ANNALES DE L'I. H. P., SECTION B

D. REVUZ

Lois du tout ou rien et comportement asymptotique pour les probabilités de transition des processus de Markov

Annales de l'I. H. P., section B, tome 19, nº 1 (1983), p. 9-24

<http://www.numdam.org/item?id=AIHPB_1983__19_1_9_0>

Ann. Inst. Henri Poincaré,
Vol. XIX, n° 1, 1983, p. 9-24.

Section B :
Calcul des Probabilités et Statistique.

Lois du tout ou rien et comportement asymptotique pour les probabilités de transition des processus de Markov

par

D. REVUZ
Université Paris-VII[e], U. E. R. de Mathématiques,
2, Place Jussieu, 75221 Paris Cedex 05

INTRODUCTION

Pour les chaînes de Markov récurrentes au sens de Harris l'étude des classes cycliques (Orey [*11*]) et le théorème de Jamison et Orey ([*8*]) donnent une description satisfaisante du comportement asymptotique des itérés de la probabilité de transition. Dans [*14*] nous avons montré comment on peut faire un exposé assez ramassé de ces questions en utilisant l'exposé de Derriennic [*4*] sur les lois du tout ou rien d'Ornstein-Sucheston.

Notre propos dans cet article est d'essayer d'étendre ces méthodes et ces résultats aux processus de Markov à temps continu. Dans [*5*] nous avions donné une condition suffisante dans le cas des processus récurrents pour que l'on ait

$$\lim_{t \to \infty} \| \nu_1 P_t - \nu_2 P_t \| = 0,$$

pour tout couple (ν_1, ν_2) de probabilités initiales. Un de nos buts est de clarifier cette question en énonçant d'autres propriétés équivalentes et en donnant une autre démonstration. Pour cela nous commencerons par énoncer les résultats de Derriennic [*4*] pour le temps continu.

Je remercie P. Bougerol de m'avoir signalé une erreur dans une première version de cet article.

I. NOTATIONS, HYPOTHÈSES ET PRÉLIMINAIRES

Dans tout cet article, nous allons considérer un espace E lusinien de tribu borélienne $\mathscr{E}$. L'objet de notre étude est le comportement asymptotique d'un semi-groupe markovien de noyaux boréliens sur $(E, \mathscr{E})$, c'est-à-dire, d'une famille P_t, $t \geq 0$, de noyaux sur $(E, \mathscr{E})$ tels que :

i) $P_0(x, .) = \varepsilon_x$ pour tout x de E ;

ii) $P_t(x, E) = 1$ pour tout $t \geq 0$ et tout x de E ;

iii) Pour tout x de E, tout $A \in \mathscr{E}$ et $s, t \geq 0$,

$$\int_E P_t(x, dy)P_s(y, A) = P_{s+t}(x, A).$$

Bien que nous n'allions pas utiliser toute la force de cette hypothèse, nous allons par commodité supposer que nous avons affaire à un *semi-groupe droit*, c'est-à-dire qu'il existe un processus de Markov *droit* X qui l'admet pour semi-groupe de transition. Nous renvoyons à [*6*] (page 50) pour les notions employées ci-dessous, mais nous allons ci-dessous faire quelques rappels pour préciser les notations qui nous seront les plus utiles.

Le processus X est défini sur l'espace Ω des applications continues à droite de $\mathbb{R}_+$ dans E. On appelle X_t les applications coordonnées. Comme d'habitude, on note $\mathscr{F}^0$ (resp. $\mathscr{F}_t^0$) la tribu sur Ω engendrée par les applications $(X_t, t \geq 0)$ (resp. $(X_s, 0 \leq s \leq t)$). Pour toute probabilité ν sur $(E, \mathscr{E})$, il existe une probabilité P_ν sur $(\Omega, \mathscr{F}^0)$ telle que $(X_t, \mathscr{F}_t^0, P_\nu)$ soit un processus de Markov de mesure initiale ν et de probabilité de transition P_t ; en d'autres termes

i) $P_\nu[X_0 \in A] = \nu(A), A \in \mathscr{E}$;

ii) $E_\nu[f(X_{t+s}) \mid \mathscr{F}_t^0] = P_t f(X_s)$ pour f borélienne bornée.

On utilisera en fait comme d'habitude les tribus $\mathscr{F}$ et $\mathscr{F}_t$ définies comme suit. On appelle $\mathscr{F}^\mu$ la complétée de $\mathscr{F}^0$ par rapport à P_μ et $\mathscr{F}_t^\mu$ la tribu engendrée par $\mathscr{F}_t^0$ et les ensembles de P_μ-mesure nulle dans $\mathscr{F}^\mu$. Puis on pose $\mathscr{F} = \cap \mathscr{F}^\mu$ et $\mathscr{F}_t^\mu = \cap \mathscr{F}_t$ où dans les deux cas, l'intersection est prise par rapport à l'ensemble de toutes les probabilités μ sur $(E, \mathscr{E})$. La propriété de Markov (*ii*) ci-dessus) reste vraie si on remplace $\mathscr{F}_t^\mu$ par $\mathscr{F}_t$. On peut de plus l'exprimer en utilisant l'opérateur de translation θ_t défini comme d'habitude par $X_s \circ \theta_t = X_{s+t}$. Si Z est $\mathscr{F}$-mesurable et positive, pour toute probabilité ν sur $(E, \mathscr{E})$ on a

$$E_\nu[Z \circ \theta_t \mid \mathscr{F}_t] = E_{X_t}[Z]P_\nu - ps.$$

On rappelle que E_ν est l'opérateur espérance mathématique associé à P_ν, que E_x est celui associé à P_x et que si Z est $\mathscr{F}$-mesurable et bornée $x \to E_x[Z]$ est universellement mesurable sur E et

$$E_\nu[Z] = \int_E \nu(dx) E_x[Z].$$

Faisons observer aussi que notre semi-groupe ayant été supposé markovien, les formules ci-dessus n'ont pas à tenir compte du temps de mort. Ceci simplifiera certains des énoncés qui vont suivre.

Nous aurons aussi à considérer le processus espace-temps $(X_t, t + s)$ à valeur dans $E \times \mathbb{R}_+$. Pour P_x ce processus part du point (x, s). Une fonction universellement mesurable f est dite *invariante* (resp. E.T-*invariante*) si

$$P_t f = f, \forall t \geq 0 \left(\text{resp} \int P_t(x, dy) f(y, t + s) = f(x, s) \forall t, s \geq 0 \right).$$

Il est classique que les processus $f(X_t)$ (resp. $f(X_t, t + s)$) sont des martingales par rapport aux tribus $\mathscr{F}_t$ pour toute probabilité P_ν sur $(\Omega, \mathscr{F})$. Il résulte aussi des hypothèses droites que ces martingales sont p. s. continues à droite, et que les fonctions f correspondantes sont *presque-boréliennes.*

Nous allons maintenant introduire des définitions et des résultats qui sont banaux pour les chaînes de Markov (cf. [*13*]) mais que nous n'avons pas trouvés explicités dans le cas des processus.

DÉFINITION 1.1. — On appelle *tribu invariante* la tribu J des événements $\Gamma \in \mathscr{F}$ tels que $\Gamma = \theta_t^{-1}(\Gamma)$ pour tout $t \geq 0$. Un tel événement est dit invariant. Une variable aléatoire Z est dite invariante si elle est J-mesurable c'est-à-dire si $Z = Z \circ \theta_t$ pour tout $t \geq 0$. Deux variables invariantes Z et Z′ sont dites équivalentes si $Z = Z'$ p. s., c'est-à-dire si $P_\nu[Z \neq Z'] = 0$ quel que soit ν.

PROPOSITION 1.2. — *Il y a une correspondance biunivoque entre les fonctions h invariantes bornées et les classes d'équivalence de variables aléatoires invariantes bornées Z donnée par*

$$h(x) = E_x[Z], \qquad Z = \lim_{t \to \infty} h(X_t) \quad \text{p. s.}$$

Démonstration. — Si h est invariante et bornée, le processus $h(X_t)$ est une martingale continue à droite, bornée, donc qui converge p. s. vers une variable aléatoire Z qui peut être évidemment choisie invariante. Du fait que h est bornée, la convergence a également lieu dans $L^1(P_x)$ pour tout x de E et il en résulte que

$$h(x) = \lim_{t\to\infty} E_x[h(X_t)] = E_x[\lim_{t\to\infty} h(X_t)] = E_x[Z].$$

Réciproquement, si Z est J-mesurable et bornée, la fonction h définie par $h(x) = E_x[Z]$ est universellement mesurable et la propriété de Markov entraîne facilement, puisque le temps de mort est p. s. infini par hypothèse, que h est invariante.

DÉFINITION 1.3. — On appelle *tribu asymptotique* la tribu

$$\mathscr{A} = \bigcap_{t=0}^{\infty} \theta_t^{-1}(\mathscr{F}).$$

On posera aussi

$$\mathscr{A}^0 = \bigcap_{t=0}^{\infty} \theta_t^{-1}(\mathscr{F}^0) = \bigcap_{t=0}^{\infty} \sigma(X_s, s \geq t).$$

une variable aléatoire Z est dite *asymptotique* si elle est $\mathscr{A}$-mesurable.

Une variable aléatoire Z est donc asymptotique si et seulement si, pour tout t, il existe une variable Z_t $\mathscr{F}$-mesurable et telle que $Z = Z_t \circ \theta_t$. Cette variable Z_t est elle-même asymptotique. En effet, pour toute trajectoire ω, $Z_t(\omega) = Z(\overline{\omega})$ où $\overline{\omega}$ est une trajectoire arbitraire telle que $\omega = \theta_t(\overline{\omega})$. Il n'y a pas d'ambiguité du fait que Z est asymptotique et ne dépend donc pas des coordonnées de $\overline{\omega}$ inférieures à t. Cette variable Z_t est donc unique et $Z_{t+s} \circ \theta_s = Z_t$. On posera

DÉFINITION 1.4. — Deux variables asymptotiques Z et Z′ sont dites équivalentes si pour chaque t on a $Z_t = Z'_t$ p. s.

Il est clair que $J \subset \mathscr{A}$ et que des variables invariantes équivalentes au sens précédent sont équivalentes au sens de 1.1. Nous allons maintenant énoncer pour la tribu asymptotique la proposition analogue à 1.2.

PROPOSITION 1.5. — *La formule*

$$h(x, t) = E_x[Z_t]$$

établit une correspondance biunivoque entre les classes d'équivalence de variables asymptotiques bornées et les fonctions E.T-invariantes bornées. On a de plus

$$Z = \lim_{t\to\infty} h(X_t, t) \quad \text{p. s.}$$

Démonstration. — Si Z est asymptotique et bornée, on a, en utilisant la propriété de Markov

$$\int \mathrm{P}_t(x, dy)h(y, t+s) = \mathrm{E}_x[\mathrm{E}_{\mathrm{X}_t}[\mathrm{Z}_{t+s}]] = \mathrm{E}_x[\mathrm{Z}_{t+s} \circ \theta_t] = h(x, s)$$

ce qui prouve que h est E.T-invariante. Réciproquement, la martingale bornée $h(\mathrm{X}_t, t)$ converge p. s. vers $\mathrm{Z} = \overline{\lim}\ h(\mathrm{X}_t, t)$ qui est mesurable puisque la martingale est p. s. continue à droite et qui est clairement asymototique. Du fait que h est bornée, le théorème de Lebesgue montre que $h(x, t) = \mathrm{E}_x[\mathrm{Z}_t]$ pour tout x et tout t.

Nous aurons également besoin de la

PROPOSITION 1.6. — *La tribu $\mathscr{A}$ est contenue dans la complétée de $\mathscr{A}^0$ par rapport à l'ensemble des probabilités* P_ν.

Démonstration. — Soit $\mathrm{A} \in \mathscr{A}$; pour tout t il existe $\mathrm{A}_t \in \mathscr{F}$ tel que $\mathrm{A} = \theta_t^{-1}(\mathrm{A}_t)$ et pour ν fixée il existe B_t et C_t dans $\mathscr{F}^0$ tels que

$$\mathrm{B}_t \subset \mathrm{A}_t \subset \mathrm{C}_t, \qquad \mathrm{P}_{\nu \mathrm{P}_t}[\mathrm{C}_t \backslash \mathrm{B}_t] = \mathrm{P}_\nu[\theta_t^{-1}(\mathrm{C}_t \backslash \mathrm{B}_t)] = 0\,.$$

Les événements $\mathrm{C}'_t = \theta_t^{-1}(\mathrm{C}_t)$ et $\mathrm{B}'_t = \theta_t^{-1}(\mathrm{B}_t)$ sont dans $\theta_t^{-1}(\mathscr{F}^0)$ et tels que

$$\mathrm{B}'_t \subset \mathrm{A} \subset \mathrm{C}'_t, \qquad \mathrm{P}_\nu[\mathrm{C}'_t \backslash \mathrm{B}'_t] = 0\,.$$

Soit $\{ t_n \}$ une suite de réels croissant vers $+\infty$; on a

$$\overline{\lim_n}\ \mathrm{B}'_{t_n} \subset \mathrm{A} \subset \underline{\lim_n}\ \mathrm{C}'_{t_n}, \qquad \mathrm{P}_\nu[(\underline{\lim_n}\ \mathrm{C}'_{t_n}) \backslash (\overline{\lim_n}\ \mathrm{B}'_t)] = 0\,.$$

Comme ces deux ensembles sont dans $\mathscr{A}_0$, on voit que $\mathscr{A}$ est contenue dans la complétée de $\mathscr{A}^0$ pour P_ν, ce qui termine la démonstration.

II. LOIS DU TOUT OU RIEN

Nous allons maintenant énoncer les lois zéro-deux d'Ornstein-Sucheston [*12*] et Winkler [*15*] en suivant l'exposé de Derriennic [*4*] pour le cas discret.

Dans ce qui suit on notera $\| \nu \|_{\mathscr{A}}$ la norme de la mesure bornée ν considérée comme mesure sur $(\Omega, \mathscr{A})$. On peut alors énoncer le résultat clé de Derriennic.

THÉORÈME 2.1. — *Pour toute mesure bornée ν sur* $(\mathrm{E}, \mathscr{E})$ *on a*

$$\lim_{t \to \infty} \| \nu \mathrm{P}_t \| = \| \mathrm{P}_\nu \|_{\mathscr{A}}\,.$$

Démonstration. — La limite existe forcément puisque $\| \nu P_t \|$ décroît avec t. On a d'ailleurs $\lim_{t \to \infty} \| \nu P_t \| = \lim_{n} \| \nu P_n \|$. Grâce à 1.6 il suffit de montrer que ceci est égal à $\| P_\nu \|_{\mathscr{A}^0}$.

Si on pose $\overline{\mathscr{A}} = \bigcap_n \sigma(X_k, k \geq n)$, on a $\overline{\mathscr{A}} \subset \mathscr{A}^0$ de manière évidente et donc $\| P_\nu \|_{\overline{\mathscr{A}}} \leq \| P_\nu^n \|_{\mathscr{A}^0}$. Soit d'autre part (A, A^c) une décomposition de Jordan-Hahn de P_ν sur $\mathscr{A}^0$. D'après la proposition 1.5, il existe une fonction E.T-invariante bornée telle que

$$1_A = \lim_{t \to \infty} h(X_t, t) P_\nu\text{-p. s.}$$

On a donc aussi $1_A = \lim_{n \to \infty} h(X_n, n) P_\nu$-p. s. et il existe donc un élément $A' \in \overline{\mathscr{A}}$ tel que $A = A' \; P_\nu$-p. s. On a donc

$$P_\nu(A) - P_\nu(A^c) = P_\nu(A') - P_\nu(A'^c) \leq \| P_\nu \|_{\overline{\mathscr{A}}}.$$

Il en résulte que $\| P_\nu \|_{\mathscr{A}^0} = \| P_\nu \|_{\overline{\mathscr{A}}}$ et comme d'après le résultat de Derriennic appliqué à la chaîne de probabilité de transition P_1 on a

$$\lim_{n \to \infty} \| \nu P_n \| = \| P_\nu \|_{\overline{\mathscr{A}}} ;$$

la démonstration est terminée.

Nous emprunterons à Winkler [*15*] notre définition suivante.

DÉFINITION 2.2. — Pour tout nombre α positif, négatif ou nul on posera

$$h_\alpha(x) = \lim_{t \to \infty} \| P_{t+\alpha}(x, .) - P_t(x, .) \| .$$

L'existence de cette limite résulte de la décroissance de $\| \nu P_t \|$ appliquée ici avec $\nu = P_{k+\alpha}(x, .) - P_k(x, .)$ où k est choisi de façon que $k + \alpha > 0$. Les fonctions h^α sont boréliennes comme cela est montré dans [*12*], mais on peut montrer le résultat plus précis suivant

LEMME 2.3. — *L'application $(\alpha, x) \to h^\alpha(x)$ est bi-mesurable.*

Démonstration. — Pour f continue à support compact ($f \in C_K$) et pour t fixé, l'application $(\alpha, x) \to P_{t+\alpha} f(x)$ est mesurable en x et continue à droite en α à cause de la continuité à droite des trajectoires, elle est donc bi-mesurable. Maintenant

$$\| P_{t+\alpha}(x, .) - P_t(x, .) \| = \sup | P_{t+\alpha} \phi_n(x) - P_t \phi_n(x) | / \| \phi_n \|$$

où $\{ \phi_n \}$ décrit une suite dense dans C_K pour la topologie de la convergence uniforme sur les compacts. L'application $(\alpha, x) \to \| P_{t+\alpha}(x, .) - P_t(x, .) \|$

est donc bi-mesurable pour t fixé et h^α qui est la limite d'une suite de telles fonctions est bi-mesurable. □

On peut énoncer pour ces fonctions la loi de zéro-deux suivante :

THÉORÈME 2.4. — *Pour tout $\alpha > 0$, ou bien $\sup_{x \in E} h^\alpha(x) = 0$ donc h^α est identiquement nulle ou bien $\sup_{x \in E} h^\alpha(x) = 2$. De plus $h^\alpha \equiv 0$ si et seulement si $\mathscr{A}$ est θ_α-invariante et dans ce cas*

$$\lim_{t \to \infty} \| \nu(P_{t+\alpha} - P_t) \| = 0$$

quelle que soit ν. En particulier, $h^\alpha \equiv 0$ pour tout α si et seulement si $\mathscr{A} = J$ p. s. et dans ce cas la convergence ci-dessus a lieu pour tout α.

Démonstration. — Pour $\alpha \neq 0$, on a

$$h^\alpha(x) = \lim_{n \to \infty} \| P_{n|\alpha|+\alpha}(x, .) - P_{n|\alpha|}(x, .) \|$$

et par suite la première phase de l'énoncé résulte immédiatement de la loi zéro-deux pour les chaînes (théorème 3 de [*4*]).

Montrons la deuxième phrase. Si $\mathscr{A}$ est invariante par θ_α, d'après la propriété de Markov, P_ν et $P_{\nu P_\alpha}$ sont égales sur $\mathscr{A}$ et par conséquent

$$\lim_{t \to \infty} \| \nu P_{t+\alpha} - \nu P_t \| = 0$$

d'après le théorème précédent.

Supposons inversement que $\mathscr{A}$ ne soit pas θ_α-invariant. Il existe alors un ensemble $F \in \mathscr{A}$ et une probabilité ν tels que $P_\nu[F \Delta \theta_\alpha^{-1}(F)] > 0$. Quitte à remplacer ν par $\sum_0^\infty 2^{-n} \nu P_\alpha$ on peut supposer que $\nu P_\alpha \ll \nu$. Si $P_\nu[F \backslash \theta_\alpha^{-1}(F)] > 0$ on pose $G = F \backslash \theta_\alpha^{-1}(F)$; l'ensemble G est dans $\mathscr{A}$, est de P_ν-mesure positive et $G \cap \theta_\alpha^{-1}(G) = (F \backslash \theta_\alpha^{-1}(F)) \cap (\theta_\alpha^{-1}(F) \backslash \theta_{2\alpha}^{-1}(F)) = \varnothing$. Sinon on a $P_\nu[\theta_\alpha^{-1}(F) \backslash F] > 0$ mais ceci vaut $P_{\nu P_\alpha}[F \backslash F_\alpha]$ et par conséquent $P_\nu[F \backslash F_\alpha] > 0$. Si on pose $G = F \backslash F_\alpha$ on a encore $G \in \mathscr{A}$, $P_\nu(G) > 0$ et

$$G \cap \theta_\alpha^{-1}(G) = (F \backslash F_\alpha) \cap (\theta_\alpha^{-1}(F) \backslash F) = \varnothing .$$

Dans chacun des cas, nous posons $Z = 1_G - 1_G \circ \theta_\alpha$ et $g_t(x) = E_x[Z_t]$. D'après la proposition 1.5 on a

$$\lim_{t \to \infty} g_t(X_t) = Z, \qquad \lim_{t \to \infty} g_{t+\alpha}(X_t) = Z_\alpha \qquad P_\nu\text{-p. s.}$$

Mais $|Z - Z_\alpha| = 2$ sur G d'après la définition de Z_α. Pour tout $\varepsilon > 0$, il existe donc un $x \in E$ et un réel s tels que

$$|g_{s+\alpha}(x) - g_s(x)| > 2 - \varepsilon.$$

Comme $\| g_t \| \leq 1$ et $P_t g_{t+s} = g_s$ il s'ensuit que

$$\begin{aligned} \| P_{t+\alpha}(x,.) - P_t(x,.) \| &\geq | P_t g_{t+s+\alpha}(x) - P_{t+\alpha} g_{t+s+\alpha}(x) | \\ &= | g_{s+\alpha}(x) - g_s(x) | > 2 - \varepsilon. \end{aligned}$$

Il en résulte que $h_\alpha(x) > 2 - \varepsilon$ ce qui termine la démonstration. □

Ceci nous conduit à la classification suivante des semi-groupes.

THÉORÈME 2.5. — *L'ensemble* $G = \{ \alpha : h^\alpha \equiv 0 \}$ *est un sous-groupe de* $\mathbb{R}$, *on a donc quatre espaces de semi-groupes suivant que*

i) $G = \mathbb{R}$;

ii) G *est dense dans* $\mathbb{R}$ *mais différent de* $\mathbb{R}$;

iii) $G = d\mathbb{Z}$ *pour un réel* $d > 0$;

iv) $G = \{ 0 \}$.

Démonstration. — Il est facile de voir que $h^\alpha = h^{-\alpha}$ et que $h^{\alpha+\beta} \leq h^\alpha + h^\beta$. Comme d'autre part $h^0 = 0$, il est clair que G est un sous-groupe de $\mathbb{R}$.

Les quatre cas peuvent se produire et nous allons en donner maintenant des exemples pour des processus ayant tous des mesures invariantes bornées, donc « récurrents » au moins au sens de la théorie ergodique ponctuelle.

1) Cas $G = \mathbb{R}$. Il y a d'innombrables exemples de ce cas, tel celui du mouvement brownien sur une variété compacte. Une étude plus approfondie de ce cas sera faite dans le paragraphe suivant.

2) Cas où G est dense sans être égal à $\mathbb{R}$. Prenons comme espace d'état le cercle de rayon 1 considéré comme un groupe. Soit μ une probabilité portée par l'ensemble des nombres rationnels et μ^n sa $n^{\text{ième}}$ puissance de convolution. On pose

$$\mu_t = \sum_{n=0}^{\infty} e^{-t} \frac{t^n}{n!} \mu^n ;$$

on définit ainsi un semi-groupe de convolution sur le cercle qui est vaguement continu en O. Si on pose $\nu_t = \varepsilon_{\bar{t}} * \mu_t$ où $\bar{t} = t (\text{mod } 1)$ on obtient encore un semi-groupe de convolution vaguement continu en O. Le processus de Feller correspondant est récurrent au sens de Harris (cf. [*2*]) car la résolvante est absolument continue par rapport à la mesure de Lebesgue (cf. [*14*], p. 91) et pour ce processus on a

$$\| P_{t+\alpha} - P_t \| < 2 \qquad \text{si } \alpha \text{ est rationnel},$$
$$\| P_{t+\alpha} - P_t \| = 2 \qquad \text{si } \alpha \text{ est irrationnel}.$$

Les fonctions h^α sont d'ailleurs constantes dans ce cas à cause de l'invariance par translation.

3) Un exemple où $G = d\mathbb{Z}$ est fourni par la translation sur le cercle ou moins trivialement par le processus de la chaleur sur le Tore.

4) Enfin un exemple où $G = \{0\}$ est fourni par la translation irrationnelle sur le tore. De façon plus précise on pose $X_t(x) = x + ty$ où y est un point dont les deux composantes sont dans un rapport irrationnel. Ceci est un exemple de processus récurrent dans les ouverts mais non récurrent au sens de Harris.

Dans le paragraphe suivant nous étudierons plus en détail certains processus récurrents. Nous allons auparavant énoncer une autre loi 0.2. Toujours grâce au théorème 2.1 on peut poser pour tout couple (x, y) de points de E

$$\beta(x, y) = \lim_{t\to\infty} \| P_t(x,.) - P_t(y,.) \| .$$

THÉORÈME 2.6. — *Ou bien* $\sup_{x,y\in E} \beta(x, y) = 0$,

ou bien $\sup_{x,y\in E} \beta(x, y) = 2$.

Le premier cas se produit si et seulement si $\mathscr{A}$ est p. s. triviale et l'on a alors

$$\lim_{t\to\infty} \| (\nu_1 - \nu_2)P_t \| = 0$$

pour tout couple (ν_1, ν_2) de probabilités initiales.

Démonstration. — Si $\mathscr{A}$ est triviale p. s. alors $\| P_{\nu_1 - \nu_2} \|_{\mathscr{A}} = 0$ et le résultat découle de 2.1.

Réciproquement, supposons qu'il existe une probabilité ν et un événement $F \in \mathscr{A}$ tels que $0 < P_\nu(F) < 1$; posons $Z = 1_F - 1_{F^c}$ et $g_t(x) = E_x[Z_t]$. Comme $\lim_{t\to\infty} g_t(X_t) = Z$ P_ν-p. s., il existe des points x et y dans E et un nombre s tels que pour un $\varepsilon > 0$ donné à l'avance on ait

$$g_s(x) > 1 - \varepsilon, \qquad g_s(y) < -1 + \varepsilon .$$

Comme dans la démonstration de 2.3, on a alors

$$\lim_{t\to\infty} \| P_t(x,.) - P_t(y,.) \| \geq | g_s(x) - g_s(y) | > 2 - 2\varepsilon$$

et la démonstration est facilement terminée.

III. PROCESSUS RECURRENTS AU SENS DE HARRIS ET APÉRIODIQUES

A partir de maintenant, nous allons supposer que le processus est *récurrent au sens de Harris*, c'est-à-dire : il existe une mesure invariante m telle que $m(A) > 0$ entraîne

$$P_x\left[\int_0^\infty 1_A(X_s)ds = \infty\right] = 1$$

pour tout x de E.

Il a été montré dans [*2*] que cette hypothèse entraîne la

PROPOSITION 3.1. — *Si f est presque-borélienne et positive et si $P_t f \leq f$ pour tout t alors f est constante m-presque partout et supérieure à cette constante partout.*

Cette propriété est en fait équivalente à l'hypothèse de récurrence au sens de Harris. Pour le montrer nous aurons besoin du noyau résolvant U_α du semi-groupe défini par

$$U_\alpha f = \int_0^\infty e^{-\alpha t} P_t f\, dt.$$

PROPOSITION 3.2. — *Si il existe une mesure ξ telle que $\alpha\xi U_\alpha \gg \xi$ pour tout α et telle que toute fonction positive, bornée, presque-borélienne f vérifiant $P_t f \leq f$ pour tout $t \geq 0$ soit constante ξ-p. s. et supérieure à cette constante partout, le processus est récurrent au sens de Harris.*

Démonstration. — Soit f une fonction borélienne bornée telle que $U_1 f \leq f$. A cause de l'équation résolvante on a aussi $\alpha U_\alpha f \leq f$ pour tout $\alpha > 0$. La fonction $\tilde{f} = \lim_{\alpha\to\infty} \alpha U_\alpha f$ vérifie $P_t \tilde{f} \leq \tilde{f}$ pour tout t (cf. [*3*], p. 116) et donc $\tilde{f} = c$ ξ-p. p. et $\tilde{f} \geq c$ partout pour une constante c. Or on a $f \geq \tilde{f}$ et si $A = \{f > \tilde{f}\}$ on a $\alpha U_\alpha(.,A) \equiv 0$ (cf. [*6*]) et donc $\xi(A) = 0$.

La chaîne de Markov Y de probabilité de transition U_1 a donc la propriété que ses fonctions surharmoniques bornées sont constantes ξ-p. s. et supérieures à cette constante partout. Soit A un ensemble tel que $\xi(A) > 0$; la probabilité d'atteinte de A par Y est une fonction surharmonique bornée par 1 et qui vaut 1 sur A, elle est donc identiquement égale à 1 et par suite Y est récurrente au sens de Harris. Il en résulte (cf. [*2*]) que le processus lui-même est récurrent au sens de Harris.

Une conséquence importante de notre hypothèse est la

PROPOSITION 3.3. — *La tribu invariante* J *est grossière p. s.*

Démonstration. — Si h est invariante bornée, il existe une constante c telle que $h = cm$-p. p. et $h \geq c$ partout. Mais en appliquant le même raisonnement à $\| h \| - h$ on voit que $h = c$ partout. Le résultat découle alors de la proposition 1.2.

Sous l'hypothèse de ce paragraphe, la loi 0.2 du théorème 2.3 se précise.

PROPOSITION 3.4. — *Pour chaque* α, *ou bien* $h^\alpha \equiv 0$ *ou bien* $h^\alpha = 2m$-p. p. *D'autre part, ou bien l'application* $(\alpha, x) \to h^\alpha(x)$ *est identiquement nulle ou bien pour m-presque tout* x *on a* $\lambda \{ \alpha : h^\alpha(x) < 2 \} = 0$ *où* λ *est la mesure de Lebesgue.*

Démonstration. — D'après 2.4 si h^α n'est pas identiquement nul, on a $\sup h^\alpha = 2$. Mais on peut montrer (cf. [*15*]) que pour tout t, on a $P_t(2 - h^\alpha) \leq 2 - h^\alpha$ d'où il résulte que $2 - h^\alpha$ est constante m-p. p. et supérieure à cette constante partout. On a donc une constante c telle que $h^\alpha = cm$-p. p. et $h^\alpha \leq c$ partout ; si $\sup h^\alpha = 2$ cette constante ne peut être égale qu'à 2 ce qui démontre la première phase de l'énoncé.

Pour démontrer la deuxième, observons que si $(\alpha, x) \to h^\alpha(x)$ n'est pas identiquement zéro le groupe G du théorème 2.5 est de mesure de Lebesgue nulle. Posons $F = \{ (\alpha, x) : h^\alpha(x) < 2 \}$; on a

$$\lambda \otimes m(F) = \int \lambda(d\alpha) \int 1_F(\alpha, x) m(dx) = \int 1_{G^c}(\alpha) \lambda(d\alpha) \int 1_F(\alpha, x) m(dx) = 0,$$

d'après la première phrase de l'énoncé. Mais d'après le théorème de Fubini ceci est aussi égal à

$$\int m(dx) \int 1_F(\alpha, x) \lambda(d\alpha)$$

et le résultat en découle donc immédiatement.

Remarque. — Pour des processus transients, les fonctions h^α peuvent prendre toute une variété de valeurs comprises entre 0 et 2. Pour s'en convaincre, on peut prendre l'exemple d'un processus consistant en le mouvement brownien sur les réels < 0 et la translation vers la droite sur les réels ≥ 0.

Nous allons maintenant nous tourner vers l'étude du cas $G = \mathbb{R}$ qui est le seul où l'on puisse avoir le théorème de convergence d'Orey. Nous allons commencer par des préliminaires valables pour tous les processus de Harris. On écrira

$$P_t(x, .) = P_t^0(x, .) + P_t^1(x, .)$$

la décomposition de Lebesgue de P_t par rapport à m. On sait que cela peut se faire de façon que P_t^0 et P_t' soient aussi des noyaux. De plus, on a l'inégalité $P_{t+s}^1 \leq P_t^1 P_s^1$ qui n'est pas difficile à montrer en utilisant l'invariance de m. Il en résulte que $P_t^1(x, E)$ décroît lorsque t croît et donc que $g(x) = \lim_{t \to \infty} P_t^1(x, E)$ existe. Le résultat suivant était dans [*5*]; nous le reproduirons pour la commodité du lecteur.

PROPOSITION 3.5. — *Ou bien* $g(x) = 0$ *pour tout* x *de* E, *ou bien pour* m*-presque tout* x *on a* $P_t^1(x, E) = 1$ *pour tout* t.

Démonstration. — On a $P_t(1 - g) \leq 1 - g$ et par suite $1 - g$ est égale à m-p. p. à une constante $1 - a$ et est supérieure à $1 - a$ partout. En appliquant le théorème de Lebesgue, on voit alors, en choisissant x tel que $g(x_0) = a$, que

$$g(x_0) \leq \lim_{s \to \infty} \int P_t^1(x_0, dy) P_s^1(y, E) = \int P_t^1(x_0, dy) g(y) \leq a P_t^1(x_0, E),$$

d'où il résulte en passant à la limite que $a \leq a^2$ donc que $a = 0$ ou que $a = 1$; dans le premier cas on a $g(x) \leq 0$ donc $g(x) = 0$ pour tout x.

Avant de passer au théorème que nous avons en vue, nous allons énoncer une propriété des chaînes de Harris.

PROPOSITION 3.6. — *Si* ν *est une mesure et* P *une probabilité de transition telle que* $\nu P \ll \nu$ *et si la contraction induite par* P *sur* $L^1(\nu)$ *est conservative* (*cf.* [*13*], *chapitre IV*) *les deux propriétés suivantes sont équivalentes*

i) P *est de Harris et apériodique ;*

ii) *Pour tout couple* (ν_1, ν_2) *de probabilités sur* $\mathscr{E}$

$$\lim_{n \to \infty} \| \nu_1 P_n(.) - \nu_2 P_n(.) \| = 0 .$$

Démonstration. — Que *i*) entraîne *ii*) est bien connue (cf. [*8*] et [*13*], chapitre VI). Nous n'allons qu'esquisser la démonstration de la réciproque. La condition *ii*) entraîne immédiatement que toute fonction harmonique bornée est constante, donc que la tribu des événements invariants de la chaîne est grossière et que l'opérateur induit sur $L^1(\nu)$ est conservatif et ergodique. La chaîne est donc ν-essentiellement irréductible (cf. [*10*] et [*13*], § 2, chapitre III) et en appliquant le théorème de dichotomie pour ces chaînes conjointement avec la propriété *ii*) il n'est pas difficile de montrer que la chaîne est Harris et apériodique. En effet, en prenant $\nu_1 \sim \nu$ et $\nu_2 = \varepsilon_x$ on voit que pour aucun x de E on ne peut avoir $\Sigma 2^{-n} P_n(x, .)$ étrangère à ν. La chaîne est donc ν-irréductible et elle ne peut être transiente à cause de la conservativité de la contraction induite par P. La chaîne est donc de

Harris, la condition *ii*) entraînant qu'il n'y a même pas d'ensemble exceptionnel à enlever. L'apériodicité est alors évidente.

Nous pouvons maintenant énoncer le résultat principal de ce paragraphe.

THÉORÈME 3.7. — *Les cinq propriétés suivantes sont équivalentes*

i) $G = \mathbb{R}$ *ou en d'autres termes* $\lim_{t\to\infty} \| P_{t+\alpha}(x,.) - P_t(x,.) \| = 0$ *pour tout* x *et tout* a ;

ii) *Pour tout couple* (ν_1, ν_2) *de probabilités sur* E *on a*

$$\lim_{t\to\infty} \| (\nu_1 - \nu_2)P_t \| = 0 ;$$

iii) *Il existe un réel* $a > 0$ *tel que la chaîne de probabilité de transition* P_a *soit de Harris par rapport à* m *et apériodique ;*

iv) *Pour tout réel* $a > 0$ *la chaîne de probabilité de transition* P_a *est de Harris par rapport à* m *et apériodique ;*

v) $\lim_{t\to\infty} P_t^1(x, E) = 0$ *pour tout* x *de* E.

Démonstration. — La propriété *ii*) entraîne *iv*) à cause de la proposition 3.6, *iv*) entraîne *iii*) est évident et *iii*) entraîne *ii*) puisque $\| (\nu_1 - \nu_2 P_t \|$ décroît avec t et que $\lim_{n\to\infty} \| (\nu_1 - \nu_2)P_{na} \| = 0$.

La propriété *i*) entraîne *ii*) car elle entraîne que $\mathscr{A} = J$ p. s. et on sait déjà que J est grossière ; il n'y a donc qu'à appliquer le théorème 2.6. Pour voir que *ii*) entraîne *v*) il suffit de prendre ν_1 équivalente à m et $\nu_2 = \varepsilon_x$.

Il reste donc à montrer que *v*) entraîne *i*). Nous allons donc supposer que G n'est pas égal à $\mathbb{R}$ et montrer que cela entraîne une contradiction. D'après 3.4 nous pouvons choisir un x tel que l'ensemble

$$H = \{ \alpha : h^\alpha(x) < 2 \}$$

soit de mesure de Lebesgue nulle. L'ensemble H est un ensemble de classes d'équivalence modulo G. En effet, si $\alpha \in H$ et $g \in G$ on a

$$\begin{aligned} h^{\alpha+g}(x) &= \lim_{t\to\infty} \| P_{t+\alpha+g}(x,.) - P_t(x,.) \| \\ &\leq \lim_{t\to\infty} \| P_{t+\alpha+g}(x,.) - P_{t+g}(x,.) \| + \lim_{t\to\infty} \| P_{t+g}(x,.) - P_t(x,.) \| \\ &= h^\alpha(x) + h^g(x) = h^\alpha(x) < 2 . \end{aligned}$$

Soit alors ν une probabilité équivalente à m. Nous appellerons $p_t(x,.)$ une densité bimesurable de $P_t^0(x,.)$ par rapport à ν et pour tout t nous poserons $E_t = \{ y : p_t(x, y) > 0 \}$. Nous choisissons un nombre $t_0 > 0$ tel que $P_{t_0}^0(x, E) > 0$; on a donc $\nu(E_s) > 0$ pour tout $s \geq t_0$. Nous pouvons alors choisir $t_1 > t_0$ tel que $t_1 - t_0 \notin H$. On a alors $\| P_{t_0}(x,.) - P_{t_1}(x,.) \| = 2$ puisque $h^{t_1 - t_0}(x) = 2$ et par suite $E_{t_0} \cap E_{t_1} = \emptyset \nu$-p. s. Comme H est de mesure nulle $\{ H + t_0 \} \cup \{ H + t_1 \}$ est encore de mesure nulle et on peut donc choisir $t_2 > t_0, t_2 \neq t_1$ tel que $t_2 - t_0$ et $t_2 - t_1$ ne soient pas dans H. Comme précédemment on aura $E_{t_2} \cap E_{t_0} = \emptyset \nu$-p. s. et $E_{t_2} \cap E_{t_1} = \emptyset \nu$-p. s.

On posera $F_0 = E_{t_0}$, $F_1 = E_{t_0} \cup E_{t_1}$, $F_2 = E_{t_0} \cup E_{t_1} \cup E_{t_2}, \ldots$ et ainsi de suite. La suite des mesures $\nu(F_i)$ est strictement croissante et bornée par 1. Nous pouvons continuer à procéder comme ceci par récurrence transfinie. Si α est un ordinal de 2^e^ espèce, on pose $F_\alpha = \bigcup_{\beta < \alpha} F_\beta$. Si α est un ordinal de 1^re^ espèce, l'ensemble $\bigcup_{\beta < \alpha} \{ H + t_\beta \}$ étant toujours de mesure nulle on peut trouver t_α tel que $t_\alpha - t_\beta$ ne soit dans H pour aucun $\beta < \alpha$ et l'on peut donc trouver E_α tel que $\nu(E_\alpha) > 0$ et $E_\alpha \cap F_{\alpha - 1} = \emptyset \nu$-p. s. On pose alors $F_\alpha = E_\alpha \cup F_{\alpha - 1}$. Comme $\nu(E) = 1$, la récurrence doit s'arrêter avant le premier ordinal non dénombrable. Il existe donc un ensemble $\{ t_j \}$ dénombrable de réels tels que les E_{t_j} soient disjoints 2 à 2 et de réunion égale à E ν-p. s. Mais $\bigcup_j \{ H + t_j \}$ est toujours de mesure nulle ; on peut donc trouver $r > t$ avec $r \notin \bigcup_j \{ H + t_j \}$ et pour tout j on aura

$$\| P_r(x,.) - P_{t_j}(x,.) \| = 2,$$

ce qui est clairement impossible puisque la partie absolument continue de $P_r(x,.)$ est non nulle.

Remarques. — 1) La démonstration précédente serait notablement plus simple si on pouvait montrer que pour presque tout x on a ou bien $h^\alpha(x) = 0$ ou bien $h^\alpha(x) = 2$ pour tout α. Ceci est en particulier le cas si le processus est un processus à accroissements indépendants sur un groupe, puisque dans ce cas h^α ne dépend pas de x. Pourrait-on montrer ceci de manière plus générale ?

2) L'équivalence des conditions *ii*) à *v*) était montrée — ou du moins sous-entendue — dans [5] par une autre méthode. Le pas difficile de [5] consistait à montrer que *v*) entraîne *iv*). Une imperfection à la fin de cette démonstration peut être corrigée en utilisant les raisonnements de la proposition 3.6.

Nous pouvons maintenant énoncer pour les processus un résultat analogue à la proposition 3.6.

PROPOSITION 3.8. — *Si ν est une mesure telle que pour tout t on ait $\nu P_t \ll \nu$ et pour tout $f \in L^1_+(\nu)$*

$$\int_0^\infty P_s f\, ds = 0 \qquad \text{ou} \qquad +\infty \ \nu\text{-p. p. } sur\ E,$$

la condition

$$\lim_{t\to\infty} \| \nu_1 P_t - \nu_2 P_t \| = 0$$

pour tout couple (ν_1, ν_2) de probabilités sur $(E, \mathscr{E})$ entraîne que le processus est récurrent au sens de Harris.

Démonstration. — D'après [*2*] il suffit de montrer que la résolvante U^1 est de Harris. Or d'une part on a $\nu U^1 \ll U^1$ et il est facile de voir que U^1 est conservative par

$$\int_0^\infty P_s f\, ds = \sum_1^\infty (U^1)^n f.$$

D'autre part, comme $(U^1)^n = \int_0^\infty e^{-s} \frac{s^n}{n\,!} P_s ds$ on a

$$\| \nu_1 (U^1)^n - \nu_2 (U^1)^n \| \leq 2 \int_0^N e^{-s} \frac{s^n}{n\,!} ds + \int_N^\infty e^{-s} \frac{s^n}{n\,!} \| \nu_1 P_s - \nu_2 P_s \| \, ds\ ;$$

en choisissant N assez grand pour que $\| P_s(x,.) - P_s(y,.) \| < \varepsilon$ pour tout $s > N$, puis n assez grand pour que la première intégrale soit $< \varepsilon$ on obtient

$$\| \nu_1 (U^1)^n - \nu_2 (U^1)^n \| \leq 2\varepsilon$$

et le résultat découle alors de la proposition 3.6.

IV. QUESTIONS OUVERTES

Dans ce paragraphe, nous faisons une liste de quelques questions soulevées par ce qui précède.

Une question générale est celle de la classification des semi-groupes suivant les diverses valeurs du groupe G ; une autre est de déterminer les processus pour lesquels on a 0 dans le théorème 2.6. De manière plus spécifique, on peut par exemple se poser la question de savoir si on a toujours zéro pour le mouvement brownien sur une variété riemannienne. La réponse à ceci est non car on obtient 2 pour tout processus à accroissements indépendants sur un groupe de Lie semi-simple, mais on peut transformer la question en : quelles sont les variétés pour lesquelles on a 0 ?

Il serait intéressant d'étudier le comportement des semi-groupes pour lesquels $G = d\mathbb{Z}$ où G est dense dans $\mathbb{R}$. Peut-on définir des « classes cycliqnes » dans ce cas et montrer que les exemples du paragraphe 2 sont typiques ?

Il existe des processus que l'on peut qualifier de « récurrents » pour lesquels $G = \{ 0 \}$; de nombreux exemples en sont donnés par l'étude des processus avec interaction. Peut-on montrer que la récurrence fine [*1*] ou la récurrence au sens de Harris excluent le cas $G = \{ 0 \}$? Y aurait-il une propriété réciproque ?

BIBLIOGRAPHIE

[*1*] J. AZEMA, M. DUFLO et D. REVUZ, Récurrence fine des processus de Markov. *Ann. Inst. Henri Poincaré*, t. **2**, 1966, p. 185-220.

[*2*] J. AZEMA, M. DUFLO et D. REVUZ, Mesure invariante sur les classes récurrentes des processus de Markov. *Z. Wahrscheinlichkeitstheorie*, t. 8, 1967, p. 157-181.

[*3*] R. M. BLUMENTHAL and R. K. GETOOR, *Markov processes and Potential Theory*, Academic Press, 1968.

[*4*] Y. DERRIENNIC, Lois « zéro-deux » pour les processus de Markov. *Ann. Inst. Henri Poincaré*, t. **12**, 1976, p. 111-129.

[*5*] M. DUFLO et D. REVUZ, Propriétés asymptotiques des probabilités de transition des processus de Markov récurrents. *Ann. Inst. Henri Poincaré*, t. **5**, N° 3, 1969, p. 233-244.

[*6*] R. K. CETOOR, Markov Processes : Ray processes and right processes. *Lecture Notes in Maths*, t. **440**, Springer-Verlag, 1975.

[*7*] N. JAIN, Some limit theorems for a general Markov process. *Z. Wahrscheinlichkeitstheorie*, t. **6**, 1966, p. 206-223.

[*8*] B. JAMISON and S. OREY, Tail σ field Markov processes recurrent in the sense of Harris. *Z. Wahrscheinlichkeitstheorie*, t. **8**, 1967, p. 18-40.

[*9*] J. NEVEU, *Bases Mathématiques du calcul des probabilités*, Masson, 1964.

[*10*] J. NEVEU, Sur l'irréductibilité des chaînes de Markov, *Ann. Inst. Henri Poincaré*, t. **8**, 1972, p. 249-254.

[*11*] S. OREY, Recurrent Markov chains. *Pacific J. of Math.*, t. **9**, 1959, p. 805-827.

[*12*] D. S. ORNSTEIN and L. SUCHESTON, An operator theorem on L^1-convergence to zero with applications to Markov kernels. *Ann. of Math. Stat.*, t. **41**, 1970. p. 1631-1639.

[*13*] D. REVUZ, *Markov chains*. North Holland, 1975.

[*14*] D. REVUZ, Sur la définition des classes cycliques des chaînes de Harris. *Israel J. of Math.*, t. **33**, 1979, p. 378-383.

[*15*] W. WINKLER, A note on continuous parameter zero-two law. *The Annals of Probability*, t. **1**, 1973, p. 341-344.

(Manuscrit reçu le 4 janvier 1982)