SÉMINAIRE DE PROBABILITÉS (STRASBOURG)

AZIZ ES-SAHIB
HENRI HEINICH

Barycentre canonique pour un espace métrique à courbure négative

Séminaire de probabilités (Strasbourg), tome 33 (1999), p. 355-370
<http://www.numdam.org/item?id=SPS_1999__33__355_0>

Barycentre canonique pour un espace métrique à courbure négative

ES-SAHIB AZIZ & HEINICH HENRI

Résumé

Pour une variable aléatoire X intégrable à valeurs dans un espace (M, d) métrique complet séparable et à courbure négative, nous définissons un barycentre de X. Ce point, $b(X)$, appartient à l'ensemble des espérances au sens de Doss de X et ne dépend que de la loi de la variable. De plus si X et Y sont deux variables intégrables, alors $d(b(X), b(Y)) \leq E[d(X, Y)]$.
Nous étudions le problème de cohérence (loi des grands nombres) pour ce barycentre et nous montrons un théorème ergodique.
Puis nous remplaçons l'espérance de Doss par celle de Herer puis par celle d'Émery et Mokobodzki.

Abstract

For X an integrable random variable with values in a complete separable metric space (M, d) with negative curvature, we define a point $b(X)$ called barycenter of X which depends only on the law of X and belongs to the set of Doss expectation of X. Moreover for two integrable variables we have $d(b(X), b(Y)) \leq E(d(X, Y))$. We study the coherence problem: strong law of larges numbers for this barycenter and an ergodic theorem is given.
In the end we change Doss expectation for the Herer one and for the Émery-Mokobodzki one.

Introduction

La notion d'espérance pour une variable aléatoire à valeurs dans un espace métrique M trouve ses premières formulations dans [8] et [5]. Les propriétés de cette espérance se développent dans [2], [6], [9],[11]. Par la suite Herer introduit en [10] et [12] une autre définition et s'intéresse aux espaces à courbure négative. Plus récemment le cas M variété a été abordé par [1] et [7].

Ces différentes approches ont en commun que l'espérance d'une variable aléatoire est un sous ensemble fermé en général non réduit à un point. Ceci soulève des difficultés, en particulier pour obtenir une loi forte des grands nombres, car il faut considérer des convergences de fermés par exemple au sens de Hausdorff ou de Wijsman *c.f.* [15]. Cette dernière référence contient aussi d'autres versions de la loi forte des grands nombres.

Lorsque $d^2(\cdot, x)$ est strictement convexe, d'autres auteurs définissent classiquement le barycentre $b(X)$ d'une variable aléatoire X intégrable, par $b(X) = \underset{a}{\text{Argmin}}\ E[d^2(a, X)]$. L'espérance est alors un singleton. Il en est de même pour la définition adoptée par [13] pour les variétés.

Notre but est d'obtenir de manière "canonique", pour une v.a X intégrable à valeurs dans M aussi général que possible, un point appelé barycentre de X (ou espérance de X).

Nous allons centrer notre travail sur l'espérance au sens de Doss du fait de sa simplicité et de l'adaptation des méthodes introduites à d'autres notions d'espérances. Rappelons, que pour un espace de Banach séparable, la notion d'espérance de Doss et celle de Bochner coïncident [3].

Plan du travail

• La première partie établit l'existence d'un barycentre canonique avec des hypothèses minimales.
- En I-1 nous rappelons les outils nécessaires.
- Dans la partie I-2 nous définissons avec la loi des grands nombres et de manière canonique, pour une v.a. X (ou une probabilité μ) intégrable, un point $b(X)$ (ou $b(\mu)$), appartenant à l'ensemble des espérances de Doss de X. Ce barycentre vérifie la propriété fondamentale :

$$d(b(X), b(Y)) \leq E[d(X, Y)] .$$

- Nous donnons dans I-3 un moyen d'obtenir le barycentre précédent sans utiliser la convergence presque sûre. Un exemple est abordé par simulation.
- La partie I-4 est consacrée au problème de cohérence sous deux aspects. Nous montrons d'abord que le barycentre obtenu à partir de la loi empirique converge p.s. vers le barycentre de la loi initiale (supposée intégrable). Nous donnons ensuite une condition nécessaire et suffisante pour que les barycentres des lois empiriques forment une martingale. Pour clore cette partie nous montrons que le théorème ergodique demeure dans ce cadre.

• Dans la partie II nous remplaçons l'espérance de Doss par celle de Herer puis par celle d'Émery-Mokobodzki.

I - 1 Définitions et notations

Dans toute la suite M est un espace métrique séparable complet muni de sa tribu borélienne. La distance de deux points x et y de M est notée xy.

Une variable aléatoire (v.a.) X définie sur un espace de probabilité $(\Omega, \mathcal{A}, P)$ et à valeurs dans M, est dite *intégrable*, si pour un point $a \in M$ (et donc pour tout point) la v.a.r. Xa est intégrable.
Pour une v.a. X intégrable, l'ensemble $\{E[X]\} = \{m | m \in M,\ am \leq E[aX],$ pour tout $a \in M\}$ est l'*espérance* de X au sens de Doss, nous écrirons aussi $\{E[\mu]\}$ pour μ probabilité (intégrable) sur M. L'ensemble $\{E[X]\}$ ne dépend que de la loi de X et est fermé.

L'espace M est dit *convexe*, respectivement *strictement convexe*, si pour toute probabilité $\mu = \frac{1}{2}(\delta_x + \delta_y)$, $\{E[\mu]\}$ est non vide, respectivement est un singleton, noté $b(x,y)$.
Nous utiliserons aussi une notion plus forte : un espace M convexe est dit à *coubure négative* si, $\forall (x_1, x_2, y_1, y_2) \in M^4$, $\forall u \in \{E[\frac{1}{2}(\delta_{x_1} + \delta_{x_2})]\}$ et tout $v \in \{E[\frac{1}{2}(\delta_{y_1} + \delta_{y_2})]\}$, on a $uv \leq \frac{1}{2}(x_1y_1 + x_2y_2)$. Un tel espace est strictement convexe (prendre $y_i = x_i$) et la condition de courbure négative s'écrit :

pour tout point $(x_1, x_2, y_1, y_2) \in M^4$, $b(x_1,x_2)b(y_1,y_2) \leq \frac{1}{2}(x_1y_1 + x_2y_2)$.

Soit $\mathcal{F}$ une sous tribu de $\mathcal{A}$, on dit qu'une v.a. Y à valeurs dans M est une *moyenne* (ou *espérance*) *conditionnelle* de X relativement à $\mathcal{F}$ si

(i) Y est $\mathcal{F}$-mesurable

(ii) $\forall a \in M$, $aY \leq E^{\mathcal{F}}[aX]$ p.s. (espérance conditionnelle d'une v.a.r.).

On note $\{E^{\mathcal{F}}[X]\}$ l'ensemble des v.a. Y vérifiant (i) et (ii).
Une suite $(Y_n)_n$ adaptée à une filtration monotone $(\mathcal{F}_n)_n$ est une *martingale* si $Y_n \in \{E^{\mathcal{F}_n}[Y_m]\}$ pour $m = n+1$ dans le cas d'une filtration croissante et pour $m = n-1$ lors d'une filtration décroissante.

I - 2 Barycentre d'une v.a. intégrable.

Dans un premier temps nous définissons un barycentre intermédiaire pour une loi uniforme sur un ensemble $\{x_1, x_2, \cdots, x_k\}$ où les x_i sont des points *non nécessairement distincts* de M.

Proposition 1 *Soit M un espace métrique complet strictement convexe. Pour tout $k > 1$, il existe une application b_k de M^k dans M telle que le point $b_k(x_1, x_2, \cdots, x_k)$, noté aussi $\sum_{i=1}^{k} \frac{1}{k} x_i$, appartient à $\{E[\sum_{i=1}^{k} \frac{1}{k}\delta_{x_i}]\}$ et vérifie*

$$\sum_{i=1}^{k} \frac{1}{k} x_i = \sum_{i=1}^{k} \frac{1}{k} \Big(\sum_{j \leq k, j \neq i} \frac{1}{k-1} x_j \Big).$$

De plus si l'espace est à courbure négative, alors

$$b_k(x_1, x_2, \cdots, x_k) b_k(y_1, y_2, \cdots, y_k) \leq \frac{1}{k} \sum_{1}^{k} x_i y_i,$$

en particulier l'application b_k est continue.

Preuve. L'existence de $b_k(\cdot)$ se fait par récurrence sur k.
Pour $k = 2$ et $\mu = \frac{1}{2}(\delta_x + \delta_y)$, comme M est strictement convexe posons $b_2(x,y) = b(x,y)$.
Par récurrence supposons que, pour tout $l < k$ ($k > 2$) et tout point $(x_1, x_2, \cdots, x_l) \in M^l$, on ait défini de manière canonique un barycentre $b_l(x_1, x_2, \cdots, x_l) = \sum_{i=1}^{l} \frac{1}{l} x_i$, vérifiant les assertions de la proposition.
Considérons alors un point $(x_1, x_2, \cdots, x_k) \in M^k$.
Définissons, par récurrence, la suite $(x_1^n, x_2^n, \cdots, x_k^n)_{n\geq 0}$, associée à $(x_1, \cdots, x_k)$, par : $x_i^0 = x_i$ et, pour $n \geq 1$, $x_i^n = \sum_{j\leq k, j\neq i} \frac{1}{k-1} x_j^{n-1}$.

Nous allons montrer que les x_i^n convergent, lorsque $n \to \infty$, vers un point indépendant de i et qui sera par définition $b_k(x_1, x_2, \cdots, x_k)$.

À cette fin, considérons l'ensemble $\mathcal{P}_f(A)$ des probabilités dont le support est contenu dans $A \subset M$, cardinal de A fini. L'enveloppe convexe de A est $E_1(A) = \bigcup_{\mu\in\mathcal{P}_f(A)} \{E[\mu]\}$. Notons $E_2(A) = E_1(E_1(A))$ et montrons que ces deux enveloppes convexes coïncident : "$A = \{x_1, \cdots x_k\} \Longrightarrow E_1(A) = E_2(A)$."

En effet soit $a \in E_2(A)$, on peut écrire $a \in \{E[\sum_{j\in J} \beta_j \delta_{m_j}]\}$ où J fini et $m_j \in \{E[\sum_{i=1}^{k} \alpha_i^j \delta_{x_i}]\}$. Posons $\gamma_i = \sum_{j\in J} \beta_j \alpha_i^j$ et $\mu = \sum_{i=1}^{k} \gamma_i \delta_{x_i}$. Pour $z \in M$ les inégalités : $az \leq \sum_j \beta_j m_j z \leq \sum_j \beta_j (\sum_i \alpha_i^j x_i z) = \sum_{i,j} \beta_j \alpha_i^j x_i z = \sum_i \gamma_i x_i z$, impliquent $a \in \{E[\mu]\}$ et donc l'inclusion $E_2(A) \subset E_1(A)$. La réciproque est triviale.
Remarquons que $E_1(A)$ est fermé.

Retour à la preuve de la proposition 1.
Avec les notations introduites à partir de la suite $(x_1, \cdots, x_k)$, considérons $A_n := \{x_1^n, x_2^n, \cdots, x_k^n\}$ et montrons que la suite $\big(E_1(A_n)\big)_n$ est décroissante et converge vers un point.

a) Décroissance

Comme $x_i^{n+1} \in E_1(A_n)$ pour tout i, on a $A_{n+1} \subset E_1(A_n)$ par conséquent $E_1(A_{n+1}) \subset E_2(A_n)$, avec l'égalité précédente, $E_1(A_{n+1}) \subset E_2(A_n) = E_1(A_n)$.

b) Montrons que les ensembles $E_1(A_n)$ et A_n ont même diamètre.

Le diamètre d'un ensemble A est, par définition, $\mathrm{diam}(A) := \sup_{x,y\in A} xy$. Soient z et u deux points de $E_1(A_n)$, $z \in \{E[\sum_i \gamma_i \delta_{x_i}]\}$, $u \in \{E[\sum_j \beta_j \delta_{y_j}]\}$ où les x_i et les y_j appartiennent à A_n. On a :

$$zu \leq \sum_i \gamma_i x_i u \leq \sup_i x_i u \leq \sup_i \sum_j \beta_j x_i y_j \leq \sup_{i,j} x_i y_j = \mathrm{diam}(A_n).$$

c) Montrons que $\mathrm{diam}(E_1(A_n))$ tend vers 0.
De manière équivalente montrons que $L^{(n)} := \sum_{i\neq j} x_i^n x_j^n \to 0$.

Comme $x_i^{n+1} \in \{E[\sum_{j\leq k, j\neq i} \frac{1}{k-1}\,\delta_{x_j^n}]\}$, on a $x_i^{n+1}x_j^{n+1} \leq (\frac{1}{k-1})^2 \sum_{p\neq i}\sum_{q\neq j} x_p^n x_q^n$.
En écrivant $\sum_{p\neq i}\sum_{q\neq j} x_p^n x_q^n = \sum_{p\neq q} x_p^n x_q^n \;-\sum_p x_p^n x_j^n - \sum_q x_i^n x_q^n + x_i^n x_j^n$, puis en sommant sur les couples $(i,j), i\neq j$, on arrive à :
$L^{(n+1)} = \sum_{i\neq j\leq k} x_i^{n+1}x_j^{n+1} \leq \frac{k^2-3k+3}{(k-1)^2}\,L^{(n)}$. L'espace M étant complet et $k>2$, nous avons la convergence de la suite $E_1(A_n)$ vers un unique point qui est, par définition, $b_k(x_1, x_2, \cdots, x_k)$.

Montrons que $b_k(x_1,x_2,\cdots,x_k) \in \{E[\frac{1}{k}\sum_{i=1}^{k} \delta_{x_i}]\}$.

On a les inclusions $\{E[\frac{1}{k}\sum_i \delta_{x_i^{n+1}}]\} \subset \{E[\frac{1}{k}\sum_i \delta_{x_i^n}]\} \subset \{E[\frac{1}{k}\sum_i \delta_{x_i}]\}$ et

$\{E[\frac{1}{k}\sum_i \delta_{x_i^n}]\} \subset E_1(A_n)$, à la limite, on obtient la relation souhaitée.

Le premier terme de la suite associée à $(x_1,\cdots,x_k)$ est $(x_1^1,\cdots,x_k^1)$ où $x_i^1 = \sum_{j\leq k, j\neq i} \frac{1}{k-1}\,x_j$ et, comme $b_k(x_1,\cdots,x_k) = b_k(x_1^1,\cdots,x_k^1)$, la relation liant $b_k(\cdot)$ et $b_{k-1}(\cdot)$ est démontrée.

Pour achever la preuve, considérons le cas d'un espace à courbure négative et montrons, par récurrence, l'inégalité recherchée.
Pour $k=2$, c'est la définition même de la courbure négative.
Admettons la relation jusqu'au rang $k-1$ et considérons deux points $(x_1,\cdots,x_k)$ et $(y_1,\cdots,y_k)$. En reprenant les notations précédentes,
$x_i^1 y_i^1 = (\sum_{j\neq i}\frac{1}{k-1}x_j)(\sum_{j\neq i}\frac{1}{k-1}y_j)$, et avec l'hypothèse de récurrence,

$x_i^1y_i^1 \leq \frac{1}{k-1}\sum_{j\neq i} x_j y_j$. Donc $\sum_{i=1}^{k} x_i^1 y_i^1 \leq \frac{1}{k-1}\sum_{i=1}^{k}\sum_{j\neq i} x_j y_j = \sum_{i=1}^{k} x_i y_i$.

En itérant, la suite $\big(\sum_{i=1}^{k} x_i^n y_i^n\big)$ est décroissante. La preuve s'achève en remarquant que $x_i^n \underset{n\to\infty}{\longrightarrow} b_k(x_1,\cdots,x_k)$ et que $y_i^n \underset{n\to\infty}{\longrightarrow} b_k(y_1,y_2,\cdots,y_k)$. $\square$

Remarques.

- Pour toute permutation σ de $\{1\cdots k\}$, $\sigma\in\Sigma(k)$, $\sum_{i=1}^{k}\frac{1}{k}x_i = \sum_{i=1}^{k}\frac{1}{k}x_{\sigma(i)}$.
- Cas d'une variable aléatoire étagée.

Pour X v.a. définie sur $[0,1]$, à valeurs dans M et de loi uniforme sur $(x_1,x_2,\cdots,x_k)\in M^k$, on écrit $b_k(X) := b_k(x_1,x_2,\cdots,x_k)$.
Lorsque $X=(X_1,\cdots,X_k)$ est une v.a. étagée à valeurs M^k, l'application :

$\omega \longrightarrow b_k(X(\omega)) = b_k(X_1(\omega), \cdots, X_k(\omega)) = \sum_{i=1}^{k} \frac{1}{k} X_i(\omega)$ est une v.a. étagée.

À partir de la suite $\left(b_n(\cdot)\right)_n$ nous allons définir un barycentre canonique, pour une variable aléatoire intégrable, en utilisant une propriété de martingale. Rappelons à cet effet le théorème de convergence p.s. des martingales selon [2] ou [6] :

"Si les boules fermées bornées de M sont compactes, alors toute martingale (Y_n) pour une filtration décroissante $(\mathcal{S}_n)$, est convergente p.s. i.e. $\lim_{n\to\infty} Y_n$ existe p.s."

Un espace dont les boules fermées et bornées sont compactes est dit *"finiment compact"*. Cette condition est en général trop exigeante pour obtenir uniquement la loi forte des grands nombres. Par exemple, elle n'est pas vérifiée dans les espaces de Banach de dimension infinie. C'est pour cela que nous introduisons la définition suivante :

L'espace M est fini-compact *si, pour tout $(x_1, \cdots, x_n) \in M^n$, l'ensemble $E_1(\{x_1, \cdots, x_n\})$ est compact.*

Énonçons maintenant le résultat central.

Théorème 2 (barycentre d'une v.a. intégrable)

Soient (X_n) une suite de v.a. i.i.d., X_1 intégrable dans M espace métrique complet séparable, fini compact à courbure négative et $\mathcal{S}_n$ la tribu des événements dépendant symétriquement de $(X_1, \cdots, X_n)$ et d'une manière quelconque de $(X_{n+1}, \cdots)$. Alors la suite $\left(Y_n = b_n(X_1, \cdots, X_n) = \sum_{i=1}^{n} \frac{1}{n} X_i\right)$ est une martingale adaptée à la filtration $(\mathcal{S}_n)$, convergeant p.s. vers une constante, $b(X_1)$, qui ne dépend que de la loi de X_1.

De plus, pour X et Y intégrables, on a $b(X) \in \{E[X]\}$ et $b(X)b(Y) \leq E[XY]$.

Preuve Montrons tout d'abord que Y_n est une v.a.

Pour cela construisons une suite de fonctions (f_k) permettant d'approcher les v.a. par des v.a. étagées. Notons (a_k) une suite dense dans M. Définissons $f_k : M \to M$ par $f_k(x) = a_i$ où i est le plus petit entier inférieur à k et tel que $a_i x \leq a_n x$, pour tout n, $1 \leq n \leq k$. On voit, sans difficulté, que $f_k(M) = \{a_1, \cdots a_k\}$ et que $\forall x \in M$, $x f_k(x)$ tend vers 0 et enfin, $x f_k(x) \leq a_1 x$.

Ainsi la suite de v.a. étagées $\left(X_n^k = f_k(X_n)\right)_k$ converge pour tout n fixé, ponctuellement et dans $\mathbb{L}^1(M)$, vers X_n, $\mathbb{L}^1(M)$ étant l'espace des v.a. à valeurs dans M et intégrables. Posons $Y_n^k(\omega) = \sum_{i=1}^{n} \frac{1}{n} X_i^k(\omega)$, Y_n^k est une v.a. étagée et la proposition 1 assure que $Y_n^k(\omega) \xrightarrow[k\to\infty]{} Y_n(\omega)$, Y_n est donc une v.a. De plus l'invariance par permutation : $\sum_1^n \frac{1}{n} x_i = \sum_1^n \frac{1}{n} x_{\sigma(i)}$, $\sigma \in \Sigma(n)$, montre que Y_n est $\mathcal{S}_n$ mesurable.

Afin d'établir que (Y_n) est une martingale, *i.e.* $Y_{n+1} \in \{E^{S_{n+1}}[Y_n]\}$, montrons la propriété suivante

(P) *Pour toute bijection,* $\sigma \in \Sigma(n+1)$, $\{E^{S_{n+1}}[\sum_1^n \frac{1}{n} X_i]\} = \{E^{S_{n+1}}[\sum_1^n \frac{1}{n} X_{\sigma(i)}]\}$.

Cette propriété une conséquence de

(P') *Pour tout* $a \in M$, $E^{S_{n+1}}[(a \sum_{i=1}^n \frac{1}{n} X_i)] = E^{S_{n+1}}[(a \sum_{i=1}^n \frac{1}{n} X_{\sigma(i)})]$ *p.s.*

Pour cela prenons une fonction $f : M^{n+1} \longrightarrow I\!R$ telle que $f(X_1, \cdots, X_{n+1})$ soit $\mathcal{S}_{n+1}$ mesurable. Notons $X = (X_1, \cdots, X_{n+1})$, $X_\sigma = (X_{\sigma(1)}), \cdots, X_{\sigma(n+1)})$ et $g(X) = a \sum_1^n \frac{1}{n} X_i$. Alors, en désignant par P_X la loi du vecteur X, $E[f(X)g(X)] = E_{P_X}[f(x)g(x)] = E[f(X_\sigma)g(X_\sigma)]$. Comme $f(X) = f(X_\sigma)$ *p.s.*, on a $E[f(X)a\sum_1^n \frac{1}{n} X_i] = E[f(X)a\sum_1^n \frac{1}{n} X_{\sigma(i)}]$. C'est la propriété désirée.

Revenons à la propriété de martingale.

La proposition 1 appliquée à la v.a. $\mathcal{S}_{n+1}$ mesurable $Y_{n+1} = \sum_1^{n+1} \frac{1}{n+1} X_i$, permet d'écrire :

$$Y_{n+1} = \sum_{i=1}^{n+1} \frac{1}{n+1} \Big(\sum_{j \leq n+1, j \neq i} \frac{1}{n} X_j \Big) \text{ et } a\, Y_{n+1} \leq \sum_{i \leq n+1} \frac{1}{n+1} \Big(a \sum_{j \leq n+1, j \neq i} \frac{1}{n} X_j \Big).$$

En prenant l'espérance conditionnelle par rapport à $\mathcal{S}_{n+1}$, il vient :

$$a\, Y_{n+1} \leq \sum_{i \leq n+1} \frac{1}{n+1} E^{S_{n+1}} \Big(a \sum_{j \leq n+1, j \neq i} \frac{1}{n} X_j \Big).$$

Avec la propriété (P'), $E^{S_{n+1}} \big(a \sum_{j \leq n+1, j \neq i} \frac{1}{n} X_j \big) = E^{S_{n+1}} \big(a \sum_{j \leq n} \frac{1}{n} X_j \big)$,

par conséquent, $aY_{n+1} \leq \sum_{i \leq n+1} \frac{1}{n+1} E^{S_{n+1}}(aY_n) = E^{S_{n+1}}(aY_n)$.

Par suite, $Y_{n+1} \in \{E^{S_{n+1}}[Y_n]\}$ et (Y_n) est bien une martingale.

Montrons la convergence presque sûre.

Pour X_1 étagée de loi $\mu = \sum_{i=1}^k \alpha_i \delta_{a_i}$ et $a \in M$, la suite $\big(aY_n(\omega)\big)$ est une sous-martingale positive convergeant p.s. Ainsi, pour $a \in D, D$ dénombrable dense dans M, on a la convergence p.s. de aY_n, Comme $Y_n \in E_1(\{a_1, \cdots, a_k\})$, l'hypothèse de compacité de cet ensemble implique alors que la suite $\big(Y_n(\omega)\big)$ converge p.s. Par le théorème de $0,1$ de Kolmogorov, la limite est presque sûrement une constante. Elle est notée provisoirement $b((X_n))$.

Montrons que cette limite ne dépend que de la loi.

On remarque que si X et Y sont intégrables, il en est de même pour XY.
Soit $(Y_n)_n$ une autre suite de v.a. i.i.d. En utilisant la propriété d'invariance par permutation de b_n pour τ et $\sigma \in \Sigma(n)$:

$$b_n(X_1(\omega),\cdots,X_n(\omega))b_n(Y_1(\omega),\cdots,Y_n(\omega)) = b_n(X_{\tau(1)}(\omega),\cdots,X_{\tau(n)}(\omega))b_n(Y_{\sigma(1)}(\omega),\cdots,Y_{\sigma(n)}(\omega)).$$

On voit, avec la proposition 1 et la convergence précédente, que

$$b((X_n))b((Y_n)) \leq \underline{\lim}_n \inf_{\sigma,\tau\in\Sigma(n)} \frac{1}{n}\sum_1^n X_{\tau(i)}(\omega)Y_{\sigma(i)}(\omega) \ p.s. \tag{1}$$

En prenant l'espérance dans l'inégalité (1), on obtient

$$b((X_n))b((Y_n)) \leq E[\underline{\lim}_n \frac{1}{n}\sum_{i=1}^n X_iY_i] \leq \underline{\lim}_n \frac{1}{n}\sum_{i=1}^n E[X_iY_i]. \tag{2}$$

Soit $\big((X_n,Y_n)\big)_n$ une suite i.i.d. de même loi que (X,Y). L'inégalité (2) donne

$$b((X_n))b((Y_n)) \leq \underline{\lim}_n \frac{1}{n}\sum_{i=1}^n E[X_iY_i] = E[XY]. \tag{3}$$

Cas des v.a.étagées. Supposons que X_1 et Y_1 ont même loi : $\mu = \sum_1^p \alpha_i\delta_{a_i}$.
À chaque n et ω, associons deux permutations τ et σ telles que $(X_{\tau(1)}(\omega),\cdots,X_{\tau(n)}(\omega)) = (a_1,\cdots,a_1,a_2,\cdots,a_{p-1},a_p,\cdots,a_p)$. Le nombre de répétitions de chaque a_i est $p_i(\omega,n) = \text{card}\{j|j\leq n, X_j(\omega)=a_i\}$. Procédons de même pour $(Y_{\sigma(1)}(\omega),\cdots,Y_{\sigma(n)}(\omega))$ avec $q_i(\omega,n) = \text{card}\{j|j\leq n, Y_j(\omega)=a_i\}$. Notons enfin $C = \sup_{i,j} a_ia_j$. Avec les relations précédentes :

$$b_n(X_1(\omega),\cdots,X_n(\omega))b_n(Y_1(\omega),\cdots,Y_n(\omega)) \leq \frac{C}{n}\sum_1^p |p_i(\omega,n)-q_i(\omega,n)|.$$

Pour presque tout ω, le terme de droite de l'inégalité tend vers 0 quand $n\to\infty$. Ainsi pour une v.a. étagée X, $b((X_n))$ ne dépend que de la loi, la notation $b(X)$ est justifiée.
De plus, avec la partie précédente, pour deux v.a. étagées, on a l'inégalité

$$b(X)b(Y) \leq E[XY]. \tag{4}$$

Cas des v.a. intégrables. Pour une suite $(X_n)_n$ i.i.d. avec X_1 intégrable, posons $X_n^k = f_k(X_n)$ où les f_k sont les fonctions définies auparavant. Pour chaque k, la suite $(X_n^k)_n$ est i.i.d. et X_1^k étagée. Nous avons vu, dans ces conditions, que la suite $\Big(b_n^k(\omega) = b_n\big(X_1^k(\omega),\cdots,X_n^k(\omega)\big)\Big)_n$ converge p.s. vers un point noté b^k.
La suite $(b^k)_k$ est convergente car l'inégalité (4) donne $b^kb^p \leq E[X_1^kX_1^p]$, ce dernier terme tend vers 0. Notons b la limite de la suite (b^k).
L'inégalité triangulaire et (4) montrent que, pour $Y_n(\omega) = \sum_{i=1}^n \frac{1}{n}X_i(\omega)$,

$$bY_n(\omega) \le bb_n^k(\omega) + b_n^k(\omega)Y_n(\omega) \le bb_n^k(\omega) + \frac{1}{n}\sum_{i=1}^{n} X_i^k(\omega)X_i(\omega).$$

La suite $(Z_i^k = X_i^k X_i)_i$ est i.i.d. réelle, Z_1^k intégrable, donc $\frac{1}{n}\sum_{i=1}^{n} Z_i^k$ converge p.s. vers $E[X_1^k X_1]$.
Ainsi presque sûrement, $\overline{\lim_n}\ bY_n(\omega) \le \overline{\lim_n}\ bb_n^k(\omega)+E[X_1^k X_1] = bb^k + E[X_1^k X_1]$.
Ceci prouve la convergence p.s. de la suite $\big(Y_n(\omega)\big)_n$ vers $b = \lim_k\ b(X_1^k)$.
De plus, en reprenant la partie précédente, on voit que les inégalités (1), (2), (3) et (4) demeurent pour les variables intégrables.
Nous avons en définitive montré que, pour toute suite i.i.d. $(X_n)_n$, X_1 intégrable, la suite $(Y_n = b_n(X_1, \cdots, X_n)_n$ converge p.s. vers une constante ne dépendant que de la loi de X_1, $b(X_1) = \lim_k\ b(X_1^k)$. Enfin, en choisissant $Y = a$ dans la relation (4), on obtient $b(X) \in \{E[X]\}$. $\square$

Remarque : Il se peut que $b_4(x,y,z,z) \neq b_8(x,x,y,y,z,z,z,z)$, et que $b_n \neq b$ pour tout n, ce qui rend difficile, a priori, l'évaluation de $b(x,y,z,z)$.

I - 3 Identification de la limite

Nous allons donner un moyen de "calculer" $b(\mu)$ pour certaines probabilités en contournant la convergence presque sûre.

Soient $\mu = \frac{1}{k}\sum_{i=1}^{k} \delta_{a_i}$, les a_i distincts, et (X_i) une suite de v.a. i.i.d. de loi μ. Définissons la suite $\mathbf{a} = (a_1, \cdots, a_k, a_1, \cdots)$ formée par répétitions successives de la suite finie $(a_1, \cdots, a_k)$ et notons $\mathbf{a}^n$ la suite formée par les n premiers termes de $\mathbf{a}$. Nous allons montrer que

$$\lim_{n\to\infty} b_n(\mathbf{a}^n) = b(\mu).$$

Fixons provisoirement un entier n et soit $n = km + k^*, 0 \le k^* < k$, son écriture modulo k. Pour tout $i \in \{1\cdots k\}$, notons p_i^n le cardinal de l'ensemble $I_i^n = \{j | 1 \le j \le n,\ X_j(\omega) = a_i\}$, $p_i^n = \#(I_i^n)$. En ordonnant de manière croissante, on a $I_i^n = \{j_1^i, \cdots, j_{p_i^n}^i\}$. Pour $J = \{1\cdots m\}$, posons $J_l = (l-1)m + J$, $\tilde{J}_l = \{1\cdots m \wedge p_l^n\}$, $\tilde{J}_l^n = (l-1)m + \tilde{J}_l$ où $l \in \{1\cdots k\}$ et $J_{k+1} = \{km+1, \cdots, n\}$. Les J_l, pour $l \in \{1\cdots k+1\}$, forment une partition de $\{1\cdots n\}$ et les $\tilde{J}_l^n$ sont disjoints pour $1 \le l \le k$.
Soit $q \in \tilde{J}_l^n$, posons $\sigma(q) = j_q^l$; on a $X_{\sigma(q)}(\omega) = a_l$ et σ est une injection de $\bigcup_{l=1}^{k} \tilde{J}_l^n$ dans $\{1\cdots n\}$ on peut donc la prolonger en une permutation, notée encore σ. Par ailleurs, $\#\Big(\bigcup_{l=1}^{k} \tilde{J}_l^n\Big) = \sum_{l=1}^{k} m \wedge p_l^n$, et si $C = \sup_{i,j} a_i a_j$, alors

$$b_n(X_1(\omega), \cdots, X_n(\omega))b_n(\mathbf{a}^n) \le \frac{C}{n}\Big(\sum_{l=1}^{n}(|m - p_l^n| + m \wedge p_l^n) + k^*\Big).$$

Ainsi, pour presque tout ω, lorsque $n \to \infty$ on obtient le résultat anoncé. □
Ce procédé s'étend naturellement à toute probabilité $\mu = \sum_{i=1}^{k} \alpha_i \delta_{a_i}$ à coefficients α_i rationnels.

Exemple Dans l'espace formé par l'arbre homogène ternaire à distance unitaire, soient x, y, z les trois sommets à distance 1 d'un sommet donné o. Une simulation montre que le barycentre $b(\mu)$ de la mesure $\mu = \frac{1}{2}\delta_x + \frac{1}{4}\delta_y + \frac{1}{4}\delta_z$ est situé sur la branche joignant o à x et $ob(\mu)$ est voisin de 0,15.

I - 4 Cohérence et théorème ergodique

Le problème soulevé ici est le suivant : le théorème 2 demeure-t-il lorsque l'on remplace $Y_n = b_n(X_1, \cdots, X_n)$ par $Y_n = b(X_1, \cdots, X_n)$? En d'autres termes le barycentre ainsi défini vérifie-t-il la loi forte des grands nombres ?
Le résultat suivant montre que la partie "convergence p.s." reste valable.

Proposition 3 (de cohérence) *Sous les conditions du théorème 2, la suite* $\left(b(\sum_{i=1}^{n} \frac{1}{n}\delta_{X_i(\omega)})\right)_n$ *converge p.s. vers* $b(X)$.

En fait l'existence de $b(\cdot)$ permet d'affaiblir les hypothèses, en particulier de supprimer l'hypothèse "fini compact". Le théorème suivant regroupe la propriété de convergence presque sûre et celle de martingale :

Théorème 4 *Soient M un espace métrique complet séparable et $\mathcal{P}_1$ l'ensemble des probabilités intégrables sur la tribu borélienne de M ($\mu \in \mathcal{P}_1 \Longrightarrow \int ax\, d\mu(x) < \infty$, $a \in M$). Supposons qu'il existe une application $b : \mathcal{P}_1 \longrightarrow M$, telle que $b(\mu)b(\nu) \leq E[XY]$, pour toutes v.a. X et Y de lois respectives μ et ν. Alors $\forall\ \mu \in \mathcal{P}_1$ et $\forall\ (X_n)$ i.i.d. de loi μ, la suite $\left(Y_n(\omega) = b(\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(\omega)})\right)$ converge presque sûrement vers $b(\mu)$.*
De plus les deux assertions suivantes sont équivalentes :

(i) $\forall\ \mu \in \mathcal{P}_1$ et $\forall (X_n)_n$ i.i.d. de loi μ, $\left(Y_n(\omega) = b(\frac{1}{n}\sum_{i=1}^{n} \delta_{X_i(\omega)})\right)$ est une martingale adaptée à $\mathcal{S}_n$.

(ii) le barycentre $b(\cdot)$ vérifie, $\forall\ a \in M$, $\forall n$ et $\forall\ (x_1, \cdots, x_{n+1}) \in M^{n+1}$,

$$ab\Big(\frac{1}{n+1}\sum_{i=1}^{n+1} \delta_{x_i}\Big) \leq \sum_{i=1}^{n+1} \frac{1}{n+1} ab\Big(\sum_{j\leq n, j\neq i} \frac{1}{n}\,\delta_{x_j}\Big).$$

Preuve Montrons au préalable que Y_n est $\mathcal{S}_n$ mesurable.
Soient $X_1, \cdots X_n$, n variables mesurables par rapport à une tribu $\mathcal{A}$. Avec les fonctions f_k précédentes posons $X_i^k = f_k(X_i)$, on obtient des v.a. étagées $\mathcal{A}$

mesurables, pour tout k et tout $i = 1 \cdots n$. Pour chaque k, il existe donc une partition Π_k finie : $A_1, \cdots, A_p$, dépendant de k et formée d'éléments de $\mathcal{A}$ telle que sur chaque A_j, $X_i^k(\omega) = x_i^k(j)$. Notons les lois empiriques $\mu_n^\omega = \sum_{i=1}^n \frac{1}{n}\delta_{X_i(\omega)}$ et $({}_k\mu)_n^\omega = \sum_{i=1}^n \frac{1}{n}\delta_{X_i^k(\omega)}$. Par hypothèse on a l'inégalité :
$b(\mu_n^\omega)b(({}_k\mu)_n^\omega) \leq \frac{1}{n}\sum_1^n X_i(\omega)X_i^k(\omega)$. Le terme de droite tend vers 0 quand $k \to \infty$, donc $b(({}_k\mu)_n^\omega) \underset{k\to\infty}{\longrightarrow} b(\mu_n^\omega)$. Or, pour $\omega \in A_j$, $b(({}_k\mu)_n^\omega) = b(\sum_{i=1}^n \frac{1}{n}\delta_{x_i^k(j)})$ et par suite, l'application $\omega \longrightarrow b(({}_k\mu)_n^\omega)$ est $\mathcal{A}$ mesurable. À la limite, $b(\mu_n^\omega)$ est $\sigma(X_1, \ldots, X_n)$ mesurable.
Remarquons alors que la probabilité μ_n^ω est invariante pour toute permutation $\tau \in \Sigma(n)$, en effet $\mu_n^\omega = \frac{1}{n}\sum_1^n \delta_{X_{\tau(i)}(\omega)}$. En reprenant le raisonnement précédent, on en déduit que $Y_n(\omega) = b(\mu_n^\omega)$ est $\mathcal{S}_n$ mesurable.

Montrons la convergence presque-sûre.
De manière classique, par exemple [14] theorem 7.1, on a, presque sûrement en ω, la convergence en loi des lois empiriques $\mu_n^\omega = \frac{1}{n}\sum_{i=1}^n \delta_{X_i(\omega)}$ vers μ. Le théorème de représentation de Skohorod, [16] [4], assure de l'existence, pour tout $\omega \notin \mathcal{N}$, $\mathcal{N}$ négligeable, d'une suite $\left(U_n^\omega(\cdot)\right)_n$ de variables aléatoires définies sur l'espace canonique $([0,1], \mathcal{B}, \lambda)$, de loi μ_n^ω et convergeant p.p. vers une v.a. U^ω de loi μ. Lorsque μ est concentrée sur une partie bornée, ce qui est le cas lorsque μ est la loi de X_1 d'une v.a. bornée, le terme de droite de l'inégalité : $b(\mu_n^\omega)b(\mu) \leq E_\lambda[U_n^\omega(\cdot)U^\omega(\cdot)]$ tend vers 0. Ainsi la première assertion du théorème est prouvée pour X_1 bornée.
Le cas général se déduit de l'inégalité :
$b(\mu)b(\mu_n^\omega) \leq b(\mu)b({}_k\mu) + b({}_k\mu)b(({}_k\mu)_n^\omega) + b(({}_k\mu)_n^\omega)b(\mu_n^\omega)$ où ${}_k\mu$ est la loi de X_1^k.
En effet, $b(({}_k\mu)_n^\omega)b(\mu_n^\omega) \leq \frac{1}{n}\sum_{i=1}^n X_i(\omega)X_i^k(\omega)$ qui tend vers $E[X_1X_1^k]$ lorsque $n \to \infty$, le terme $b(\mu)b({}_k\mu)$ est inférieur à $E[X_1X_1^k]$ qui tend vers 0 quand k infini et enfin, nous venons de le voir, $b({}_k\mu)b(({}_k\mu)_n^\omega)$ tend vers 0 pour tout k.

Établissons maintenant l'équivalence.
Comme dans la propriété (P), pour toute permutation $\tau \in \Sigma(n+1)$, on établit l'égalité
$$\{E^{\mathcal{S}_{n+1}}[b(\mu_n(\omega))]\} = \{E^{\mathcal{S}_{n+1}}[b(\frac{1}{n}\sum_{i=1}^n \delta_{X_{\tau(i)}(\omega)})]\}.$$
Avec la notation $Y^i(\omega) = b(\sum_{j\leq n+1, j\neq i} \frac{1}{n}\delta_{X_j(\omega)})$ alors, pour tout $a \in M$,
$$(*) \qquad E^{\mathcal{S}_{n+1}}[aY_n] = E^{\mathcal{S}_{n+1}}[aY^i].$$

Montrons que la condition est nécessaire.
Si $(Y_n)_n$ est une martingale alors pour tout $a \in M$,

$$aY_{n+1} \leq E^{\mathcal{S}_{n+1}}[aY_n] = E^{\mathcal{S}_{n+1}}[aY^i]. \text{ D'où } aY_{n+1} \leq \sum_1^{n+1} \frac{1}{n+1} E^{\mathcal{S}_{n+1}}[aY^i].$$

La v.a.r. $\sum_1^{n+1} \frac{1}{n+1} aY^i = \sum_{i=1}^{n+1} \frac{1}{n+1} ab\Big(\sum_{j \leq n+1, j \neq i} \frac{1}{n} \delta_{X_j}\Big)$ étant clairement $\mathcal{S}_{n+1}$ mesurable, on obtient $aY_{n+1} \leq \sum_1^{n+1} \frac{1}{n+1}\ aY^i$. C'est la relation cherchée en prenant $\mu = \frac{1}{n} \sum_1^n \delta_{x_i}$.

Montrons que la condition est suffisante.
La v.a. $Y_{n+1} = b(\sum_1^{n+1} \frac{1}{n+1} \delta_{X_i})$ vérifie alors $aY_{n+1} \leq \sum_1^n \frac{1}{n+1}\ aY^i$ p. s. Donc $aY_{n+1} \leq \sum_1^{n+1} \frac{1}{n+1} E^{\mathcal{S}_{n+1}}[aY^i]$. Avec la relation $(*)$, il vient $aY_{n+1} \leq E^{\mathcal{S}_{n+1}}[aY_n]$ qui est la propriété de martingale. La preuve du théorème est donc achevée. $\square$

Comme application nous allons prouver un théorème ergodique. À cette fin nous supposerons que l'espace de probabilité est l'espace de Lebesgue et, par conséquent, pour toute sous tribu $\mathcal{C}$ de la tribu borélienne $\mathcal{B}$, il existe une version régulière de la probabilité conditionnelle sachant $\mathcal{C}$, *i.e.* un noyau N défini sur $\Omega \times \mathcal{B}$ tel que $P(\cdot) = \int N(\omega, \cdot) dP(\omega)$ et que $N(\cdot, B)$ soit une v.a.r. $\mathcal{C}$-mesurable pour tout $B \in \mathcal{B}$. Ceci permet d'introduire la définition suivante :

Définition

Avec les conditions du théorème 4, soient X une variable aléatoire intégrable à valeurs dans M et $\mathcal{C}$ une sous tribu. *L'espérance conditionnelle de X sachant $\mathcal{C}$* est la variable aléatoire $b^{\mathcal{C}}(X)(\omega) = b_{N(\omega,\cdot)}(X)$, barycentre de X par rapport au noyau de conditionnement N relatif à $\mathcal{C}$.

Il est aisé de vérifier que l'application $\omega \longrightarrow b^{\mathcal{C}}(X)(\omega)$ est $\mathcal{C}$ mesurable et $b^{\mathcal{C}}(X) \in \{E^{\mathcal{C}}[X]\}$. De plus pour deux variables X et Y intégrables, on a : $b^{\mathcal{C}}(X)(\omega) b^{\mathcal{C}}(Y)(\omega) \leq E^{\mathcal{C}}[XY](\omega)$ *p.s.*

Théorème 5 (Théorème ergodique)
Soit (X_n) une suite stationnaire de v.a. intégrables à valeurs dans M. Alors, sous les conditions du théorème 4 et, en désignant par $\mathcal{I}$ la tribu des événements invariants, $b(\mu_n^\omega = \frac{1}{n} \sum_1^n \delta_{X_i(\omega)})$ converge p.s. vers $b^{\mathcal{I}}(X_1)(\omega)$.

Preuve Considérons d'abord le cas où X_1 est étagée. Le théorème de Birkhoff montre que pour presque tout ω et pour toute fonction numérique f continue bornée $\frac{1}{n} \sum_1^n f(X_i(\omega))$ converge vers $E^{\mathcal{I}}[f(X_1)](\omega) = \int_M f(x)\ d\mu_{\mathcal{I}}^\omega(x)$. Le

théorème de représentation de Skohorod assure de l'existence d'une suite de v.a. $U_n^\omega(\cdot)$ de loi μ_n^ω, convergeant p.s. vers une v.a. $U^\omega(\cdot)$ de loi $\mu_{\mathcal{I}}^\omega$.
Le théorème de convergence dominée de Lebesgue montre que

$$b^{\mathcal{I}}(X_1)(\omega)\sum_1^n \frac{1}{n}X_i(\omega) = b(U^\omega)b(U_n^\omega) \leq E[U_n^\omega(\cdot)U^\omega(\cdot)] \underset{n\to\infty}{\longrightarrow} 0.$$

Le cas général se fait de manière similaire avec les fonctions f_k. □

II Autres notions d'espérances

II - 1 Sur l'espérance au sens de Herer

Le but de cette partie est d'obtenir des résultats similaires en remplaçant l'espérance au sens de Doss, $\{E[X]\}$, par celle de Herer [12], notée $\mathcal{E}_{\mathcal{H}}(X)$. L'espace de probabilité de référence est, ici, l'espace canonique $([0,1],\mathcal{B},\lambda)$. Soient $a_1,\cdots,a_k$ des points distincts de M, affectés de masses α_i, $\sum_1^n \alpha_i = 1$. Une suite de points $(x_1,\cdots,x_n)$, $n \geq k$ affectés de masses respectives $\beta_i, i = 1\cdots n$ est dite issue de $\mu = \sum_1^k \alpha_i\delta_{a_i}$ si

$\{x_1,\cdots,x_n\} = \{a_1,\cdots,a_k\}$, et, pour tout i, $\alpha_i = \sum_{j|x_j=a_i} \beta_j$; *i.e.* $\mu = \sum_1^n \beta_j\delta_{x_j}$.

À partir d'une suite de points $(x_i)_{i=1\cdots n}$ affectés de masses β_i, on construit un arbre binaire (chaque sommet intermédiaire joint deux points) issu d'un sommet et dont les extrémités finales sont les x_j. En remontant à partir des extrémités finales, on affecte à chaque sommet intermédiaire la masse somme des masses de ses deux extrémités (la masse du sommet initial est 1). De la même façon, en remontant, le sommet joignant deux points m_l et m_p de masses respectives γ_l et γ_p est appelé $\dfrac{\gamma_l}{\gamma_l+\gamma_p}m_l + \dfrac{\gamma_p}{\gamma_l+\gamma_p}m_p$.
Le nom du sommet initial dépend de la forme de l'arbre et est identifié à un point de M.

On note $\mathcal{H}(\mu)$ l'ensemble des points de M, obtenus comme sommets initiaux d'un arbre binaire du type précédent, lorsque l'on fait varier les paramètres suivants : n, la suite $(x_1,\cdots x_n)$ avec les masses $(\alpha_i)_{i=1\cdots n}$ issue de μ, et la forme de l'arbre.

L'espérance de μ au sens de Herer est, par définition, $\mathcal{E}_{\mathcal{H}}(\mu) := \overline{\mathcal{H}(\mu)}$. On écrit aussi $\mathcal{H}(X)$ et $\mathcal{E}_{\mathcal{H}}(X)$ si X est de loi μ.

Soient $m_1,\cdots,m_p$ p points de $\mathcal{H}(\mu)$. Chaque m_i est le sommet initial d'un arbre binaire dont les n_i extrémités sont des points que nous notons, pour simplifier, $x_{n_{i-1}+1},\cdots,x_{n_{i-1}+n_i}$, avec la convention $n_0 = 0$. La suite $(x_j)_{n_{i-1}+1\leq j\leq n_{i-1}+n_i}$ affectée de masses $(\beta_j^i)_j$ est issue de μ. Construisons un

arbre binaire partant d'un point m et dont les extrémités sont les $m_i, 1 \leq i \leq p$. On obtient naturellement, par prolongement, un arbre binaire joignant m aux points $x_1, \cdots, x_k$ où $k = \sum_1^p n_i$. Affectons chaque m_i de la masse γ_i, $\sum_i^p \gamma_i = 1$, et chaque x_j, pour $n_{i-1} + 1 \leq j \leq n_{i-1} + n_i$, est affecté de la masse $\gamma_i \beta_j^i$. On obtient alors une suite issue de μ. Le point m s'interprète comme un point de $\mathcal{H}(\mu)$ et aussi comme un point de $\mathcal{H}(\sum_1^p \gamma_i \delta_{m_i})$. Ceci montre le résultat suivant :
"Soient μ une probabilité discrète sur M et ν une autre probabilité discrète dont le support est contenu dans $\mathcal{H}(\mu)$, alors $\mathcal{H}(\nu) \subset \mathcal{H}(\mu)$."
Ou encore
"Soient X et Y deux v.a. étagées, si Y prend ses valeurs dans $\mathcal{H}(X)$, alors $\mathcal{H}(Y) \subset \mathcal{H}(X)$."

On obtient alors la propriété suivante qui, grâce à la proposition 1, sert de point de départ de l'étude du barycentre canonique relatif à l'espérance de Doss. En adaptant les notations de la partie I-1, posons $\mathcal{H}_1(A) = \bigcup_{\mu \in \mathcal{P}(A)} \mathcal{E}_{\mathcal{H}}(\mu)$ et $\mathcal{H}_2(A) = \mathcal{H}_1(\mathcal{H}_1(A))$. On a : *si $A = \{x_1, \cdots, x_n\}$, alors* $\mathcal{H}_1(A) = \mathcal{H}_2(A)$.
En effet l'inclusion $\mathcal{H}_1(A) \subset \mathcal{H}_2(A)$ est triviale en prenant pour ν une masse de Dirac, et l'inclusion inverse est une conséquence des propriétés ci-dessus.

L'adaptation de la proposition 1 permet d'obtenir l'existence de $b_n(x_1 \cdots x_n)$ comme un point de $\mathcal{E}_{\mathcal{H}}(\mu = \sum_1^n \frac{1}{n} \delta_{x_i})$ et, de là, une version du théorème 2 relative à l'espérance au sens de Herer.

II - 2 Sur l'espérance au sens d'Émery et Mokobodzki

Suivant [7] on considère, à la place de l'espace M précédent, un espace V qui est à courbure négative et est un compact géodésiquement convexe d'une variété riemanniene.
Un point $x \in V$ est un *barycentre convexe* de μ probabilité sur V, si pour toute fonction f convexe définie sur un voisinage de V on a $f(x) \leq \mu(f)$, rappelons que f est convexe si sa restriction à toute géodésique est convexe. On note $\{E_{E-M}[\mu]\}$ l'ensemble des barycentres convexes de μ.
L'égalité entre les ensembles $\mathcal{E}_1(A = \{x_1, \cdots, x_n\}) = \bigcup_{\mu \in \mathcal{P}(A)} \{E_{E-M}[\mu]\}$ et $\mathcal{E}_2(A) = \mathcal{E}_1(\mathcal{E}_1(A))$ résulte de la proposition 1 de [7].

Montrons que le barycentre $b(\mu)$ appartient à $\{E_{E-M}[\mu]\}$.
Établissons d'abord la propriété pour une probabilité discrète $\mu = \frac{1}{k} \sum_1^k \delta_{x_i}$ et pour $b_k(\mu)$.
Pour $k = 2$, l'exemple d) de [7] montre que $b_2(x, y) \in \{E_{E-M}[\frac{1}{2}(\delta_x + \delta_y)]\}$.

Admettons par récurrence, que $f(b_l(x_1,\cdots,x_l)) \leq \frac{1}{l}\sum_1^l f(x_i)$ pour $l \leq k-1$ et $k > 2$. Il faut établir $f(\sum_1^k \frac{1}{k}x_i) \leq \frac{1}{k}\sum_1^k f(x_i)$. Avec les notations de la proposition 1 : $x_i^{n+1} = \sum_{j\neq i}\frac{1}{k-1}x_j^n$, on obtient $f(x_i^{n+1}) \leq \frac{1}{k-1}\sum_{j\neq i} f(x_j^n)$. D'où, en notant $S^n = \sum_{j=1}^k f(x_j^n)$, $S^{n+1} \leq \frac{1}{k-1}(kS^n - \sum_{i=1}^k f(x_i^n)) = S^n$, la suite (S^n) est décroissante. Par continuité, pour tout $\epsilon > 0$, on a pour n suffisament grand, $f(b_k(x_1,\cdots,x_k)) \leq \frac{1}{k}S^n \leq \epsilon + \frac{1}{k}\sum_{j=1}^k f(x_j)$, c'est la relation cherchée : $b_k(\mu) \in \{E_{E-M}[\mu]\}$.

Le passage à $b(\cdot)$ peut se faire par (voir I-3) : $b(\mu = \frac{1}{k}\sum_1^k \delta_{x_i}) = \lim_n b_n(\mathbf{x}^n)$.

Nous avons, pour f convexe continue et bornée et $n = kp$,

$$f(b(\mu)) = \lim_n f(b_n(\mathbf{x}^n)) \leq \lim_n \frac{1}{n}\sum_1^n f(x_i^n) = \frac{1}{k}\sum_1^k f(x_i).$$

Ce qui exprime que $b(\mu) \in \{E_{E-M}[\mu]\}$. On peut aussi, plus rapidement, utiliser la fermeture de $\{E_{E-M}[\mu]\}$.

Pour le cas général, considérons une suite (X_n) i.i.d. de loi μ intégrable. Le théorème 2 montre la convergence p.s. de $b(\frac{1}{n}\sum_1^n \delta_{X_i(\omega)})$ vers $b(\mu)$, ce qui donne la relation cherchée. $\square$

Remarquons aussi que les conditions a) et b) de la définition 1.2 de [13] vérifient la seconde partie du théorème 4.

Pour conclure, examinons le cas d'un exemple simple.

On considère la loi uniforme $\mu = \frac{1}{3}(\delta_x + \delta_y + \delta_z)$ où x, y et z sont trois points du plan muni de sa structure euclidienne usuelle. Le barycentre ordinaire coïncide avec le barycentre exponentiel. Si on modifie la structure en introduisant une perturbation dans un (petit) voisinage d'un point situé par exemple au milieu du segment $[x, y^1], y^1$ étant, conformément à nos notations, le milieu de $[x, z]$, le barycentre exponentiel ne change pas. Par contre le milieu du segment $[x, z]$ change et le barycentre canonique de μ (pour la nouvelle structure) diffère alors du barycentre exponentiel.

Références

[1] M. ARNAUDON *Barycentres convexes et approximations des martingales continues dans les variétés.* Séminaire de Probabilités XXIX, Lecture Notes in Mathematics **1613** 70-85, 1997.
[2] V. BENÈS *Martingales on metric spaces.* Theor. Veroyatnost. i Primenen **7**, 81-82 1962.
[3] B. BRU, H. HEINICH, J.C. LOOGIETER *Distance de Lévy et extensions des théorèmes de la limite centrale et de Glivenko-Cantelli.* Pub. Inst. Stat. Univ. de Paris **37**, 29-42, 1993.
[4] J. A. CUESTA, C. A. MATRÁN *Strong convergence of weighted sums of random element through the equivalence of sequences of distributions.* J. Multivariate Anal. **25**, 311-322, 1988.
[5] S. DOSS *Sur la moyenne d'un élément aléatoire dans un espace distancié* Bull. Sci. Math. **73**, 48-72, 1949
[6] S. DOSS *Moyennes conditionnelles et martingales dans un espace métrique* C. R. Acad. Sci. Paris Série I, t.**254**, 3630-3632, 1962
[7] M. ÉMERY, G. MOKOBODZKI *Sur le barycentre d'une probabilité dans une variété* Sémi. Prob. XXV, Lect. Notes in Math. **1485**, 220-233, 1991
[8] M. FRECHET *Les éléments aléatoires de nature quelconque* Ann. Inst. H. Poincaré **14**, 215-310, 1948.
[9] W. HERER *Espérance mathématique au sens de Doss d'une variable aléatoire dans un espace métrique* C. R. Acad. Sci. Paris Série I, t.**302**, 131-134, 1983.
[10] W. HERER *Espérance mathématique d'une variable aléatoire à valeurs dans un espace métrique à courbure négative* C. R. Acad. Sci. Paris Série I, t.**306**, 681-684, 1988 .
[11] W. HERER *Mathematical expectation and martingales of random subsets of a metric spaces* Prob. and Math. Stat. **11**, 291-304, 1991.
[12] W. HERER *Mathematical expectation and strong law of large numbers for random variables with values in a metric space of negative curvative* Prob. and Math. Stat. **13**, 59-70, 1992.
[13] J. PICARD *Barycentres et martingales sur une variété* Ann. Inst. H. Poincaré **30**, 647-702, 1994.
[14] K. R. PARTHASARATHY *Probability Measures on Metric Spaces.* Academic Press 1967.
[15] P. RAYNAUD de FITTE *Théorème ergodique ponctuel et lois fortes des grands nombres pour des points aléatoires d'un espace métrique à courbure négative.* À paraitre dans Annals of Probability.
[16] A. V. SKOHOROD *Limit theorems for stochastic processes* Theory Probab. Appl. **1** 261-290, 1956.