Analyse harmonique dyadique : une introduction

Aline Bonami
aline.bonami@univ-orleans.fr

Ces notes sont une incitation, pour les probabilistes, à s'intéresser à l'analyse réelle. Pour les analystes, leur but est de montrer à quel point les probabilités jouent un rôle central dans l'analyse réelle, qu'il s'agisse de construire des contre-exemples avec des méthodes aléatoires ou d'y trouver les raisons pour lesquelles on s'est intéressé à tel ou tel problème. L'analyse harmonique dyadique joue un rôle évident de passerelle. Le développement prodigieux des inégalités à poids autour de la conjecture A_2 donne la possibilité d'illustrer ces convergences entre les deux domaines avec des résultats très récents.

Il y a de nombreux textes qui permettent aux spécialistes de trouver une présentation exhaustive (par exemple le texte de Cristina Pereyra dans Contemporary Maths. Ce n'est pas le but de celui-ci, qu'il faut plutôt voir comme une introduction à la théorie de Calderón-Zygmund à poids entraînant le lecteur dans une sorte de promenade qui lui permet de s'initier à quelques méthodes d'analyse réelle dans des situations simples.

Il s'agit d'une version provisoire, mise en ligne après l'école de recherche d'Abidjan. Elles seront complétés ultérieurement, en particulier par une bibliographie

1. Fonction maximale et convergence presque partout

Notre point de départ est le théorème de Hardy-Littlewood, relatif à la fonction maximale de Hardy-Littlewood, définie dès que f est localement intégrable sur $\mathbb{R}^d$ par

$$Mf(x) = \sup_{x\in Q} \frac{1}{|Q|} \int_Q |f(y)|\, dy$$

où le supremum est pris sur tous les cubes (à côtés parallèles aux axes) contenant x.

La fonction maximale intervient naturellement dans beaucoup de problèmes. Le premier exemple d'application est le théorème de différentiation de Lebesgue, qui permet d'affirmer que si f est localement intégrable alors, presque partout,

$$f(x) = \lim_{r\to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy. \tag{1}$$

Ici $B(x,r)$ désigne la boule de centre x et de rayon r.

La propriété cruciale de l'opérateur M est l'inégalité faible L^1 du théorème maximal.

Théorème 1.1. *(théorème maximal) [Hardy–Littlewood–Wiener]*
M est de type faible $(1,1)$ c'est-à-dire

$$|\{Mf > \lambda\}| \leq \frac{c\|f\|_{L^1}}{\lambda}.$$

De plus M est de type fort (p,p), $1 < p < \infty$. Plus précisément,

$$\|Mf\|_{L^p} \leq \frac{cp}{p-1}\|f\|_{L^p}. \tag{2}$$

Ici $|A|$ désigne la mesure de l'ensemble mesurable A pour la mesure de Lebesgue.

Les constantes c dépendent de la dimension. La détermination des meilleures constantes est encore en partie inconnue. Voir le blog de Tao http ://terrytao.wordpress.com/2011/05/21/steins-spherical-maximal-theorem/

Exercice 1. *Montrer que, quelle que soit la fonction ϕ positive sur l'espace mesuré (X,μ), on a la formule*

$$\int_X \phi^p \, d\mu = p\int_0^\infty t^p \mu(\{x \in X : \phi(x) > t\}) \, \frac{dt}{t}.$$

Exercice 2. *Montrer que l'inégalité faible $(1,1)$ entraîne l'inégalité forte (p,p). Indication : on pourra écrire $f = g + h$ avec $g = f\mathbb{1}_{|f|>\lambda/2}$ et en déduire que*

$$\{Mf > \lambda\} \subset \{Mg > \lambda/2\}.$$

Exercice 3. *Montrer que si f est la fonction indicatrice de la boule unité centrée en 0, alors $Mf(x) \geq c|x|^{-d}$ à l'infini. En déduire qu'il n'y a pas d'inégalité forte $(1,1)$. En déduire que la plus petite constante dans l'inégalité*

$$\|Mf\|_{L^p} \leq A\|f\|_{L^p} \tag{3}$$

est supérieure à $\frac{c}{p-1}$ (les constantes c peuvent changer d'une ligne à l'autre).

L'exercice 2 donne un cas particulier du théorème de Marcinkiewicz, que nous énonçons maintenant.

Théorème 1.2. *Si T est un opérateur (linéaire, sous-linéaire) qui est de type faible (p_0, q_0) et de type faible (p_1, q_1), avec $1 \leq p_0 \leq q_0 \leq \infty$ et*

$1 \leq p_1 \leq q_1 \leq \infty$, alors T est de type fort (p, q) pour tout $p \in (p_0, p_1)$ si p et q sont tels que

$$\frac{1}{p} = \frac{t}{p_0} + \frac{1-t}{p_1} \qquad\qquad \frac{1}{q} = \frac{t}{q_0} + \frac{1-t}{q_1}.$$

Rappelons que les hypothèses et conclusions peuvent encore s'écrire
Hypothèses : $\mu(\{x \in X : |Tf(x)| > \lambda\}) \leq \frac{c_0\|f\|_{p_0}^{q_0}}{\lambda^{q_0}}$;

$$\mu(\{x \in X : |Tf(x)| > \lambda\}) \leq \frac{c_1\|f\|_{p_1}^{q_1}}{\lambda^{q_1}}.$$

Conclusion : $\|f\|_q \leq c\|f\|_p$.

Nous allons montrer le théorème de différentiation de Lebesgue à partir du théorème maximal, que nous démontrerons plus tard. Il s'agit d'une démonstration modèle, qui sert de modèle à de nombreuses démonstrations.

Démonstration du théorème de différentiation de Lebesgue

Soit f une fonction localement intégrable. Nous allons montrer qu'il existe un ensemble L_f dit **ensemble de Lebesgue** dont le complémentaire est de mesure 0 et tel que

$$f(x) = \lim_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} f(y)\, dy \qquad x \in L_f \tag{4}$$

Nous allons même montrer un résultat plus fort, l'existence de L_f tel que

$$\lim_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy = 0 \qquad x \notin L_f. \tag{5}$$

Comme la conclusion est locale (elle ne dépend que des valeurs de f autour du point considéré), nous pouvons supposer f intégrable. Nous définissons

$$A_\lambda = \{x : \limsup_{r \to 0} \frac{1}{|B(x,r)|} \int_{B(x,r)} |f(y) - f(x)|\, dy > \lambda\} \qquad \lambda \geq 0.$$

On veut montrer que $|A_0| = 0$. Comme $A_0 = \cup_j A_{1/j}$ il suffit de montrer que $|A_\lambda| = 0$ pour tout $\lambda > 0$. Il est important de remarquer que (5) est satisfait quand f est continue. Or le sous-espace des fonctions continues est dense dans L^1. Si $f \in L^1(\mathbb{R}^d)$ est une fonction quelconque, pour $\varepsilon > 0$ donné on peut l'écrire $f = g + h$ avec g continue et $\|h\|_{L^1(\mathbb{R}^d)} < \varepsilon$. Comme la limite supérieure d'une somme est inférieure ou égale à la somme des limites supérieures, et comme la limite pour g vaut 0,

$$\limsup_{r\to 0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |f(y)-f(x)|\,dy$$
$$\leq \limsup_{r\to 0} \tfrac{1}{|B(x,r)|}\int_{B(x,r)} |h(y)-h(x)|\,dy.$$

Nous avons immédiatement, pour $x \in A_\lambda$,

$$\lambda \;<\; \sup_{r>0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |h(y)-h(x)|\,dy \tag{6}$$

$$\leq\; |h(x)| + \sup_{r>0} \frac{1}{|B(x,r)|}\int_{B(x,r)} |h(y)|\,dy. \tag{7}$$

Le second terme est défini comme la fonction maximale, mais avec des boules centrées en x à la place des cubes contenant x. Or la boule $B(x,r)$ est incluse dans un cube qui centré en x et de volume $(2r)^d = \kappa_d|B(x,r)|$. Ce second terme est donc majoré par $\kappa_d Mh(x)$. On en déduit que

$$|A_\lambda| \leq |\{x\in\mathbb{R}^d : \kappa_d Mh(x) > \lambda/2\}| + |\{x\in\mathbb{R}^d : |h(x)| > \lambda/2\}|$$

On utilise ensuite l'inégalité maximale pour le nouveau premier terme, l'inégalité de Tchebychev pour le second :

$$|A_\lambda| \leq \frac{2(\kappa_d c+1)}{\lambda}\int_{\mathbb{R}^d} |h(x)|\,dx \leq \frac{2(\kappa_d c+1)}{\lambda}\varepsilon$$

En faisant tendre ε vers 0 on obtient que $|A_\lambda| = 0$.

2. Cubes dyadiques, théorème maximal dyadique

Un intervalle dyadique de $\mathbb{R}$ est un intervalle $[2^{-k}n, 2^{-k}(n+1))$, avec $n,k\in\mathbb{Z}$. Le paramètre k (ou, suivant les cas, 2^{-k} est appelé paramètre d'échelle. On note $\mathcal{D}$ l'ensemble des intervalles dyadiques et $\mathcal{D}_k$ le sous ensemble constitué des intervalles $I\in\mathcal{D}_k$ tels que $|I| = 2^{-k}$.

Les propriétés suivantes sont très faciles à montrer :
- A l'échelle k les intervalles dyadiques forment une partition de $\mathbb{R}$.
- $\mathcal{D}$ est invariant par la dilatation $x\mapsto 2x$. Autrement dit, si I est un intervalle dyadique, alors $2I$ est aussi un intervalle dyadique, et réciproquement.
- Chaque intervalle dyadique I est réunion disjointe de deux intervalles de l'échelle moitié.
- Si I, I' sont deux intervalles dyadiques, alors l'un est inclus dans l'autre ou bien les deux intervalles sont disjoints.

Ces propriétés impliquent le lemme suivant.

Lemme 2.1. *Soit $\mathcal{L}$ une collection d'intervalles dyadiques I de taille bornée. Si $\mathcal{L}_{max}$ désigne la sous-collection des intervalles dyadiques maximaux pour l'inclusion, alors ceux-ci sont deux à deux disjoints et tout intervalle de $\mathcal{L}$ est inclus dans un intervalle de $\mathcal{L}_{max}$.*

On définit dans $\mathbb{R}^d$ les cubes dyadiques comme des produits d'intervalles dyadiques de la même longueur. Il est facile de voir que les propriétés précédentes s'étendent à plusieurs dimensions.

On définit la fonction maximale dyadique par

$$M^{\mathcal{D}} f(x) = \sup_{x \in Q} \frac{1}{|Q|} \int_Q |f(y)|\, dy$$

où le supremum est pris sur tous les cubes dyadiques Q contenant x.

Cette fonction maximale est a priori plus petite, mais on peut contrôler la fonction maximale en termes de celle-ci grâce à la propriété suivante (qui n'a été observée sous cette forme qu'assez récemment !). Notons tout d'abord $\mathcal{D}_\alpha$, pour $\alpha \in \{0, 1/3\}^d$, la collection des cubes $2^{-k}([0,1)^d + j + (-1)^k \alpha)$, avec $k \in \mathbb{Z}$ et $j \in \mathbb{Z}^d$. La famille $\mathcal{D}_\alpha$ possède toutes les propriétés que nous avons énoncées pour la famille des cubes dyadiques.

Lemme 2.2. *Tout cube Q d'arête ℓ est inclus dans un cube Q' d'échelle k appartenant à une des familles $\mathcal{D}_\alpha$ et tel que $2^k \leq 6\ell$.*

Démonstration. Nous allons faire uniquement la preuve en dimension 1. Soit k tel que $2^{-k-1} \leq 3\ell < 2^{-k}$. Si l'intervalle I ne contient aucun point de la forme $2^{-k}j$, alors il est tout entier contenu dans un cube dyadique classique de taille $2^{-k} < 6\ell$. Supposons que I contienne un point de la forme $2^{-k}j$. Alors il ne contient aucun point de la forme $2^{-k}(j \pm 1/3)$. Il est donc contenu dans un intervalle de la seconde famille. □

Nous pouvons maintenant démontrer une inégalité qui est meilleure que l'inégalité maximale dyadique, et pour laquelle la constante vaut 1 :

$$\lambda |\{M^{\mathcal{D}} f > \lambda\}| \leq \int_{M^{\mathcal{D}} f > \lambda} |f(x)| dx. \tag{8}$$

Démonstration. Soit $A = \{M^{\mathcal{D}} f > \lambda\}$. On peut aussi écrire A comme l'union des cubes dyadiques Q tels que

$$\frac{1}{|Q|} \int_Q |f(x)| dx > \lambda.$$

Ces cubes sont de volume borné par $\frac{\int |f(x)| dx}{\lambda}$, donc bornés, et on peut utiliser le lemme 2.1. Autrement dit A est l'union disjointe de cubes dyadiques Q_j qui satisfont la même propriété et donc pour lesquels

$$|Q_j| < \frac{1}{\lambda}\int_{Q_j} |f(x)|dx.$$

On conclut en faisant la somme en j. $\square$

Exercice 4. *Déduire de* (8) *que*

$$\int |M^{\mathcal{D}} f|^p dx \leq \frac{p}{p-1}\int |M^{\mathcal{D}} f|^{p-1}|f|dx.$$

On utilisera l'exercice 1.

En déduire l'inégalité de type fort (p,p) *avec la constante* $\frac{p}{p-1}$.

Le premier corollaire du théorème maximal dyadique est le théorème maximal grâce au lemme 2.2, qui permet d'écrire la fonction maximale comme somme de deux fonctions maximales dyadiques.

D'autre part on voit que démonstration est géométrique et n'utilise pas les propriétés spécifiques de la mesure de Lebesgue. Elle s'étend à une mesure de Radon quelconque. Il suffit, pour cela, de commencer par s'intéresser à la fonction

$$M_N^{\mathcal{D}} f(x) = \sup_{x\in Q, |Q|\leq 2^{dN}} \frac{1}{|Q|}\int_Q |f(y)|\, dy$$

puis de faire tendre N vers $+\infty$.

3. Martingales dyadiques et autres, convergence des martingales

Donnons tout d'abord quelques notations. Si Q est un cube dyadique, on note $\mathcal{D}(Q)$ (resp. $\mathcal{D}_k(Q)$) la sous-famille des cubes dyadiques inclus dans Q (resp. des cubes dyadiques de l'échelle 2^{-k}). On parle aussi de la k-ième génération.

Pour reconnaître une martingale, on va se limiter à considérer les fonctions définies sur un cube dyadique. Par commodité on se limitera à la dimension 1 et on prendra l'intervalle $I_0 = [0, 1)$. On appelle $\mathcal{F}_n$ la tribu engendrée par les intervalles de $\mathcal{D}_n(I_0)$. Les tribus $\mathcal{F}_n$ sont croissantes et engendrent la tribu borélienne. On va reconnaître en les moyennes d'une fonction f sur les intervalles dyadiques l'espérance conditionnelle de la v. a. identifiée à f par rapport à la tribu $\mathcal{F}_n$. Plus précisément,sur un intervalle I de la n-ième génération,

$$\mathbb{E}(f|\mathcal{F}_n) = \frac{1}{|I|}\int_I f(x)dx.$$

On rappelle que $\mathbb{E}(g|\mathcal{F}_n)$ est $\mathcal{F}_n$-mesurable et possède la propriété caractéristique, parmi les fonctions $\mathcal{F}_n$-mesurables, que, quelle que soit la fonction h $\mathcal{F}_n$-mesurable

$$\mathbb{E}(\mathbb{E}(g|\mathcal{F}_n)h) = \mathbb{E}(gh).$$

On reconnaît en la fonction maximale dyadique (restreinte à l'intervalle I_0) la fonction maximale au sens des martingales :

$$\sup_{x\in I, I\in\mathcal{D}(I_0)} \frac{1}{|I|}\int_I f(y)\,dy = \sup_n |\mathbb{E}(f|\mathcal{F}_n)|(x) = f^*(x).$$

(Attention, les deux notions ne coïncident que pour les fonctions positives).

L'inégalité que nous avons démontrée correspond à l'inégalité de Doob pour f positive :

$$\lambda\mathbb{P}(f^* > \lambda) \le \int_{f^*>\lambda} f. \tag{9}$$

La démonstration classique en est la suivante : on considère le temps d'arrêt $\tau = \inf\{n; \mathbb{E}(f|\mathcal{F}_n) > \lambda\}$ de sorte que l'évènement $f^* > \lambda$ coïncide avec l'évènement $\tau < \infty$. Nous allons utiliser les notations classiques de la théorie des martingales : $\mathbb{E}(f|\mathcal{F}_n) = f_n$, et $f_\tau = f_n$ lorsque $\tau = n$. Il en résulte que $f_\tau \ge \lambda$ lorsque $\tau < \infty$. Pour conclure il suffit de montrer que $\int_{\tau<\infty} f_\tau = \int_{\tau<\infty} f$, ce qui découle de la propriété caractéristique de l'espérance conditionnelle que nous avons rappelée.

On peut voir la technique utilisée pour les cubes dyadiques de sélection des cubes maximaux comme une technique de temps d'arrêt.

Plus généralement, si on a une suite croissante de tribus $\mathcal{F}_n$ sur un espace de probabilité, on dit que la suite (f_n) définit une martingale intégrable si chaque f_n est $\mathcal{F}_n$ mesurable, si $\mathbb{E}(f_{n+1}|\mathcal{F}_n) = f_n$, et si $\sup_{n\ge 0}\mathbb{E}|f_n| < \infty$. Un cas particulier de martingale est donné par une fonction f intégrable telle que $f_n = \mathbb{E}(f|\mathcal{F}_n)$ où f est une fonction intégrable. On peut alors démontrer que la martingale converge de manière analogue à ce que nous avons fait pour le théorème de différentiation de Lebesgue. Mais une martingale n'est pas en général donné ainsi (il est nécessaire et suffisant que la martingale soit équi-intégrable). Le cas dyadique est donné par les moyennes d'une mesure qui est éventuellement non absolument continue par rapport à la mesure de Lebesgue, qui converge p.p. vers sa dérivée de Radon Nikodym par rapport à celle-ci.

Exercice 5. *Montrer que si f est dans L^p, avec $1 < p < \infty$, alors f_n tend vers f dans L^p. (On pourra utiliser le théorème de Lebesgue, le majorant étant donné par la fonction maximale).*

4. Variables aléatoires de Bernoulli, fonctions de Haar et inégalité de Khintchine

Les martingales dyadiques peuvent être vues d'une autre manière. Rappelons tout d'abord qu'une **v.a. de Bernoulli** (centrée) est une v. a. qui prend les valeurs -1 et 1 avec probabilité $1/2$. Il y a plusieurs manières de considérer une suite infinie de v. a. de Bernoulli indépendantes.

- Le groupe $\Omega = \{-1, +1\}^{\mathbb{N}}$ est un groupe abélien compact, muni de la mesure produit $d\omega$. Si on considère pour X_j la j-ième fonction coordonnée, alors les X_j sont indépendantes et sont des v. a. de Bernoulli. On ajoute $X_0 = 1$.
- Les **fonctions de Rademacher**, qui sont définies sur $[0, 1)$ (muni de la mesure de Lebesgue) par $r_n(t) = \operatorname{sgn}\sin(2^n\pi t)$, sont orthogonales et sont de Bernoulli. On peut voir qu'elles sont indépendantes. On ajoute $r_0 = 1$.
- Soit une suite (Y_n) de v. a. gaussiennes centrées réduites indépendantes. Alors la suite $X_n = \operatorname{sgn} Y_n$ forme une suite de Bernoulli indépendantes.

On peut encore définir $r_n(t)$ à partir du développement en base 2 de t : si $t = \sum_{n>0} \frac{\varepsilon_n}{2^n}$, les développements avec uniquement des 1 à partir d'un certain rang étant exclus, alors $r_n(t) = 1$ si $\varepsilon_n = 0$, alors que $r_n(t) = -1$ sinon.

Le développement dyadique d'un nombre permet de passer de $[0, 1)$ à $\{-1, +1\}^{\mathbb{N}}$ et de faire se correspondre les fonctions coordonnées d'une part, les fonctions de Rademacher d'autre part. Ainsi **tout énoncé relatif aux martingales dyadiques peut s'interpréter en termes de v. a. de Bernoulli.**

La tribu $\mathcal{F}_n$ peut aussi être vue comme la tribu engendrée par $X_0, \cdots, X_n$.

A partir des fonctions de Rademacher on peut construire une base de $L^2([0, 1))$ qui présente des analogies avec la base de Fourier : si $n = \sum_{j=1}^{\infty} \varepsilon_j 2^j$ avec $\varepsilon_j \in \{0, 1\}$ (il y a évidemment un nombre fini de 1), alors la **fonction de Walsh** w_n prend pour valeur le produit des r_j pour lesquels ε_j vaut 1.

Une autre base a des propriétés remarquables, la **base de Haar**. Pour tout intervalle dyadique I, on note I_- et I_+ les deux intervalles dyadiques moitié qui le composent. Alors

$$h_I(x) = \frac{1}{\sqrt{|I|}} \begin{cases} 1 & x \in I_- \\ -1 & x \in I_+ \\ 0 & x \notin I \end{cases}.$$

Théorème 4.1. *Les fonctions de Haar $(h_I)_{I\in\mathcal{D}}$ forment une base de $L^2(\mathbb{R})$. Les fonctions de Haar $(h_I)_{I\in\mathcal{D}(I_0)}$, complétées par $\mathbb{1}_{I_0}$, forment une base de $L^2(I_0)$.*

Démonstration. Le fait que ce soit des systèmes orthonormaux est élémentaire. La complétion est une conséquence du théorème de différentiation de Lebesgue dyadique (ou du théorème de convergence des martingales). Montrons-le pour $L^2(I_0)$. Soit I un sous-intervalle dyadique de I_0. On calcule immédiatement les produits scalaires

$$\langle \mathbb{1}_I, \mathbb{1}_{I_0}\rangle = |I|, \qquad \langle \mathbb{1}_I, h_{I'}\rangle = \begin{cases} & \frac{|I|}{\sqrt{|I'|}} & I \subset I'_- \\ & -\frac{|I|}{\sqrt{|I'|}} & I \subset I'_- \\ 0 \quad \text{sinon} & & \end{cases}.$$

On vérifie que pour ces fonctions particulières de $L^2(I_0)$ on a bien

$$f = \langle f, \mathbb{1}_{I_0}\rangle + \sum_{I'} \langle f, h_{I'}\rangle h_{I'}.$$

C'est également vrai pour toute combinaison linéaire de celles-ci, donc pour toutes les fonctions $\mathcal{F}_n$ mesurables. Mais toute fonction f est limite dans L^2 des f_n, ce qui permet de conclure. □

La base de Haar est la base d'ondelettes la plus simple. Elle est très utilisée en traitement du signal ou de l'image.

Du fait que h_I est orthogonal à toutes les fonctions $\mathcal{F}_n$ mesurables, on déduit que

$$f_{n+1} - f_n = \sum_{I\in\mathcal{D}_n(I_0)} \langle f, h_I\rangle h_I.$$

Ainsi **tout énoncé relatif aux martingales dyadiques peut s'interpréter en termes de la base de Haar.**

Nous allons terminer cette section par l'inégalité de Khintchine :

Théorème 4.2. *Soit X_j une suite de v. a. de Bernoulli indépendantes, $X_0 = 1$. Alors, si $\sum a_j^2 < \infty$, la somme $\sum a_j X_j$, qui converge dans L^2 est aussi convergente dans L^p pour tout $p < \infty$ et*

$$\mathbb{E}|\sum a_j X_j|^p \le C p^{p/2} (\sum a_j^2)^{p/2}.$$

Il y a beaucoup de démonstrations de cet énoncé fondamental. La démonstration qui donne la meilleure constante est due à Haagerup. Nous allons ici le déduire du cas gaussien.

Démonstration. Nous commençons par démontrer cette inégalité sur les sommes finies. Soit Y_j une suite de v. a. gaussiennes centrées réduites indépendantes, $Y_0 = 1$. On peut supposer que $X_n = \operatorname{sgn} Y_n$. On peut même vérifier aisément que $\mathbb{E}(Y_n|\mathcal{F}) = \sqrt{\frac{\pi}{4}} X_n$, où $\mathcal{F}$ est la tribu engendrée par les X_j. Du fait de l'inégalité de Jensen, valable pour les espérances conditionnelles, il suffit de montrer l'inégalité pour la somme $\sum_0^N a_j Y_j$. Mais celle-ci est une gaussienne centrée de variance $\sum_0^N a_j^2$. Or le moment d'ordre p d'une gaussienne centrée réduite est égal à

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} |t|^p e^{-t^2/2} dt.$$

On le calcule aisément en fonction de la fonction Γ et on trouve une majoration en $Cp^{p/2}$.

Pour passer aux sommes infinies on utilise le fait que si une suite converge dans L^2 vers f et est majorée en norme par C dans L^p, alors f est dans L^p, de norme majorée par C. $\square$

Exercice 6. *Montrer que la limite n'est pas bornée en général. Soit X la somme infinie. En utilisant la même démonstration, montrer que $\exp \lambda X^2$ est intégrable si $\lambda \sum a_j^2 < 1/\sqrt{\pi}$.*

Corollaire 4.3. *Soit X_j une suite de v. a. de Bernoulli indépendantes, $X_0 = 1$. Alors*

$$(\sum a_j^2)^{1/2} \leq c \sup_N \mathbb{E} \left| \sum_0^N a_j X_j \right|.$$

Pour la démonstration il suffit de considérer les sommes finies. On utilise ensuite la log-convexité des moments (conséquence de l'inégalité de Hölder) :

$$\mathbb{E}Z^2 \leq (\mathbb{E}Z)^\alpha (\mathbb{E}Z^4)^{1-\alpha}$$

si on choisit α de sorte que $2 = \alpha + 4(1-\alpha)$ donc $\alpha = 2/3$. On utilise ensuite l'inégalité de Khintchine.

Nous allons tout de suite donner une application de l'inégalité de Khintchine.

Théorème 4.4. *Soit T un opérateur borné sur $L^P(X, d\mu)$, de norme $\|T\|$. Alors l'opérateur s'étend en un opérateur borné sur $L^p(X, \mathbb{C}^N)$ avec une norme indépendante de N par la formule*

$$\widetilde{T}(f_1, \cdots, f_N) = (T(f_1), \cdots, T(f_N)).$$

Démonstration. Il s'agit de montrer que

$$\|(\sum |Tf_j|^2)^{1/2}\|_p \leq C\|T\|\|(\sum |f_j|^2)^{1/2}\|_p$$

avec C une constante qui ne dépend que de p. Mais

$$\sum |f_j(x)|^2)^{p/2} \simeq \int_0^1 |\sum r_j(t) f_j(x)|^p$$

du fait de l'inégalité de Khintchine (ou son corollaire). Donc

$$\int_X |\sum r_j(t) Tf_j(x)|^p d\mu \leq \|T\|^p \int_X \int_0^1 |\sum r_j(t) f_j(x)|^p d\mu(x).$$

On intègre en t, on change l'ordre des intégrations et on utilise les inégalités de Khintchine pour conclure. □

5. Transformée de Hilbert et opérateurs de Calderón-Zygmund

La transformée de Hilbert est le prototype d'un opérateur intégral singulier. Il est donné formellement sur $\mathbb{R}$ par

$$Hf(x) = \text{v. p.}\frac{1}{\pi}\int \frac{f(y)}{x-y}dy = \lim_{\varepsilon\to 0 \frac{1}{\pi}} \int_{|x-y|>\varepsilon} \frac{f(y)}{x-y}dy.$$

C'est la convolution avec la distribution v. p.$\frac{1}{\pi x}$, qui est bien définie lorsqu'on la teste sur des fonctions de classe $\mathcal{C}^1$ à support compact. Le lemme suivant est fondamental quant à son étude :

Lemme 5.1. *La distribution* v. p.$\frac{1}{\pi x}$ *a pour transformée de Fourier la fonction* $-i\,\mathrm{sgn}$. *Comme conséquence,*

$$\widehat{Hf}(\xi) = -i\,\mathrm{sgn}\,\xi \widehat{f}(\xi). \tag{10}$$

Une manière de démontrer le lemme est d'écrire que $x \times$v. p.$\frac{1}{\pi x} = \frac{1}{\pi}$, dont la transformée de Fourier est $\frac{\delta_0}{\pi}$. Comme la dérivation côté Fourier provient d'une multiplication par $-2\pi i x$, la transformée de Fourier de v. p.$\frac{1}{\pi x}$ a pour dérivée $-2i\delta_0$. On conclut en remarquant que v. p.$\frac{1}{\pi x}$, et donc aussi sa transformée de Fourier.

Exercice 7. *Montrer que la transformée de Hilbert est un opérateur unitaire sur L^2. Montrer que l'image d'une fonction de $L^1 \cap L^2$ n'est pas dans L^1 si l'intégrale de f n'est pas nulle.*

On aurait pu définir plus facilement la transformée de Hilbert à partir de (10) puisque cette formule permet de définir la transformée de Hilbert d'une fonction de L^2 (dont la transformée de Fourier est dans

L^2 de même norme du fait de l'identité de Plancherel. Mais avec la notion d'opérateur intégral singulier c'est la première définition que nous souhaitons généraliser. Dans beaucoup d'applications les opérateurs ne sont pas invariants par translation, ils ne se présentent pas comme des convolutions.

On appelle noyau de Calderón-Zygmund toute fonction K mesurable dans $\mathbb{R}^d \times \mathbb{R}^d$ en dehors de la diagonale pour laquelle il existe deux constantes positives C et δ telles que

$$\begin{aligned} |K(x,y)| &\leq \frac{C}{|x-y|^d}, \\ |K(x,y)-K(x',y)| &\leq \frac{C|x-x'|^\delta}{|x-y|^{d+\delta}} \\ |K(x,y)-K(x,y')| &\leq \frac{C|y-y'|^\delta}{|x-y|^{n+\delta}} \end{aligned}$$

lorsque $x \neq y$ et la condition $\max(|x-x'|, |y-y'| \leq \frac{1}{2}|x-y|, \leq \frac{1}{2}|x-y|$ est réalisée.

Exercice 8. *Montrer que le noyau de la transformée de Hilbert est de Calderón-Zygmund.*

Un opérateur est dit **opérateur de Calderón-Zygmund** de noyau K s'il est borné dans L^2 et s'il est associé au noyau K au sens suivant :

$$\int g(x)T(f)(x)\,dx = \iint g(x)K(x,y)f(y)\,dy\,dx$$

lorsque f et g sont des fonctions régulières dont les supports sont disjoints.

Le théorème principal, dont un des ingrédients est le théorème maximal, est le suivant :

Théorème 5.2. *Soit T un opérateur de Calderón-Zygmund. Alors T est de type faible $(1,1)$ et de type fort (p,p).*

La démonstration est tout à fait classique. Nous donnerons une démonstration analogue dans le cas des martingales dyadiques.

Nous allons montrer très rapidement comment la transformée de Hilbert intervient dans l'étude des processus gaussiens stationnaires et a mené Szegö à s'intéresser aux inégalités à poids.

Relations avec les processus stationnaires : on considère un processus Gaussien stationnaire discret, c'est-à-dire une suite de v. a. gaussiennes centrées Y_n telles que les covariances $\mathbb{E}(Y_jY_k)$ ne dépendent que de la différence $j-k$. La suite r, telle que

$$\mathbb{E}(Y_jY_k) = r(j-k)$$

est symétrique définie positive : quelle que soit la suite finie (ξ_j), la quantité $\sum_{j,k} r(j-k)\xi_j\xi_k$ est positive. D'après le théorème de Bochner (ou de Herglotz), il existe une mesure $d\mu$ positive 2π périodique telle que

$$r(n) = \int_0^{2\pi} e^{-int} d\mu(t).$$

On veut savoir s'il y a une prédiction possible. Comme

$$\mathbb{E}(Y_j Y_k) = \int_0^{2\pi} e^{-ijt} e^{ikt} d\mu(t)$$

au lieu de projeter sur le passé dans l'espace des gaussiennes on projette sur l'espace des fonctions 2π périodiques ne faisant intervenir que des fréquences négatives dans l'espace $L^2(d\mu)$. Mais cette projection est bornée dans $L^2(d\mu)$ si et seulement si l'opérateur H^{per} l'est, où celui-ci est défini par le fait que les coefficients de Fourier de f et $H^{per}(f)$ sont liés par

$$c_n(H^{per}(f)) = -i \operatorname{sgn}(n) c_n(f).$$

C'est l'analogue de la transformée de Hilbert dans le cadre périodique. On aurait pu aussi bien se placer dans le cadre de processus continus et trouver alors la transformée de Hilbert. Helson et Szegö ont donné en 1965 une étonnante caractérisation des mesures $d\mu$ à l'aide de méthodes d'analyse complexe : $d\mu = wdx$ et $w = exp(u + H^{per}v)$, où u et v sont bornés, $|u| < \pi/2$. A partir des années 70, on a trouvé d'autres méthodes et d'autres types de caractérisation.

6. Inégalités à poids précisées pour la fonction maximale

Soit w une fonction positive localement intégrable dans $\mathbb{R}^d$. On dit que w est un poids de la classe A_p, $1 < p < \infty$ si

$$[w]_{A_p} = \sup_Q \left(\frac{1}{|Q|} \int_Q w(y)\, dy \right) \left(\frac{1}{|Q|} \int_Q w(y)^{-\frac{1}{p-1}}\, dy \right)^{p-1} < \infty \tag{11}$$

On dit que w est un poids de la classe A_1 s'il existe une constante K telle que

$$\frac{1}{|Q|} \int_Q w(y)\, dy \leq K \inf_Q w, \tag{12}$$

et nous noterons dans ce cas $[w]_{A_1}$ la plus petite constante possible.

Exercice 9. *Montrer (inégalité de Jensen) que*

$$1 \leq p \leq q \quad \Longrightarrow \quad [w]_{A_q} \leq [w]_{A_p} \quad i.e. \quad A_p \subset A_q.$$

Le théorème suivant a été une avancée majeure. Il est dû sous sa première forme à Muckenhoupt (1972). Avec les constantes données ici il est dû à Buckley (1993). La démonstration que nous en donnons a été trouvée par Lerner (2008). Pour simplifier les notations, on note $w(Q)$ la mesure de Q relative à la mesure à poids wdx. L'espace L^p pour la mesure à poids wdx est noté $L^p(w)$.

Théorème 6.1. *Soit $1 < p < \infty$. Les conditions suivantes sont équivalentes :*

(a) $w \in A_p$.
(b) *Il existe une constante K telle que pour tout cube Q et toute fonction f positive*

$$\left(\frac{1}{|Q|}\int_Q f(y)\,dy\right)^p w(Q) \leq [w]_{A_p}\int_Q f(y)^p\,w(y)dy \tag{13}$$

(c) *M est de type faible (p,p) dans $L^p(w)$. De plus*

$$\lambda^p \int_{Mf>\lambda} w(x)dx \leq c_d[w]_{A_p}\|f\|_{L^p(w)}. \tag{14}$$

(d) *M est borné sur $L^p(w)$. De plus*

$$\|M\|_{L^p(w)} \leq \frac{c_n p}{p-1}\,[w]_{A_p}^{\frac{1}{p-1}}, \tag{15}$$

Démonstration. Il est clair que (b) est équivalent à (a) si la constante est supposée égale à 1 dans (13). En effet, si (13) est satisfait avec la constante A à la place de $[w]_{A_p}$, en prenant $f = \mathbb{1}_Q w^{-\frac{1}{p-1}}$, on voit que w est dans A_p et $[w]_{A_p} \leq A$. Réciproquement, si w est dans A_p, il résulte de l'inégalité de Hölder que

$$\begin{aligned}\left(\frac{1}{|Q|}\int_Q f\,dy\right)^p w(Q) &= \left(\frac{1}{|Q|}\int_Q f w^{1/p}\,w^{-1/p}dy\right)^p w(Q)\\ &\leq [w]_{A_p}\int_Q f(y)^p\,w(y)dy.\end{aligned}$$

Les autres conditions nécessaires s'obtiennent facilement si on n'essaie pas d'optimiser les constantes : si M est de type fort, alors M est de type faible. De plus, si M est de type faible, comme, sur Q, on a l'inégalité $Mf > \frac{1}{|Q|}\int_Q f\,dy$, l'inégalité traduisant le fait que l'opérateur est de type faible (1, 1) donne immédiatement l'inégalité (13) avec même constante. On montre donc non seulement ainsi toutes les conditions nécessaires, mais aussi que le fait que la croissance en $[w]_{A_p}$ de la constante de (14) est optimale.

Montrons (c) et (d). Nous avons vu qu'il suffisait de considérer le cas dyadique. On suppose donc maintenant que la condition A_p porte uniquement sur les cubes dyadiques et que la fonction maximale a été remplacée par la fonction maximale dyadique. Le point clé est le fait qu'on a aussi bien des théorèmes maximaux lorsque la mesure de Lebesgue est remplacée par une mesure à poids. On notera $\sigma(x) = w(x)^{-\frac{1}{p-1}}$ et M_w, M_σ les deux opérateurs maximaux correspondants. Plus précisément, par exemple,

$$M_w f(x) = \sup_{x\in Q} \frac{1}{w(Q)} \int_Q |f(y)| w(y) dy.$$

Nous ne prenons ici que les cubes dyadiques.

Montrons l'inégalité faible (14) pour $f > 0$. L'inégalité (16) peut encore se lire

$$\left(\frac{1}{|Q|}\int_Q f\, dy\right)^p \leq [w]_{A_p} \frac{1}{w(Q)} \int_Q f(y)^p\, w(y) dy \leq [w]_{A_p} M_w(f(x)^p)$$

lorsque $x \in Q$. Comme ceci a lieu pour tout Q contenant x, on en déduit que

$$M^{\mathcal{D}} f(x)^p \leq [w]_{A_p} M_w(f(x)^p).$$

L'inégalité faible découle immédiatement de l'inégalité forte pour M_w.

Montrons l'inégalité forte (15). On commence par écrire

$$\frac{1}{|Q|}\int_Q f \leq [w]_{A_p}^{\frac{1}{p-1}} \left\{ \frac{|Q|}{w(Q)} \Big(\frac{1}{\sigma(Q)} \int_Q f\Big)^{p-1}\right\}^{\frac{1}{p-1}}.$$

Quel que soit $x \in Q$, on a

$$\frac{1}{\sigma(Q)}\int_Q f \leq M_\sigma(f\sigma^{-1})(x),$$

d'où il découle que

$$\begin{aligned}
\frac{1}{|Q|}\int_Q f &\leq [w]_{A_p}^{\frac{1}{p-1}} \left\{ \frac{|Q|}{w(Q)} \inf_Q \Big(M_\sigma(f\sigma^{-1})\Big)^{p-1}\right\}^{\frac{1}{p-1}} \\
&\leq [w]_{A_p}^{\frac{1}{p-1}} \left\{ \frac{1}{w(Q)} \int_Q \Big(M_\sigma(f\sigma^{-1})\Big)^{p-1}\right\}^{\frac{1}{p-1}} \\
&\leq [w]_{A_p}^{\frac{1}{p-1}} \left\{ M_w \left[w^{-1} \Big(M_\sigma(f\sigma^{-1})\Big)^{p-1}\right]\right\}^{\frac{1}{p-1}} (x).
\end{aligned}$$

Comme cette inégalité a lieu quel que soit le cube contenant x, il en résulte que

$$M^{\mathcal{D}}f \leq [w]_{A_p}^{\frac{1}{p-1}} \left\{ M_w \left[w^{-1} \Big(M_\sigma(f\sigma^{-1}) \Big)^{p-1} \right] \right\}^{\frac{1}{p-1}}.$$

Pour conclure on commence par utiliser le théorème maximal pour M_w avec l'exposant de Lebesgue $p' = \frac{p}{p-1}$: on en déduit que

$$\|M^{\mathcal{D}}f\|_{L^p(w)} \leq [w]_{A_p}^{\frac{1}{p-1}} p^{\frac{1}{p-1}} \left(\int \left[w^{-1} \Big(M_\sigma(f\sigma^{-1}) \Big)^{p-1} \right]^{\frac{p}{p-1}} w dx \right)^{\frac{1}{p}}.$$

On reconnait σ en $w^{-\frac{p}{p-1}}w$ et on utilise le théorème maximal pour M_σ et pour l'exposant p. On conclut aisément que

$$\|M^{\mathcal{D}}f\|_{L^p(w)} \leq \frac{p}{p-1}[w]_{A_p}^{\frac{1}{p-1}} p^{\frac{1}{p-1}} \|f\|_{L^p(w)} \leq \frac{ep}{p-1}[w]_{A_p}^{\frac{1}{p-1}} \|f\|_{L^p(w)}.$$

Au facteur e près, c'est optimal pour l'inégalité sans poids. □

7. Optimalité du théorème de Buckley et méthode d'extrapolation de Rubio de Francia

Nous allons suivre Luque, Pérez et Rela pour montrer que le théorème (6.1) est optimal. Plus précisément, nous allons montrer que

Proposition 7.1. *Soit $p > 1$ donné et soit $\psi : [1, \infty) \to (0, \infty)$ une fonction croissante telle que, quel que soit le poids $w \in A_p$ et quel que soit la fonction f positive,*

$$\|Mf\|_{L^p(w)} \leq \psi([w]_{A_p})\|f\|_{L^p(w)}.$$

Alors il existe des constantes c, c' telle que $\psi(t) \geq ct^{\frac{1}{p-1}}$ pour $t > c'$.

On peut définir la fonction ψ comme

$$\psi(t) = \sup_{w \in A_p\,;\, [w]_{A_p} \leq t} \left(\sup_{f \in L^p(w)\,;\, \|f\|_{L^p(w)}=1} \|Mf\|_{L^p(w)} \right).$$

Démonstration. Pour simplifier les notations nous allons supposer que $p = 2$, mais la démonstration générale est la même. Nous supposons donc que, quel que soit le poids $w \in A_p$ et quel que soit la fonction f positive,

$$\|Mf\|_{L^2(w)} \leq \psi([w]_{A_2})\|f\|_{L^2(w)}. \tag{16}$$

Nous allons maintenant utiliser la lettre p pour un exposant variable inférieur à 2 et que nous ferons tendre vers 1. Rappelons que la plus petite constante dans l'inégalité maximale (3), que nous notons $\|M\|_p$, peut être minorée par $\frac{c}{p-1}$. On en déduit qu'il existe pour chaque $p > 1$ une fonction f positive de norme 1 dans L^p et telle que $\|Mf\|_p \geq \frac{c}{p-1}$.

Nous allons maintenant utiliser l'algorithme d'itération de Rubio de Francia en posant, pour h dans L^p,

$$R(h) = \sum_{k=0}^{\infty} \frac{1}{2^k} \frac{M^k(h)}{\|M\|_p^k}.$$

Il est clair qu'on a les propriétés suivantes :

$$h \leq R(h) \qquad \|R(h)\|_{L^p} \leq 2\,\|h\|_{L^p},$$

$$[R(h)]_{A_1} \leq 2\,\|M\|_p.$$

On prend pour poids la fonction $w = (Rf)^{p-2})$, où f est la fonction choisie au préalable, et on calcule

$$\|Mf)\|_{L^p} = \Big(\int \Big((Mf)(Rf)^{\frac{p}{2}-1})\Big)^p \,(Rf)^{p(1-\frac{p}{2})}dx\Big)^{\frac{1}{p}}.$$

Après utilisation de l'inégalité de Hölder pour les exposants conjugués $\frac{2}{p}$ et $(1-\frac{p}{2})^{-1}$, on trouve la majoration

$$\|Mf)\|_{L^p} \leq \Big(\int (Mf)^2\,(Rf)^{p-2}dx\Big)^{\frac{1}{2}} \Big(\int (Rf)^p dx\Big)^{1-\frac{p}{2}}.$$

Après utilisation de (16) pour le poids w choisi, on obtient

$$\|Mf)\|_{L^p} \leq \psi([w]_{A_2})\Big(\int f^2\,(Rf)^{p-2}dx\Big)^{\frac{1}{2}} \Big(\int (Rf)^p dx\Big)^{1-\frac{p}{2}}.$$

Les propriétés de Rf et le fait que f soit de norme 1 entraînent que

$$\psi([w]_{A_2}) \geq \frac{c}{p-1}.$$

Maintenant, en utilisant l'inégalité de Jensen puisque $2-p < 1$, on trouve que

$$\begin{aligned}
[w]_{A_2} &= \sup_Q \Big(\frac{1}{|Q|}\int_Q (Rf)^{p-2}dx\Big)\Big(\frac{1}{|Q|}\int_Q (Rf)^{2-p}dx\Big) \\
&\leq \sup_Q \Big(\frac{1}{|Q|}\int_Q (Rf)^{-1}dx\Big)^{2-p}\Big(\frac{1}{|Q|}\int_Q (Rf)dx\Big)^{2-p} \\
&\leq [Rf]_{A_2}^{2-p} \leq [Rf]_{A_2} \leq 2\|M\|_p \leq \frac{C}{p-1}.
\end{aligned}$$

Finalement, si on pose $t = \frac{C}{p-1}$, on a $\psi(t) \geq c't$. $\square$

La méthode de Rubio de Francia est très efficace. Elle permet en particulier de montrer que pour les opérateurs de Calderón- Zygmund les inégalités L^p sont conséquences des inégalités à poids pour un seul exposant de Lebesgue p_0, ceci à condition de les avoir pour tout poids de la classe A_{p_0}.

La méthode utilisée dans le dernier théorème permet d'obtenir de nombreuses bornes inférieures, voir [?].

8. Retour aux martingales

Nous allons tout d'abord montrer que la méthode de Lerner pour démontrer le théorème de Buckley s'adapte au cas des martingales. Nous considérons uniquement le cas $p = 2$ pour simplifier, et l'inégalité forte. Le théorème maximal en toute généralité pour les martingales a été démontré en 1986 par Jawerth. La démonstration de Lerner est plus simple en même temps qu'elle donne la constante optimale.

On suppose donnée une suite croissante de tribus $\mathcal{F}_n$ sur un espace de probabilité $(\Omega, \mathbb{P})$. La condition A_2 pour un poids Z s'écrit dans ce contexte :

$$[Z]_{A_2} = \sup_n \mathbb{E}(Z|\mathcal{F}_n)\mathbb{E}(Z^{-1}|\mathcal{F}_n) < \infty. \tag{17}$$

On a alors la proposition suivante :

Proposition 8.1. *Soit* $(f_n = \mathbb{E}(f|\mathcal{F}_n)$ *une martingale positive,* f^* *sa fonction maximale. Alors*

$$\mathbb{E}((f^*)^2 Z) \leq 16[Z]_{A_2}^2 \mathbb{E}((f^2 Z).$$

Démonstration. Nous l'esquissons seulement. Si on reprend la preuve du théorème de Buckley, ce qu'il nous faut c'est une majoration ponctuelle de la fonction maximale f^* en termes des deux fonctions maximales M_Z et M_U, où on a remplacé la loi de probabilité initiale par la loi $Zd\mathbb{P}$ et la loi $Ud\mathbb{P}$, avec $U = Z^{-1}$. Il s'agit donc de montrer que

$$f^* \leq [Z]_{A_2} M_Z \left[U\Big(M_U(fZ)\Big)\right],$$

la démonstration étant ensuite identique. Mais de la définition de l'espérance conditionnelle (avec des notations évidentes pour les nouvelles probabilités) on déduit que

$$\mathbb{E}(f|\mathcal{F}_n) = \mathbb{E}(U|\mathcal{F}_n)\mathbb{E}_{\mathbb{U}}(fZ|\mathcal{F}_n)$$

et, pour les mêmes raisons,

$$1 = \mathbb{E}(Z|\mathcal{F}_n)\mathbb{E}_{\mathbb{Z}}(U|\mathcal{F}_n).$$

On en déduit que

$$\mathbb{E}(f|\mathcal{F}_n) = \mathbb{E}(U|\mathcal{F}_n)\mathbb{E}(Z|\mathcal{F}_n)\mathbb{E}_{\mathbb{Z}}(U\mathbb{E}_{\mathbb{U}}(fZ|\mathcal{F}_n)|\mathcal{F}_n).$$

On a donc la majoration

$$\mathbb{E}(f|\mathcal{F}_n) \leq [Z]_{A_2}\mathbb{E}_{\mathbb{Z}}(UM_U(fZ)|\mathcal{F}_n).$$

On passe au supremum pour conclure. $\square$

Nous allons maintenant définir les **transformées de martingales**, à la fois en général et plus spécifiquement pour les martingales dyadiques. Etant donnée une suite v_n de fonctions $\mathcal{F}_{n-1}$ mesurables (suite prévisible) la fonction v_0 étant constante, on définit l'opérateur T_v comme l'opérateur qui à la martingale $(f_n)_{n\geq 0}$ fait correspondre la martingale (g_n) définie par les égalités

$$g_0 = v_0 f_0 \qquad g_n - g_{n-1} = v_n(f_n - f_{n-1}), \qquad n \geq 1.$$

La suite $d_n = f_n - f_{n-1}$ est appelée suite des différences de la martingale. On pose $d_0 = f_0$. Le fait que (f_n) est une martingale se traduit par le fait que

$$\mathbb{E}(d_n|\mathcal{F}_{n-1}) = 0.$$

Avec ces notations on voit immédiatement que la n-ième différence de $T_v f$ est donnée par $v_n d_n$.

On a le lemme suivant :

Lemme 8.2. *L'opérateur T_v est borné dans L^2 si (et seulement si) la suite v_n est bornée.*

C'est une conséquence immédiate du fait que

$$\|f\|_2^2 = \sup \|f_n\|_2^2 = \sum_{n\geq 0} \mathbb{E}(d_n^2).$$

On démontre tout d'abord la formule lorsque $d_n = 0$ à partir d'un certain rang. C'est alors une conséquence du fait que les d_n sont orthogonaux :

$$\mathbb{E}(d_n d_{n+k}) = \mathbb{E}(\mathbb{E}(d_n d_{n+k}|\mathcal{F}_n)) = \mathbb{E}(d_n\mathbb{E}(d_{n+k}|\mathcal{F}_n)) = 0.$$

Exercice 10. *Montrer que la condition est nécessaire pour les martingales dyadiques.*

Pour les martingales dyadiques de l'intervalle I_0, il est facile de voir les transformées de martingales dans la base de Haar : soit I un intervalle de la n-ième génération. La fonction v_n est constante sur I. Si α_I

est sa valeur sur I, on a $T_v(h_I) = \alpha_I h_I$. Pour une fonction f dans L^2 quelconque,

$$T_v(f) = \sum_I \alpha_I \langle f, h_I \rangle h_I + \alpha_0 \mathbb{1}_{I_0}.$$

Cette transformation est ce qu'on appelle un **multiplicateur de Haar**. Son nom est dû au fait qu'elle se diagonalise dans la base de Haar. On définit de la même manière un multiplicateur de Haar sur $\mathbb{R}$ entier. Nous noterons plutôt T_α la transformation dans ce cas. A nouveau le théorème principal, dont un des ingrédients est le théorème maximal, est le suivant :

Théorème 8.3. *Soit T_v une transformée de martingale, avec $\sup_n |v_n| \leq 1$. Alors T_v est de type faible $(1,1)$ et de type fort (p,p).*

Ce théorème est essentiellement dû à Burkholder. Il suffit de montrer l'inégalité faible (1, 1). La démonstration générale comporte des difficultés liées aux sauts de la martingale dans le cas général. Nous allons le montrer dans le cas dyadique. Plus précisément nous allons considérer un multiplicateur de Haar dans $\mathbb{R}$ donné par la suite α_I telle que $|\alpha_I| \leq 1$.

Démonstration. Il suffit de montrer que

$$|\{T_\alpha f \geq \lambda\}| \leq C\|f\|_{L^1}$$

lorsque f possède seulement un nombre fini de termes non nuls dans la base de Haar. La propriété spécifique de T_α est la suivante :

Lemme 8.4. *Si ϕ est à support dans I et d'intégrale nulle, alors $T_\alpha(\phi) = \sum_{I' \subset I} \alpha_{I'} h_{I'}$ est à support dans I et de norme L^2 majorée par celle de ϕ.*

Le point-clé est la **décomposition de Calderón-Zygmund** de la fonction f. On suppose donné $\lambda > 0$. On rappelle que l'ensemble des points $x \in I$ pour lesquels la fonction maximale dyadique de f est supérieure à λ est la réunion des intervalles dyadiques maximaux I_j pour lesquels

$$\int_{I_j} w > \lambda |I_j|.$$

Mais, par maximalité, on a aussi que

$$\int_{I_j} w \leq 2\lambda |I_j|.$$

Donc, si on pose $f = g + b$ (ou encore $f = f_\tau + (f - f_\tau)$) avec

$$g(x) = \left\{ \begin{array}{ll} f(x) & \text{si } x \quad \text{n'appartient à aucun des } I_j \\ \frac{1}{|I_j|}\int_{I_j} f(y)dy & \text{si } x \in I_j \end{array} \right. .$$

Alors $|g| \leq 2\lambda$ et $T_\alpha b$ est à support dans la réunion des I_j, c'est-à-dire dans l'ensemble où $M^{\mathcal{D}}f > \lambda$. Donc

$$|\{T_\alpha f \geq \lambda\}| \leq |\{M^{\mathcal{D}} f \geq \lambda\}| + |\{T_\alpha g \geq \lambda\}| \leq \frac{1}{\lambda}\|f\|_1.$$

On a utilisé le fait que

$$\|\{T_\alpha g\|_2 \leq 2\lambda\|g\|_1 \leq \lambda\|f\|_1.$$

$\square$

En utilisant le théorème de Khintchine on en déduit immédiatement que

$$\|(\sum_{n\geq 0} \mathbb{E}(d_n^2))^{1/2}\|_p \leq C_p\|f\|_p.$$

La quantité $(\sum_{n\geq 0} \mathbb{E}(d_n^2))^{1/2}$ est la **fonction carrée de la martingale**.

9. Retour sur les conditions A_p

Le but de cette section est de donner un aperçu d'autres aspects de la théorie de Calderón-Zygmund. Une des propriétés importantes des poids A_p est le fait qu'un tel poids w satisfait à des inégalité de Hölder inverses : il existe $r > 1$ tel que

$$\left(\frac{1}{|Q|}\int_Q w^r\right)^{1/r} \leq \frac{c}{|Q|}\int_Q w.$$

C'est une propriété fondamentale, due à Gehring, qui joue en particulier un rôle important dans la théorie des fonctions quasi-conformes. Nous allons montrer une propriété de ce type, avec une valeur de r explicite, quand w est dans la classe A_1, qui est, rappelons-le, plus petite. Nous suivons Lerner, Ambrosi et Pérez (2008) qui donnent des constantes très précises. Ces résultats ont depuis été améliorés avec des versions A_p.

Proposition 9.1. *Soit w un poids sur $\mathbb{R}^d$ tel que*

$$M^{\mathcal{D}} w \leq [w]_{A_1} w.$$

Alors, si $r = 1 + 1/(2^{d+1}[w]_{A_1})$, on a l'inégalité pour tout cube dyadique

$$\left(\frac{1}{|Q|}\int_Q w^r\right)^{1/r} \leq \frac{2[w]_{A_1}}{|Q|}\int_Q w.$$

Démonstration. Nous considérerons pour simplifier uniquement le cas $d = 1$ mais la démonstration est la même en dimension supérieure. Nous revenons au théorème maximal pour la mesure wdx et pour la fonction w dans l'intervalle dyadique I. On suppose que λ est plus grand que la moyenne $\langle w\rangle_I$ de w sur I. On rappelle que l'ensemble des points $x \in I$ pour lesquels la fonction maximale dyadique de w (restreinte à I) est supérieure à λ est la réunion des intervalles dyadiques maximaux I_j pour lesquels

$$\int_{I_j} w > \lambda |I_j|.$$

Mais, par maximalité, on a aussi que

$$\int_{I_j} w \leq 2\lambda |I_j|.$$

Donc

$$\int_{M_I^{\mathcal{P}} w > \lambda} wdx \leq 2|M_I^{\mathcal{P}} w > \lambda|.$$

On calcule

$$\begin{aligned} \int_I (M_I^{\mathcal{P}} w)^\delta w dx &= \delta \int_0^\infty \lambda^{\delta-1} \left(\int_{M_I^{\mathcal{P}} > \lambda} w \right) d\lambda \\ &\leq \frac{1}{|I|^\delta} (\int_I wdx)^{\delta+1} + \frac{2\delta}{\delta+1} \int_I (M_I^{\mathcal{P}} w)^{\delta+1} dx. \end{aligned}$$

Avec le choix proposé pour δ on peut faire passer le second terme du membre de droite à gauche et conclure. □

A compléter