Séminaire d'analyse fonctionnelle
École Polytechnique

B. Maurey

# Système de Haar

*Séminaire d'analyse fonctionnelle (Polytechnique)* (1974-1975), exp. nº 1, p. 1-11
<http://www.numdam.org/item?id=SAF_1974-1975____A1_0>

ECOLE POLYTECHNIQUE

CENTRE DE MATHEMATIQUES

*17, rue Descartes*
*75230 Paris Cedex 05*

S E M I N A I R E M A U R E Y - S C H W A R T Z 1 9 7 4 - 1 9 7 5

# SYSTEME DE HAAR

par B. MAUREY

Exposé N$^{o}$ I

6 Novembre 1974

Cet exposé est consacré à une revue rapide des notions fondamentales de la théorie des martingales. Les propriétés du système de Haar seront démontrées dans le second exposé, comme cas particulier des propriétés générales des martingales.

1. <u>Espérance conditionnelle</u>.

Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $\mathcal{B}$ une sous-tribu de $\mathcal{A}$. On désignera par $L^o(\mathcal{B})$ l'ensemble des classes de fonctions mesurables $f \in L^o(\Omega, \mathcal{A}, P)$ telles qu'un représentant au moins de f soit $\mathcal{B}$-mesurable.
(C'est-à-dire qu'il existe une fonction g appartenant à la classe de f, et telle que $g^{-1}(B) \in \mathcal{B}$ pour tout borélien B de $\mathbb{R}$). Si la tribu $\mathcal{B}$ est complète, c'est-à-dire contient tous les ensembles négligeables, ou bien tous les représentants de f sont $\mathcal{B}$-mesurables, ou bien aucun représentant de f n'est $\mathcal{B}$-mesurable.

On posera $L^p(\mathcal{B}) = L^o(\mathcal{B}) \cap L^p(\Omega, \mathcal{A}, P)$, $0 \leq p \leq \infty$.

On notera $\|f\|_p$ la norme de f dans $L^p(\Omega, \mathcal{A}, P)$. (quasi-norme pour $p < 1$).

Si $f \in L^1(\Omega, \mathcal{A}, P)$, on définit l'espérance conditionnelle de f par rapport à $\mathcal{B}$, notée $E^{\mathcal{B}} f$, comme un élément de $L^1(\mathcal{B})$ caractérisé par le théorème suivant :

<u>Théorème 1</u> : 1) Pour tout $f \in L^1(\Omega, \mathcal{A}, P)$, il existe un élément unique $E^{\mathcal{B}} f \in L^1(\mathcal{B})$ tel que :

$$\forall g \in L^\infty(\mathcal{B}) \qquad \int g E^{\mathcal{B}} f \, dP = \int g f \, dP .$$

En fait, $E^{\mathcal{B}} f$ est déjà caractérisé par :

$$E^{\mathcal{B}} f \in L^1(\mathcal{B}), \text{ et } \forall B \in \mathcal{B} \qquad \int_B E^{\mathcal{B}} f \, dP = \int_B f \, dP .$$

2) L'opérateur linéaire $f \to E^{\mathcal{B}} f$ est positif, idempotent (i.e. $f \in L^1(\mathcal{B}) \Leftrightarrow f = E^{\mathcal{B}} f$). De plus, la restriction de $E^{\mathcal{B}}$ à $L^p(\Omega,\mathcal{A},P)$, $1 \leq p \leq \infty$, est continue de norme $\leq 1$ à valeurs dans $L^p(\mathcal{B})$ :

$$\forall\, f \in L^p(\Omega,\mathcal{A},P) \qquad \|E^{\mathcal{B}} f\|_p \leq \|f\|_p \ .$$

3) Si h est réelle et $\mathcal{B}$-mesurable, et si f et hf sont intégrables :

$$E^{\mathcal{B}}(hf) = h\, E^{\mathcal{B}} f \ .$$

Soit maintenant F un espace de Banach. Désignons par $L^o(\Omega,\mathcal{A},P,F)$ l'espace des (classes de) variables aléatoires mesurables à valeurs dans F. Si $\mathcal{B}$ est une sous-tribu de $\mathcal{A}$, on notera $L^o(\mathcal{B},F)$ l'ensemble des classes $f \in L^o(\Omega,\mathcal{A},P,F)$ admettant au moins un représentant $\mathcal{B}$-mesurable, et on posera encore :

$$L^p(\mathcal{B},F) = L^o(\mathcal{B},F) \cap L^p(\Omega,\mathcal{A},P,F), \qquad 0 \leq p \leq \infty \ .$$

On notera $\|f\|_p$ la quasi-norme dans $L^p(\Omega,\mathcal{A},P,F)$, $\left( \|f\|_p = \left(\int \|f(\omega)\|_F^p \, dP(\omega)\right)^{1/p} \right)$.

On sait que $L^1(\Omega,\mathcal{A},P,F) = L^1(\Omega,\mathcal{A},P) \hat{\otimes}_\pi F$. Il est donc possible de définir un opérateur $E_F^{\mathcal{B}} = E^{\mathcal{B}} \hat{\otimes}\, \mathrm{Id}$ de $L^1(\Omega,\mathcal{A},P,F)$ dans lui-même. Cet opérateur est l'opérateur d'espérance conditionnelle par rapport à $\mathcal{B}$ pour les variables aléatoires intégrables à valeurs dans F. Nous le noterons simplement $E^{\mathcal{B}}$, et nous aurons :

<u>Théorème 1 bis</u> : 1) Pour tout $f \in L^1(\Omega,\mathcal{A},P,F)$ il existe un élément unique $E^{\mathcal{B}} f \in L^1(\mathcal{B},F)$ tel que :

$$\forall\, g \in L^\infty(\mathcal{B}) \qquad \int g\, E^{\mathcal{B}} f \, dP = \int g f \, dP \ .$$

En fait, $E^{\mathcal{B}} f$ est déjà caractérisé par :

$$E^{\mathcal{B}} f \in L^1(\mathcal{B},F) \quad \text{et} \quad \forall\, B \in \mathcal{B} \quad , \quad \int_B E^{\mathcal{B}} f\, dP = \int_B f\, dP \quad .$$

2) L'opérateur $f \to E^{\mathcal{B}} f$ est idempotent (i.e. $f \in L^1(\mathcal{B},F) \Leftrightarrow f = E^{\mathcal{B}} f$). De plus, la restriction de $E^{\mathcal{B}}$ à $L^p(\Omega,\mathcal{A},P,F)$, $1 \leq p \leq \infty$, est continue de norme $\leq 1$ à valeurs dans $L^p(\mathcal{B},F)$ :

$$\forall\, f \in L^p(\Omega,\mathcal{A},P,F) \qquad \|E^{\mathcal{B}} f\|_p \leq \|f\|_p \, .$$

3) Si h est réelle (resp. à valeurs dans F) et $\mathcal{B}$-mesurable, et f intégrable à valeurs dans F (resp. intégrable réelle), et si hf est intégrable :

$$E^{\mathcal{B}}(hf) = h\, E^{\mathcal{B}} f \, .$$

4) Si $f \in L^1(\Omega,\mathcal{A},P,F)$ :

$$\|E^{\mathcal{B}} f\| \leq E\, \|f\| \, , \quad \text{et} \quad \forall\, \xi \in F' \qquad \langle \xi, E^{\mathcal{B}} f \rangle = E^{\mathcal{B}} \langle \xi, f \rangle \, .$$

<u>Exemple 1</u> : L'exemple le plus simple de calcul d'espérance conditionnelle est le cas où $\mathcal{B}$ est engendrée par une partition finie $B_1, \dots, B_n$ de $\Omega$, avec $B_i \in \mathcal{A}$. Dans ce cas, si $f \in L^1(\Omega,\mathcal{A},P,F)$ :

$$E^{\mathcal{B}} f = \sum_{i=1}^{n} \left( \frac{1}{P(B_i)} \int_{B_i} f\, dP \right) 1_{B_i}$$

(En désignant par $1_A$ la fonction indicatrice d'une partie $A \subset \Omega$)

Dans la suite nous noterons indifféremment Ef ou $\int f\,dP$ l'intégrale d'une fonction f réelle ou vectorielle (Ef se lit : espérance de f).

## 2. Martingales.

Soient $(\Omega, \mathcal{A}, P)$ un espace de probabilité et $\mathcal{B}_o \subset \mathcal{B}_1 \subset \ldots \subset \mathcal{B}_n \subset \ldots$ une suite croissante de sous-tribus de $\mathcal{A}$.
On notera $\mathcal{B}_\infty$ la tribu $\overset{\infty}{\underset{n=0}{\mathrm{V}}} \mathcal{B}_n$ engendrée par les tribus $(\mathcal{B}_n)$.

Soit F un espace de Banach. Nous dirons qu'une suite $(X_n)_{n \in \mathbb{N}}$ d'élements de $L^1(\Omega, \mathcal{A}, P, F)$ est une <u>martingale intégrable</u> si :

1) $X_n$ est $\mathcal{B}_n$-mesurable pour tout n. (Autrement dit : il existe un représentant $Y_n$ de la classe $X_n$ tel que $Y_n^{-1}(B) \in \mathcal{B}_n$ pour tout borélien B de F.)

2) Pour tout n : $\quad E^{\mathcal{B}_n} X_{n+1} = X_n$ .

On voit que la propriété 2) implique $E^{\mathcal{B}_n} X_m = X_n$ pour $m \geq n$, et on voit que cette propriété 2) est réalisée si et seulement si :

$$\forall\, B \in \mathcal{B}_n \qquad \int_B X_{n+1}\, dP = \int_B X_n\, dP \quad .$$

En particulier si $\mathcal{B}_n$ est engendrée par une partition finie $(B_1^n, \ldots, B_k^n)$, il suffit que :

$$\forall\, j = 1, \ldots, k, \qquad \int_{B_j^n} X_{n+1}\, dP = \int_{B_j^n} X_n\, dP \ .$$

Soit maintenant $(Y_n)$ une suite d'éléments de $L^1(\Omega, \mathcal{A}, P)$. Nous dirons que $(Y_n)$ est une <u>sous-martingale intégrable si</u> :

1) $Y_n$ est $\mathcal{B}_n$-mesurable pour tout n.

2) Pour tout n : $E^{\mathcal{B}_n} Y_{n+1} \geq Y_n$ .

(Ici encore, la propriété 2) implique $E^{\mathcal{B}_n} Y_m \geq Y_n$ pour tout $m \geq n$ et cette propriété 2) est réalisée dès que :

$$\forall B \in \mathcal{B}_n \qquad \int_B X_{n+1}\, dP \geq \int_B X_n\, dP \text{ . )}$$

<u>Proposition 1</u> : Si $(X_n)$ est une martingale intégrable à valeurs dans F, $Y_n(\omega) = \|X_n(\omega)\|$ définit une sous-martingale intégrable $(Y_n)$.

La démonstration est immédiate : en effet, $Y_n$ est $\mathcal{B}_n$-mesurable et :

$$E^{\mathcal{B}_n} \|X_{n+1}\| \geq \|E^{\mathcal{B}_n} X_{n+1}\| = \|X_n\| \text{ .}$$

<u>Exemple 2</u> : Un exemple important de martingale vectorielle est fourni par les "arbres" (cf. [1] ou [2], exposé 13).

Rappelons qu'un arbre dans un espace de Banach consiste en la donnée d'un point $x_o \in F$, et pour tout entier $k \geq 1$ de points $x_{\varepsilon_1,\ldots,\varepsilon_k}$, $\varepsilon_j = \pm 1$, tels que :

$$x_o = \frac{x_1 + x_{-1}}{2} \quad ; \quad \text{et pour tout } k \geq 1 \quad :$$

$$x_{\varepsilon_1,\ldots,\varepsilon_k} = 1/2\ (x_{\varepsilon_1,\ldots,\varepsilon_k,1} + x_{\varepsilon_1,\ldots,\varepsilon_k,-1}) \text{ .}$$

(On peut toujours supposer que k varie de 1 à $+\infty$. En effet, si les $x_{\varepsilon_1,\ldots,\varepsilon_k}$ sont définis seulement pour $k \leq n$, on posera pour $m > n$ :

$$x_{\varepsilon_1,\ldots,\varepsilon_n,\varepsilon_{n+1},\ldots,\varepsilon_m} = x_{\varepsilon_1,\ldots,\varepsilon_n} \text{ . )}$$

Il est facile d'interpréter les points de l'arbre comme l'ensemble des valeurs d'une martingale vectorielle. A cet effet posons $\Omega = \{-1,+1\}^{\mathbb{N}}$, $P = \bigotimes_{i=1}^{\infty} [1/2\ (\delta_1 + \delta_{-1})]_i$ (c'est la puissance tensorielle de la probabilité sur $\{-1,+1\}$ qui donne la masse 1/2 à chaque point. On peut voir aussi que c'est la probabilité de Haar du groupe compact $\Omega$.)

Posons $\mathcal{B}_o = \{\emptyset, \Omega\}$, et désignons par $\mathcal{B}_n$ la σ-algèbre engendrée par les n premières applications coordonnées. On voit que $\mathcal{B}_n$ est engendrée par la partition $B_{\varepsilon_1^o,\dots,\varepsilon_n^o}$, $\varepsilon_i^o = \pm 1$, où :

$$B_{\varepsilon_1^o,\dots,\varepsilon_n^o} = \{ (\varepsilon_i)_{1 \le i \le \infty} \in \Omega \quad ; \quad \varepsilon_i = \varepsilon_i^o \quad i = 1,2,\dots,n\}.$$

Supposons donné un arbre $(x_o, x_{\varepsilon_1,\dots,\varepsilon_k}, k \ge 1)$. Définissons une suite de variables aléatoires $(X_n)$ sur $\Omega$ par :

$$X_o = x_o, \text{ et pour } n \ge 1 \quad :$$

$$X_n(\varepsilon) = x_{\varepsilon_1^o,\dots,\varepsilon_n^o} \quad \text{si } \varepsilon \in B_{\varepsilon_1^o,\dots,\varepsilon_n^o} .$$

On voit facilement que $X_n$ est $\mathcal{B}_n$-mesurable ; d'autre part $B_{\varepsilon_1^o,\dots,\varepsilon_n^o} = B_{\varepsilon_1^o,\dots,\varepsilon_n^o,1} \cup B_{\varepsilon_1^o,\dots,\varepsilon_n^o,-1}$,

et $P(B_{\varepsilon_1^o,\dots,\varepsilon_n^o,1}) = P(B_{\varepsilon_1^o,\dots,\varepsilon_n^o,-1})$, donc :

$$\int_{B_{\varepsilon_1^o,\dots,\varepsilon_n^o}} X_{n+1}\, dP = \frac{1}{2} P(B_{\varepsilon_1^o,\dots,\varepsilon_n^o}) \, [x_{\varepsilon_1^o,\dots,\varepsilon_n^o,1} + x_{\varepsilon_1^o,\dots,\varepsilon_n^o,-1}] = \int_{B_{\varepsilon_1^o,\dots,\varepsilon_n^o}} X_n\, dP$$

ce qui prouve que $(X_n)$ est une martingale d'après les remarques qui suivent la définition.

Rappelons le théorème de convergence des martingales réelles, ou plus généralement des sous-martingales :

<u>Théorème 2</u> : Soit $(Y_n)$ une sous-martingale intégrable (relative à la suite croissante $(\mathcal{B}_n)$). Posons $Y_n^+ = \sup(Y_n, 0)$.

a) Si $\sup_n E\, Y_n^+ < \infty$, la suite $(Y_n)$ converge presque sûrement vers une variable aléatoire intégrable $Y$ (qui est donc $\mathcal{B}_\infty$-mesurable).

b) La convergence a lieu dans $L^1$ si et seulement si la suite $(Y_n)$ est équi-intégrable.

Dans le cas d'une martingale $(X_n)$ à valeurs dans un espace de Banach, la condition $\sup_n E\,\|X_n\| < \infty$ ne suffit pas en général à assurer la convergence presque sûre de $(X_n)$. Cette question sera étudiée dans l'exposé 4.

Nous allons définir les transformations de martingales, qui seront l'outil essentiel de l'exposé N° 2.

Soit $(X_n)_{n \geq 0}$ une suite de variables aléatoires intégrables à valeurs dans un espace de Banach. Posons $d_o = X_o$, et $d_n = X_n - X_{n-1}$ pour $n \geq 1$. On voit immédiatement que $(X_n)$ est une martingale si et seulement si :

a) $d_n$ est $\mathcal{B}_n$-mesurable pour tout $n \geq 0$.

b) $E^{\mathcal{B}_n}\, d_{n+1} = 0$ pour tout $n \geq 0$.

Supposons donc que $(X_n)_{n \geq 0}$ soit une martingale intégrable vectorielle (resp. réelle) et soit $(U_n)_{n \geq 0}$ une suite de variables aléatoires réelles (resp. vectorielles) telle que :

1) La suite $(U_n)$ est prévisible, c'est-à-dire que $U_n$ est $\mathcal{B}_{n-1}$-mesurable pour tout $n \geq 1$.

2) Pour tout $n \geq 0$, $d_n U_n$ est intégrable.

Définissons une suite de variables aléatoires $(U * X)_n$ par :

$$(U * X)_n = \sum_{k=0}^{n} U_k \, d_k \, .$$

On voit que $(U * X)_n$ est une martingale. En effet, si $\Delta_o = U_o d_o$, et $\Delta_n = (U * X)_n - (U * X)_{n-1}$ pour $n \geq 1$, on voit que $\Delta_n$ est $\mathcal{B}_n$-mesurable pour tout n, et d'autre part :

$$E^{\mathcal{B}_n} \Delta_{n+1} = E^{\mathcal{B}_n} U_{n+1} \, d_{n+1} = U_{n+1} \, E^{\mathcal{B}_n} d_{n+1} = 0$$

($U_{n+1}$ est $\mathcal{B}_n$-mesurable, et on peut appliquer le théorème 1 bis 3).)

La martingale $(U * X)_n$ est appelée martingale transformée de $(X_n)$ par la suite $(U_n)$. Un cas particulier très important est celui des martingales transformées par un temps d'arrêt.

Rappelons qu'un temps d'arrêt (relatif à la suite des tribus $(\mathcal{B}_n)$) est une application mesurable T de $\Omega$ dans l'ensemble des entiers $\geq 0$, et telle que :

$$(*) \qquad \forall \, n \geq 0 \qquad \{T = n\} \in \mathcal{B}_n \, .$$

(Il est équivalent de dire que $\{T \leq n\} \in \mathcal{B}_n$ pour tout $n \geq 0$).

Si T est un temps d'arrêt, définissons une suite de variables aléatoires $(U_n)$ par :

$$U_n(\omega) = 1 \qquad \text{si } n \leq T(\omega), \quad 0 \text{ sinon.}$$

On voit que $U_n$ est $\mathcal{B}_{n-1}$-mesurable pour $n \geq 1$, puisque $\{U_n = 0\} = \{T < n\} = \{T \leq n-1\} \in \mathcal{B}_{n-1}$. Si $(X_n)$ est une martingale vectorielle, la martingale transformée $(U * X)_n$ vaut :

$$(U * X)_n(\omega) = \sum_{k=0}^{T(\omega)\wedge n} d_k(\omega) = X_{T(\omega)\wedge n}(\omega)$$

(nous notons $a \wedge b$ l'inf de a et de b).

La martingale ci-dessus est simplement notée $(X_{T\wedge n})_n$ ; c'est la martingale "arrêtée au temps T".

Nous allons rappeler maintenant les inégalités de Doob, concernant la fonction maximale d'une martingale, ou plus généralement d'une sous-martingale :

<u>Lemme</u> : Soit $(Y_n)$ une sous-martingale intégrable $\geq 0$.
Posons $Y_n^*(\omega) = \sup_{0 \leq k \leq n} Y_k(\omega)$. Pour tout $a > 0$, et pour tout n, on a :

$$a\, P(Y_n^* > a) \leq \int_{\{Y_n^* > a\}} Y_n \, dP \quad .$$

<u>Démonstration</u> : Posons

$$T(\omega) = \inf \{k \,;\, Y_k(\omega) > a\} \quad (T(\omega) = +\infty \text{ si } \sup_k Y_k(\omega) \leq a).$$

On voit facilement que T est un temps d'arrêt. Par ailleurs on voit que :

$$\{Y_n^* > a\} = \{T \leq n\}.$$

En utilisant le fait que $Y_k > a$ sur $\{T = k\}$, puis en utilisant l'inégalité des sous-martingales, on obtient :

$$a\, P(T \leq n) = a \sum_{k=0}^{n} P\,(T = k) \quad \leq \quad \sum_{k=0}^{n} \int_{\{T = k\}} Y_k \, dP$$

$$\leq \sum_{k=0}^{n} \int_{\{T = k\}} Y_n \, dP = \int_{\{T \leq n\}} Y_n \, dP$$

<u>Théorème 3</u> : Soit $(X_n)$ une martingale vectorielle, cu une sous-martingale positive, bornée dans $L^p$, $1 < p < \infty$. Posons $X^*(\omega) = \sup_n \|X_n(\omega)\|$.

On a :

$$\|X^*\|_p \leq \frac{p}{p-1} \sup_n \|X_n\|_p .$$

<u>Démonstration</u> : Posons $Y_n(\omega) = \|X_n(\omega)\|$. La suite $(Y_n)$ est une sous-martingale positive. Il suffit donc de faire la démonstration dans le cas d'une sous-martingale positive $(Y_n)$. Il suffit de montrer que :

$$\forall\, n, \quad \|Y_n^*\|_p \leq \frac{p}{p-1} \|Y_n\|_p \quad .$$

D'après le lemme, on a pour tout $a > 0$ :

$$a\, P(Y_n^* > a) \quad \leq \quad \int Y_n \, 1_{\{Y_n^* > a\}} \, dP .$$

En multipliant l'inégalité par $pa^{p-2}$ et en intégrant par rapport à a, on obtient :

$$\int (Y_n^*)^p dP \leq p \int Y_n \left[ \int 1_{\{Y_n^* > a\}} \, a^{p-2} \, da \right] dP$$

$$= \frac{p}{p-1} \int Y_n \, (Y_n^*)^{p-1} \, dP$$

$$\leq \frac{p}{p-1} \left(\int Y_n^p\right)^{1/p} \left(\int (Y_n^*)^{p'(p-1)} \, dP\right)^{1/p'} ,$$

avec $1/p + 1/p' = 1$, d'où $p'(p-1) = p$, d'où le résultat en divisant les deux membres par $(\int Y_n^{*})^p \, dP)^{1/p'}$.

<u>Corollaire</u> : Soit $p$ tel que $1 < p < \infty$. Si $(X_n)$ est une martingale réelle telle que $\sup_n \|X_n\|_p < \infty$, la suite $(X_n)$ converge presque sûrement et dans $L^p$ vers une variable $X$.

<u>Démonstration</u> : D'après le théorème 3, $X^*(\omega) = \sup_n \|X_n(\omega)\|$ est dans $L^p$, et $X_n \to X$ p.s, d'après le théorème 2, avec $|X_n| \leq X^*$ : il suffit donc d'appliquer le théorème de Lebesgue dans $L^p$.

---

## BIBLIOGRAPHIE

[1] R.C. James : Some self dual properties of normed linear spaces. Ann. Math. Studies n°69 p.159-175.

[2] Séminaire Maurey-Schwartz 1973-1974.

[3] J. Neveu : Martingales à temps discret (Masson).

---