B. DELYON

Mélange : exemples, applications

Publications de l'Institut de recherche mathématiques de Rennes, 1986, fascicule 1
« Probabilités », , p. 42-60
<http://www.numdam.org/item?id=PSMIR_1986___1_42_0>

MELANGE : EXEMPLES, APPLICATIONS

B.Delyon

Abstract : We recall briefly some recent results about mixing and enlarge the class of mixing processes in the domain of Markov chains.

Introduction

Lorsque l'on veut obtenir des théorèmes du genre "central-limit" sur des suites de variables aléatoires X_n , non nécessairement indépendantes, on est amené à définir une condtion, dite de mélange (due à Kolmogorov), exprimant qu'en un certain sens, les v.a. X_n pour n petit sont "quasiment" indépendantes des X_n pour n grand. Le but de cet article est d'étendre la classe des processus possédant une telle propriété, processus dit complètement réguliers, aux chaînes de Markov dont le noyau de transition vérifie des conditions voulues ici peut restrictives en regard des conclusions désirées tout en étant assez faciles à vérifier. On commencera par donner brièvement dans un premier paragraphe quelques résultats obtenus récemment sur ces processus.

I) Lemmes de mélange et théorèmes "central-limit"

Tous les résultats annoncés ici sont démontrés dans [D].

a) Définition des processus mélangeants et lemmes de mélange.

Soient $\mathcal{F}$ et $\mathcal{G}$ tribus sur un même espace Ω, on définira leur taux de mélange par :

$$\beta = \frac{1}{2} \sup \sum_{i,j} |P(A_i \cap B_j) - P(A_i)\, P(B_j)|$$

où le sup est pris sur les familles finies finies A_i et B_j vérifiant:

$$A_i \in \mathcal{F}, \quad B_j \in \mathcal{G}$$
$$P(A_i \cap A_j) = P(B_i \cap B_j) = 0 \qquad i \neq j$$

On peut alors montrer que cette condition est équivalente à l'existence d'une v.a. $\Delta=\Delta(\mathcal{F},\mathcal{G})$, $\mathcal{F}$ -mesurable, $0 \leq \Delta \leq 1$, telle que:

$$\text{(C)} \qquad \begin{array}{l} | P(A\cap B) - P(A)\, P(B) | \leq E[\Delta\, 1_A] \qquad \text{pour tous} A \in \mathcal{F}, \quad B \in \mathcal{G} \\ \text{et} \qquad E[\Delta] = \beta \end{array}$$

Un processus complètement régulier sera défini de la façon suivante:

Définition : *Le processus* $(X_t)_{t\geq 0}$ *est dit complètement régulier si en posant :*

$$\mathcal{F}_s^{\,t} = \sigma(X_u,\ s \leq u \leq t)$$

$$\Delta_{s,t} = \Delta(\mathcal{F}_0^{\,s}, \mathcal{F}_t^{\,\infty}) \qquad s \leq t$$

il existe une fonction $\beta(t)$ *tendant vers* 0 *à l'infini telle que :*

$$E[\Delta_{s,t}] \leq \beta(t-s).$$

Remarques :

(1) Le rôle de (X_t) est secondaire, c'est en fait une propriété de $(\Omega, \mathcal{F}_s^{\,t}, P)$, de sorte que pour toute application mesurable f, le processus $f(X_t)$ sera complètement regulier.

(2) En pratique on a besoin de propriétés d'intégrabilité de β telles que $\beta(t) = o(1)$. On trouvera ici : $\beta(t) \leq c^t$ avec $0 \leq c < 1$.

On peut alors obtenir, entre autres, les lemmes de mélanges suivants :

Lemme 1 : *Si* X *et* Y *sont deux variables aléatoires réelles dont l'une est centrée et respectivement* $\mathcal{F}_0^{\,s}$ *et* $\mathcal{F}_t^{\,\infty}$ *mesurables, on a alors, si le premier membre est défini :*

$$|E(XY)| < 2\,\beta(t-s)^{1/p}\,\|X\|_q\,\|Y\|_r \qquad \text{où :} \qquad p^{-1}+q^{-1}+r^{-1}=1$$

Lemme 2 : *Soient* $(X_n)_{n\in\mathbb{N}}$ *centrées avec* $X_n \in \mathcal{F}_t^{\,\infty}$ *pour tout n, alors, si s < t :*

$$|E[\,\sup_n E[X_n \mid \mathcal{F}_0^{\,s}]\,]| \leq 2\,\|\sup_n |X_n|\,\|_p\,\beta(t-s)^{1/q}$$

$$\text{où :} \qquad p^{-1}+q^{-1}=1$$

Il est alors possible à l'aide de ces deux lemmes d'obtenir le résultat qui va suivre.

b) Théorème "central-limit" pour les algorithmes stochastiques.

Il s'agit d'une généralisation du théorème "central-limit" habituel.

Soit :

$$F : \ \mathbb{R}^d \times \mathbb{R}^m \to \mathbb{R}^d$$

une application mesurable et $Y_n \in \mathbb{R}^d$ un processus stationnaire mélangeant au sens de la définition précédente. On se propose d'annuler en θ la fonction :

$$\bar{F}(\theta) = E[\,F(\theta, Y_n)\,]$$

Typiquement on sera dans la situation :

$$F(\theta, Y) = \frac{\partial H}{\partial \theta}(\theta, Y)$$

où H est une application réelle dont on cherche à minimiser (en θ) la moyenne pour la loi de Y, cas qui correspond à l'algorithme dit "de gradient stochastique".
On est alors amené à construire l'algorithme à pas constant (où ε est un petit réel) :

$$\theta^{\varepsilon}_{n+1} = \theta^{\varepsilon}_{n} + \varepsilon F(\theta^{\varepsilon}_{n}, Y_n)$$
$$\theta^{\varepsilon}_{0} = \theta_0$$

où $\theta_0 \in \mathbb{R}^d$ est fixé . On suppose que $F(\theta, Y_n)$ est stationnaire à l 'ordre 2 à θ fixé (condition en fait plus faible que la condition de stationnarité de Y_n demandée plus haut), et que les tribus $\mathcal{F}_n^p = \sigma\{ Y_i , n \le i \le p\}$ forment une famille complètement régulière pour la probabilité P sur Ω, avec des taux de mélange $\beta(p-n)$. Posons :

$$x^{\varepsilon}_t = \theta^{\varepsilon}_{t/\varepsilon} \qquad \text{où } t \to \theta^{\varepsilon}_t \text{ est l'extension afine par morceaux de } n \to \theta^{\varepsilon}_n$$

et en définissant la trajectoire moyenne x_t par :

$$\frac{dx_t}{dt} = \bar{F}(x_t) \qquad \text{où} \quad \bar{F}(x) = E[\, F(x, Y_n)\,]$$
$$x_0 = \theta_0$$

posons également :

$$y^{\varepsilon}_t = \frac{1}{\sqrt{\varepsilon}} (x^{\varepsilon}_t - x_t)$$

On peut alors énoncer le

<u>*Théorème*</u> : *Soit* T>0, *supposons qu' existent* p>4 *et un voisinage* V *de la trajectoire* $(x_t)_{0 \le t \le T}$, *et pour tout* M>0 , *une constante* c_M *tels que :*

(H1) $\quad E[\, \sup_{|\theta|<M} | F(\theta, Y_n)|^p \,] < c_M$

(H2) $\quad \theta \to \bar{F}(\theta)$ est une application $C^1(V)$

(H3) $\quad \sum_{n=1}^{\infty} n\, \beta(n)^{1/q} < \infty \qquad$ où $\quad \frac{4}{p} + \frac{1}{q} = 1$

(H4) $\quad E[\, F^*_\alpha(\theta, Y_n)\,] < c_M\, \alpha \qquad$ pour tous $\theta \in V$ et $n > 0$ $\quad$ où :

$$F^*_\alpha(\theta, Y) = \sup_{|h|<\alpha} | F(\theta+h, Y) - F(\theta, Y) |$$

alors, les processus $(y_t^\varepsilon)_{0\leq t\leq T}$ *tendent en loi vers la solution de :*

$$y_t = u_t + \int_0^t \bar{F}'(x_s) y_s \, ds$$

où u_t *est un processus gaussien à accroissements indépendants vérifiant :*

$$E[\, u_t u_t^T \,] = \sum_{n=0}^{\infty} \int_0^t E[\, \tilde{F}(x_s, Y_0)\tilde{F}(x_s, Y_n)^T + \tilde{F}(x_s, Y_n)\tilde{F}(x_s, Y_0)^T \,] \, ds$$

$$\text{où :} \quad \tilde{F}(x,Y) = F(x,Y) - \bar{F}(x,Y)$$

<u>Remarques</u> :

(1) Le théorème "central-limit" habituel correspond au cas où F est indépendant de x.

(2) Toutefois la force de ce théorème vient de ce que l'on y autorise F à être discontinue, ce qui est souvent le cas dans la pratique. C'est précisément ce cas qui utilise dans sa toute sa richesse l'hypothèse de mélange faite ici.

(3) Le même résultat existe en temps continu, i.e. où l'on remplace l'algoritme stochastique par un équation différentielle aléatoire :

$$\frac{dx_t^\varepsilon}{dt} = F(x_t^\varepsilon, Y_t)$$

II) <u>Propriétés mélangeantes des suites gaussiennes et des chaînes de markov</u>

Le problème de la mise en évidence de processus vérifiant la condition de mélange a été résolu en 1959 dans [VR] pour les suites gaussiennes, avec le résultat suivant :

<u>Théorème</u> : soit X_n *une suite gaussienne stationnaire de fonction de corrélation* $r(n)$. *Si la densité spectrale de* X_n *est strictement positive et s'il existe* $\gamma > 2$ *tel que :*

$$r(k) = O(\, k^{-\gamma}\,)$$

alors:

$$\beta(k) = O(k^{-\gamma})$$

Au vu de l'hypothèse (H3) du théorème cité plus haut, il sera bon d'avoir $\gamma > 2$.

On se propose ici de montrer le résultat suivant :

Pour toute fonction $f \in C_b(\mathbb{R}^d)$ (i.e. f continue bornée) on note :

$$\|f\| = \sup_x |f(x)|$$

$$\|f\|_\eta = \sup_{|x-y|<\eta} |f(x) - f(y)| \qquad \eta > 0$$

Le théorème s'énonce alors ainsi :

<u>Théorème</u> : *Soit* Y_n *une chaîne de Markov construite à l'aide d'un noyau de transition Fellerien* $\Pi(x,dy)$ *et d'une mesure de départ* μ. *On suppose que la seule valeur propre de module* 1 *de* Π, *considéré comme opérateur sur* $C_b(\mathbb{R}^d)$, *est* 1 *avec comme sous-espace propre associé l'ensemble des fonctions constantes.*

On suppose de plus qu'existent a, ã, n, η, γ, p, Q, b, *et une fonction* ϕ *sur* $\mathbb{R}^d$ *de la forme* $\phi(x)=\phi_0(|x|)$ *o ù* ϕ_0 *est croissante,* >1, *et tend vers* 0 *à l'infini tels que :*

(1) $\mu(\phi) < \infty$

(2) $\| \Pi^n f \|_\eta \le a \|f\| + \tilde{a} \|f\|_\eta$ *pour toute* f *continue bornée et* $a+\tilde{a}<1$

(3) $\Pi^p \phi(x) \le \gamma\phi(x) + Q$ *et* $\gamma < 1$

(4) $\Pi\phi(x) \le b\phi(x)$

Alors la chaîne Y_n *est mélangeante avec des constantes de mélange* $\beta(n)$ *vérifiant :*

$$\beta(n) \le c\beta^n$$

pour un certain $c>0$ *et un certain* $\beta<1$.

La mesure invariante sera notée plus bas π.

Dans l'hypothèse (2), la partie importante du second terme est $a\|f\|$ (on se ramènera assez facilement dans la démonstrationn au cas ã=0, il suffira de prendre n un peut plus grand). Remarquons également que si Π s'écrit $\Pi = \alpha\Pi_1 + \beta\Pi_2$ avec $\alpha+\beta=1$, $\alpha>0$, $\beta>0$, où Π_1 et Π_2 sont deux opérateurs de transition dont l'un vérifie l'hypothèse (2) alors on montre sans difficulté que Π vérifie également cette hypothèse.

Les hypothèses (1) à (4) sont en pratique, comme on le verra plus bas, d'une vérification assez élémentaire.

Ramenons nous, dans un premier temps, à la condition (4) énoncée plus bas.

Soit $(Y_p)_{p\geq 0}$ une réalisation de la chaîne avec une probabilité de départ μ. La condition (C) du (I.a) s'écrit :

pour tout $F \in \mathcal{F}_o^n$ et $G \in \mathcal{F}_p^\infty$, $n < p$:

(1) $$E_\mu[F(G-E_\mu[G])] \leqslant \sup(|G|)E_\mu[|F| \Delta_{n,p}]$$

$$E[\Delta_{n,p}] \leqslant \alpha\ (p-n)$$

où $n \to \alpha(n)$ est une fonction décroissant assez vite vers 0 à l'infini (ici, la décroissance sera géométrique).

Comme $E_\mu[G | \mathcal{F}_o^p] = g(Y_p)$ pour une certaine fonction g vérifiant $\sup|g| \leqslant \sup|G|$, on se ramène assitôt à montrer (1) pour les G s'écrivant $G = g(Y_p)$; on obtient donc la condition :

$$E_\mu[F (\Pi^{p-n} g(Y_n) - E_\mu[g(Y_p)])]$$
$$\leq \sup|g| \; E_\mu[|F| \Delta_{n,p}] .$$

Si l'on pose :

$$\bar{\mu} = \sum_{n=0}^{\infty} \frac{1}{2^{n+1}} \Pi^n \mu$$

il suffit d'avoir l'existence d'applications $\Delta_{n,p}(y)$ telles que pour tout $g \in L^\infty(\pi)$, on ait $\bar{\mu}$. p.s :

$$|\Pi^{p-n} g(y) - E_\mu[g(Y_p)]| \leq ||g||_\infty \Delta_{n,p}(y)$$

$$\text{et} \quad E_\mu[\Delta_{n,p}(Y_n)] \leq \alpha(p-n)$$

il suffit donc qu'il existe des applications Δ_q telles que :

$$(2) \quad \begin{aligned} &|\Pi^q g(y) - \pi g| \leq ||g||_\infty \Delta_q(y) \qquad \bar{\mu} - \text{ps} \\ &E_\mu[\Delta_q(Y_n)] \leq \alpha(q) \end{aligned}$$

car alors

$$|E_\mu[g(Y_p)] - \pi g| \leq \int |\Pi^p g(y) - \pi g| \mu(dy)$$
$$\leq ||g||_\infty \alpha(p) \leq ||g||_\infty \alpha(p-n)$$

Soit Φ l'ensemble des applications ϕ de R^d dans R qui peuvent s'écrire :

$$\phi(y) = \phi_o(|y|)$$

où ϕ_o est croissante et ≥ 1.

On va en fait chercher une condition pour avoir l'existnece de constantes M et ν et d'une application ϕ telles que :

$$(3)\quad \begin{aligned}&0 < v < 1 \quad , M > 0 \quad , \quad \phi \in \Phi \qquad \text{et}\\ &|\Pi^n g(y) - \pi g| \leqslant M v^n \phi(y) \sup|g|\\ &\text{pour tous } n \geqslant 0 \text{ et } g \text{ continue bornée et :}\\ &\qquad \sup_n \mu(\Pi^n \phi) < \infty.\end{aligned}$$

Mais (3) est impliqué par :

$$(4)\quad \begin{aligned}&\sup_{y\in R^d} \left| \frac{\Pi^n g(y) - \pi g}{\phi(y)} \right| \leqslant M\, v^n \sup_{y\in R^d} \left| \frac{g(y)}{\phi(y)} \right|\\ &\text{et} \qquad \mu(\phi) < \infty\end{aligned}$$

pour toute g continue bornée.

Posons pour $\phi \in \Phi$:

$$C_\phi = \{f \text{ continue sur } R^d, \ \frac{f}{\phi} \text{ bornée}\}$$

$$C^o_\phi = \{f \in C_\phi, \ \lim_{|x|\to\infty} \frac{f}{\phi}(x) = 0\}$$

$$||f||_\phi = \sup \left| \frac{f}{\phi} \right|$$

alors, la première ligne de (4) équivaut à dire que le rayon spectral de $\Pi-\pi$, défini sur C_ϕ, est strictement inférieur à 1. Il suffit même que cette propriété soit vérifiée sur C^o_Φ, à condition que ce dernier espace soit stable par Π.

Pour cela, on utilisera le théorème suivant dû à Yosida et Kakutani ([YK]).

Théorème : Soit U un opérateur quasi-compact sur un espace de Banach tel que $||U^n|| \leqslant K$, $n \in N$, alors :

Les valeurs propres λ_i de module 1 de U sont en nombre fini avec un ordre de multiplicité fini et il existe des projecteurs U_i sur des sous-espaces *de dimension finie et un opérateur* V *de rayon spectral strictement inférieur à 1 tels que :*

$$U^n = \sum_{i=1}^{p} \lambda_i^n U_i + V^n \qquad n > 0$$

$$U_i U_j = 0 \text{ si } i \neq j \qquad U_i^2 = U_i$$

$$U_i V = V U_i = 0.$$

Le résultat reste vrai si on remplace l'hypothèse

U est quasi complact

par *U* est quasi compact*

Rappelons qu'un opérateur U est dit quasi-compact s'il existe $m > 0$ et T opérateur compact tels que :

$$||U^m - T|| < 1$$

Démontrons la dernière assertion qui n'est pas donnée dans l'énoncé de Yoshida-Kakutani :

Si U^* est quasi-compact et $\sup_n ||U^n|| < \infty$, alors U^{**} vérifie les hypothèses du théorème original et on peut écrire :

$$U^{**} = \sum \lambda_i U_i + V$$

L'inclusion $B \subset B^{**}$ peut permettre d'en déduire la décomposition :

$$U = \sum \lambda_i U_i + V$$

à condition que l'on ait montré :

$$U_i(B) \subset B \quad \text{pour tout } i.$$

Mais on a pour tout $x \in B^{**}$:

$$U_i(x) = \lim \frac{1}{N} \sum_{n=1}^{N} \bar{\lambda}_i^n \quad U^{**}(x)$$

ce qui prouve que $U_i(B) \subset B$.

Il n'est pas toujours commode de vérifier qu'un opérateur est quasi-compact et il est par conséquent difficile d'appliquer le théorème précédent. Le lemme suivant a pour but de mettre en évidence une classe assez générale d'opérateurs de transition rentrant dans ce cadre.

Lemme : Soit U un opérateur sur un espace de Banach B tel qu'il existe une famille finie d'applications linéaires continues $(\ell_i)_{1 \leq i \leq N}$, $n > 0$ et $\alpha < 1$ tels que :

$$(5) \qquad ||U^n x|| \leq \alpha ||x|| + \sum_{i=1}^{N} |\ell_i(x)| \qquad x \in B$$

Alors U^* est quasi-compact.

Démonstration : On peut supposer que les ℓ_i sont indépendantes. Soit alors $(\ell_i^o)_{1 \leq i \leq N}$ une base normale de $[\ell_1, \ldots, \ell_N]$, i.e., telle que $||\ell_i^o|| = 1$ et que les applications coordonnées soient de norme 1 (l'existence d'une telle base est toujours vérifiée en dimension finie, voir [5] p. 257 théorème 2.2) ; alors le théorème de Hahn-Banach permet de prolonger ces applications coordonnées à tout B^*, et, en les notant $x_1, x_2 \ldots x_n$ $(x_i \in B^{**})$ on définit un projecteur sur $[\ell_1, \ldots \ell_N]$ par :

$$P\ell = \Sigma\, \ell(x_i)\ell_i^o$$

ou encore :

$$P\ell(x) = \ell(\Sigma\, x_i\, \ell_i^o(x))$$

avec : $\qquad ||\ell_i^o|| = ||x_i|| = 1 \qquad 1 \leq i \leq N$

Ecrivons (en posant $U^n = U_n$) :

$$U_n^* = P\,U_n^* + (I-P)U_n^*$$

PU_n^* est compact ; calculons la norme de $(I-P)U_n^*$. Pour cela, on a besoin de vérifier que (5) a lieu sur B^{**}. En effet, si $x \in B$ et $||x|| \leqslant 1$, pour toute application linéaire ℓ de B^*, on aura :

$$|U_n^*\,\ell x| \leqslant ||\ell||\;(\alpha + \Sigma|\ell_i(x)|)$$

Grâce au théorème de densité de Goldstine ([BZ]), cette relation se prolonge à la boule unité de B^{**} et donc (5) aussi.

Soit donc $\ell \in B^*$ et $x \in B$, on a :

$$|(I-P)U_n^*\,\ell(x)| = |U_n^*\ell\,(x - \sum_{i=1}^{N} x_i\;\ell_i^{\,o}(x))|$$

$$\leqslant \alpha||\ell||\;||x - \sum_{i=1}^{N} x_i\;\ell_i^{\,o}(x)||$$

$$\leqslant \alpha(N+1)||\ell||\;||x||$$

donc :

$$||(I-P)U_n^*|| \leqslant \alpha(N+1)$$

Si $\alpha < \frac{1}{N+1}$ alors on a montré que U^* est quasi-compact. Sinon, remarquons que :

$$||U^{kn}x|| \leqslant \alpha^k||x|| + \sum_{\substack{1\leqslant i\leqslant N\\ 0\leqslant j<k}} |\ell_i(U^{nj}x)|$$

et donc il suffit de choisir k tel que :

$$\alpha^k(kN+1) < 1$$

et de raisonner comme précédemment mais cette fois-ci sur U^*_{nk}.

CQFD

On peut maintenant appliquer le théorème de Yosida-Kakutani ; définissons avant pour toute application continue f :

$$||f|| = \sup_x |f(x)|$$
$$||f||_\eta \quad \sup_{|x-y|<\eta} |f(x)-f(y)| \qquad \eta > 0.$$

Théorème : Soit Π *un noyau Markovien, Fellerien tel qu'il existe* $\alpha, \overline{\alpha}, \eta, n, b, \beta, p, Q$, *et une fonction* ϕ *sur* R^d *de la forme* $\phi(x) = \phi_o(|x|)$ *où* ϕ_θ *est croissante,* $\geqslant 1$, *et tend à l'infini vers l'infini tels que :*

(6) $\quad ||\Pi^n f||_\eta \leqslant \alpha ||f|| + \overline{\alpha} ||f||_\eta \qquad \alpha + \overline{\alpha} < 1$ *pour toute* f *continue*

(7) $\quad \Pi^p \phi(x) < \beta\, \phi(x) + Q \qquad \beta < 1$

(8) $\quad \Pi\phi(x) \leqslant b\phi(x)$.

Alors, pour tout $0 < \varepsilon \leqslant 1$, $C^o_\phi\varepsilon$ *est stable par* Π *et il existe* $0 < \varepsilon \leqslant 1$ *tel que* Π *défini sur* $(C^o_\phi\varepsilon)$ *vérifie les hypothèses du théorème de Yosida-Kateutani.*

Démonstration : On va montrer successivement :

. $C^o_\phi\varepsilon$ est stable par Π

. Il existe K tel que $||\Pi^k||_\phi\varepsilon < K$, $k > 0$

. Le lemme précédent est applicable avec $B = C^o_\phi\varepsilon$ pour un $\varepsilon > 0$

Commençons par les deux premiers points qui sont faciles :

si $f \in C^o_\phi\varepsilon$, on peut, en posant pour tout $M > 0$:

$$\theta_M(x) = 1_{|x|<M}$$

écrire :

$$\left| \frac{\Pi f(x)}{\phi^{\varepsilon}(x)} \right| = \frac{|\Pi(f\theta_M)(x)|}{\phi^{\varepsilon}(x)} + \frac{|\Pi(f(1-\theta_M))(x)|}{\phi^{\varepsilon}(x)}$$

$$\leqslant \frac{1}{\phi^{\varepsilon}(x)} \sup_{|y|<M} |f(y)| + \sup_{|y|>M} \left| \frac{f}{\phi^{\varepsilon}}(y) \right| \frac{\Pi\phi^{\varepsilon}(x)}{\phi^{\varepsilon}(x)}$$

donc :

$$\limsup_{|x|\to\infty} \left| \frac{\Pi f(x)}{\phi^{\varepsilon}(x)} \right| \leqslant \sup_{|y|>M} \left| \frac{f}{\phi^{\varepsilon}}(y) \right| b^{\varepsilon}$$

pour tout M ; alors, en faisant tendre M vers l'infini, on obtient :

$$\lim_{|x|\to\infty} \frac{|\Pi f(x)|}{\phi^{\varepsilon}(x)} = 0$$

soit : $\Pi f \in C^{o}_{\phi^{\varepsilon}}$

Pour le deuxième point, on déduit de (7) que pour tout $k > 0$:

(10) $$\Pi^{kp}\phi(x) < \beta^{k}\phi(x) + \frac{1-\beta^{k}}{1-\beta} Q$$

et donc :

$$\Pi^{kn} \phi^{\varepsilon}(x) \leqslant \beta^{k\varepsilon} \phi^{\varepsilon}(x) + \frac{Q^{\varepsilon}}{(1-\beta)^{\varepsilon}} \qquad \text{car} \qquad \varepsilon \leqslant 1$$

$$\frac{\Pi^{kn} \phi^{\varepsilon}(x)}{\varphi^{\varepsilon}(x)} \leqslant \beta^{k\varepsilon} + \frac{Q^{\varepsilon}}{(1-\beta)^{\varepsilon}}$$

d'où, avec (8), l'existence d'un $K > 0$ tel que :

$$\frac{\Pi^{k} \phi^{\varepsilon}(x)}{\phi^{\varepsilon}(x)} \leqslant K \qquad \text{pour tous } k,x$$

$$||\Pi^{k} \phi^{\varepsilon}||_{\phi^{\varepsilon}} \leqslant K$$

et donc, pour tout $f \in C^o_{\phi^\varepsilon}$:

$$||\Pi^k f||_{\phi^\varepsilon} \leqslant ||f||_{\phi^\varepsilon} \, ||\Pi^k \phi^\varepsilon||_{\phi^\varepsilon}$$

$$\leqslant ||f||_{\phi^\varepsilon} K$$

ce qui achève de prouver le deuxième point.

Pour le troisième point, commençons par nous ramener à :

$$n = p = 1 \qquad \bar{\alpha} = 0.$$

Remarquons que (6) entraîne que pour tout $k > 0$, on a :

$$||\Pi^{kn} f||_\eta \leqslant \bar{\alpha}^k ||f||_\eta + \frac{1-\bar{\alpha}^k}{1-\bar{\alpha}} \alpha ||f||$$

$$\leqslant (2\bar{\alpha}^k + \frac{\alpha}{1-\bar{\alpha}}) \, ||f||$$

et si k est choisi assez grand :

$$2\bar{\alpha}^k + \frac{\alpha}{1-\bar{\alpha}} < 1$$

on s'est donc ramené, en remplaçant n par kn, à $\bar{\alpha} = 0$; au vu de (10), on peut aussi se ramener à $p = n$.

Quitte maintenant à poser $\Pi' = \Pi^n$ et à déduire la quasi-compacité de Π^* de celle de $(\Pi')^*$, on peur supposer $n = 1$. En résumé, on est ramené à :

(6') $\qquad ||\Pi f||_\eta \leqslant \alpha ||f|| \qquad \alpha < 1$

(7') $\qquad \Pi\phi(x) < \beta\phi(x) + Q \qquad \beta < 1$

Donnons-nous $M > 0$, $0 < \varepsilon \leqslant 1$, $k > 0$ que l'on choisira plus tard de manière adéquate. On a toujours :

$$\Pi^k \phi^\varepsilon(x) \leqslant (\Pi^k \phi(x))^\varepsilon$$

$$\leqslant (\beta^k \phi(x) + \frac{Q}{1-\beta})^\varepsilon$$

$$\leqslant \beta^{k\varepsilon} \phi^\varepsilon(x) + (\frac{Q}{1-\beta})^\varepsilon$$

Soit maintenant $f \in C^{o}_{\phi^\varepsilon}$, posons :

$$\Pi^k f = g+h$$

où :

$$g(x) = \Pi^k f(x) 1_{|x|<M}$$

$$h(x) = \Pi^k f(x) 1_{|x|\geqslant M}$$

alors :

$$|g(x)| \leqslant ||f||_{\phi^\varepsilon} (\beta^{k\varepsilon} C^\varepsilon + (\frac{Q}{1-\beta})^\varepsilon) \qquad \text{où} \qquad C = \phi(M)$$

et

$$|h(x)| \leqslant ||f||_{\phi^\varepsilon} \phi^\varepsilon(x) (\beta^{k\varepsilon} + (\frac{Q}{(1-\beta)C})^\varepsilon)$$

pour la première inégalité, on a utilisé :

$$\phi(x) \leqslant C \quad \text{si} \quad |x| < M$$

et pour la seconde :

$$\frac{1}{\phi(x)} \leqslant \frac{1}{C} \qquad \text{si} \qquad |x| > M$$

Soient alors x et y dans R^d tels que :

$$|x-y| < \eta$$

$$|x| < M \quad , \qquad |y| < M$$

on peut écrire :

$$|\Pi^{2k} f(x) - \Pi^{2k} f(y)| \leqslant$$

$$\leqslant |\Pi^k g(x) - \Pi^k g(y)| + |\Pi^k h(x)| + |\Pi^k h(y)|$$

$$\leqslant \alpha ||g|| + ||f||_{\phi^\varepsilon} [\beta^{k\varepsilon} + (\frac{Q}{(1-\beta)c})^\varepsilon] \{\Pi^k \phi^\varepsilon(x) + \Pi^k \phi^\varepsilon(y)\}$$

$$\leqslant \alpha ||f||_{\phi^\varepsilon} [\beta^{k\varepsilon} C^\varepsilon + (\frac{Q}{1-\beta})^\varepsilon] +$$

$$+ 2||f||_{\phi^\varepsilon} [\beta^{k\varepsilon} (\frac{Q}{(1-\beta)c})^\varepsilon] [\beta^{k\varepsilon} C^\varepsilon + (\frac{Q}{1-\beta})^\varepsilon]$$

pour tout $\lambda > 0$, on peut choisir :

.. ε tel que $\left(\frac{Q}{1-\beta}\right)^{\varepsilon} < 1+\lambda$ $\qquad \varepsilon > 0$

. puis M tel que $\left(\frac{Q}{(1-\beta)c}\right)^{\varepsilon} < \lambda$

. puis k tel que $\beta^{k\varepsilon} C^{\varepsilon} < \lambda$

et alors :

$$|\Pi^{2k} f(x) - \Pi^{2k} f(y)| \leqslant ||f||_{\phi^{\varepsilon}} \{\alpha(1+2\lambda) + 2(2\lambda)(1+2\lambda)\}$$

posons :

$$\tilde{\alpha} = \alpha(1+2\lambda) + 2(2\lambda)(1+2\lambda)$$

si λ a été choisi assez petit, on a $\tilde{\alpha} < 1$.

En résumé, il existe k, η, M, ε et $\tilde{\alpha} < 1$ tels que :

$$(11) \quad \left[\begin{array}{l} \text{si } |x-y| < \eta \text{ et } |x| < M,\ |y| < M \\ \text{alors :} \\ |\Pi^{2k} f(x) - \Pi^{2k} f(y)| \leqslant \tilde{\alpha} ||f||_{\phi^{\varepsilon}}. \end{array}\right.$$

Mais si $|x| > M$, alors :

$$\begin{aligned} |\Pi^{2k} f(x)| &\leqslant ||f||_{\phi^{\varepsilon}} \Pi^{2k} \phi^{\varepsilon}(x) \\ &\leqslant ||f||_{\phi^{\varepsilon}} \{\beta^{2k\varepsilon} \phi^{\varepsilon}(x) + \left(\frac{Q}{1-\beta}\right)^{\varepsilon}\} \\ &\leqslant ||f||_{\phi^{\varepsilon}} \phi^{\varepsilon}(x) \{\beta^{2k\varepsilon} + \left(\frac{Q}{(1-\beta)c}\right)^{\varepsilon}\} \\ &\leqslant 2\lambda\, \phi^{\varepsilon}(x) ||f||_{\phi^{\varepsilon}} \\ &\leqslant \tilde{\alpha} ||f||_{\phi^{\varepsilon}} \phi^{\varepsilon}(x) \end{aligned}$$

Soit maintenant $(x_i)_{1 \leqslant i \leqslant N}$ un η-recouvrement de la boule de rayon M et de centre 0, on a alors :

$$|\Pi^{2k} f(x)| \leqslant \sum_{i=1}^{N} |\Pi f(x_i)| + \tilde{\alpha}\|f\|_{\phi^\varepsilon} \qquad \text{si} \quad |x| \leqslant M$$

$$|\Pi^{2k} f(x)| \leqslant \tilde{\alpha}\|f\|_{\phi^\varepsilon}\, \phi^\varepsilon(x) \qquad \text{si} \quad |x| > M$$

d'où pour tout x :

$$\frac{|\Pi^{2k} f(x)|}{\phi^\varepsilon(x)} \leqslant \sum_{i=1}^{N} |\Pi f(x_i)| + \tilde{\alpha}\|f\|_{\phi^\varepsilon}$$

ce qui est exactement la relation (5) du lemme écrite dans l'espace séparable $C^0_{\phi^\varepsilon}$.

<u>CQFD</u>

Pour pouvoir maintenant déduire de ce théorème celui annoncé au début de ce paragraphe, il s'agit de motrer que Π n'admet pas de valeur propre de module 1 autre que 1 dans l'espace C^0_ψ (en ayant posé $\psi = \phi^\varepsilon$). Sachant que cette propriété est vraie sur $C_b(\mathbb{R}^d)$. On peut écrire :

$$\Pi = \sum_{i=1}^{n} \lambda_i U_i + V$$

alors :

$$U_i = \lim_N \sum_{j=1}^{N} \sum_{i=1}^{n} \overline{\lambda}_i^{\,n}\, \Pi^n$$

et donc :

$$C_b \supset U_i(C_b)$$

or par hypothèse :

$$U_i(C^0_\psi) \cap C_b = \begin{cases} \{\text{constantes}\} & \text{si } \lambda_i = 1 \\ \{0\} & \text{si } \lambda_i \neq 1 \end{cases}$$

donc :

$$U_i(C_b) = \begin{cases} \{\text{constantes}\} & \text{si } \lambda_i = 1 \\ \{0\} & \text{si } \lambda_i \neq 1 \end{cases}$$

d'où par densité du second espace dans le premier :

$$U_i(C^0_\psi) = \begin{cases} \{\text{constantes}\} & \text{si } \lambda_i = 1 \\ \{0\} & \text{si } \lambda_i \neq 1 \end{cases}$$

ce qui est ce que l'on voulait montrer.

A titre d'exemple, on a le résultat suivant :

Théorème : Soit $(X_n)_{n\geq 0}$ une chaîne de Markov gérée par un opérateur de transition Π, avec une mesure de départ μ, tels qu'existent $A, p, a > 0$ vérifiant:

$$\Pi f(x) = \int f(Ax+y)p(dy)$$

i) Les valeurs propres de la matrice A sont de module strictement inférieur à 1.

ii) La mesure de probabilité p n'est pas singulière par rapport à la mesure de Lebesgue.

iii)
$$\int |x|^a \, p(dx) < \infty$$
$$\int |x|^a \, \mu(dx) < \infty$$

Alors la condition (4) est vérifiée et le processus est complètement régulier.

Remarque : Tout processus de la forme $Y_n = f(X_n)$ pour un tel X_n sera donc également complètement régulier (en vertu d'une remarque de la fin du Chapitre I).

Démonstration : Il est clair que X_n peut se représenter par :

$$X_n = W_n + AW_{n-1} + \ldots + A^{n-1} W_1 + A^n W_0$$

où les variables indépendantes W_i ont pour loi p si $i > 0$ et μ si $i=0$. Il s'en suit que pour toute f continue bornée, $\Pi^n f(x)$ tend pour tout x vers πf où π est la loi de :

$$W_1 + AW_2 + A^2 W_3 + \ldots.$$

Pour tout $n > 0$, on peut décomposer :

$$\Pi^n f(x) = \int f(A^n x+y)p_n(y)dy + \int f(A^n x+y)q_n(dy)$$

avec : $\int p_n(y)dy > 0$

Soit η tel que $||A^n|| < 1$, alors :

$$||\Pi^n f||_\eta \leq ||f|| \sup_{|x-y|<\eta} \int |p_n(z-A^n x)-p_n(z-A^n y)|dz$$

$$+ \int \sup_{|x-y|<\eta} |f(A^n x+z)-f(A^n y+z)|q_n(dz)$$

$$||\Pi^n f||_\eta \leq ||f|| \sup_{|x|<\eta} \int |p_n(z-A^n x)-p_n(z)|dz$$

$$+ ||f||_\eta \int q_n(dz)$$

comme $\int q_n(ds) < 1$, il suffit de choisir η assez petit et (6) est vérifié.

On peut supposer $a \leq 1$, ce qui entraîne :

$$|x+y|^a \leq |x|^a + |y|^a$$

d'autre part, il existe C tel que :

$$\int \sum_{i\geq 0} |A^i y|^a \, p(dy) < C$$

car les A^i décroissant exponentiellement vite. On peut alors vérifier (7) et (8) avec $\phi(x) = 1 + |x|^a$; pour tout $n > 0$:

$$\Pi^n \phi(x) = \int \phi(A^n x + A^{n-1} y_n + \ldots + y_1) p(dy_1) \ldots p(dy_n)$$

$$\leq \phi(A^n x) + \sum_{i=1}^{n} \int |A^{i-1} y_i|^a \, p(dy_i)$$

$$\leq 1 + ||A^n||^a \, \phi(x) + c$$

ce qui prouve (8), en prenant $n=1$ et (7) en prenant n tel que $||A^n|| < 1$.

<u>CQFD</u>.

Il va de soit que l'exemple ci-dessus est assez restrictif et que la classe des Π qui vérifient (6) (7) (8) est bien plus grande ; chaque cas doit être examiné à part.

<u>References</u> :

[BZ] Brezis : "Analyse fonctionnelle : théorie et applications".
Masson, 1983.

[D] B.Delyon : "Un théorème de limite centrale pour certains algorithmes stochastiques avec discontinuités".
Irisa, Publication interne 312.

[S] I.Singer : "Bases in banach spaces".
Springer Verlag, vol.154.

[YK] Yoshida et Kakutani : "Operator theoretical treatment of Markoff process and mean ergodic theorem".
Ann.Math. 42, pp. 188-228.