PAUL-ANDRÉ MEYER

Convergence faible et compacité des temps d'arrêt, d'après Baxter et Chacón

Séminaire de probabilités (Strasbourg), tome 12 (1978), p. 411-423
<http://www.numdam.org/item?id=SPS_1978__12__411_0>

Université de Strasbourg
Séminaire de Probabilités 1976/77

CONVERGENCE FAIBLE ET COMPACITE DES TEMPS D'ARRET D'APRES BAXTER ET CHACON

par P.A.Meyer

J.R.Baxter et R.V.Chacon viennent d'écrire un article, intitulé "Compactness of stopping times", qui contient deux résultats d'un très grand intérêt. Le premier est un théorème de compacité de l'ensemble des temps d'arrêt "randomisés" (je les appellerai ici "temps d'arrêt flous' afin de ne pas écrire franglais, même si je le parle) pour une topologie parfaitement raisonnable. Le second concerne les passages à la limite, dans cette topologie, sur les processus quasicontinus à gauche. Par rapport au travail original de Chacon et Baxter, cet exposé lève les restrictions de séparabilité, et remplace la quasicontinuité à gauche par une condition plus faible de «régularité».

NOTATIONS. VARIABLES ALEATOIRES ET TEMPS D'ARRET FLOUS

Nous considérons un espace probabilisé filtré $(\Omega,\underline{\underline{F}},P\ ;\ (\underline{\underline{F}}_t)_{t\geq 0})$ satisfaisant aux conditions habituelles.

Nous appelons <u>variable aléatoire floue</u> (randomized r.v.) sur $(\Omega,\underline{\underline{F}},P)$ toute mesure de probabilité μ sur $[0,\infty]\times\Omega$ muni de $\underline{\underline{B}}([0,\infty])\times\underline{\underline{F}}$, dont la projection sur Ω est égale à P. Ainsi, si X est un processus mesurable borné admettant pour ensemble des temps $[0,\infty]$, nous savons calculer l'intégrale $\mu(X)=\int X_s(\omega)\mu(ds,d\omega)$. Par exemple, toute v.a. S sur Ω, à valeurs dans $[0,\infty]$, détermine une v.a. floue μ_S par

$$\mu_S(X) = E[X_S]$$

Comment construit on toutes les v.a. floues ? D'après le théorème de désintégration des mesures , il existe un processus croissant (A_t), tel que $A_{0-}=0$, pouvant présenter un saut à l'infini, et tel que

(1) $\qquad \mu(X) = E[\int_{[0,\infty]} X_s dA_s]$ pour tout processus positif X

Ce processus est unique (à ensemble évanescent près) si l'on en prend comme d'habitude la version continue à droite. Le fait que la projection de μ sur Ω est exactement P signifie que $A_\infty=1$ p.s., et la tradition - que nous suivrons ici - est d'utiliser plutôt le processus <u>décroissant</u> $M_t=1-A_t$, tel que $M_{0-}=1$, $M_\infty=0$, de sorte que $\mu(X)=E[-\int X_s dM_s]$.

On dit qu'une v.a. floue est un <u>temps d'arrêt flou</u> si M_t est $\underline{\underline{F}}_t$-mesurable pour tout t.

On peut interpréter les v.a. floues et t.a. flous comme de véritables v.a. ou t.a. de la manière suivante. Soit $\overline{\Omega}=[0,1]\times\Omega$, muni des tribus $\overline{\underline{\underline{F}}}^{\circ} = \underline{\underline{B}}([0,1])\times\underline{\underline{F}}$, $\overline{\underline{\underline{F}}}^{\circ}_t=\underline{\underline{B}}([0,1])\times\underline{\underline{F}}_t$, et de la loi $\overline{P}=\lambda\otimes P$, où λ est la mesure de Lebesgue. Après complétion et adjonction de tous les ensembles de mesure nulle aux $\overline{\underline{\underline{F}}}^{\circ}_t$, nous avons une nouvelle famille $(\overline{\underline{\underline{F}}}_t)$, qui satisfait aux conditions habituelles - mais nous ne démontrons pas ici ce petit résultat, car nous n'en aurons pas besoin.

Nous convenons d'identifier toute v.a. $H(\omega)$ sur Ω à la v.a. $(u,\omega)\mapsto H(\omega)$ sur $\overline{\Omega}$, que nous désignons par la même lettre. Alors $\underline{\underline{F}},\underline{\underline{F}}_t$ apparaissent comme des sous-tribus de $\overline{\underline{\underline{F}}},\overline{\underline{\underline{F}}}_t$.

Soit μ une v.a. floue, et soit (M_t) le processus décroissant associé Introduisons la v.a. sur $\overline{\Omega}$

(2) $\quad S(\overline{\omega}) = S(u,\omega) = \inf\{\ s\ :\ M_s(\omega)\leq u\ \}$

Noter que $S(.,\omega)$ est une fonction décroissante et continue à droite, et que
$$M_t(\omega) = \inf\{\ u\ :\ S(u,\omega)\leq t\}$$
$$(t<S(u\ ,\omega)) \iff (M_t(\omega)>u)$$
L'ensemble $\{(u,\omega)\ :\ S(u,\omega)\neq S(u-,\omega)\}$ est à coupes dénombrables au dessus de ω, donc (Fubini) négligeable pour $\overline{P}$. Nous avons donc
$$\overline{P}\{S(u,\omega)>t\,|\,\underline{\underline{F}}\} = \overline{P}\{S(u-,\omega)>t\,|\,\underline{\underline{F}}\}=\overline{P}\{u<M_t(\omega)\,|\,\underline{\underline{F}}\}=M_t(\omega)$$
et il en résulte que pour tout processus X sur Ω, mesurable et borné

(3) $\quad \mu(X) = \overline{E}[X_S]$

Si μ est un t.a. flou, l'ensemble $\{(u,\omega)\ :\ S(u,\omega)>t\}=\{(u,\omega)\ :\ M_t(\omega)>u\}$ est $\overline{\underline{\underline{F}}}^{\circ}_t$-mesurable, donc S est un t.a. de la famille $(\overline{\underline{\underline{F}}}^{\circ}_t)$, et a fortiori de la famille $(\overline{\underline{\underline{F}}}_t)$.

Inversement, soit $\Sigma(\overline{\omega})$ une v.a. réelle $\overline{\underline{\underline{F}}}$-mesurable à valeurs dans $[0,\infty]$; définissons un processus (M_t) sur Ω par $M_{0-}=1$, $M_{\infty}=0$, et pour $0\leq t<\infty$

($\overline{4}$) $\quad M_t = \overline{P}\{\Sigma>t\,|\,\underline{\underline{F}}\}$ (version décroissante et continue à dr.)

Si Σ est un t.a. de la famille $(\overline{\underline{\underline{F}}}_t)$, $I_{\{\Sigma>t\}}$ est $\overline{\underline{\underline{F}}}_t$-mesurable. Soit $H(\omega)$ une v.a. bornée orthogonale à $\underline{\underline{F}}_t$ sur Ω ; H est alors orthogonale à $\overline{\underline{\underline{F}}}_t$ sur $\overline{\Omega}$ et nous avons
$$E[HM_t] = \overline{E}[HI_{\{\Sigma>t\}}]=0$$
Par conséquent M_t est $\underline{\underline{F}}_t$-mesurable. Lorsqu'on part de Σ, que l'on construit (M_t) par (4), puis S par (2), on ne retombe pas sur Σ en général : S est une sorte de "réarrangement décroissant" de Σ. Cependant, si $\Sigma(.,\omega)$ est décroissante et continue à droite, le raisonnement fait plus haut montre que M est la fonction inverse de Σ, et l'on a $\Sigma=S$ p.s.

Il y a donc bijection entre les v.a.floues (les t.a. flous) sur Ω, les processus décroissants (décroissants adaptés) sur Ω, compris entre

0 et 1 et continus à droite, et finalement les fonctions $\Sigma(u,\omega)$ (les t.a.) sur $[0,1]\times\Omega$, décroissantes et continues à droite en u. Bien entendu, lorsqu'il est question de bijections, il faut identifier les processus indistinguables ou les v.a. égales $\bar{P}$-p.s..

Dans cette correspondance, une vraie v.a. $S(\omega)$ est associée au processus décroissant $M_t(\omega)=I_{\{t\geq S(\omega)\}}$, et à la v.a. $S(u,\omega)=S(\omega)$.

L'ensemble des (classes de) v.a. $S(u,\omega)$, décroissantes et continues à droite en u, est muni d'une relation d'ordre naturelle (la relation $S\leq T$ se lit simplement $M_t\leq N_t$ sur les processus décroissants associés). Il est stable pour les inf finis, les sup et les sommes dénombrables, ce qui permet si on le désire - mais nous ne donnerons pas de détails ici - de définir l'inf d'une famille finie de v.a. floues, le sup d'une famille dénombrable ou le sup essentiel d'une famille quelconque de v.a. floues (on peut aussi définir les inf dénombrables ou essentiels, au prix d'une régularisation de la limite).

TOPOLOGIE FAIBLE ET COMPACITE

Nous appellerons processus continus les processus mesurables bruts (i.e. non nécessairement adaptés) $X=(X_t)_{0\leq t\leq+\infty}$, dont toutes les trajectoires $X_{\cdot}(\omega)$ sont continues bornées sur $[0,\infty]$. De même, nous appellerons processus càdlàg. les processus $X=(X_t)_{0\leq t\leq+\infty}$ dont toutes les trajectoires sont càdlàg. bornées sur $[0,\infty]$. Le processus X est dit borné s'il est uniformément borné sur $[0,\infty]\times\Omega$ entier. Nous posons comme d'habitude $X^*(\omega) = \sup_t|X_t|$.

Rappelons un théorème établi dans le séminaire X, p.383 .

THEOREME 1. Soit $\underline{\underline{D}}$ l'espace des processus càdlàg. bruts bornés, et soit H une forme linéaire sur $\underline{\underline{D}}$, telle que $|H(X)|\leq E[X^*]$ pour tout $X\in\underline{\underline{D}}$. Il existe alors deux processus (A_t) et (B_t) non adaptés, à variation intégrable, tels que

$$A_{0-} = B_{0-}=0 \ ; \quad A_\infty=A_{\infty-} \ ; \ B_0=B_{0-} \ ; \quad \int_0^\infty|dA_s|+\int_0^\infty|dB_s| \leq 1$$

et

$$H(X) = E[\int_{[0,\infty[} X_s dA_s + \int_{]0,\infty]} X_{s-}dB_s] \quad \text{pour } X\in\underline{\underline{D}} . \tag{5}$$

Noter que si $H(1)=1$, les deux processus sont nécessairement croissants et l'on a $A_\infty+B_\infty=1$ p.s..

Soit $\underline{\underline{C}}\subset\underline{\underline{D}}$ l'espace des processus bruts continus bornés, et soit H une forme linéaire sur $\underline{\underline{C}}$ telle que $|H(X)|\leq E[X^*]$, $H(1)=1$. Prolongeons la (th. de Hahn-Banach) en une forme sur $\underline{\underline{D}}$ possédant les mêmes propriétés . Revenant à $X\in\underline{\underline{C}}$, on a $X_s=X_{s-}$, et on obtient simplement :

THEOREME 2. Toute forme linéaire H sur $\underline{\underline{C}}$, telle que $|H(X)|\leq E[X^*]$ et $H(1)=1$, s'écrit de manière unique

(6) $$H(X) = \mu(X)$$

où μ est une v.a. floue.

L'unicité résulte aussitôt de ce que les processus X de la forme $X_t(\omega)=a(\omega)f(t)$, a bornée, $f\in\underline{\underline{C}}[0,\infty]$, appartiennent à l'espace $\underline{\underline{C}}$ et engendrent toute la tribu par classes monotones.

Nous posons donc la définition suivante :

DEFINITION. La topologie faible sur l'ensemble des v.a. floues est la topologie la moins fine rendant continues les applications $\mu\mapsto\mu(X)$, $X\in\underline{\underline{C}}$.

Nous avons aussitôt le théorème suivant :

THEOREME 3. L'ensemble de toutes les v.a. floues et l'ensemble des t.a. flous sont compacts.

La première assertion résulte du th.2. Pour la seconde, nous vérifions que l'ensemble des t.a. flous est fermé pour la topologie faible. Or soit μ une v.a. floue, et soit (M_t) le processus décroissant associé. Pour vérifier que (M_t) est $(\underline{\underline{F}}_t)$-adapté , il nous suffit de vérifier que $\mu(X)=0$ pour tout processus $X\in\underline{\underline{C}}$ de la forme

$$X_s(\omega)=a(\omega)b(s) \quad \text{où a est bornée, orthogonale à } \underline{\underline{F}}_t$$
$$\text{b est continue, à support dans } [0,t[$$

et cette propriété passe à la limite faible.

Ce théorème ne servirait à rien si l'on n'avait pas un résultat d' extraction dénombrable. Le voici.

THEOREME 4. Si la tribu $\underline{\underline{F}}$ est séparable mod.P (ou encore , si $L^1(\underline{\underline{F}})$ est séparable), la topologie faible est métrisable. Dans tous les cas, de toute suite de v.a. floues on peut extraire une suite faiblement convergente.

DEMONSTRATION. Si la tribu $\underline{\underline{F}}$ est séparable mod.P, ou ce qui revient au même si $L^1(\underline{\underline{F}})$ est séparable, nous pouvons trouver une algèbre de Boole dénombrable $\underline{\underline{A}}$ engendrant $\underline{\underline{F}}$ mod.P. D'autre part, soit $\underline{\underline{J}}$ un ensemble dénombrable dense dans $\underline{\underline{C}}([0,\infty])$. Soit $\underline{\underline{X}}$ l'ensemble des processus $X_t(\omega)=I_A(\omega)f(t)$, où A parcourt $\underline{\underline{A}}$ et f parcourt $\underline{\underline{J}}$. Les fonctions continues $\mu\mapsto\mu(X)$ pour $X\in\underline{\underline{X}}$,en infinité dénombrable, séparent les points de l'espace des v.a. floues, et celui-ci est donc compact métrisable.

Soit μ_n une suite de v.a. floues, et soient (M^n_t) les processus décroissants correspondants. Soit $\underline{\underline{G}}$ une sous-tribu séparable de $\underline{\underline{F}}$, telle que toutes les v.a. M^n_t ($n\in\mathbb{N}$, t rationnel) soient $\underline{\underline{G}}$-mesurables.Pour tout processus continu X borné, soit $\overline{X}$ une projection bien-mesurable de X sur la famille de tribus constante et égale à $\underline{\underline{G}}$; $\overline{X}$ est indistinguable d'un processus continu borné, nous le choisissons effectivement continu et borné partout. Nous avons alors $\mu_n(X)=\mu_n(\bar{X})$ pour tout $X\in\underline{\underline{C}}$. La tribu $\underline{\underline{G}}$ étant séparable, nous pouvons trouver une sous-suite n_k telle que $\mu_{n_k}(\bar{X})$ converge pour tout $X\in\underline{\underline{C}}$. Mais alors $\mu_{n_k}(X)$ converge pour tout $X\in\underline{\underline{C}}$ et nous avons gagné.

PROCESSUS CÀDLÀG. REGULIERS

La convergence faible d'une suite (μ_n) de t.a. flous vers un t.a. flou μ entraîne la convergence de $\mu_n(X)$ vers $\mu(X)$ pour quantité de processus bornés non continus. C'est peut être là le point le plus remarquable du travail de Baxter et Chacon.

DEFINITION. <u>Soit</u> $X=(X_t)_{0\leq t\leq+\infty}$ <u>un processus càdlàg. borné</u> (non nécessairement adapté). <u>Nous dirons que</u> X <u>est</u> régulier <u>si</u>

<u>Pour tout t.a. prévisible</u> (<u>ordinaire</u>) T <u>on a</u> $E[X_T-X_{T-}|\underline{\underline{F}}_{T-}]=0$ <u>p.s.</u>

Exemples : les martingales ou les surmartingales régulières ; les processus quasi-continus à gauche ; les projections optionnelles de processus continus. La condition de régularité équivaut à la suivante :

<u>Pour toute suite</u> (T_n) <u>de t.a.</u> (<u>ordinaires</u>) <u>telle que</u> $T_n\uparrow T$, <u>on a</u> $E[X_{T_n}]\to E[X_T]$

THEOREME 3. <u>Soit</u> (μ_n) <u>une suite de t.a. flous qui converge faiblement vers</u> μ. <u>Alors</u> $\mu_n(X)\to\mu(X)$ <u>pour tout processus</u> X <u>càdlàg. régulier borné</u>.

DEMONSTRATION. Il nous suffit de prouver que pour tout ultrafiltre $\underline{u}$ sur $\mathbb{N}$, qui converge vers $+\infty$, on a $\lim_{\underline{u}} \mu_n(X)=\mu(X)$.

Considérons la forme linéaire $\lim_{\underline{u}} \mu_n(X)$ pour $X\in\underline{\underline{D}}$, que nous noterons $H(X)$. Comme elle satisfait à $|H(X)|\leq E[X^*]$, $H(1)=1$, il existe deux processus <u>croissants</u> intégrables (A_t) et (B_t), satisfaisant aux conditions du théorème 1 , tels que

$$\text{pour } X\in\underline{\underline{D}}\ ,\ H(X) = E[\textstyle\int X_s dA_s + \int X_{s-}dB_s]$$

Soient respectivement A^o et B^p la projection duale optionnelle de A et la projection duale prévisible de B , et soit λ la mesure bornée sur $[0,\infty]\times\Omega$

$$\lambda(X) = E[\textstyle\int X_s(dA^o_s+dB^p_s)]$$

Noter qu'on n'a pas a priori $A^o_\infty+B^p_\infty=1$, de sorte que λ n'est pas une v.a. floue.

Comme les μ_n sont des t.a. flous, les processus décroissants associés étant adaptés, nous avons pour tout $X\in\underline{\underline{D}}$, de projection optionnelle X^o

$$\mu_n(X)=\mu_n(X^o) \quad\text{et de même } \mu(X)=\mu(X^o),\quad \lambda(X)=\lambda(X^o).$$

Ensuite, soit $X\in\underline{\underline{D}}$, càdlàg. régulier. Nous avons

$$\mu_n(X) = \mu_n(X^o) \xrightarrow[\underline{u}]{} E[\textstyle\int X^o_s dA_s + \int X^o_{s-}dB_s] = E[\int X^o_s dA^o_s + \int X^o_{s-}dB^p_s]$$

Le processus X^o est càdlàg. régulier, donc sa projection prévisible est le processus X^o_{s-} , et comme B^p est prévisible

$$E[\int X^o_{s-}dB^p_s] = E[\int X^o_s dB^p_s]$$

Ainsi

$$\mu_n(X) \xrightarrow[\underline{u}]{} E[\int X^o_s(dA^o_s+dB^p_s)] = \lambda(X^o)=\lambda(X)$$

lorsque X est càdlàg régulier. En particulier, lorsque X est continu $\mu_n(X) \xrightarrow[\underline{u}]{} \lambda(X)$ et $\lambda(X)= \lim_n \mu_n(X)=\mu(X)$. Cela entraîne que $\lambda=\mu$ et le théorème est établi.

EXTENSION AUX PROCESSUS CROISSANTS INTEGRABLES

Nous avons employé le langage des v.a. floues, mais le théorème suivant se démontre à peu près de la même manière :

THEOREME 6. <u>Soit</u> (A^n_t) <u>une suite de processus croissants adaptés, telle que les v.a.</u> A^n_∞ <u>soient uniformément intégrables. Il existe alors un processus croissant intégrable</u> (A_t) <u>et une suite</u> $(n_k)\uparrow+\infty$ <u>tels que l'on ait, pour tout processus</u> X <u>càdlàg. régulier borné</u>

(7) $$\lim_k E[\int X_s dA^{n_k}_s] = E[\int X_s dA_s]$$

(Noter que X doit être càdlàg. sur $[0,\infty]$, et que A^n,A peuvent présenter des sauts en 0 et à l'infini).

Sous cette forme, on reconnaît le vieil argument de convergence faible qui intervient dans la théorie de la décomposition des surmartingales. Ce qui est nouveau, c'est qu'on ne procède pas par limite faible pour t fixe et régularisation, mais directement par convergence faible sur $\underline{\underline{C}}$, et qu'on obtient la convergence sur une classe de processus bien plus riche qu'autrefois.

D'autre part, soit X un processus càdlàg. régulier sur $[0,\infty[$, borné, ne possédant pas nécessairement de limite à gauche à l'infini. Soit X^N le processus $(X_{t\wedge N})$, qui est càdlàg. régulier sur $[0,\infty]$. La formule (7) nous donne - en supposant, pour la simplicité des notations, que la suite (A^k_t) toute entière converge vers (A_t)

$$\lim_k E[\int X^N_s dA^k_s] = E[\int X^N_s dA_s]$$

On peut en déduire le même résultat sur X lui même si l'on sait que

$$E[\ A^k_\infty - A^k_N\ ] \xrightarrow[N\to\infty]{} 0 \qquad \text{uniformément en } k$$

ce qui a lieu par exemple si les potentiels $E[A^k_\infty - A^k_t|\underline{\underline{F}}_t]$ sont majorés par un même potentiel (ξ_t) [si les (A^k_t) sont prévisibles, et (ξ_t) appartient à la classe (D), cela entraîne aussi l'intégrabilité uniforme des v.a. A^k_∞ , qui figure dans nos hypothèses ci-dessus]

Baxter et Chacon étudient aussi des propriétés de convergence sur des processus X bornés, quasi-continus à gauche jusqu'à un certain temps d'arrêt fixé ζ, par analogie avec les "processus standard" de la théorie des processus de Markov[1].

UNE APPLICATION A LA THEORIE DU POTENTIEL

Considérons un semi-groupe droit (P_t) sur un espace d'états E, et le processus de Markov (Y_t) correspondant. Nous supposons que le noyau potentiel U est propre.

Soit λ une loi de probabilité sur E. Rappelons qu'une mesure positive μ est une balayée de λ si $\mu(f)\leqq\lambda(f)$ pour toute fonction excessive f, ce qui revient - le noyau U étant propre - à écrire que $\mu U\leqq\lambda U$. La théorie de Skorokhod pour les processus de Markov nous dit que

$$(\mu \text{ est une balayée de } \lambda) \iff (\text{il existe un temps d'arrêt flou } T \text{ tel que } \mu = \lambda P_T)$$

en convenant que $f(Y_t)=f(\partial)=0$ si $t\geqq\zeta$ pour toute f sur E. D'autre part, soit f=Ug un potentiel borné. Le processus $X_t=f\circ Y_t$ est càdlàg. régulier borné. Le théorème de compacité de Baxter-Chacon entraîne donc

PROPOSITION. <u>De toute suite</u> (μ_n) <u>de balayées de</u> λ <u>on peut extraire une suite</u> (μ_{n_k}) <u>qui converge vers une balayée</u> μ <u>de</u> λ <u>au sens suivant</u>

1) <u>Pour tout potentiel borné</u> Ug , $< \mu_{n_k},Ug > \rightarrow < \mu,Ug >$

2) <u>Pour toute fonction excessive</u> f, $<\mu,f> \leqq \liminf_k < \mu_{n_k},f >$.

La propriété 2) est une conséquence immédiate de 1), puisque toute fonction excessive est limite d'une suite croissante de potentiels bornés. Noter que plus généralement, si f est un potentiel de la classe (D) borné <u>régulier</u> (potentiel d'une fonctionnelle additive continue), on a encore $\mu(f) = \lim_k \mu_{n_k}(f)$.

Cette notion de limite n'est pas très satisfaisante, car on a ignoré ce qui se passe à l'instant ζ,en convenant que $f(\partial)=0$ pour toute f sur E. La théorie de la compactification de Martin permet de "subdiviser le point ∂ " en ajoutant une frontière, et en redéfinissant le processus (Y_t) pour $t\geqq\zeta$ - et en particulier pour $t=+\infty$ - de manière que le nombre des processus $(f\circ Y_t)$ réguliers sur $[0,\infty]$ augmente. Dans ce cas, la topologie peut être améliorée, il n'y a plus de perte de masse à l'infini, mais la mesure limite μ est portée par l'espace de Martin et non par E en général.

1. Une version paléolithique des résultats de compacité de Baxter-Chacon se trouve dans les Annales de l'Institut Fourier, tome XII, 1962, p. 195 -212 (avec pas mal d'erreurs), pour le cas des fonctionnelles additives.

AUTRES RESULTATS DE CONVERGENCE

Nous revenons à la situation du début : une suite (μ_n) de v.a. floues qui converge faiblement vers une v.a. floue μ. Cette convergence est tout à fait analogue à la convergence étroite des mesures de probabilité, et nous nous proposons ici de suivre cette analogie.

THEOREME 7. <u>Soit</u> X <u>un processus mesurable positif</u>, <u>dont presque toutes les trajectoires sont s.c.i. sur</u> $[0,\infty]$. <u>On a alors</u>

$$\mu(X) \leqq \liminf_n \mu_n(X)$$

DEMONSTRATION. Soit X^p le processus

$$X^p = \sum_{k=1}^{n} 2^{-p} I_{\{X>k2^{-p}\}}$$

On a $X^p \uparrow X$ lorsque $p \to \infty$, donc il suffit d'établir le résultat pour chaque X^p. Comme X^p est une somme finie d'indicatrices d'ouverts aléatoires (non adaptés), il suffit de l'établir pour une telle indicatrice I_H . On peut alors représenter H comme une réunion d'intervalles stochastiques $]\!]S_n,T_n[\![$ (où les S_n et les T_n sont des v.a. positives, non des temps d'arrêt en général), et il suffit de traiter le cas d'un intervalle $]\!]S,T[\![$. Or il est immédiat de représenter $I_{]\!]S,T[\![}$ comme limite d'une suite croissante de processus (non adaptés) à trajectoires continues.

COROLLAIRE. <u>Soit</u> X <u>un processus mesurable borné</u>. <u>Si l'ensemble des</u> (t,ω) <u>tels que</u> $X_.(\omega)$ <u>ne soit pas continue en</u> t <u>est</u> μ-<u>négligeable</u>, <u>on a</u> $\mu(X) = \lim_n \mu_n(X)$.

DEMONSTRATION. Supposons X compris entre 0 et 1, et définissons les deux processus (mesurables : cf. Probabilités et potentiels, p.225)

$$\overline{X}_t = \limsup_{s\to t} X_s \ , \quad \underline{X}_t = \liminf_{s\to t} X_s \quad ; \ \underline{X} \leqq X \leqq \overline{X} \ .$$

D'après le théorème précédent appliqué à $\underline{X}$ et à $1-\overline{X}$, nous avons

$$\mu(\underline{X}) \leqq \liminf_n \mu_n(\underline{X}) \leqq \liminf_n \mu_n(X)$$

$$\mu(\overline{X}) \geqq \limsup_n \mu_n(\overline{X}) \geqq \limsup_n \mu_n(X)$$

L'hypothèse sur X revient à dire que $\mu(\overline{X})=\mu(\underline{X})=\mu(X)$. D'où aussitôt la propriété cherchée.

REMARQUE. Comme dans Probabilités et Potentiels p.115 (th. III.58), on peut démontrer le théorème plus fort suivant : si les μ_n et μ sont portés par un ensemble aléatoire A, et si X est continu <u>sur A</u> et borné, alors $\mu_n(X) \to \mu(X)$.

REMARQUE. Le corollaire permet d'éclaircir un peu, en la rendant plus élémentaire, la convergence introduite par Baxter-Chacon. Introduisons les processus décroissants (M_t^n), (M_t) associés à μ_n et à μ , et choisissons un t tel que $P\{M_t \neq M_{t-}\}=0$. Soit aussi $A \in \underline{\underline{F}}$. Le processus

$$X_s(\omega) = I_A(\omega) I_{]t,\infty]}(s)$$

a un ensemble de points de discontinuité μ-négligeable, et on a par conséquent

$$\int_A M_t^n P = \mu_n(X) \rightarrow \mu(X) = \int_A M_t P$$

Autrement dit, $M_t^n \rightarrow M_t$ pour la topologie faible de L^1. Il est peut être intéressant de noter aussi que, si T est un t.a. totalement inaccessible, le processus

$$X_s(\omega) = I_A(\omega) I_{\{s>T\}}$$

est quasi-continu à gauche, et que (si les μ_n sont des t.a. flous) on a donc $\lim_n M_T^n = M_T$ pour la topologie faible de L^1.

TEMPS D'ARRET FLOUS ET PROCESSUS NON BORNES

Cette section a été rajoutée en Novembre 1977, sur la suggestion de C. Dellacherie. Celui-ci a en effet exposé au séminaire une démonstration d'un remarquable théorème de Bismut (énoncé plus loin) qui repose sur une extension du th.5 à des processus non bornés.

Soit X un processus mesurable, indexé par $[0,\infty]$. Nous poserons

(8) $$\|X\|_1 = \sup_T E[|X_T|]$$

T parcourant l'ensemble de tous les temps d'arrêt (ordinaires). Si Φ est une fonction convexe croissante nulle en 0, telle que $\lim_t \Phi(t)/t = +\infty$ lorsque $t \to \infty$ (fonction d'Orlicz), on pose pour toute v.a. Z

$$\|Z\|_\Phi = \inf \{ t : E[\Phi(|Z|/t)] \leqq 1 \}$$

et pour tout processus X comme ci-dessus

(9) $$\|X\|_\Phi = \sup_T \|X_T\|_\Phi$$

D'après le lemme de La Vallée-Poussin, un processus X appartient à la classe (D) si et seulement s'il existe une fonction d'Orlicz Φ telle que $\|X\|_\Phi < \infty$.

Soient maintenant μ un temps d'arrêt flou, (M_t) le processus décroissant adapté associé à μ ; nous posons $S_u = \inf \{ t : M_t < u \}$ et, pour tout processus X comme ci-dessus

(10) $$X_\mu = -\int_{[0,\infty]} X_s dM_s = \int_0^1 X_{S_u} du$$

Nous notons les propriétés immédiates suivantes . D'abord

$$E[|X_\mu|] \leqq \int_0^1 E[|X_{S_u}|]du \leqq \|X\|_1$$

On peut donc définir $\mu(X) = E[X_\mu]$ pour tout processus X tel que $\|X\|_1 < \infty$, prolongeant ainsi la définition donnée plus haut pour les processus bornés. Puis, soit Φ une fonction d'Orlicz ; on a par convexité $\Phi(X_\mu) \leqq \int_0^1 \Phi(X_{S_u})du$, et donc

$$\|X_\mu\|_\Phi \leqq \|X\|_\Phi$$

Il en résulte que si X appartient à la classe (D), toutes les v.a. X_μ correspondant aux temps d'arrêt flous sont uniformément intégrables, comme c'était le cas pour les temps d'arrêt non flous (cela peut se voir aussi en remarquant que les X_μ appartiennent à l'enveloppe convexe fermée des X_T dans L^1). Enfin, notons le lemme suivant :

LEMME. Soit (X^n) une suite de processus, dominés en valeur absolue par un processus Y de la classe (D), et tels que $(X^n)^*$ converge simplement vers 0 . Alors $\mu(X^n) \to 0$ uniformément sur l'ensemble des t.a. flous.

DEMONSTRATION. On peut supposer les X^n positifs et écrire

$$\mu(X^n) = \int_0^1 E[X^n_{S_u}]du \leqq \int_0^1 E[X^n_{S_u} \wedge c]du + \int_0^1 E[X^n_{S_u} I_{\{X^n_{S_u}>c\}}]du$$
$$\leqq E[(X^n)^* \wedge c] + \sup_T E[Y_T I_{\{Y_T>c\}}]$$

et la conclusion est immédiate.

Maintenant, nous étendons le théorème 5 . Pour un processus X de la classe (D), la définition de la régularité s'étend sans modification (si T est un temps prévisible, annoncé par une suite T_n , les v.a. X_{T_n} convergent vers X_{T-} , X étant toujours supposé càdlàg. jusqu'à l'infini ; l'intégrabilité uniforme s'étend donc aux v.a. X_{T-}). Noter que le théorème s'applique en particulier aux processus continus de la classe (D), non nécessairement adaptés.

THEOREME 8. Soit (μ_n) une suite[1] de t.a. flous qui converge faiblement vers μ. Alors $\mu_n(X) \to \mu(X)$ pour tout processus X càdlàg. régulier de la classe (D).

DEMONSTRATION. Soit $\underline{\underline{R}}$ l'ensemble des processus càdlàg. (non nécessairement adaptés) de la classe (D) ; $\underline{\underline{R}}$ est un espace vectoriel réticulé, contenant tous les processus càdlàg. bornés, et en particulier les constantes. Soit $\underline{u}$ un ultrafiltre sur $\mathbb{N}$; posons pour tout $X \in \underline{\underline{R}}$

$$H(X) = \lim_{\underline{u}} \mu_n(X) .$$

H(X) est fini : pour le voir, on peut considérer par exemple une fonction d'Orlicz Φ telle que $\|X\|_\Phi < \infty$, et remarquer que $|\mu_n(X)| \leqq \|X\|_\Phi$ pour tout n . Tout revient, comme dans la démonstration du théorème 5, à démontrer que H admet pour $X \in \underline{\underline{R}}$ une représentation

$$H(X) = E[\int X_s dA_s + \int X_{s-} dB_s] .$$

Il suffit pour cela (séminaire X, p.383) de montrer que si des $X^n \in \underline{\underline{R}}$ positifs tendent en décroissant vers 0, de telle manière que $(X^n)^* \to 0$ simplement, alors $H(X^n) \to 0$. Mais ceci résulte aussitôt du lemme ci-dessus.

1. Nous nous limitons aux suites uniquement pour la simplicité.

Mais le théorème 8 ne suffit pas pour atteindre le théorème de Bismut. Il faut sortir un peu des temps d'arrêt flous. Soit $\underline{\underline{C}}$ l'espace des processus[1] X (bruts) continus, tels que $E[X^*]<\infty$, avec la norme $\|X\|=E[X^*]$. Le dual $\underline{\underline{M}}$ de $\underline{\underline{C}}$ s'identifie à l'espace des mesures sur $\overline{\mathbb{R}}_+\times\Omega$ de la forme

$$\mu(X) = E[\int_{[0,\infty]} X_s dA_s]$$

où (A_t) est un processus à variation intégrable (brut), pouvant présenter un saut en 0 et un saut à l'infini, tel que $\int_{[0,\infty]} |dA_s| \in L^\infty$, muni de la norme $\|\mu\| = \| \int |dA_s| \|_\infty$. Nous désignons par $\underline{\underline{M}}_a$ le sous-espace fermé de $\underline{\underline{M}}$ constitué par les μ dont le processus associé A est adapté (ce qui revient à dire que $\mu(X)=\mu(X^o)$ pour tout processus borné X). Voici la version du théorème 8 qui nous est nécessaire :

THEOREME 8'. <u>Soient</u> (μ_n) <u>une suite</u>[2] <u>d'éléments de la boule unité de</u> $\underline{\underline{M}}$, <u>et</u> $\mu\in\underline{\underline{M}}$; <u>les propriétés suivantes sont équivalentes</u>

1) $\mu_n(X) \to \mu(X)$ <u>pour tout</u> X <u>continu borné</u> (<u>brut</u>).

2) $\mu_n(X) \to \mu(X)$ <u>pour tout</u> $X\in\underline{\underline{C}}$ (i.e. $\mu_n\to\mu$ pour $\sigma(\underline{\underline{M}},\underline{\underline{C}})$).

<u>De plus, si les</u> μ_n <u>appartiennent à</u> $\underline{\underline{M}}_a$, <u>il en est de même de</u> μ, <u>et les propriétés ci-dessus équivalent encore à</u>

3) $\mu_n(X) \to \mu(X)$ <u>pour tout</u> X <u>càdlàg. régulier (brut) de la classe</u> (D).

La démonstration est la même que celle des théorèmes 5 et 8, il faut seulement remplacer l'ensemble des t.a. flous par la boule unité de $\underline{\underline{M}}_a$, i.e. étendre le lemme de la page précédente à la boule unité. C'est immédiat, car un élément de la boule unité est différence de deux éléments positifs.

Il nous reste à énoncer le théorème de Bismut, dont nous espérons rendre compte dans le prochain volume : <u>Tout processus optionnel càdlàg. de la classe</u> (D) <u>régulier est projection optionnelle d'un processus brut</u> X, <u>continu et tel que</u> $E[X^*]<\infty$. C'est vraiment un résultat remarquable.

UN EXEMPLE DE CONVERGENCE FAIBLE

Pour conclure cet exposé, nous voudrions donner (d'après Chacon-Baxter) un exemple de temps d'arrêt ordinaires convergeant vers un temps d'arrêt flou.

Soit (B_t) un mouvement brownien plan issu de 0, et soit pour $0<r\leq 1$ $U_r = \inf\{t : |B_t|=r\}$. Puis soit pour $s\in[0,r]$

$$V_{r,s} = \inf\{ t\geq U_r : |B_t|=s \text{ ou } |B_t|=1 \}$$

On a $V_{r,s} \leq U_1$, donc les $V_{r,s}$ forment un ensemble relativement compact pour la topologie faible. Lorsque s varie de 0 à r, $P\{|B_{V_{r,s}}|=s\}$ varie

1. Toujours indexés par $[0,\infty]$. 2. Extension immédiate aux filtres.

de 0 à 1. Choisissons donc $s=s(r)$ tel que cette probabilité soit égale à 1/2 et posons $W_r=V_{r,s(r)}$. Nous avons

$$W_r \leqq U_1 \qquad P\{|B_{W_r}| \leqq r\} \geqq 1/2 \qquad , \ P\{|B_{W_r}|=1\} \geqq 1/2$$

Toute valeur d'adhérence faible W des W_r lorsque $r \to 0$ doit satisfaire à

$$W \leqq U_1 \ , \ P\{|B_W|=0\} \geqq 1/2 \ , \ P\{|B_W|=1\} \geqq 1/2$$

Mais ceci est impossible pour un temps d'arrêt non flou. En effet, $\{0\}$ est polaire, donc $(|B_W|=0) \Rightarrow (W=0)$ p.s.. D'après la loi de tout ou rien , pour un temps d'arrêt non flou $P\{W=0\}>0 \Rightarrow P\{W=0\}=1$, et cela contredit la dernière propriété.

<u>Références</u>

[1]. J.R. Baxter et R.V. Chacon. Compactness of stopping times. Z.f.W. 40, 1977, p.169-182.

APPENDICE : PROCESSUS QUASI-CONTINUS A GAUCHE

Je voudrais exposer ici une très jolie caractérisation des processus càdlàg. quasi-continus à gauche, qui figure dans le travail de Baxter et Chacon, mais que nous n'avons pas utilisée dans l'exposé lui même puisque nous avons remplacé la quasi-continuité à gauche par la " régularité" .

Soit X un processus càdlàg. sur $[0,\infty]$ - il est indifférent que X soit adapté. Nous dirons que X est <u>quasi-continu à gauche</u> si

<u>Pour toute suite</u> (T_n) <u>de t.a. telle que</u> $T_n \uparrow T$ <u>on a</u> $X_{T_n} \to X_T$ <u>p.s.</u>

(ou, ce qui revient au même, en P.). Le théorème de Baxter-Chacon dit alors :

X <u>est uniformément continu en probabilité sur les t.a.</u>: <u>pour la distance usuelle d définissant la convergence en P. des v.a. réelles</u>, <u>on a la propriété suivante</u>, <u>où</u> S <u>et</u> T <u>désignent des t.a</u>.

$$\forall a>0 \quad \exists b>0 \quad : (d(S,T)<b) \Rightarrow (d(X_S,X_T)<a)$$

Inversement, il est clair que cette propriété entraîne la quasi-continuité à gauche.

Voici une démonstration rapide de ce résultat, qui utilise un peu plus de "théorie générale des processus" que celle de Baxter-Chacon. Nous pouvons nous ramener au cas où X est compris entre 0 et 1. Quitte à ramener l'intervalle $[0,\infty]$ sur $[0,1]$, puis à prolonger le processus par la v.a. X_1 sur $[1,\infty[$, nous pouvons supposer que X est arrêté à l'instant 1. Si $s<t$, nous notons $Y_{st}(\omega)$ l'oscillation de $X_{\cdot}(\omega)$ sur

l'intervalle $]s,t]$ (aucune difficulté de mesurabilité puisque X est càdlàg.). En particulier, posons pour m entier

$$Y_t^m = Y_{t,t+1/m}$$

On vérifie aussitôt que c'est un processus càdlàg.. D'autre part, $Y^m \downarrow 0$ lorsque $m \to \infty$.

Soit Z^m la <u>projection optionnelle</u> de Y^m. Comme Y^m est càdlàg., Z^m l'est aussi ; comme $Y^m \downarrow 0$, on a $Z^m \downarrow 0$. La quasi-continuité à gauche de X va entraîner le lemme suivant, qui est très proche du lemme de Shur sur les fonctions excessives régulières (Prob. et Pot. 1e éd. th. VII.36).

LEMME. <u>Soit</u> $R_m = \inf \{ t : Z_t^m \geq a \}$. <u>Alors</u> $P\{R_m<\infty\} \to 0$ <u>lorsque</u> $m \to \infty$.

En d'autres termes, $Z^m_\cdot$ converge p.s. vers 0 <u>uniformément</u>. Il en résulte que $\sup_S E[Z_S^m] = \sup_S E[Y_S^m] \to 0$ lorsque $m \to \infty$, S parcourant l'ensemble de tous les t.a..

DEMONSTRATION. Nous remarquons que R_m croît avec m ; soit R sa limite. On a $Z_{R_m}^m \geq a$ sur $\{R_m<\infty\}$, donc $E[Y_{R_m}^m I_{\{R_m<\infty\}}]=E[Z_{R_m}^m I_{\{R_m<\infty\}}] \geq aP\{R_m<\infty\}$. D'autre part $Y_{R_m}^m \leq Y_{R_m,R} + Y_R^m$; $Y_{R_m,R}$ tend vers 0 d'après la régularité de X, le second terme aussi tend vers 0. Tout étant borné $E[Y_{R_m}^m] \to 0$ et le lemme est établi.

Démontrons le théorème. Etant donnée la situation à laquelle nous nous sommes ramenés, nous pouvons raisonner sur des t.a. ≤ 1 et prendre comme distances $d(S,T)=E[|S-T|]$ et $d(X_S,X_T)=E[|X_S-X_T|]$. Supposons que l'énoncé soit faux. Il existe alors des (S_n,T_n) tels que $E[|S_n-T_n|] \to 0$ et $E[|X_{S_n}-X_{T_n}|]$ reste $\geq a > 0$. Quitte à remplacer S_n,T_n par $S_n \wedge T_n$ et $S_n \vee T_n$ nous pouvons supposer $S_n \leq T_n$. Nous avons alors

$$E[|X_{S_n}-X_{T_n}|] \leq 2P\{T_n \geq S_n + 1/m\} + E[Y_{S_n}^m]$$

Nous majorons le dernier terme par $\sup_S E[Y_S^m]$, majoré par a/3 pour m assez grand d'après le lemme. Puis m étant ainsi choisi on a $2P\{\ldots\} \leq a/3$ pour n assez grand et on a obtenu une contradiction.

On peut commenter ainsi ce résultat : si X est un processus càdlàg. <u>régulier</u>, l'application $T \mapsto E[X_T]$ est continue pour la topologie de Chacon-Baxter sur l'ensemble des temps d'arrêt (flous, et a fortiori non flous), d'après le théorème 5 si X est borné, ou le théorème 8 si X appartient à la classe (D). On trouvera plus loin une note de Dellacherie montrant que, pour les t.a. non flous, cela revient à une continuité pour la convergence en probabilité. Si maintenant on remplace l'application $T \mapsto E[X_T]$ par $T \mapsto X_T$, le résultat analogue caractérise, non plus les processus réguliers, mais les processus quasi-continus à gauche (et n'exige plus de restriction d'intégrabilité).