ISRAEL JOURNAL OF MATHEMATICS, Vol. 33, Nos. 3–4, 1979

SUR LA DEFINITION DES CLASSES CYCLIQUES DES CHAINES DE HARRIS

PAR
DANIEL REVUZ

Cet article est dédié à la mémoire de Shlomo Horowitz

ABSTRACT

We give an alternative proof of the existence of the cyclically moving sub-classes of a Harris chain relying on the so-called zero-two laws and avoiding the use of the differentiation lemma of Orey's original method.

Depuis la démonstration originale publiée dans [4] le théorème de convergence d'Orey pour les chaînes de Harris a connu plusieurs autres démonstrations dues à Ornstein–Sucheston, Papangelou, Derriennic et plus récemment Athreya et Ney. Ce résultat suppose que la chaîne soit apériodique et il faut donc avoir fait au préalable une étude des classes cycliques pour ces chaînes. Cette étude jusqu'à présent ne se fait que par la méthode originale d'Orey [4] qui comporte la démonstration d'un lemme relativement peu agréable (lemme VI.1.1 de [8]) ou par la méthode exposée dans [3] qui comporte également des difficultés.

Le but de cette note, dont l'intérêt est principalement "pédagogique", est de montrer que les résultats de Derriennic [2] dont l'exposé est déjà très ramassé permettent sans effort nouveau d'obtenir aussi la définition des classes cycliques. On obtient en fait d'un seul coup les résultats de périodicité et le théorème de convergence. Cette partie de la théorie des chaînes de Harris devient alors assez rapide à exposer et ne comporte plus aucune difficulté technique.

Une première version de cette note a été rédigée lorsque l'auteur bénéficiait de l'hospitalité de l'Université de Colombie Britanique à Vancouver.

I. Notations et preliminaires

Dans tout ce qui suit on considère une chaîne de Markov X sur un espace

Received September 1, 1978

d'état de type dénombrable $(E, \mathscr{E})$; on la suppose récurrente au sens de Harris et l'on note P sa probabilité de transition et m sa mesure invariante. Les notations qui ne sont pas explicitement rappelées sont celles de [8], notamment des chapitres III et VI.

La décomposition de Lebesgue par rapport à m des puissances P_n de P sera notée

$$P_n(x, \cdot) = P_n^0(x, \cdot) + P_n^1(x, \cdot)$$

où $P_n^0(x, \cdot)$ est absolument continue et $P_n^1(x, \cdot)$ singulière par rapport à m. Pour tout n on peut choisir une fonction p_n qui soit $\mathscr{E} \otimes \mathscr{E}$-mesurable et telle que

$$P_n^0(x, A) = \int_A p_n(x, y) m(dy).$$

Enfin on peut facilement montrer que (cf. [8] p. 81)

$$\lim_{n \to \infty} P_n^1(x, E) = 0.$$

Notre démonstration va utiliser les fonctions introduites ci-dessous (voir [9]).

Définition. Pour tout entier relatif k on pose

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

La limite existe car la quantité sur la droite décroit avec n. Pour tout k, la fonction h^k est positive et bornée par 2. Elle est de plus mesurable; ceci est démontré dans [6] et se fait facilement à partir du lemme I.5.3 de [8].

Proposition. *Pour tout entier relatif k on a*

ou bien $\quad h^k(x) = 0 \quad$ *pour tout $x \in E$,*

ou bien $\quad h^k(x) = 2 \quad$ *pour m-presque tout x de E.*

Démonstration. On a $Ph^k \geqq h^k$. En effet si l'on pose

$$\gamma_n(x, \cdot) = P_{n+k}(x, \cdot) - P_n(x, \cdot), \qquad h_n(x) = \| \gamma_n(x, \cdot) \|,$$

on a, pour tout $A \in \mathscr{E}$,

$$h_n(x) \geqq \gamma_n(x, A) - \gamma_n(x, A^c).$$

On a donc

$$Ph_n(x) \geqq \int P(x, dy)(\gamma_n(y, A) - \gamma_n(y, A^c))$$

$$= \gamma_{n+1}(x, A) - \gamma_{n+1}(x, A^c).$$

En prenant pour (A, A^c) une décomposition de Jordan–Hahn de $\gamma_{n+1}(x, \cdot)$ nous voyons que $Ph_n \geqq h_{n+1}$ ce qui entraîne $Ph^k \geqq h^k$. Comme h^k est bornée supérieurement par 2 il résulte de la constance m-presque-partout des fonctions surharmoniques d'une chaîne de Harris ([8] p. 83) que h^k est constante m-presque-partout et est inférieure à cette constante partout.

D'autre part le théorème 3 de [2] appliqué à la probabilité de transition $P_{|k|}$ entraine que l'on a ou bien $\sup_x h^k(x) = 2$ ou $\sup_x h^k(x) = 0$. La proposition en résulte donc immédiatement.

II. Théorème d'Orey et périodicité

Nous allons maintenant montrer

THÉORÈME. *Si X est une chaîne de Harris, il n'y a que deux cas possibles:*
(1) *Ou bien pour tout couple (ν_1, ν_2) de probabilités sur E on a*

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

(2) *Ou bien il existe un entier $d > 1$ et une partition $C_0, C_1, \cdots, C_{d-1}, F$ de E telle que $m(F) = 0$ et $P(\cdot, C_i) = 1$ sur C_{i-1}, $P(\cdot, C_0) = 1$ sur C_{d-1}.*

DÉMONSTRATION. Il est facile de voir que $h^{k+l} \leqq h^k + h^l$. Il en résulte que l'ensemble $G = \{k : h^k \equiv 0\}$ est un sous-groupe du groupe $\mathbf{Z}$ des entiers relatifs et par suite est égal soit à $\{0\}$, soit à $\mathbf{Z}$ lui-même, soit à $d\mathbf{Z}$ pour un entier $d > 1$. Nous allons montrer que le premier cas est impossible et que les deux autres correspondent à l'alternative du théorème.

Supposons donc d'abord que $G = \{0\}$; en enlevant au besoin un ensemble de mesure nulle on peut donc supposer que

$$\| P_{n+k}(x, \cdot) - P_n(x, \cdot) \| = 2$$

pour tout x de E et tous entiers $n, k > 0$. Soit x un point de E et n un entier tel que $P_n^0(x, E) > 0$; posons $A = \{y : p_n(x, y) > 0\}$. Soit B un ensemble de mesure nulle tel que $\Sigma_0^\infty P_n^1(x, B^c) = 0$. A cause de l'hypothèse de recurrence il existe un

$m > n$ tel que $P_m(x, A \setminus B) > 0$, donc tel que $p_m(x, \cdot)$ soit strictement positif sur un sous-ensemble de $A \setminus B$ chargé par m. Il en découle que

$$\|P_n(x, \cdot) - P_m(x, \cdot)\| < 2$$

ce qui contredit l'hypothèse. Le cas $G = \{0\}$ est impossible.

Si $G = \mathbf{Z}$, nous avons en particulier

$$\lim_n \|P_{n+1}(x, \cdot) - P_n(x, \cdot)\| = 0$$

pour tout x de E. Par le théorème 3 de Derriennic [2] la tribu asymptotique de la chaîne est presque surement égale à sa tribu invariante; or cette dernière tribu est grossière dans le cas des chaînes de Harris (proposition III.3.6 de [8]). La tribu asymptotique est donc grossière et il est alors classique (proposition VI.2.4 de [8] ou théorème 4 de [2]) que cela entraîne la propriété de convergence de l'énoncé.

Considérons maintenant le dernier cas, à savoir $G = d\mathbf{Z}$ avec $d > 1$. En rejetant éventuellement un ensemble de mesure nulle, qui sera joint à la fin à l'ensemble F, on peut supposer que pour tout x de E on a

$$\lim_n \|P_{n+d}(x, \cdot) - P_n(x, \cdot)\| = 0, \qquad \|P_{n+j}(x, \cdot) - P_n(x, \cdot)\| = 2$$

pour tout entier n positif et tout j qui n'est pas un multiple de d.

Fixons un point x_0 arbitraire dans E et posons pour $j = 0, 1, \cdots, d-1$,

$$B_j = \bigcup_{n=0}^{\infty} \{y : p_{nd+j}(x_0, y) > 0\}.$$

Pour toute paire (i, j) on a $m(B_i \cap B_j) = 0$; en effet, on pourrait sinon trouver n_1 et n_2 tels que

$$m\{y : p_{n_1 d+i}(x_0, y) > 0 \text{ et } p_{n_2 d+j}(x_0, y) > 0\} > 0$$

et il en résulterait que

$$\|P_{n_1 d+i}(x_0, \cdot) - P_{n_2 d+j}(x_0, \cdot)\| < 2,$$

ce qui est impossible.

Par ailleurs la reunion des B_j est égale à E à un ensemble de mesure nulle près. Ceci résulte facilement de la propriété de récurrence et du fait que $\sum_1^\infty P_n^1(x_0, \cdot)$ est étrangère à m. On peut donc supposer que les B_j sont disjoints et que le complémentaire $\tilde{F}$ de leur réunion a pour mesure zéro.

Posons alors, pour $j \neq i+1$ modulo d,

$$D = \{y \in B_i,\ P(y, B_j) > 0\}$$

et supposons $m(D) > 0$. Il existe un entier n tel que $P^0_{nd+i}(x_0, D) > 0$ et donc tel que

$$P^0_{nd+i+1}(x_0, B_j) \geqq \int_D P^0_{nd+i}(x_0, dy) P(y, B_j) > 0.$$

On a donc $m(B_{i+1} \cap B_j) > 0$ ce qui contredit le fait que les B_j sont disjoints. On a donc $P(\cdot, B_{i+1}) = 1$ m-presque-partout sur B_i. En définissant $C_i = B_i \setminus \{y \in B_i : P(y, B_{i+1}) < 1\}$ et $F = (\bigcup C_i)^c$ nous obtenons la situation décrite dans la partie (2) de l'énoncé ce qui termine la démonstration.

La deuxième situation décrite par l'énoncé précédent est celle dite du cas *périodique*. Le nombre d qui intervient dans l'énoncé est la *période* de la chaîne et les ensembles C_i sont les *classes cycliques*. Dans ce cas la tribu asymptotique $\mathfrak{A}_\infty$ de la chaîne est égale à la tribu des évènements invariants par θ_d comme cela est montré dans [2]. Ces tribus sont atomiques avec d atomes comme cela est énoncé dans le corollaire 2.7 chapitre VI de [8], mais de manière inadéquate comme cela nous a été montré par R. Durrett. Nous allons préciser ceci ci-dessous.

PROPOSITION. *Dans le cas périodique, si ν_1 et ν_2 sont deux probabilités telles que $\nu_1(C_i) = \nu_2(C_i)$ pour tout i, et $\nu_1(F) = \nu_2(F) = 0$ on a*

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

D'autre part la tribu $\mathfrak{A}_\infty$ est atomique et toute variable aléatoire Z asymtotique et bornée peut s'écrire

$$Z = \sum_{i=1}^{d} \alpha_i \, 1_{\underline{\lim}\{X_{nd} \in C_i\}}$$

où les α_i sont des nombres réels arbitraires.

DÉMONSTRATION. La première phrase est facile à montrer à partir de ce qui précède. Pour la seconde nous utilisons les fonctions harmoniques bornées pour la chaîne "espace-temps" associée, c'est-à-dire les fonctions h sur $E \times \mathbf{N}$ telles que

$$h(x, n) = \int P(x, dy) h(y, n+1).$$

De l'inégalité

$$|h(x,n)-h(x',n)| \leqq \|h\| \|P_k(x,\cdot)-P_k(x',\cdot)\|$$

et de la première phrase de l'énoncé, il résulte que ces fonctions ont constantes sur les ensembles $C_i \times \{k\}$, i et k fixés, donc sur les ensembles $C_i \times \{\mathbf{N}d\}$.

Soit Z une variable aléatoire asymptotique et bornée; il existe une fonction h comme ci-dessus telle que ([8] p. 165)

$$Z = \lim_{n\to\infty} h(X_n, n) \qquad \text{p.s.}$$

Comme F est transient et si h vaut α_i sur $C_i \times \{\mathbf{N}d\}$, on peut alors écrire que presque-surement

$$Z = \lim_{n\to\infty} \sum_{i=1}^{d} \alpha_i 1_{C_i\times\{\mathbf{N}d\}}(X_n, n)$$

$$= \lim_{n\to\infty} \sum_{i=1}^{d} \alpha_i 1_{C_i}(X_{nd}).$$

Comme d'autre part on peut trouver Z telle que tous les α_i sont nuls sauf un, ceci prouve que $\lim_{n\to\infty} 1_{C_i}(X_{nd})$ existe p.s. et est donc p.s. égale à la variable asymptotique $\underline{\lim}\, 1_{C_i}(X_{nd})$. La variable Z précédente peut donc s'écrire

$$Z = \sum_{i=1}^{d} \alpha_i 1_{\underline{\lim}\{X_{nd}\in C_i\}}.$$

Bibliographie

1. K. B. Athreya and P. Ney, *A new approach to the limit theory of recurrent Markov chains*, to appear.

2. Y. Derriennic, *Lois "zéro ou deux" pour les processus de Markov. Applications aux marches aléatoires.* Ann. Inst. H. Poincaré **12** (1976), 111–129.

3. S. R. Foguel, *The Ergodic Theory of Markov Processes*, Van Nostrand, New York, 1969.

4. B. Jamison and S. Orey, *Tail σ-field of Markov processes recurrent in the sense of Harris*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **8** (1967), 41–48.

5. S. Orey, *Recurrent Markov chains*, Pacific J. Math. **9** (1959), 805–827.

6. D. S. Ornstein and L. Sucheston, *An operator theorem on L^1-convergence to zero with applications to Markov kernels*, Ann. Math. Statist. **41** (1970), 1631–1639.

7. G. Papangelou, *A martingale approach to the convergence of iterates of a transition function*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **37** (1977), 211–226.

8. D. Revuz, *Markov Chains*, North-Holland Publ. Co., 1975.

9. W. Winkler, *A note on continuous parameter zero-two law*, Ann. Probability **1** (1973), 341–344.

Université de Paris VII