hypothèse (H_2).

On note B_{oo} le sous-espace de $\underline{\underline{C}}_b(\mathfrak{X})$ formé des fonctions f telles que

$$\lim_{t \to 0} \sup_{x \in \mathfrak{X}} \int |f(y)-f(x)| P_t(x,dy) = 0$$

où $P_t(x,dy)$ est la probabilité de transition qui correspond au semi-groupe. On remarque dans [2] que B_{oo} est un sous-espace vectoriel fermé de B_o , et que si f et f^2 sont dans B_o , alors f est dans B_{oo} . Cette propriété résulte immédiatement de l'inégalité de Schwarz. On peut en fait prouver le résultat plus fort :

<u>Proposition 11</u> . <u>Si</u> $f \in B_{oo}$ <u>et si</u> $g \in B_o$, <u>le produit</u> fg <u>est dans</u> B_o .
B_{oo} <u>est la plus grande sous-algèbre de</u> $\underline{\underline{C}}_b(\mathfrak{X})$ <u>contenue dans</u> B_o, <u>et l'on a</u>

(10) $\qquad f \in B_{oo} \iff (f \in B_o \text{ et } f^2 \in B_o)$.

<u>Démonstration</u> : On écrit

$$\begin{aligned} |(P_t fg - fg)(x)| &= \left|\int (fg(y)-fg(x)) P_t(x,dy)\right| \\ &\leq \left|\int (f(y)-f(x)) g(y) P_t(x,dy)\right| + |f(x)| \cdot \left|\int (g(y)-g(x)) P_t(x,dy)\right| \\ &\leq \|g\| \int |f(y)-f(x)| P_t(x,dy) + \|f\| \, \|P_t g - g\| \end{aligned}$$

Il en résulte que si $f \in B_{oo}$, $g \in B_o$, alors fg appartient à B_o .

L'implication $\Rightarrow$ de (10) est alors immédiate en prenant f=g ; l'implication $\Leftarrow$ de (10) est établie dans [2], et il en résulte, comme $fg = \frac{1}{2}((f+g)^2 - f^2 - g^2)$, que B_{oo} est une sous-algèbre de $\underline{\underline{C}}_b(\mathfrak{X})$ contenue dans B_o, la plus grande d'après (10).

On va montrer que si B_{oo} définit la topologie de $\mathfrak{X}$, l'hypothèse (H_2) de [2] est vérifiée . Explicitement

<u>Théorème 12</u>. <u>Si</u> B_{oo} <u>définit la topologie de</u> $\mathfrak{X}$, <u>la topologie étroite sur</u> $\underline{\underline{M}}(\mathfrak{X})$ <u>coïncide avec la topologie</u> $\sigma(\underline{\underline{M}}(\mathfrak{X}), B_{oo})$ <u>de la convergence sur</u> B_{oo}.

<u>Démonstration</u> . Puisque B_{oo} définit la topologie de $\mathfrak{X}$, $\mathfrak{X}$ est un sous-espace du spectre $\mathfrak{S}$ de l'algèbre B_{oo} . L'espace $\mathfrak{S}$ est une compactification de $\mathfrak{X}$, et $\overline{B_{oo}}$ s'identifie à l'espace des fonctions continues sur $\mathfrak{S}$.
La topologie $\sigma(\underline{\underline{M}}(\mathfrak{X}), B_{oo})$ est donc la topologie induite sur $\underline{\underline{M}}(\mathfrak{X})$ par la

topologie étroite de $\underline{\underline{M}}(\mathfrak{S})$. Il est alors bien connu que cette topologie coïncide avec la topologie étroite de $\underline{\underline{M}}(\mathfrak{X})$ (Bourbaki, Intégr. chap.IX, §5, n°3, prop.8).

V. Continuité à droite sur les trajectoires.

Il est aisé de voir que si $\mathfrak{X}$ n'est pas localement compact, il peut se faire que le processus associé à un semi-groupe de Feller n'admette pas de modification continue à droite. Il suffit pour cela de restreindre l'espace d'états du mouvement brownien sur $\mathbb{R}$ à l'espace polonais $\mathbb{R}\backslash\mathbb{Q}$ qui porte les probabilités de transition, alors que presque toute trajectoire traverse $\mathbb{Q}$ sur tout intervalle de temps de longueur non nulle.

Néanmoins, une transposition de la démonstration de P.A.Meyer (cf. [3]) permet de prouver :

<u>Théorème 13</u> . <u>Si</u> $f\in B_0$, <u>le processus</u> $(f\circ X_t)$ <u>admet une modification continue à droite</u>.

En ajoutant une constante à f , on peut supposer $f\geqq 0$. Pour tout $p>0$, le processus $(e^{-ps}U_p f\circ X_s)$, où U_p est l'opérateur $\int_0^\infty e^{-pt}P_t dt$ de la résolvante, est une surmartingale, donc possède presque sûrement des limites à droite suivant $\mathbb{Q}$ en tout $s\geqq 0$.

Comme $f\in B_0$, $pU_p f$ converge uniformément vers f sur $\mathfrak{X}$ lorsque $p\to\infty$. Le processus $(f\circ X_t)$ a donc aussi presque sûrement des limites à droite suivant $\mathbb{Q}$ en tout $t\geqq 0$. En définissant, pour presque tout $\omega\in\Omega$, le processus (g_t) par

$$g_t(\omega) = \lim_{s\to t_+ ,\ s\in\mathbb{Q}} f\circ X_s(\omega)$$

on obtient un processus continu à droite, et il reste à montrer que pour toute loi initiale μ et tout t, on a $g_t=f\circ X_t$ presque sûrement pour P^μ. Or on a , les espérances E étant relatives à la loi P^μ

$$E[g_t^2] = \lim_{s\to t_+ ,s\in\mathbb{Q}} E[f^2\circ X_s] = \lim_{h\to 0} E[P_h f^2\circ X_t] = E[f^2\circ X_t]$$

puisque $P_h f^2\to f^2$ simplement. De même

$$E[f\circ X_t . g_t] = \lim_{s\to t_+ ,s\in\mathbb{Q}} E[f\circ X_t . f\circ X_s] = \lim_{h\to 0} [f\circ X_t . P_h g\circ X_s]$$
$$= E[f^2\circ X_t]$$

et enfin

$$E[(f\circ X_t - g_t)^2] = E[f^2\circ X_t]-2E[f\circ X_t . g_t]+E[g_t^2] = 0 .$$

REFERENCES

[1] M.D. Donsker et S.R.S. Varadhan - Asymptotic Evaluations... I . Comm. Pure Appl. Math. 28, 1975, p.1-47.

[2] M.D. Donsker et S.R.S. Varadhan - Asymptotic Evaluations... III. Comm. Pure Appl. Math. 29, 1976, p. 389-461.

[3] P.A. Meyer. Processus de Markov. Lect. Notes in M. 26, 1967. Springer-Verlag.

J. Saint Raymond
Université Paris VI
Mathématiques, Tour 46
4 Place Jussieu
75230 Paris Cedex 05