CHAÎNES DE MARKOV INDEXÉES PAR $-\mathbf{N}$: existence et comportement

Jean BROSSARD et Christophe LEURIDAN

Résumé

Dans cet article, nous donnons des conditions pour qu'il existe une chaîne de Markov indexée par $-\mathbf{N}$ dont le noyau de transition Π est donné. Lorsque le noyau Π est irréductible et lorsque de telles chaînes existent, nous décrivons leurs comportements extrémaux. Nous montrons qu'il n'y a que deux types de comportements possibles : un comportement de type stationnaire, et un comportement de type transient, où le temps est une fonction déterministe de la position jusqu'à un instant aléatoire strictement supérieur à $-\infty$. Nous donnons des exemples illustrant ces situations.

1. Introduction

Soient $(E,\mathscr{E})$ un espace polonais muni de sa tribu borélienne, et Π un noyau de transition markovien sur $(E,\mathscr{E})$. L'existence d'une chaîne de Markov (indexée par $\mathbf{N}$) de noyau de transition Π est assurée par le théorème de Kolmogorov, et le comportement asymptotique d'une telle chaîne (récurrence, transience...) est bien classifié. Il en va tout autrement si on étudie les chaînes de Markov indexées par $-\mathbf{N}$, simplement parce qu'on ne peut plus parler de loi initiale.

D'après le théorème de Kolmogorov, l'existence d'une chaîne de Markov indexée par $-\mathbf{N}$ de noyau Π est équivalente à l'existence d'une loi d'entrée formée de probabilités, c'est-à-dire d'une famille de probabilités $(\mu_l)_{l\in-\mathbf{N}}$ telle que $\mu_l\Pi = \mu_{l+1}$ pour tout $l \in -\mathbf{N}^*$. Si on excepte le cas bien connu où le noyau Π possède une probabilité invariante, l'existence d'une telle famille de probabilités est loin d'être automatique, et est même plutôt rare.

Bien que de nombreux travaux de Dynkin, Maisonneuve, Fitzsimmons, Fourati,... portent sur les processus de Markov à naissance et mort aléatoire et les processus de Markov indexés par $\mathbf{R}$, la question de l'existence d'une chaîne de Markov indexée par $-\mathbf{N}$ pour un noyau de transition donné n'a curieusement pas été posée, et encore moins résolue. Notons que le problème de l'existence de processus de Markov indexés par $\mathbf{R}$ pour un semi-groupe donné se ramème immédiatement à celui de l'existence d'une chaîne de Markov indexée par $-\mathbf{N}$, puisque c'est un problème de loi d'entrée.

Dans la deuxième partie, nous donnons – entre autres résultats – une condition nécessaire et suffisante d'existence, sous l'hypothèse que les probabilités de transition sont absolument continues par rapport à une même mesure ν (cette hypothèse est discutée dans la cinquième partie).

Classification math. : 60J05.
Mots-clés : chaînes de Markov indexées par -$\mathbf{N}$, existence, lois d'entrée, comportement asymptotique.

Dans la troisième partie, nous supposons en plus que le noyau Π est irréductible, et nous décrivons alors les comportements « extrémaux » (c'est-à-dire rendant triviale la tribu asymptotique $\mathscr{F}_{-\infty}$) des Π-chaînes de Markov indexées par $-\mathbf{N}$. Nous montrons qu'il n'y a que deux types de comportements possibles : un comportement stationnaire aux phénomènes de périodicité près, et un comportement transient dans lequel le temps est une fonction déterministe de la position jusqu'à un instant aléatoire strictement supérieur à $-\infty$. Le caractère rigide de ce comportement entraîne sa rareté, mais nous montrons par des exemples qu'il peut se produire avec certains noyaux de transition, même s'ils possèdent une probabilité invariante. L'exemple 2 montre qu'il est même possible de construire une chaîne irréductible indexée par $\mathbf{Z}$, déterministe avant l'instant 0, et stationnaire ensuite.

Dans la quatrième partie, nous supposons que les probabilités de transition $\Pi(x,\cdot)$ dépendent continûment de x. Nous donnons une condition suffisante pour l'existence d'une Π-chaîne de Markov indexée par $-\mathbf{N}$. Cette condition, plus simple à formuler que la condition nécessaire et suffisante de la seconde partie est *grosso modo* que les probabilités $\Pi^n(x,\cdot)$ ne convergent pas uniformément vers la mesure nulle (pour la convergence vague), et que la chaîne ne peut pas arriver trop vite de l'infini.

Signalons enfin que les résultats des deuxième et quatrième parties, sont encore valables *mutatis mutandis* pour les chaînes de Markov inhomogènes car leur étude se ramène (pour les questions abordées dans ces parties) à celle de chaînes homogènes sur l'espace-temps $E \times (-\mathbf{N})$. Pour ne pas alourdir inutilement l'exposé et les notations, nous nous sommes cependant limités aux chaînes homogènes.

2. Remarques générales et résultats préliminaires

Dans cette partie $(E,\mathscr{E})$ est un espace polonais, muni de sa tribu borélienne. On se donne un noyau de transition markovien Π sur $(E,\mathscr{E})$ tel que les probabilités de transition $\Pi(x,\cdot)$ sont absolument continues par rapport à une même mesure ν. D'après le lemme 5.3 du chapitre 1 de [4], on peut trouver une application $\mathscr{E}\otimes\mathscr{E}$ - mesurable $p : E\times E \to \mathbf{R}_+$ telle que pour tout $x\in E$, la probabilité $\Pi(x,\cdot)$ admette l'application $p(x,\cdot)$ comme densité par rapport à la mesure ν. Plus généralement, pour tout $n\in\mathbf{N}^*$ et tout $x\in E$, la probabilité $\Pi^n(x,\cdot)$ admet comme densité l'application $p_n(x,\cdot)$ donnée par $p_1 = p$, et $p_{n+1}(x,y) = \int_E \Pi(x,dz)\,p_n(z,y)$ pour $n\in\mathbf{N}^*$ et $x,y\in E$.

Notons $(X_n)_{n\in-\mathbf{N}}$ le processus canonique sur $\Omega = E^{-\mathbf{N}}$, défini par $X_n(\omega) = \omega(n)$ pour $n\in-\mathbf{N}$ et $\omega\in\Omega$, et $(\mathscr{F}_n)_{n\in-\mathbf{N}}$ la filtration engendrée par le processus $(X_n)_{n\in-\mathbf{N}}$. Nous allons étudier l'ensemble $\mathscr{C}$ des probabilités sur Ω sous lesquelles le processus $(X_n)_{n\in-\mathbf{N}}$ est une chaîne de Markov de noyau Π dans la filtration canonique $(\mathscr{F}_n)_{n\in-\mathbf{N}}$. Commençons par rappeler quelques faits classiques, que on rencontre dans un contexte similaire dans [5] ou [6] :

* l'ensemble $\mathscr{C}$ est un convexe (éventuellement vide) de l'espace vectoriel des mesures signées sur Ω ;
* deux probabilités de $\mathscr{C}$ sont égales si et seulement si leurs restrictions à $\mathscr{F}_{-\infty}$ coïncident (ce point se démontre de façon analogue à la proposition 1 ci-dessous, en utilisant la propriété de Markov et le théorème de convergence pour les martingales à rebours) ;
* les points extrémaux de $\mathscr{C}$ sont les probabilités pour lesquelles la tribu asymptotique $\mathscr{F}_{-\infty}$ est triviale.

L'idée d'utiliser le théorème de convergence des martingales est naturelle pour ce type de problème (voir par exemple [3], ou même [2] dans une situation un peu différente). Nous cherchons principalement à savoir si l'ensemble $\mathscr{C}$ est vide ou non, et lorsqu'il ne l'est pas, à décrire le comportement de la chaîne $(X_n)_{n\in-\mathbf{N}}$ sous les lois extrémales de $\mathscr{C}$. Pour cela, un résultat-clé est le suivant :

Proposition 2.1 *Soit P une probabilité de $\mathscr{C}$. Alors pour P-presque tout $\omega \in \Omega$, la suite de fonctions $(p_n(\omega(-n),\cdot))_{n\geqslant 1}$ converge ν-presque partout et dans $L^1(\nu)$ vers une fonction $f_0(\omega,\cdot)$, où $(f_0(\omega,\cdot))_{\omega\in\Omega}$ est une version régulière de la densité par rapport à ν de X_0 sachant $\mathscr{F}_{-\infty}$. En particulier, si P est extrémale dans $\mathscr{C}$, la suite $(p_n(\omega(-n)),\cdot))_{n\geqslant 1}$ converge ν-presque partout et dans $L^1(\nu)$ vers la densité de X_0 par rapport à ν.*

Démonstration. On vérifie facilement que pour tout $y \in E$, la suite de variables aléatoires $(p_n(X_{-n},y))_{n\geqslant 1}$ est une martingale à rebours. Elle converge P-presque sûrement et dans $L^1(P)$ vers une variable aléatoire L_y. Grâce au théorème de Fubini, on voit que pour P-presque tout $\omega \in \Omega$, la suite de fonctions $(p_n(X_{-n}(\omega),\cdot))_{n\geqslant 1}$ converge ν-presque partout vers la fonction $y \mapsto L_y(\omega)$. Par positivité des fonctions, on a d'après le lemme de Fatou :

$$\int_E L_y(\omega)\nu(dy) \leqslant \liminf_{n\to+\infty}\int_E p_n(X_{-n}(\omega),y)\nu(dy) = 1.$$

Pour montrer la convergence dans $L^1(\nu)$, il suffit de vérifier que cette inégalité est en fait une égalité, ce qu'on fait en écrivant :

$$\mathbf{E}[\int_E L_y\nu(dy)] = \int_E \mathbf{E}[L_y]\nu(dy) = \int_E \mathbf{E}[p_1(X_{-1},y)]\nu(dy) = \mathbf{E}[\int_E p_1(X_{-1},y)\nu(dy)] = 1.$$

Il reste à montrer que la famille de fonctions $f_0(\omega,\cdot) : y \mapsto L_y(\omega)$ est une version régulière de la densité par rapport à ν de X_0 sachant $\mathscr{F}_{-\infty}$. Soit donc $B \in \mathscr{E}$. Pour tout $n \in \mathbf{N}^*$, on a presque sûrement : $P[X_0 \in B|\mathscr{F}_{-n}] = \int_B p_n(X_{-n},y)\nu(dy)$. Par passage à la limite, on a donc presque sûrement : $P[X_0 \in B|\mathscr{F}_{-\infty}] = \int_B L_y\nu(dy)$, ce qu'il fallait démontrer. □

Plus généralement, notons C l'événement formé des $\omega \in \Omega$ tels que pour tout $k \in \mathbf{N}$, la suites de fonctions $(p_{n-k}(\omega(-n),\cdot))_{n>k}$ converge dans $L^1(\nu)$. Alors la proposition 2.1 entraîne immédiatement que $P(C) = 1$ pour toute probabilité $P \in \mathscr{C}$. Le résultat suivant constitue une sorte de réciproque :

Lemme 2.2 *Si $\omega \in C$, alors il existe une probabilité et une seule dans $\mathscr{C}$ sous laquelle pour tout $k \in \mathbf{N}$, la variable X_{-k} admette comme densité par rapport à ν la fonction f_{-k}, où f_{-k} est la limite dans $L^1(\nu)$ de la suite de fonctions $(p_{n-k}(\omega(-n),\cdot))_{n>k}$.*

Démonstration. L'unicité est évidente, la loi de $(X_{-k}, X_{-k+1}, \cdots, X_0)$ étant imposée pour tout $k \in \mathbf{N}$. L'existence repose sur le théorème de Kolmogorov, en remarquant que la famille de probabilités $(f_l.\nu)_{l\in-\mathbf{N}}$ est une loi d'entrée, au sens où $(f_l.\nu)\Pi = f_{l+1}.\nu$ pour tout $l \in -\mathbf{N}^*$. □

On déduit immédiatement de la proposition 2.1 et du lemme 2.2 une condition nécessaire et suffisante (mais peu pratique) pour l'existence d'une chaîne de Markov de noyau Π indexée par $-\mathbf{N}$:

Corollaire 2.3 *Pour que l'ensemble $\mathscr{C}$ soit non-vide, il faut et il suffit que l'événement C soit non-vide.*

Pour $\omega \in C$, notons Q^ω la probabilité définie par le lemme 2.2, et lorsque $\omega \in \Omega\backslash C$, prenons Q^ω égale à une probabilité fixée sur Ω (non nécessairement dans $\mathscr{C}$). Nous pouvons alors énoncer un résultat de décomposition de toute probabilité de $\mathscr{C}$ en probabilités de la forme Q^ω. Les corollaires 2.4 et 2.5 sont dus à Dynkin (voir le paragraphe « décomposition into ergodic measures », pages 269 à 275 de [3]), mais nous en donnons ici une démonstration plus concise.

COROLLAIRE 2.4 *Soit $P \in \mathscr{C}$. Alors la famille de probabilités $(Q^\omega)_{\omega\in\Omega}$ est une version régulière de la loi conditionnelle de P par rapport à $\mathscr{F}_{-\infty}$.*

Démonstration. Soient $k \in \mathbf{N}$ et $f_0, \cdots, f_k : E \mapsto \mathbf{R}$ des applications $\mathscr{E}$-mesurables bornées. Alors pour tout $n \geqslant k$,

$$\mathbf{E}[\, f_0(X_{-k}) \cdots f_k(X_0)\,|\,\mathscr{F}_{-n}] = \int_E p_{n-k}(X_{-n}, y) g(y) \nu(dy),$$

où $g = f_0(\Pi(f_1\Pi(f_2 \cdots \Pi(f_{k-1}\Pi f_k) \cdots)))$. En faisant tendre n vers l'infini, on obtient pour tout $\omega \in C$,

$$\mathbf{E}[\, f_0(X_{-k}) \cdots f_k(X_0)\,|\,\mathscr{F}_{-\infty}](\omega) = Q^\omega[\, f_0(X_{-k}) \cdots f_k(X_0)].$$

$\square$

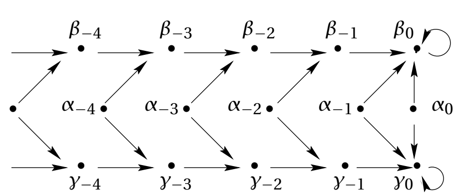
On déduit du corollaire 2.4 les faits suivants :

COROLLAIRE 2.5 *Toute probabilité extrémale de $\mathscr{C}$ est de la forme Q^ω. Réciproquement, si $P \in \mathscr{C}$, alors pour P-presque tout $\omega \in C$, la probabilité Q^ω est extrémale (car la tribu $\mathscr{F}_{-\infty}$ est Q^ω-triviale).*

Démonstration. Le seul point non évident est que pour P-presque tout $\omega \in C$, la tribu $\mathscr{F}_{-\infty}$ est Q^ω-triviale (bien qu'il soit facile de montrer que pour tout $A \in \mathscr{F}_{-\infty}$, $Q^\omega(A) \in \{0\,;1\}$ pour P-presque tout $\omega \in C$, l'interversion du « pour tout A » et du « pour presque tout ω » pose problème car la tribu $\mathscr{F}_{-\infty}$ n'est pas séparable). Nous allons le montrer en appliquant le corollaire 2.4 aux probabilités Q^ω, ce qui est possible car Q^ω appartient à $\mathscr{C}$ pour $\omega \in C$.

Soit donc $P \in \mathscr{C}$. Une simple application du théorème des classes monotones montre que pour toute application $\mathscr{F}_{-\infty} \otimes \mathscr{F}_0$-mesurable bornée $f : E \times E \mapsto \mathbf{R}$, la variable aléatoire $\omega \mapsto \int_\Omega f(\omega, \omega') Q^\omega(d\omega')$ est une espérance conditionnelle (sous la probabilité P) par rapport à $\mathscr{F}_{-\infty}$ de la variable aléatoire $\omega \mapsto f(\omega, \omega)$. En appliquant ce résultat à la fonction $f : (\omega, \omega') \mapsto |Q^{\omega'}(A) - Q^\omega(A)|$ pour $A \in \mathscr{F}_0$, on obtient grâce à la séparabilité de $\mathscr{F}_0$ que pour P-presque tout $\omega \in C$ et pour Q^ω-presque tout $\omega' \in \Omega$, $Q^{\omega'} = Q^\omega$. Mais comme pour un tel $\omega \in C$, la famille $(Q^{\omega'})_{\omega'\in\Omega}$ est une version régulière de loi conditionnelle de Q^ω par rapport à $\mathscr{F}_{-\infty}$, on en déduit que la tribu $\mathscr{F}_{-\infty}$ est Q^ω-triviale. $\square$

Il se peut cependant qu'une probabilité de la forme Q^ω ne soit pas extrémale, comme le montre l'exemple de la « fermeture-éclair » :



Dans le schéma ci-dessus, les flèches horizontales et en boucle indiquent des probabilités de transition égales à 1, et les autres des probabilités de transition égales à 1/2. Les probabilités Q^β et Q^γ sont extrémales, mais $Q^\alpha = 1/2(Q^\beta + Q^\gamma)$ ne l'est pas.

3. Description des lois extrémales des chaînes irréductibles

On garde dans cette partie les notations et les hypothèses de la partie précédente (en fait, il n'est pas nécessaire de supposer que $(E, \mathscr{E})$ est un espace polonais : il suffit que la tribu $\mathscr{E}$ soit séparable, car les résultats de cette partie n'utilisent que la proposition 2.1, qui ne fait pas appel au théorème de Kolmogorov). On suppose en outre que le noyau Π est ν-irréductible, c'est-à-dire que pour tout $x \in E$, la mesure ν est absolument continue par rapport à la probabilité $\sum_{n\in\mathbf{N}^*} 2^{-n}\Pi^n(x, \cdot)$. D'après les théorèmes 2.6, 2.7 et 2.10 du chapitre 3 de [4], il y a deux possibilités :

* soit il existe une suite croissante $(B_m)_{m\in\mathbf{N}}$ de parties mesurables de $(E, \mathscr{E})$ dont la réunion est égale à E, et telle que pour tout $m \in \mathbf{N}$, la fonction $x \mapsto \sum_{n\in\mathbf{N}} \Pi^n(x, B_m)$ est bornée ; nous dirons alors que le noyau Π est **transient** ;
* soit il existe une mesure σ-finie non nulle m sur $(E, \mathscr{E})$, invariante pour le noyau Π, et pour laquelle tout ensemble de mesure non nulle est récurrent. Cette mesure est unique à un facteur multiplicatif près ; nous dirons alors que le noyau Π est récurrent au sens de Harris, **positivement récurrent** si $m(E) < +\infty$ et **nul-récurrent** si $m(E) = +\infty$.

Notations. Nous noterons d la période lorsque le noyau est récurrent (voir le théorème 1.3 et la définition 2.4 du chapitre 6 de [4]). Lorsque le noyau est transient, nous poserons $d = 1$. Si μ_1 et μ_2 sont deux probabilités sur $(E, \mathscr{E})$, on notera $\|\mu_2 - \mu_1\|$ leur **distance en variation totale**, définie par : $\|\mu_2 - \mu_1\| = |\mu_2 - \mu_1|(E)$.

Nous pouvons alors énoncer un résultat simple :

LEMME 3.1 *Soit μ une probabilité sur $(E, \mathscr{E})$. Si $\mu\Pi^d \neq \mu$, alors* $\inf\{\|\mu\Pi^n - \mu\| \,;\, n \in \mathbf{N}^*\} > 0$.

Démonstration. Si le noyau Π est transient ou nul-récurrent, on a alors $\|\mu\Pi^n - \mu\| > 0$ pour tout $n \in \mathbf{N}^*$, et $\|\mu\Pi^n - \mu\| \to 2$ quand $n \to +\infty$, donc $\inf\{\|\mu\Pi^n - \mu\| \,;\, n \in \mathbf{N}^*\} > 0$.

Si le noyau Π est positivement récurrent, alors d'après le théorème d'Orey (voir [4], chapitre 6, théorème 2.8 et corollaire 2.9), les suites $(\mu\Pi^{nd})_{n\in\mathbf{N}}$, $(\mu\Pi^{nd+1})_{n\in\mathbf{N}}$, $\cdots$, $(\mu\Pi^{nd+d-1})_{n\in\mathbf{N}}$ convergent vers des mesures m_0 , m_1 , $\cdots$, m_{d-1} invariantes pour le noyau Π^d. Donc

$$\liminf_{n\to+\infty} \|\mu\Pi^n - \mu\| = \min\{\|m_0 - \mu\| \,,\, \|m_1 - \mu\|, \cdots , \|m_{d-1} - \mu\|\} > 0.$$

Par ailleurs, on ne peut pas avoir $\mu\Pi^k = \mu$ pour un certain $k \in \mathbf{N}^*$, car cela entraînerait l'égalité $\mu\Pi^{nkd} = \mu$ pour tout $n \in \mathbf{N}^*$, d'où à la limite $m_0 = \mu$, ce qui contredirait l'hypothèse $\mu\Pi^d \neq \mu$. Donc $\inf\{\|\mu\Pi^n - \mu\| \,;\, n \in \mathbf{N}^*\} > 0$. $\square$

Nous pouvons maintenant énoncer et démontrer le théorème décrivant les comportements possibles du processus canonique sous les lois extrémales de $\mathscr{C}$.

THÉORÈME 3.2 *Soit Q une loi extémale de $\mathscr{C}$. Alors de deux choses l'une :*

* *ou bien la chaîne $(X_{dn})_{n\in-\mathbf{N}}$ est stationnaire ;*

* *ou bien il existe une suite $(E_n)_{n\in\mathbf{N}}$ de parties mesurables de $(E,\mathscr{E})$, deux-à-deux disjointes, telles que Q-presque sûrement, $X_{-n}\in E_n$ à partir d'un certain rang.*

Remarques.

* Il est évident que le premier comportement (de type stationnaire) ne peut être observé que si le noyau Π est positivement récurrent. Un fait plus surprenant est que le second comportement (de type transient) puisse se produire même pour un noyau Π positivement récurrent. Nous verrons dans l'exemple 2 qu'il est possible de construire une chaîne de Markov indexée par $\mathbf{Z}$ de type transient vers l'instant $-\infty$ mais stationnaire à partir de l'instant 0.

* Le second cas (comportement de type transient) entraîne que pour Q-presque tout ω, les valeurs de la suite $\omega(-n)_{n\in\mathbf{N}}$ sont toutes distinctes à partir d'un certain rang. Plus précisément, la chaîne $(X_n)_{n\in-\mathbf{N}}$ possède une horloge interne, au sens où le temps est une fonction déterministe de la position jusqu'à un certain temps d'arrêt presque sûrement différent de $-\infty$. Ce déterminisme de la transience montre que l'existence d'une chaîne de Markov indexée par $-\mathbf{N}$ est un phénomène rare, si on excepte le cas bien connu des chaînes de type stationnaire.

Démonstration. Supposons que sous la probabilité Q, la chaîne $(X_{dn})_{n\in-\mathbf{N}}$ n'est pas stationnaire, et montrons alors qu'elle est de type transient. On peut trouver $n_0\in-\mathbf{N}$ tel que la loi de la variable X_{n_0} n'est pas invariante pour le noyau Π^d. Quitte à considérer la chaîne $(X_{n_0+n})_{n\in-\mathbf{N}}$ (dont la loi est encore une probabilité extrémale de $\mathscr{C}$) au lieu de $(X_n)_{n\in-\mathbf{N}}$, on peut supposer que $n_0=0$.

Notons alors μ_0 la loi de la variable X_0 et f_0 la densité de μ_0 par rapport à ν. D'après le lemme 3.1, le réel $\alpha=\inf\{\|\mu_0\Pi^n-\mu_0\|;\, n\in\mathbf{N}^*\}$ est strictement positif. On définit alors une suite $(E_n)_{n\in\mathbf{N}}$ de partie mesurables de $(E,\mathscr{E})$ par :

$$E_n=\{x\in E : \|\Pi^n(x,\cdot)-\mu_0\|<\alpha/2\}.$$

On sait d'après la proposition 2.1 que pour Q-presque tout $\omega\in\Omega$,

$$\|\Pi^n(\omega(-n),\cdot)-\mu_0\|=\|p_n(\omega(-n),\cdot)-f_0\|_{L^1(\nu)}\to 0 \text{ quand } n\to+\infty,$$

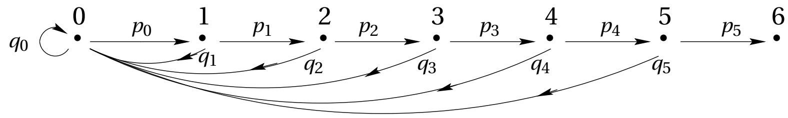
ce qui entraîne que $\omega(-n)\in E_n$ à partir d'un certain rang.

Il reste à montrer que les ensembles E_n sont deux-à-deux disjoints. S'il n'en était pas ainsi, on pourrait trouver un point x appartenant à deux ensembles E_n et E_{n+k}, avec $n\in\mathbf{N}$ et $k\in\mathbf{N}^*$. On aurait alors d'une part $\|\Pi^{n+k}(x,\cdot)-\mu_0\|<\alpha/2$, et d'autre part

$$\|\Pi^{n+k}(x,\cdot)-\mu_0\Pi^k\|=\|\Pi^n(x,\cdot)\Pi^k-\mu_0\Pi^k\|\leqslant\|\Pi^n(x,\cdot)-\mu_0\|<\alpha/2,$$

d'où $\|\mu_0\Pi^k-\mu_0\|<\alpha$, ce qui contredirait la définition de α. □

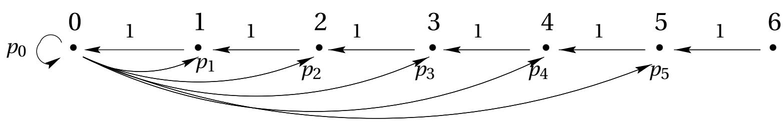
Exemple 1. Sur $\mathbf{N}$, on considère le noyau de transition Π donné par $p(n,n+1) = p_n$ et $p(n,0) = q_n = 1 - p_n$ pour tout $n \in \mathbf{N}$ où $p_n \in]0\,;1[$ pour tout $n \in \mathbf{N}$.



Le noyau Π est irréductible apériodique récurrent. On montre facilement que la mesure invariante m vérifiant $m(\{0\}) = 1$ est donnée par $m(\{n\}) = p_0 \cdots p_{n-1}$ pour tout $n \in \mathbf{N}$.

Si l'ensemble $\mathscr{C}$ n'est pas vide, considérons une probabilité extrémale Q de $\mathscr{C}$. Alors pour Q-presque tout $\omega \in \Omega$, on a $\omega(n - \omega(n)) = 0$ pour tout $n \in -\mathbf{N}$, ce qui entraîne que la trajectoire ω passe une infinité de fois en 0. D'après le théorème 3.2, la chaîne $(X_n)_{n\in-\mathbf{N}}$ est stationnaire, ce qui entraîne que la mesure invariante m est finie. Par conséquent, si $\sum_{n\in\mathbf{N}} p_0 \cdots p_{n-1} < +\infty$, alors l'ensemble $\mathscr{C}$ est le singleton formé de la loi stationnaire ; sinon l'ensemble $\mathscr{C}$ est vide.

Exemple 2. Sur $\mathbf{N}$, on considère le noyau de transition Π donné par $p(n,n-1) = 1$ pour tout $n \in \mathbf{N}^*$ et $p(0,n) = p_n$ pour tout $n \in \mathbf{N}$, les réels p_n étant strictement positifs de somme 1.



Le noyau Π est irréductible apériodique récurrent. On montre facilement que la mesure invariante m vérifiant $m(\{0\}) = 1$ est donnée par

$$m(\{n\}) = 1 - \sum_{k=0}^{n-1} p_k = \sum_{k=n}^{+\infty} p_k \text{ pour tout } n \in \mathbf{N}.$$

En remarquant que $\sum_{n\in\mathbf{N}} m(\{n\}) = \sum_{k\in\mathbf{N}} (k+1)\,p_k$, on voit que le noyau Π est positivement récurrent si et seulement si la somme $\sum_{k\in\mathbf{N}} k\,p_k$ est finie. Pour $k \in \mathbf{Z}$, notons ω_k la trajectoire définie par $\omega_k(-n) = \max(k+n, 0)$ pour tout $n \in \mathbf{N}$. On montre grâce au théorème 3.2 et au corollaire 2.5 que les lois extrémales de $\mathscr{C}$ sont les lois Q^{ω_k} auxquelles il faut ajouter la loi stationnaire si $\sum_{k\in\mathbf{N}} k\,p_k < +\infty$.

Remarques sur l'exemple 2.

* On peut prendre les probabilités de transition $p(n,n-1)$ différentes de 1 pourvu que $\prod_{n\in\mathbf{N}^*} p(n,n-1) > 0$;

* Sous la probabilité $Q^{\omega_{-1}}$, on a $X_{-n} = n-1$ presque sûrement pour tout $n \in \mathbf{N}^*$, et la variable X_0 suit la loi $\sum_{n\in\mathbf{N}} p_n\delta_n$. Si on prend $p_n = (1-r)r^n$ pour tout $n \in \mathbf{N}$, avec $r \in]0\,;1[$ fixé, on voit que sous la probabilité $Q^{\omega_{-1}}$, la loi de X_0 est la probabilité invariante bien que la chaîne $(X_n)_{n\in-\mathbf{N}^*}$ soit déterministe !

4. Une condition suffisante d'existence

Nous avons vu dans la seconde partie (corollaire 2.3) une condition nécessaire et suffisante pour qu'il existe une chaîne de Markov indexée par $-\mathbf{N}$ de noyau Π donné. Cette condition étant complexe dans sa formulation, nous allons voir une condition plus simple.

Dans cette partie $(E, \mathscr{E})$ est un espace polonais muni de sa tribu borélienne, et Π un noyau de transition markovien sur $(E, \mathscr{E})$. On note toujours $\mathscr{C}$ l'ensemble des probabilités sur $\Omega = E^{-\mathbf{N}}$ sous lesquelles le processus canonique $(X_n)_{n\in-\mathbf{N}}$ est une chaîne de Markov de noyau Π.

Une obstruction fréquente à l'existence d'une Π-chaîne de Markov indexée par $-\mathbf{N}$ est le fait que quand n tend vers $+\infty$, les probabilités de transition $\Pi^n(x,\cdot)$ convergent vaguement vers la mesure nulle *uniformément par rapport à* $x \in E$. Ce fait est l'objet du lemme suivant :

LEMME 4.1 *Supposons que pour tout compact K de E, la suite de fonctions $x \mapsto \Pi^n(x,K)$ converge uniformément vers la fonction nulle sur E. Alors l'ensemble $\mathscr{C}$ est vide : il n'existe pas de Π-chaîne de Markov indexée par $-\mathbf{N}$.*

Démonstration. Si $\mathscr{C}$ contenait une probabilité P, on aurait $P[X_0 \in K] = \mathbf{E}[\Pi^n(X_{-n}, K)]$ pour tout compact K de E et pour tout $n \in \mathbf{N}$. En faisant tendre n vers $+\infty$, on obtiendrait $P[X_0 \in K] = 0$ pour tout compact K, ce qui est absurde car la loi de X_0 est tendue. □

On déduit du lemme qu'une condition *nécessaire* à l'existence d'une Π-chaîne de Markov indexée par $-\mathbf{N}$ est qu'il existe un compact K tel que la suite de fonctions $x \mapsto \Pi^n(x,K)$ ne converge pas uniformément vers la fonction nulle sur E. Nous allons voir que cette condition est *suffisante* sous les hypothèses supplémentaires assez générales suivantes :

(1) les probabilités de transition $\Pi(x,\cdot)$ dépendent continûment de $x \in E$ pour la convergence étroite ;

(2) pour tout compact K de E, $\Pi(x,K)$ tend vers 0 quand x tend vers l'infini dans le compactifié d'Alexandrov de E.

THÉORÈME 4.2 *Si les hypothèses (1) et (2) sont satisfaites et s'il existe un compact K tel que la suite de fonctions $x \mapsto \Pi^n(x,K)$ ne converge pas uniformément vers la fonction nulle sur E, alors l'ensemble $\mathscr{C}$ n'est pas vide : il existe donc une Π-chaîne de Markov indexée par $-\mathbf{N}$.*

Pour démontrer le théorème, nous aurons besoin du lemme suivant :

LEMME 4.3 *Les hypothèses (1) et (2) entraînent que pour tout $n \in \mathbf{N}^*$, tout compact K de E, et tout $\epsilon > 0$, l'ensemble $\{x \in E : \Pi^n(x,K) \geqslant \epsilon\}$ est compact.*

Démonstration. Les ensembles $\{x \in E : \Pi^n(x,K) \geqslant \epsilon\}$ sont fermés par semi-continuité supérieure des fonctions $x \mapsto \Pi^n(x,K)$, puisque les probabilités $\Pi^n(x,\cdot)$ dépendent continûment de x. On démontre par récurrence sur $n \in \mathbf{N}^*$ que pour tout compact K de E et tout $\epsilon > 0$, l'ensemble $\{x \in E : \Pi^n(x,K) \geqslant \epsilon\}$ est relativement compact. Cette propriété est vraie au rang 1 d'après l'hypothèse (2). Pour passer du rang n au rang $n+1$, on écrit que pour tout $x \in E$,

$$\Pi^{n+1}(x,K) = \int_E \Pi^n(x,dy)\Pi(y,K) \leqslant \frac{\epsilon}{2} + \Pi^n(x,K'),$$

où $K' = \{y \in E : \Pi(y, K) \geqslant \epsilon/2\}$ est compact, d'où :

$$\{x \in E : \Pi^{n+1}(x, K) \geqslant \epsilon\} \subset \{x \in E : \Pi^n(x, K') \geqslant \epsilon/2\}.$$

□

Le théorème se démontre alors en deux étapes, qui font l'objet des lemmes ci-dessous :

LEMME 4.4 *Sous les hypothèses du théorème, il existe une suite strictement croissante d'entiers naturels $(s(n))_{n\in\mathbf{N}}$ et une suite de compacts non vides $(K_n)_{n\in\mathbf{N}}$ telles que le produit infini $\prod_{n\in\mathbf{N}} r_n$ converge, où $r_n = \inf_{x\in K_{n+1}} \Pi^{s(n+1)-s(n)}(x, K_n)$.*

Démonstration. Soit K un compact tel que la suite de fonctions $x \mapsto \Pi^n(x, K)$ ne converge pas uniformément vers la fonction nulle sur E. Posons alors

$$\alpha = \limsup_{n\to+\infty} \big(\sup_{x\in E} \Pi^n(x, K)\big) > 0.$$

Fixons une suite strictement décroissante $(\epsilon_n)_{n\in\mathbf{N}}$ de réels tendant vers 0 telle que le produit infini $\prod_{n\in\mathbf{N}} \frac{\epsilon_n-\epsilon_{n+1}}{\epsilon_n+\epsilon_{n+1}}$ converge (on peut prendre par exemple $\epsilon_n = 2^{-n^2}$ pour tout $n \in \mathbf{N}$). On peut alors trouver une suite strictement croissante d'entiers naturels $(s(n))_{n\in\mathbf{N}}$ telle que pour tout $n \in \mathbf{N}$,

$$\alpha(1 - \epsilon_{n+1}) \leqslant \sup_{x\in E} \Pi^{s(n)}(x, K) \leqslant \alpha(1 + \epsilon_{n+1}).$$

Posons alors $K_n = \{x \in E : \Pi^{s(n)}(x, K) \geqslant \alpha(1 - \epsilon_n)\}$. Comme $\alpha(1 - \epsilon_{n+1}) > \alpha(1 - \epsilon_n)\}$, le compact K_n n'est pas vide. Par ailleurs, pour tout $n \in \mathbf{N}$ et tout $x \in K_{n+1}$:

$$\begin{aligned}
\alpha(1 - \epsilon_{n+1}) &\leqslant \Pi^{s(n+1)}(x, K) \\
&= \int_{K_n} \Pi^{s(n+1)-s(n)}(x, dy)\Pi^{s(n)}(y, K) + \int_{K_n^c} \Pi^{s(n+1)-s(n)}(x, dy)\Pi^{s(n)}(y, K) \\
&\leqslant \alpha(1 + \epsilon_{n+1})\Pi^{s(n+1)-s(n)}(x, K_n) + \alpha(1 - \epsilon_n)\Pi^{s(n+1)-s(n)}(x, K_n^c) \\
&= \alpha(1 - \epsilon_n) + \alpha(\epsilon_n + \epsilon_{n+1})\Pi^{s(n+1)-s(n)}(x, K_n).
\end{aligned}$$

Donc

$$\epsilon_n - \epsilon_{n+1} \leqslant (\epsilon_n + \epsilon_{n+1})\Pi^{s(n+1)-s(n)}(x, K_n),$$

d'où :

$$\Pi^{s(n+1)-s(n)}(x, K_n) \geqslant \frac{\epsilon_n - \epsilon_{n+1}}{\epsilon_n + \epsilon_{n+1}},$$

ce qui achève la preuve. □

LEMME 4.5 *Supposons que les probabilités de transition $\Pi(x,\cdot)$ dépendent continûment de $x \in E$ (pour la convergence étroite), et qu'il existe une suite strictement croissante d'entiers naturels $(s(n))_{n\in\mathbf{N}}$ et une suite de compacts non vides $(K_n)_{n\in\mathbf{N}}$ telles que le produit infini $\prod_{n\in\mathbf{N}} r_n$ converge, où $r_n = \inf_{x\in K_{n+1}} \Pi^{s(n+1)-s(n)}(x, K_n)$. Alors l'ensemble $\mathscr{C}$ n'est pas vide.*

Démonstration. A l'aide des suites $(s(n))_{n\in\mathbf{N}}$ et $(K_n)_{n\in\mathbf{N}}$, nous allons construire une loi d'entrée formée de probabilités sur $(E, \mathscr{E})$. Tout d'abord, une récurrence immédiate montre que pour tous entiers $n \geqslant N \geqslant 0$, et tout point $x \in K_n$, $\Pi^{s(n)-s(N)}(x, K_N) \geqslant \prod_{k\in[N\cdots n-1]} r_k$.

Choisissons pour tout $n \in \mathbf{N}$ un point $a_n \in K_n$, et posons $\mu_{-l,n} = \Pi^{s(n)-l}(a_n, \cdot)$ pour l, n entiers tels que $n \geqslant l \geqslant 0$ (d'où $s(n) \geqslant l$). Alors pour tout $l \in \mathbf{N}$, la suite de probabilités $(\mu_{-l,n})_{n \geqslant l}$ est tendue.

En effet, pour tout $\epsilon > 0$, on peut trouver un entier naturel $N \geqslant l$ tel que $\prod_{n \geqslant N} r_n \geqslant 1 - \epsilon$. Pour tout entier $n \geqslant N$, on a alors :

$$\Pi^{s(n)-s(N)}(a_n, K_N) \geqslant \prod_{k \in [N \cdots n-1]} r_k \geqslant \prod_{k \geqslant N} r_k \geqslant 1 - \epsilon.$$

Or la famille de probabilités $(\Pi^{s(N)-l}(x, \cdot))_{x \in K_N}$ est compacte pour la topologie de la convergence étroite, par continuité de l'application $x \mapsto \Pi^{s(N)-l}(x, \cdot))$. Elle est donc tendue d'après le critère de Prokhorov (voir[1], page 37). On peut donc trouver un compact K' tel que pour tout $x \in K_N$, $\Pi^{s(N)-l}(x, K') \geqslant 1 - \epsilon$. Pour tout entier $n \geqslant N$, on a ainsi :

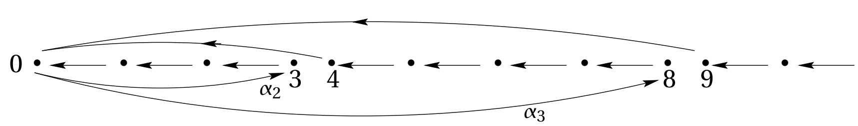
$$\mu_{-l,n}(K') \geqslant \int_{K_N} \Pi^{s(n)-s(N)}(a_n, dx)\Pi^{s(N)-l}(x, K') \geqslant (1-\epsilon)\Pi^{s(n)-s(N)}(a_n, K_N) \geqslant (1-\epsilon)^2,$$

ce qui montre que la suite de probabilités $(\mu_{-l,n})_{n \geqslant l}$ est tendue.

Par le procédé diagonal, on construit une suite strictement croissante $(\phi(n))_{n \in \mathbf{N}}$ d'entiers naturels telle pour tout $l \in \mathbf{N}$, la suite de probabilités $(\mu_{-l,\phi(n)})_{n \geqslant l}$ converge étroitement vers une probabilité μ_{-l}. On vérifie que la famille de probabilités $(\mu_k)_{k \in -\mathbf{N}}$ est une loi d'entrée pour le noyau Π en passant à la limite dans les égalités $\mu_{k+1,n} = \mu_{k,n}\Pi$ (pour le second membre, on utilise la continuité des probabilités $\Pi(x, \cdot)$ vis-à-vis de $x \in E$). $\square$

Remarques concernant les hypothèses (1) et (2). Le théorème 4.2 ne donne qu'une condition *suffisante* pour l'existence d'une Π-chaîne de Markov indexée par $-\mathbf{N}$. Cette condition suffisante est formée d'une condition *nécessaire* à laquelle on a ajouté les hypothèses (1) et (2). Nous montrons que les conditions (1) et (2) ne sont pas superflues (à défaut d'être nécessaires) au moyen de deux contre-exemples, dans lesquels seule la condition (1) (exemple 3) ou la condition (2) (exemple 4) n'est pas satisfaite.

Exemple 3. On prend $E = \mathbf{N}$, et on pose $\Pi(x, \cdot) = \delta_{x-1}$ si x n'est pas un carré parfait, $\Pi(x, \cdot) = \delta_0$ si x est un carré parfait non nul, et $\Pi(0, \cdot) = \sum_{n \geqslant 2} \alpha_n \, \delta_{n^2-1}$, où les réels α_n sont strictement positifs et de somme 1.



Le noyau Π est irréductible et récurrent. Il est nul-récurrent si $\sum_{n \geqslant 2} n\alpha_n = +\infty$. Dans ce cas, on démontre qu'il n'existe pas de Π-chaîne de Markov indexée par $-\mathbf{N}$ comme dans l'exemple 1 de la troisième partie. Pourtant, les probabilités $\Pi(x, \cdot)$ dépendent continûment de x, et $\Pi^{2n-1}(n^2 - 1, \{0\}) = 1$ pour tout $n \geqslant 2$, ce qui montre que les fonctions $x \mapsto \Pi^n(x, \{0\})$ ne convergent pas uniformément vers 0.

Exemple 4. On prend $E = [0\,;1]$, et on pose $\Pi(x,\cdot) = \delta_{g(x)}$ pour tout $x \in [0\,;1]$, où $g(x) = x/2$ si $x > 0$ et $g(0) = 1$. On vérifie facilement qu'il n'existe pas de « suite » $(x_n)_{n\in-\mathbf{N}}$ vérifiant la relation de récurrence $x_{n+1} = g(x_n)$ pour tout $n \in -\mathbf{N}^*$, ce qui entraîne qu'il n'existe pas de Π-chaîne de Markov indexée par $-\mathbf{N}$, bien que $\Pi^n(x,[0\,;1]) = 1$ pour tout $n \in \mathbf{N}$ et tout $x \in [0\,;1]$.

Voici pour finir un exemple simple illustrant les théorèmes précédents :

Exemple 5. On prend $E = \mathbf{Z}$, et on pose $\Pi(x,\cdot) = p_x\delta_{x+1} + q_x\delta_{x-1} + r_x\delta_x$ pour $x \in \mathbf{Z}$, où p_x, q_x, r_x sont des réels positifs de somme 1 tels que $p_xq_x > 0$. Le théorème 3.2 montre alors que si Q est une loi extrémale de $\mathscr{C}$ sous laquelle la chaîne $(X_{2n})_{n\in-\mathbf{N}}$ n'est pas stationnaire, alors il existe $\epsilon \in \{-1\,;1\}$ et $a \in \mathbf{Z}$ tel qu'on ait Q-presque sûrement, $X_{-n} = \epsilon n + a$ à partir d'un certain rang. En utilisant le théorème 4.2 on montre ainsi qu'une condition nécessaire et suffisante pour l'existence d'une Π-chaîne de Markov indexée par $-\mathbf{N}$ de type transient est que $\prod_{x\in\mathbf{N}} q_x > 0$ ou que $\prod_{x\in-\mathbf{N}} p_x > 0$.

5. Remarques sur l'hypothèse d'absolue continuité de la seconde partie

Dans la seconde partie, nous avons supposé que les probabilités de transition $\Pi(x,\cdot)$ sont absolument continues par rapport à une même mesure ν. Sous cette hypothèse, nous avons démontré (proposition 2.1) que si $(X_n)_{n\in\mathbf{N}}$ est une Π-chaîne de Markov indexée par $-\mathbf{N}$, alors presque sûrement, la suite de probabilités $\Pi^n(X_{-n},\cdot)$ converge en variation totale. Le contre-exemple suivant montre que ce résultat n'est plus vrai sans l'hypothèse d'absolue continuité.

Exemple. Soient $E = \mathbf{R}/\mathbf{Z}$, et $\alpha \in E$ un irrationnel fixé. Posons pour tout $x \in E$, $\Pi(x,\cdot) = 1/2(\delta_{(x+\alpha)}+\delta_{(x-\alpha)})$. Soit Q la probabilité sur $E^{-\mathbf{N}}$ faisant du processus canonique une Π-chaîne de Markov stationnaire. Alors pour tout $\omega \in \Omega$, on peut construire un borélien $B(\omega)$ tel que la suite $(\Pi^n(\omega(-n), B(\omega)))_{n\in\mathbf{N}}$ n'ait pas de limite.

Démonstration. Le point-clé est que si on pose $\mu_n = \Pi^n(\omega(-n),\cdot)$, alors chaque mesure μ_n est portée par un ensemble fini, mais $\mu_n(F) \to 0$ quand $n \to +\infty$ pour toute partie finie F de E. On construit alors par récurrence une suite strictement croissante d'entiers $(s(n))_{n\in\mathbf{N}}$ et une suite croissante $(F_n)_{n\in\mathbf{N}}$ de parties finies de E : on pose $s(0) = 0$, $F_{-1} = \varnothing$; une fois construits les entiers $s(0) < s(1) < \cdots < s(n)$, on prend F_n égale à la réunion des supports de $\mu_{s(0)}, \mu_{s(1)}, \cdots, \mu_{s(n)}$, et on choisit $s(n+1) > s(n)$ de telle sorte que $\mu_{s(n+1)}(F_n) \leqslant 1/3$. En posant $B = \bigcup_{n\in\mathbf{N}}(F_{2n}\backslash F_{2n-1})$, on a alors $\mu_{s(n)}(B) \geqslant \mu_{s(n)}(F_n\backslash F_{n-1}) \geqslant 2/3$ lorsque n est pair et $\mu_{s(n)}(B) \leqslant \mu_{s(n)}(F_{n-1}) \leqslant 1/3$ lorsque n est impair, ce qui montre que la suite $(\mu_n(B))_{n\in\mathbf{N}}$ diverge. □

Toutefois, si $(E,\mathscr{E})$ est un espace localement compact à base dénombrable, alors presque sûrement, les probabilités $\Pi^n(X_{-n},\cdot)$ convergent étroitement. En effet, soit $(f_m)_{m\in\mathbf{N}}$ une suite totale de fonctions continues à support compact de E dans $\mathbf{R}$. En utilisant la convergence presque sûre des martingales à rebours $(\Pi^n f_m(X_{-n}))_{n\in\mathbf{N}}$, on montre la convergence vague des probabilités $\Pi^n(\omega(-n),\cdot)$ vers une mesure $\mu(\omega,\cdot)$ pour P-presque tout $\omega \in \Omega$. Il reste à montrer que pour P-presque tout $\omega \in \Omega$, la sous-probabilité $\mu(\omega,\cdot)$ est une probabilité. Pour cela, on utilise une suite croissante $(K_m)_{m\in\mathbf{N}}$ de compacts dont la réunion est E, et on écrit :

$$\begin{aligned}
\mathbf{E}[\ \mu(\cdot, E)\ ] &= \lim_{m\to+\infty} \mathbf{E}[\ \mu(\cdot, K_m)\ ] \\
&\geqslant \lim_{m\to+\infty} \mathbf{E}[\ \limsup_{n\to+\infty}\ \Pi^n(X_{-n}, K_m)\ ] \\
&\geqslant \lim_{m\to+\infty} \limsup_{n\to+\infty}\ \mathbf{E}[\ \Pi^n(X_{-n}, K_m)\ ] \\
&= \lim_{m\to+\infty} P[X_0 \in K_m] = 1.
\end{aligned}$$

En supposant alors que les probabilités de transition $\Pi(x, \cdot)$ dépendent continûment de x pour la convergence étroite, et en adaptant les démonstrations, on obtient des énoncés similaires au lemme 2.2 et aux corollaires 2.3, 2.4 et 2.5, où les convergences en variation (ou les convergences dans $L^1(\nu)$ des densités) sont remplacées par des convergences étroites.

Bibliographie

[1] BILINGSLEY P., *Convergence of probability measures*, Wiley, New-York (1968).

[2] DUBINS L.E., SCHWARZ G., *On extremal martingale distributions*, Proc. 5th Berkeley symp. math. stat. prob., University of California 1965/1966, **2**, part. 1 (1967) 295–299.

[3] DYNKIN E.B., *The initial and final behaviour of trajectories of Markov processes*, Russ. Math. Surv., **26**-4 (1971), 165–185.

[4] REVUZ D., *Markov chains*, North-Holland Mathemetical Library (1975).

[5] YOR M., *De nouveaux résultats sur l'équation de Tsirel'son*, C. R. Acad. Sci. Paris Sér. I Math., **309** Série I (1989), 511–514.

[6] YOR M., *Tsirel'son's equation in discrete time*, Probability Theory and Related Fields, **91**-2 (1992), 135–152.

Jean BROSSARD et Christophe LEURIDAN
INSTITUT FOURIER
Laboratoire de Mathématiques
UMR5582 (UJF-CNRS)
BP 74
38402 St MARTIN D'HÈRES Cedex (France)