Université de Franche-Comté

Mémoire de Master 1

Une démonstration probabiliste du théorème de Dvoretzky

FOVELLE Audrey

Encadrant : M. LANCIEN Gilles

Année 2018

Table des matières

Dans l'unité d'espaces fonctionnels de L3, nous avons vu que dans un espace de Hilbert H, pour tout sous-espace vectoriel fermé F de H, il existe une projection $P : H \to F$ linéaire continue, ce qui signifie que tout sous-espace vectoriel fermé de H est complémenté dans H.

Un théorème dû à Lindenstrauss et et Tzafriri dit quant à lui que si X est un espace de Banach de dimension infinie dans lequel tout sous-espace vectoriel fermé est complémenté, alors X est isomorphe à un espace de Hilbert. Ce théorème permet de faciliter la classification des espaces de Banach, une des motivations de la géométrie de ces espaces. Sa preuve fait appel au théorème de Dvoretzky, démontré par Aryeh Dvoretzky au début des années 1960.

Ce théorème de Dvoretzky permet également de répondre par l'affirmative à la question de Grothendieck, posée en 1956, qui se demandait si tout espace de Banach de dimension infinie admettait un sous-espace de dimension finie dont la distance à un espace de Hilbert était au plus d'ordre constant, et dont la dimension pouvait être arbitrairement grande.

C'est ce dernier théorème que nous allons prouver ici, en reprenant la démonstration de Milman, datant des années 1970, et qui utilise la théorie des probabilités.

Pour cela, nous commencerons tout d'abord par introduire l'ellipsoïde de John d'un espace de Banach de dimension finie, et étudier certaines de ses propriétés. Nous démontrerons ensuite dans une deuxième partie un théorème connu sous le nom de "phénomène de concentration de la mesure", qui dit qu'une fonction lipschitzienne définie sur la sphère euclidienne de dimension n se rapproche de plus en plus de sa moyenne quand n augmente.
Pour finir, nous démontrerons le théorème de Dvoretzky.

Notons que tous les espaces vectoriels seront pris réels, même si la plupart des théorèmes énoncés peuvent être étendus au cas complexe.

1 Ellipsoïde de John

Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n et soit $\|.\|_E$ une norme euclidienne sur X, *i.e* une norme induite par un produit scalaire sur X.

Définition 1. *On appelle ellipsoïde de X toute boule unité fermée de l'espace vectoriel normé $(X, \|.\|)$, $\mathcal{E}_n := \{x \in X; \|x\| \leq 1\}$, où $\|.\|$ est une norme euclidienne sur X.*

Commençons par donner une caractérisation un peu plus pratique d'un ellipsoïde.

Proposition 1. *Soit $A \subset X$. A est un ellipsoïde de X si et seulement si il existe une application $S : \ell_2^n \to X$ linéaire bijective telle que $A = S(B_{\ell_2^n})$.*

Démonstration. Soit $\mathcal{E}_n$ un ellipsoïde de X, dont on note la norme euclidienne associée $\|.\|_E$.
Montrons qu'il existe une application $S : \ell_2^n \to X$ linéaire bijective telle que $\mathcal{E}_n = S(B_{\ell_2^n})$.
Soit $(e_i)_{i=1}^n \in (\ell_2^n)^n$ la base de vecteurs unitaires de ℓ_2^n. On choisit une base $(x_i)_{i=1}^n \in X^n$ de X, orthonormée par rapport au produit scalaire induisant $\|.\|_E$. On peut alors définir une application linéaire bijective S :

$$S : \begin{cases} \ell_2^n & \to \quad X \\ e_i & \mapsto \quad x_i \text{ pour tout } i \in \{1, \cdots, n\} \end{cases}$$

On pose $B_{\ell_2^n} := \{y = (y_i)_{i=1}^n \in \mathbb{R}^n; \sum_{i=1}^n |y_i|^2 \leq 1\}$.
Montrons que $S(B_{\ell_2^n}) = \mathcal{E}_n$. Soit $y = (y_i)_{i=1}^n \in X$.

$$\begin{aligned} y \in S(B_{\ell_2^n}) &\Leftrightarrow \exists (\lambda_i)_{i=1}^n \in \mathbb{R}^n; y = S(\sum_{i=1}^n \lambda_i e_i), \sum_{i=1}^n |\lambda_i|^2 \leq 1 \\ &\Leftrightarrow \exists (\lambda_i)_{i=1}^n \in \mathbb{R}^n; y = \sum_{i=1}^n \lambda_i x_i, \sum_{i=1}^n |\lambda_i|^2 \leq 1 \\ &\Leftrightarrow y \in \mathcal{E}_n \end{aligned}$$

Réciproquement, soit $S : \ell_2^n \to X$ une application linéaire bijective. Montrons que $S(B_{\ell_2^n})$ est un ellipsoïde de $(X, \|.\|_X)$.
On note $\langle , \rangle$ le produit scalaire de ℓ_2^n.

$$(,) : \left\{ \begin{array}{lll} X \times X & \to & \mathbb{R} \\ (x, y) & \mapsto & \langle S^{-1}x, S^{-1}y \rangle \end{array} \right.$$

est un produit scalaire sur X, dont on note $\|.\|_E$ la norme associée.
Montrons que $S(B_{\ell_2^n}) = \{x \in X; \|x\|_E \leq 1\}$.
Soit $x \in X$. On a :

$$\|x\|_E^2 = (x, x) = \langle S^{-1}x, S^{-1}x \rangle = \|S^{-1}x\|_{\ell_2^n}^2$$

donc $x \in \{x \in X; \|x\|_E \leq 1\}$ si et seulement si $x \in S(B_{\ell_2^n})$. On a donc montré que tout $A \subset X$ est un ellipsoïde de X si et seulement si $A = S(B_{\ell_2^n})$ où $S : \ell_2^n \to X$ est linéaire bijective. $\square$

On veut désormais montrer qu'il existe un ellipsoïde de volume maximal contenu dans $B_X := \{x \in X; \|x\|_X \leq 1\}$. Plus précisément, si $S : \mathbb{R}^n \to X$ est une application linéaire bijective, le volume d'un borélien A de X est naturellement mesuré par $vol_S(A) = \lambda^{(n)}(S^{-1}(A))$, où l'on a noté $\lambda^{(n)}$ la mesure de Lebesgue sur $\mathbb{R}^n$. On dit alors que $vol_S : \mathscr{B}(X) \to [0, \infty]$ est une mesure de Lebesgue sur X.
La quantité $vol_S(A)$ dépend évidemment de l'application S mais puisque nous cherchons seulement à comparer ces valeurs, et non à les calculer, il suffit d'observer, pour deux boréliens A et B de X, le quotient $\frac{vol_S(A)}{vol_S(B)}$ qui est indépendant de S. En effet, si $S_1 : \ell_2^n \to X$ et $S_2 : \ell_2^n \to X$ sont deux isomorphismes, on a pour tout borélien A de X, à l'aide d'un changement de variable affine :

$$\begin{aligned} vol_{S_1}(A) &= \int_{\mathbb{R}^n} \mathbb{1}_{S_1^{-1}(A)} d\lambda^{(n)} \\ &= \int_{\mathbb{R}^n} |\det(S_1^{-1}S_2)| \mathbb{1}_{S_1^{-1}(A)} \circ S_1^{-1}S_2 d\lambda^{(n)} \\ &= |\det(S_1^{-1}S_2)| vol_{S_2}(A) \end{aligned}$$

car $\mathbb{1}_{S_1^{-1}(A)} \circ S_1^{-1}S_2 = \mathbb{1}_{S_2^{-1}(A)}$.
Donc si A et B sont deux boréliens de X, on a :

$$\frac{vol_{S_1}(A)}{vol_{S_1}(B)} = \frac{|\det(S_1 S_2^{-1})| vol_{S_2}(A)}{|\det(S_1 S_2^{-1})| vol_{S_2}(B)} = \frac{vol_{S_2}(A)}{vol_{S_2}(B)}$$

dès que cela a un sens.
On note que deux mesures de Lebesgue sur un espace vectoriel normé X de dimension finie diffèrent d'une constante (multiplicative).

Théorème 1. *Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n. Il existe un ellipsoïde de dimension n et de volume maximal contenu dans B_X.*

Démonstration. Il existe un ellipsoïde $\mathcal{E}_n$ contenu dans B_X. Il existe une application linéaire $S : \ell_2^n \to X$ telle que $\mathcal{E}_n = S(B_{\ell_2^n})$. On note $\|.\|_E$ la norme euclidienne induite sur X par S (cf prop 1).
Soit $\mathcal{E}'_n$ un ellipsoïde de X. $\mathcal{E}'_n$ est l'image de $\mathcal{E}_n$ par un isomorphisme $T \in \mathscr{L}(X)$.
De plus $\mathcal{E}'_n \subset B_X$ si et seulement si $\|T\|_{E\to X} \leq 1$. En effet, supposons que $\mathcal{E}'_n \subset B_X$. Alors :

$$\begin{aligned}\|T\|_{E\to X} &= \sup\{\|T(x)\|_X, x \in \mathcal{E}_n\} \\ &\leq \sup\{\|y\|_X, y \in \mathcal{E}'_n\} \\ &\leq \sup\{\|y\|_X, y \in B_X\} \\ &\leq 1\end{aligned}$$

Réciproquement, supposons que $\|T\|_{E\to X} \leq 1$. Soit $y \in \mathcal{E}'_n$. Il existe $x \in \mathcal{E}_n$ tel que $y = T(x)$. On a alors :

$$\|y\|_X = \|T(x)\|_X \leq \|T\|_{E\to X} \leq 1$$

D'où $\mathcal{E}'_n \subset B_X$ si et seulement si $\|T\|_{E\to X} \leq 1$.
Or :

$$\begin{aligned}vol_S(\mathcal{E}'_n) = |\det(T)| vol_{T\circ S}(\mathcal{E}'_n) &= |\det(T)|\lambda^{(n)}(S^{-1}(T^{-1}(\mathcal{E}'_n))) \\ &= |\det(T)|\lambda^{(n)}(S^{-1}(\mathcal{E}_n)) = |\det(T)|\lambda^{(n)}(B_{\ell_2^n})\end{aligned}$$

donc il suffit de montrer l'existence d'un maximum pour la fonction $|\det|$ sur la boule unité fermée de l'espace de Banach $(\mathscr{L}(X), \|.\|_{E\to X})$. Or $(\mathscr{L}(X), \|.\|_{E\to X})$ est de dimension finie donc sa boule unité fermée est compacte et $|\det|$ est continue donc on a bien le résultat recherché. □

On pourrait également montrer qu'il n'existe qu'un unique ellipsoïde de volume maximal contenu dans B_X mais nous n'avons besoin que de son existence dans la suite. Ceci justifie la définition suivante :

Définition 2. *L'ellipsoïde de volume maximal contenu dans la boule unité fermée B_X est appelé ellipsoïde de John.*

On note alors $\mathcal{E}_X$ l'ellipsoïde de John de X, $\|.\|_{\mathcal{E}_X}$ la norme euclidienne qu'il induit sur X et $\langle,\rangle$ le produit scalaire associé. On sait qu'il existe une application linéaire $S : \ell_2^n \to X$ telle que $\mathcal{E}_X = S(B_{\ell_2^n})$.
$\mathcal{E}_X \subset B_X$ donc $\|I\|_{\mathcal{E}_X\to X} \leq 1$ et pour toute $T \in \mathscr{L}(X)$ vérifiant $\|T\|_{\mathcal{E}_X\to X} \leq 1$, on a :
$|\det(T)| \leq 1$.
En effet, soit $T \in GL(X)$ (le résultat est clair si T n'est pas inversible) vérifiant $\|T\|_{\mathcal{E}_X\to X} \leq 1$. Puisque $\|T\|_{\mathcal{E}_X\to X} \leq 1$, $T(\mathcal{E}_X)$ est un autre ellipsoïde de X. Or :

$$vol_S(T(\mathcal{E}_X)) = |\det(T)|\lambda^{(n)}(B_{\ell_2^n}) = |\det(T)| vol_S(\mathcal{E}_X)$$

d'où $|\det(T)| \leq 1$ par définition de $\mathcal{E}_X$.
On en déduit que si $dim(X) = n$, pour toute $T \in GL(X)$, on a :

$$|\det(T)| = \left|\det\left(\|T\|_{\mathcal{E}_X\to X} \times \frac{T}{\|T\|_{\mathcal{E}_X\to X}}\right)\right| = \|T\|^n_{\mathcal{E}_X\to X}\left|\det\left(\frac{T}{\|T\|_{\mathcal{E}_X\to X}}\right)\right| \leq \|T\|^n_{\mathcal{E}_X\to X}$$

On note que cette inégalité reste vraie si T n'est pas inversible.
On note également que $\det(I) = 1$ donc $\|I\|_{\mathcal{E}_X \to X} = 1$, ce qui signifie qu'il existe $x \in \mathcal{E}_X$ tel que $\|x\|_X = 1$ car la boule unité fermée de X est compacte.

Nous noterons $(X, \|.\|_{X^*})$ l'espace dual de $(X, \|.\|_X)$ où, pour tout $x \in X$, on note :

$$\|x\|_{X^*} = Sup\{|\langle x, y\rangle|; y \in B_X\}$$

Nous cherchons désormais à estimer la "distance entre X et ℓ_2^n". Pour cela, nous avons besoin de plusieurs résultats, énoncés ci-dessous.

Lemme 1. *On a :* $\|I\|_{X^* \to \mathcal{E}_X} \leq 1$.

Démonstration. Soit $x \in X\backslash\{0\}$. On a :

$$\begin{aligned}\|x\|_{\mathcal{E}_X} &= \frac{1}{\|x\|_{\mathcal{E}_X}}\langle x, x\rangle \leq \frac{1}{\|x\|_X}\langle x, x\rangle \\ &\leq |\langle x, \frac{1}{\|x\|_X}x\rangle| \leq \sup\{|\langle x, y\rangle|, y \in B_X\} \\ &\leq \|x\|_{X^*}\end{aligned}$$

Donc pour tout $x \in X$, $\|x\|_{\mathcal{E}_X} \leq \|x\|_{X^*}$.
D'où $\|I\|_{X^* \to \mathcal{E}_X} \leq 1$. □

On a alors :

$$\forall x \in X, \|x\|_X \leq \|x\|_{\mathcal{E}_X} \leq \|x\|_{X^*}$$

Lemme 2. *Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n. Soit $x \in X$ tel que $\|x\|_X = \|x\|_{\mathcal{E}_X} = 1$.*
On a alors $\|x\|_{X^*} = 1$.

Démonstration. Soit $y \in B_X$. D'après le lemme 1, il suffit de montrer que $\langle x, y\rangle \leq 1$.
Pour tout $t > 0$, on a :

$$1 \leq (1+t)\|x\|_X - t\|y\|_X \leq \|(1+t)x - ty\|_X \leq \|(1+t)x - ty\|_{\mathcal{E}_X}$$

car $(1+t)\|x\|_X - t\|y\|_X = 1 + t(1 - \|y\|_X) \geq 1$. D'où, pour tout $t > 0$:

$$\begin{aligned}1 &\leq \|(1+t)x\|_{\mathcal{E}_X}^2 + \|ty\|_{\mathcal{E}_X}^2 - 2\langle(1+t)x, ty\rangle \\ &\leq (1+t)^2 + t^2\|y\|_{\mathcal{E}_X}^2 - 2t(1+t)\langle x, y\rangle\end{aligned}$$

On a alors, pour tout $t > 0$:

$$0 \leq 2 + t + t\|y\|_{\mathcal{E}_X}^2 - 2(1+t)\langle x, y\rangle$$

i.e pour tout $t > 0$:

$$2\langle x, y\rangle \leq 2 + t + t\|y\|_{\mathcal{E}_X}^2 - 2t\langle x, y\rangle$$

On a donc le résultat en passant à la limite quand $t \to 0^+$. □

Lemme 3. *On a, pour tout $H \in \mathcal{M}_n(\mathbb{R})$, $d\det(I)H = tr(H)$.*

Démonstration. Pour tout $(i,j) \in [|1,n|]^2$, on a :

$$1 + \lambda\delta_{i,j} = \det(I + \lambda E_{i,j}) = 1 + \lambda d\det(I)E_{i,j} + \circ_{\lambda\to 0}(\lambda)$$

donc $d\det(I)E_{i,j} = \delta_{i,j}$, d'où le résultat. □

Lemme 4. *Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n, et $T \in \mathscr{L}(X)$.*
Il existe $x \in X$ tel que $\|x\|_X = \|x\|_{\mathcal{E}_X} = 1$ et $|tr(T)| \leq n\|T(x)\|_X$.

Démonstration. Pour tout $k \in \mathbb{N}^*$, il existe $x_k \in X$ tel que $\|x_k\|_{\mathcal{E}_X} = 1$ et

$$\left\|x_k + \frac{1}{k}T(x_k)\right\|_X = \left\|I + \frac{1}{k}T\right\|_{\mathcal{E}_X\to X}$$

car X est de dimension finie et $I + \frac{1}{k}T$ est linéaire.
Pour tout $k \in \mathbb{N}^*$, $x_k \in \mathcal{E}_X$ et $\mathcal{E}_X$ est compact donc il existe $(k_l)_{l=1}^{\infty} \subset \mathbb{N}^*$ strictement croissante et $x \in \mathcal{E}_X$ tels que $\lim\limits_{l\to+\infty} x_{k_l} = x$.
On note que $\|x\|_{\mathcal{E}_X} = 1$ et, par passage à la limite, $\|x\|_X = \|I\|_{\mathcal{E}_X\to X} = 1$.
De plus, pour tout $l \in \mathbb{N}^*$:

$$\det\left(I + \frac{1}{k_l}T\right) \leq \left\|I + \frac{1}{k_l}T\right\|_{\mathcal{E}_X\to X} = \left\|x_{k_l} + \frac{1}{k_l}T(x_{k_l})\right\|_X^n \leq \left(1 + \frac{1}{k_l}\|T(x_{k_l})\|_X\right)^n$$

D'où :

$$tr(T) = \lim_{t\to 0^+} \frac{det(I + tT) - det(I)}{t} \leq \lim_{l\to+\infty} \frac{(1 + \frac{1}{k_l}\|T(x_{k_l})\|_X)^n - 1}{\frac{1}{k_l}} = n\|T(x)\|_X$$

□

Théorème 2 (Hahn-Banach)**.** *Soit $(X, \|.\|)$ un $\mathbb{R}$-espace vectoriel normé, E un sous-espace vectoriel de X, $f \in E^*$.*
Il existe $F \in X^*$ telle que $F|_E = f$ et $\|F\|_{X^*} = \|f\|_{E^*}$.

Remarque 1. *On note que l'application J est une isométrie linéaire bijective, où J est définie par :*

$$J : \left\{\begin{array}{lcl} (X, \|.\|_{X^*}) & \to & (X^*, \|.\|_{X^*}) \\ x & \mapsto & \left\{\begin{array}{lcl} X & \to & \mathbb{R} \\ y & \mapsto & \langle x, y\rangle \end{array}\right. \end{array}\right.$$

Lemme 5. *Soit $u \in X$. Il existe $y \in X$ tel que $\|y\|_{X^*} = 1$ et $\|u\|_X = \langle u, y\rangle$.*

Démonstration. On pose $E := \mathbb{R}u$ et $f : \left\{\begin{array}{lcl} E & \to & \mathbb{R} \\ \lambda u & \mapsto & \lambda\|u\|_X \end{array}\right.$. f est une forme linéaire continue sur E vérifiant $\|f\|_{E^*} = 1$.
D'après le théorème de Hahn-Banach, il existe $F \in X^*$ telle que $F|_E = f$ et $\|F\|_{X^*} = 1$. D'après la remarque 1, il existe $y \in X$ tel que pour tout $x \in X$, $\langle x, y\rangle = F(x)$. On note alors que $\|u\|_X = F(u) = \langle u, y\rangle$ et :

$$\|y\|_{X^*} = \sup\{|\langle x, y\rangle|, x \in B_X\} = \sup\{|F(x)|, x \in B_X\} = \|F\|_{X^*} = 1$$

□

Théorème 3. *Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n.*
On a $\pi_2(I_{X\to\mathcal{E}_X}) \le \sqrt{n}$ où $\pi_2(I_{X\to\mathcal{E}_X})$ est la plus petit constante $C > 0$ telle que pour tout $k \in \mathbb{N}^*$, pour tout $(x_1, \cdots, x_k) \in X^k$, on ait

$$\left(\sum_{i=1}^{k} \|x_i\|_{\mathcal{E}_X}^2\right)^{\frac{1}{2}} \le C \sup\left\{\left(\sum_{i=1}^{k} |x^*(x_i)|^2\right)^{\frac{1}{2}} ; x^* \in B_{X^*}\right\}$$

Démonstration. Soit $k \in \mathbb{N}^*$, $(x_1, \cdots, x_k) \in X^k$. On note $(u_1, \cdots, u_n)$ une base orthonormée de $(X, \|.\|_{\mathcal{E}_X})$.
On note T l'application linéaire définie par :

$$T : \left\{\begin{array}{lll} X & \to & X \\ u & \mapsto & \sum_{i=1}^{k} \langle x_i, u\rangle x_i \end{array}\right.$$

On a :

$$\begin{aligned} tr(T) &= \sum_{j=1}^{n} \langle T(u_j), u_j\rangle = \sum_{j=1}^{n} \langle \sum_{i=1}^{k} \langle x_i, u_j\rangle x_i, u_j\rangle \\ &= \sum_{j=1}^{n}\sum_{i=1}^{k} \langle x_i, u_j\rangle\langle x_i, u_j\rangle = \sum_{i=1}^{k}\langle x_i, \sum_{j=1}^{n}\langle x_i, u_j\rangle u_j\rangle \\ &= \sum_{i=1}^{k}\langle x_i, x_i\rangle = \sum_{i=1}^{k}\|x_i\|_{\mathcal{E}_X}^2 \end{aligned}$$

D'après les lemmes 2 et 4, il existe $x \in X$ tel que $\|x\|_X = \|x\|_{\mathcal{E}_X} = \|x\|_{X^*} = 1$ et $tr(T) \le n\|T(x)\|_X$.
D'après le lemme 5, il existe $y \in X$ tel que $\|y\|_{X^*} \le 1$ et $\|T(x)\|_X = \langle T(x), y\rangle$.
On a alors, d'après l'inégalité de Cauchy-Schwarz :

$$\begin{aligned} \sum_{i=1}^{k}\|x_i\|_{\mathcal{E}_X}^2 = tr(T) \le n\|T(x)\|_X &= n\langle T(x), y\rangle \\ &\le n\langle \sum_{i=1}^{k}\langle x_i, x\rangle x_i, y\rangle = n\sum_{i=1}^{k}\langle x_i, x\rangle\langle x_i, y\rangle \\ &\le n\left(\sum_{i=1}^{k}|\langle x_i, x\rangle|^2\right)^{\frac{1}{2}}\left(\sum_{i=1}^{k}|\langle x_i, y\rangle|^2\right)^{\frac{1}{2}} \\ &\le n \max\left\{\sum_{i=1}^{k}|\langle x_i, u\rangle|^2, \|u\|_{X^*} \le 1\right\} \end{aligned}$$

Or, d'après la remarque 1, on a :

$$\max\left\{\sum_{i=1}^{k}|\langle x_i, u\rangle|^2, \|u\|_{X^*} \le 1\right\} = \sup\left\{\left(\sum_{i=1}^{k}|x^*(x_i)|^2\right)^{\frac{1}{2}} ; x^* \in B_{X^*}\right\}$$

D'où $(\pi_2(I_{X\to\mathcal{E}_X}))^2 \le n$ *i.e* $\pi_2(I_{X\to\mathcal{E}_X}) \le \sqrt{n}$. □

Afin d'estimer la "distance entre X et ℓ_2^n", nous allons avoir besoin d'introduire la distance de Banach-Mazur et de certaines de ses propriétés.

Définition 3. *Soit X et Y deux espaces vectoriels normés de même dimension finie.*
La distance de Banach-Mazur entre X et Y, noté $d(X,Y)$ est définie par :

$$d(X,Y) := \inf\{\|T\|\|T^{-1}\|; T \in GL(X,Y)\}$$

Proposition 2. *Soit E, F et G trois espaces vectoriels normés de même dimension finie. On a :*
(i) $d(E,F) \geq 1$;
(ii) $d(E,F) = d(F,E)$;
(iii) $d(E,F) = \min\{\|T\|\|T^{-1}\|; T \in GL(E,F)\}$;
(iv) E et F sont isométriques si et seulement si $d(E,F) = 1$;
(v) $d(E,G) \leq d(E,F) \times d(F,G)$.

Démonstration. (i) Pour tout $T \in GL(E,F)$, on a :

$$1 = \|I_F\| = \|T \circ T^{-1}\| \leq \|T\|\|T^{-1}\|$$

donc $d(E,F) \geq 1$.
(ii) On a :

$$d(E,F) = \inf\{\|T\|\|T^{-1}\|; T \in GL(E,F)\} = \inf\{\|T\|\|T^{-1}\|; T^{-1} \in GL(F,E)\}$$

car $T \to T^{-1}$ est une bijection de $GL(E,F)$ dans $GL(F,E)$. D'où :

$$d(E,F) = \inf\{\|(T^{-1})^{-1}\|\|T^{-1}\|; T^{-1} \in GL(F,E)\} = \inf\{\|T^{-1}\|\|T\|; T \in GL(F,E)\} = d(F,E).$$

(iii) Il suffit de montrer qu'il existe $S \in GL(E,F)$ tel que $\|S\|\|S^{-1}\| = d(E,F)$.
Soit $C > d(E,F)$. On pose :

$$K := \{T \in \mathscr{L}(E,F); \forall x \in E, \|x\|_E \leq \|T(x)\|_F \leq C\|x\|_E\}$$

K est un ensemble fermé borné, non vide puisque $C > d(E,F)$, de $\mathscr{L}(E,F)$ qui est de dimension finie donc K est un compact de $\mathscr{L}(E,F)$, ce qui assure alors l'existence de $S \in K$ telle que $\|S\| = \min\{\|T\|; T \in K\}$. Pour tout $x \in E$, $\|x\|_E \leq \|S(x)\|_F \leq \|S\|\|x\|_E$ donc S est injective. Or $dim(E) = dim(F) < \infty$ donc $S \in GL(E,F)$.
Montrons que $\|S\|\|S^{-1}\| = d(E,F)$. Sinon, il existe $U \in GL(E,F)$ tel que $\|U\| < \|S\|\|S^{-1}\| \leq \|S\|$ et $\|U^{-1}\| = 1$. On note alors que $U \in K$, ce qui contredit la définition de S.
(iv) Supposons E et F isométriques. Il existe $S \in \mathscr{L}(E,F)$ telle que pour tout $x \in E$, $\|S(x)\|_F = \|x\|_E$. On a alors $S \in GL(E,F)$ et $\|S\|\|S^{-1}\| = 1 \times 1 = 1$ donc, $d(E,F) \leq 1$ et avec (i), $d(E,F) = 1$.
Réciproquement, supposons $d(E,F) = 1$. D'après (iii), il existe $S \in GL(E,F)$ telle que $\|S\|\|S^{-1}\| = 1$. Montrons que $T := \frac{1}{\|S\|}S$ est une isométrie :

$$\forall x \in E, \|T(x)\|_F \leq \|x\|_E = \|T^{-1}(T(x))\|_E \leq \|T(x)\|_F$$

car $\|T\| = 1$ et, comme $T^{-1} = \|S\|S^{-1}$, $\|T^{-1}\| = 1$.
(v) D'après (iii), il existe $U \in GL(E,F)$ et $V \in GL(F,G)$ telles que $d(E,F) = \|U\|\|U^{-1}\|$, $d(F,G) = \|V\|\|V^{-1}\|$. On pose $W := V \circ U \in GL(E,G)$. On a :

$$d(E,G) \leq \|W\|\|W^{-1}\| \leq \|V\|\|U\|\|U^{-1}\|\|V^{-1}\| = d(E,F) \times d(F,G).$$

□

Nous pouvons désormais établir une inégalité concernant $d(X, \ell_2^n)$:

Théorème 4. *Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n.*
On a $d(X, \ell_2^n) \leq \sqrt{n}$.

Démonstration. On a $d(X, \ell_2^n) \leq d(X, (X, \|.\|_{\mathcal{E}_X})) \times d((X, \|.\|_{\mathcal{E}_X}), \ell_2^n)$, $d((X, \|.\|_{\mathcal{E}_X}), \ell_2^n) = 1$, et $d((X, \|.\|_{\mathcal{E}_X}), X) \leq \|I\|_{\mathcal{E}_X \to X} \|I\|_{X \to \mathcal{E}_X}$. Or $\|I\|_{\mathcal{E}_X \to X} = 1$ et $\pi_2(I_{X \to \mathcal{E}_X}) \leq \sqrt{n}$ donc il suffit de montrer que $\|I\|_{X \to \mathcal{E}_X} \leq \pi_2(I_{X \to \mathcal{E}_X})$. Or, pour tout $x \in X$, on a :

$$\|x\|_{\mathcal{E}_X} \leq \pi_2(I_{X \to \mathcal{E}_X}) \sup\{|y^*(x)|, y^* \in B_{X^*}\} \leq \pi_2(I_{X \to \mathcal{E}_X}) \sup\{1 \times \|x\|_X, y^* \in B_{X^*}\} \leq \pi_2(I_{X \to \mathcal{E}_X}) \|x\|_X$$

D'où le résultat. □

La proposition suivante montre qu'on ne peut pas trouver de meilleur estimation.

Définition 4. *Une suite de Rademacher est une suite de variables aléatoires $(\varepsilon_n)_{n \in \mathbb{N}}$ mutuellement indépendantes définies sur un espace de probabilité $(\Omega, \mathscr{T}, \mathbb{P})$ telle que pour tout $n \in \mathbb{N}$, $\mathbb{P}(\varepsilon_n = 1) = \mathbb{P}(\varepsilon_n = -1) = \frac{1}{2}$.*

Proposition 3. *Si $X = \ell_\infty^n$, alors $d(X, \ell_2^n) = \sqrt{n}$.*

Démonstration. Soit $S : \ell_\infty^n \to \ell_2^n$ un isomorphisme linéaire. On pose $D = \|S\|_{\ell_\infty^n \to \ell_2^n}$ et $C = \|S^{-1}\|_{\ell_2^n \to \ell_\infty^n}$. Pour tout $x \in \ell_\infty^n$, on a :

$$\frac{1}{C} \|x\|_\infty \leq \|S(x)\|_2 \leq D \|x\|_\infty$$

Pour tout $(\varepsilon_i)_{i=1}^n \in \{-1, 1\}^n$, on pose : $U_{\varepsilon_1, \cdots, \varepsilon_n} : \left\{ \begin{array}{lcl} \ell_\infty^n & \to & \ell_\infty^n \\ x & \mapsto & (\varepsilon_1 x_1, \cdots, \varepsilon_n x_n) \end{array} \right.$, qui est une isométrie, ce qui implique que pour tout $x \in \ell_\infty^n$, on a également $\frac{1}{C} \|x\|_\infty \leq \|SU_{\varepsilon_1, \cdots, \varepsilon_n}(x)\|_2 \leq D \|x\|_\infty$.
Soit $(\varepsilon_1, \cdots, \varepsilon_n)$ une suite finie de Rademacher définie sur un espace de probabilité $(\Omega, \mathscr{T}, \mathbb{P})$.
On pose :

$$T : \left\{ \begin{array}{lcl} \ell_\infty^n & \to & L^2(\Omega, \mathbb{P}; \ell_2^n) \\ x & \mapsto & \left\{ \begin{array}{lcl} (\Omega, \mathbb{P}) & \to & \ell_2^n \\ \omega & \mapsto & SU_{\varepsilon_1(\omega), \cdots, \varepsilon_n(\omega)}(x) \end{array} \right. \end{array} \right.$$

Pour tout $x \in \ell_\infty^n$, on a :

$$(\frac{1}{C} \|x\|_\infty)^2 \leq \|T(x)\|_{L^2(\mathbb{P})}^2 = \int_\Omega \|T(x)(\omega)\|_2^2 d\mathbb{P}(\omega) \leq (D \|x\|_\infty)^2$$

d'où, pour tout $x \in \ell_\infty^n$, $\frac{1}{C} \|x\|_\infty \leq \|T(x)\|_{L^2(\mathbb{P})} \leq D \|x\|_\infty$.
Or, d'après l'identité du parallélogramme itérée, comme les variables sont mutuellement indé-

pendantes, pour tout $x \in \ell_\infty^n$, on a :

$$\begin{aligned}
\left\| T\left(\sum_{i=1}^n x_i e_i\right) \right\|_{L^2(\mathbb{P})}^2 &= \mathbb{E}\left(\left\| \sum_{i=1}^n x_i \varepsilon_i S(e_i) \right\|_2^2 \right) \\
&= \sum_{(u_1,\cdots,u_n)\in\{-1,1\}^n} \mathbb{E}\left(\left\| \sum_{i=1}^n u_i x_i S(e_i) \right\|_2^2 \mathbb{1}_{(\varepsilon_1,\cdots,\varepsilon_n)=(u_1,\cdots,u_n)} \right) \\
&= \sum_{(u_1,\cdots,u_n)\in\{-1,1\}^n} \left\| \sum_{i=1}^n u_i x_i S(e_i) \right\|_2^2 \mathbb{P}(\varepsilon_1 = u_1)\cdots\mathbb{P}(\varepsilon_n = u_n) \\
&= \frac{1}{2^n} \sum_{(u_1,\cdots,u_n)\in\{-1,1\}^n} \left\| \sum_{i=1}^n u_i x_i S(e_i) \right\|_2^2 \\
&= \sum_{i=1}^n \|x_i S(e_i)\|_2^2 \\
&= \sum_{i=1}^n |x_i|^2 \|S(e_i)\|_2^2
\end{aligned}$$

On pose $x = \sum_{i=1}^n e_i$. On a alors :

$$D = D\|x\|_\infty \geq \|T(x)\|_{L^2(\mathbb{P})} = \left(\sum_{i=1}^n \|S(e_i)\|_2^2\right)^{\frac{1}{2}} \geq \left(\sum_{i=1}^n \left(\frac{1}{C}\times 1\right)^2\right)^{\frac{1}{2}} = \frac{1}{C}\sqrt{n}$$

d'où $\sqrt{n} \leq CD$.
On en déduit que pour tout $S \in GL(\ell_\infty^n, \ell_2^n)$, $\|S\|_{\ell_\infty^n \to \ell_2^n} \|S^{-1}\|_{\ell_2^n \to \ell_\infty^n} \geq \sqrt{n}$ d'où $d(\ell_\infty^n, \ell_2^n) \geq \sqrt{n}$.
Le théorème précédent assure alors que $d(\ell_\infty^n, \ell_2^n) = \sqrt{n}$. $\square$

Montrons désormais un autre théorème, indépendant du reste mais intéressant en lui-même, qui dit que dans un espace de Banach X, tout sous-espace vectoriel de dimension finie est complémenté dans X avec la définition suivante :

Définition 5. *Soit X un espace de Banach, E un sous-espace vectoriel de X.*
On dit que E est complémenté dans X si et seulement si il existe une projection linéaire continue $P : X \to E$.

Pour démontrer ce théorème, nous aurons besoin d'un résultat d'Auerbach :

Lemme 6 (Auerbach)**.** *Soit $(E, \|.\|)$ un $\mathbb{R}$-espace vectoriel normé de dimension n.*
Il existe une base de E, $(e_1, \cdots, e_n)$, vérifiant : pour tout $i \in [|1,n|]$, $\|e_i\| = 1$ telle que sa base duale, $(e_1^*, \cdots, e_n^*)$, vérifie également : pour tout $i \in [|1,n|]$, $\|e_i^*\|_{E^*} = 1$.

Démonstration. Il existe une base $\mathscr{B} = (b_1, \cdots, b_n)$ de E, et comme E est de dimension finie, sa sphère unité $\mathcal{S} = \{x \in E; \|x\| = 1\}$ est un compact et $\mathcal{S}^n$ est un compact de (E^n, N) où on a noté N une norme sur l'espace produit. On pose :

$$\Delta : \left\{ \begin{array}{lcl} E^n & \to & \mathbb{R}^+ \\ (x_1, \cdots, x_n) & \mapsto & |\det_{\mathscr{B}}(x_1, \ldots, x_n)| \end{array} \right.$$

On a les propriétés suivantes :

— Δ est non identiquement nulle car $\Delta(\mathscr{B}) = 1$;

— Δ est continue (car det est continue) ;

— Sur le compact $\mathcal{S}^n$, comme Δ est continue, Δ est bornée sur $\mathcal{S}^n$ et il existe $(e_1, \cdots, e_n) \in \mathcal{S}^n$ tel que pour tout $(x_1, \cdots, x_n) \in \mathcal{S}^n$, $\Delta((x_1, \cdots, x_n)) \leq \Delta((e_1, \cdots, e_n))$;

— $\Delta((e_1, \cdots, e_n)) > 0$ puisque Δ est non identiquement nulle, donc la famille $(e_1, \cdots, e_n)$ est libre, c'est donc une base de E.

On a alors que $(e_1, \cdots, e_n)$ est une base de E vérifiant pour tout $i \in [|1,n|]$, $\|e_i\| = 1$.
On note $(e_1^*, \cdots, e_n^*)$ la base duale de $(e_1, \cdots, e_n)$.
Soit $k \in [|1,n|]$, $x \in \mathcal{S}$. Comme $x = \sum_{i=1}^n e_i^*(x) e_i$, on a :

$$|e_k^*(x)|\Delta((e_1, \cdots, e_n)) = \Delta((e_1, \cdots, e_{k-1}, x, e_{k+1}, \cdots, e_n)) \leq \Delta((e_1, \cdots, e_n))$$

Donc pour tout $x \in \mathcal{S}$, $|e_k^*(x)| \leq 1$ *i.e* $\|e_k^*\|_{E^*} \leq 1$. Or $|e_k^*(e_k)| = 1$ d'où, pour tout $k \in [|1,n|]$, $\|e_k^*\|_{E^*} = 1$.
On a donc bien le résultat. □

Théorème 5. *Soit $(X, \|.\|)$ un espace vectoriel normé, E un sous-espace vectoriel de X de dimension n.*
Il existe une projection $P : X \to E$ linéaire continue vérifiant $\|P\| \leq n$.

Démonstration. D'après le lemme 6, il existe une base normée $(e_1, \cdots, e_n)$ de E dont la base duale, notée $(e_1^*, \cdots, e_n^*)$, est également normée.
Pour tout $i \in [|1,n|]$, $e_i^* \in E^*$ donc, d'après le théorème de Hahn-Banach, il existe $f_i \in X^*$ telle que $f_i|_E = e_i^*$ et $\|f_i\|_{X^*} = \|e_i^*\|_{E^*} = 1$.
On pose $P : \left\{ \begin{array}{lcl} X & \to & E \\ x & \mapsto & \sum_{i=1}^n f_i(x) e_i \end{array} \right.$.
P est une application linéaire vérifiant :

— Pour tout $x \in E$, $P(x) = \sum_{i=1}^n f_i(x)e_i = \sum_{i=1}^n e_i^*(x)e_i = x$;

— Pour tout $x \in X$, $\|P(x)\| = \|\sum_{i=1}^n f_i(x)e_i\| \leq \sum_{i=1}^n \|f_i\|_{X^*}\|x\|\|e_i\| \leq \sum_{i=1}^n \|x\| = n\|x\|$, donc P est continue et $\|P\| \leq n$.

On a donc bien le résultat recherché. $\square$

Remarque 2. *Un résultat de Kadets et Snobar, qui se démontre à l'aide du théorème 3, affirme que l'on peut même trouver une projection $P : X \to E$ continue vérifiant $\|P\| \leq \sqrt{n}$.*

2 Le phénomène de concentration de la mesure

Pour démontrer le théorème de Dvoretzky, nous allons avoir besoin d'un principe connu sous le nom de "Phénomène de concentration de la mesure", qui affirme qu'une fonction lipschitzienne sur la sphère euclidienne de $\mathbb{R}^n$ est de plus en plus proche de sa moyenne quand n augmente. Nous le démontrerons dans cette partie à l'aide de Gaussiennes, en suivant une approche dûe à Maurey et Pisier, plutôt qu'à partir de l'inégalité isopérimétrique de Lévy.
Pour cela, on considère $\mathbb{R}^n$ muni de sa norme euclidienne canonique $\|.\|_2$.
On note σ_n la mesure sur $\mathbb{R}^n$ normalisée sur la sphère $\mathcal{S}^{n-1} := \{\xi \in \mathbb{R}^n ; \|\xi\|_2^2 = \sum_{j=1}^n \xi_j^2 = 1\}$ (voir l'annexe pour plus d'informations).
Introduisons également un vecteur gaussien $G = (g_i)_{i=1}^n$ où les g_i sont des varaiables aléatoires réelles mutuellement indépendantes de loi normale centrée réduite définies sur un même espace de probabilité $(\Omega, \mathscr{F}, \mathbb{P})$. Le vecteur G admet alors pour densité la fonction :

$$\left\{ \begin{array}{lll} \mathbb{R}^n & \to & \mathbb{R} \\ \xi & \mapsto & \dfrac{1}{(2\pi)^{\frac{n}{2}}} \exp\left(-\frac{\|\xi\|_2^2}{2}\right) \end{array} \right.$$

Commençons tout d'abord par faire deux remarques.

Remarque 3. *On rappelle que si $f : \mathbb{R}^n \to E$ avec E un Banach, f est de classe $\mathcal{C}^1$ sur $\mathbb{R}^n$ si et seulement si elle admet des dérivées partielles continues sur $\mathbb{R}^n$.*

Remarque 4. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application de classe $\mathcal{C}^1$ et $1-$lipschitzienne. On note que, pour tout $x \in \mathbb{R}^n$, on a $\|\nabla f(x)\|_2 \leq 1$.*
En effet, l'application $\varphi : \left\{ \begin{array}{lll} \ell_2^n & \to & (\ell_2^n)^* \\ x & \mapsto & \left\{ \begin{array}{lll} \ell_2^n & \to & \mathbb{R} \\ y & \mapsto & \langle x, y \rangle \end{array} \right. \end{array} \right.$ est une isométrie linéaire et pour tout $x \in \mathbb{R}^n$, pour tout $h \in \mathbb{R}^n$, $df(x)h = \langle \nabla f(x), h \rangle$, donc $\|\nabla f(x)\|_2 = \|df(x)\|_{(\ell_2^n)^*} \leq 1$

Pour démontrer le phénomène de concentration de la mesure, nous allons avoir besoin des deux lemmes énoncés ci-dessous, qui nous permettront par la suite d'approcher uniformément une fonction lipschitzienne par des fonctions lipschitziennes de classe $\mathcal{C}^1$.

Lemme 7. *Soit $f : (\mathbb{R}^n, \|.\|) \to (\mathbb{R}, |.|)$ une application lipschitzienne, $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^+)$ une fonction positive à support compact vérifiant $\int_{\mathbb{R}^n} \varphi d\lambda = 1$. On pose :*

$$g : \left\{ \begin{array}{lll} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & \int_{\mathbb{R}^n} f(x-u)\varphi(u) d\lambda(u) = f * \varphi(x) \end{array} \right.$$

On a :
(i) g est lipschitzienne et $Lip(g) \leq Lip(f)$;
(ii) Pour tout $(x, u) \in \mathbb{R}^n \times \mathbb{R}^n$, la dérivée directionnelle $\partial_u g(x)$ existe et $\partial_u g : \mathbb{R}^n \to \mathbb{R}$ est continue.

Démonstration. Notons tout d'abord que g est bien définie.

(i) Soit $(x,y) \in \mathbb{R}^n \times \mathbb{R}^n$. On a :

$$\begin{aligned}
|g(x)-g(y)| &= \Big| \int_{\mathbb{R}^n} (f(x-u)-f(y-u))\varphi(u)d\lambda(u) \Big| \\
&\leq \int_{\mathbb{R}^n} |f(x-u)-f(y-u)|\varphi(u)d\lambda(u) \\
&\leq \int_{\mathbb{R}^n} Lip(f)\|x-u-(y-u)\|\varphi(u)d\lambda(u) \\
&=\leq Lip(f)\|x-y\| \int_{\mathbb{R}^n} \varphi d\lambda \\
&= Lip(f)\|x-y\|
\end{aligned}$$

Donc g est lipschitzienne et $Lip(g) \leq Lip(f)$.
(ii) Soit $(x,u) \in \mathbb{R}^n \times \mathbb{R}^n$. Pour tout $t \in \mathbb{R}^*$, on a :

$$\begin{aligned}
\frac{g(x+tu)-g(x)}{t} &= \frac{1}{t}\left(\int_{\mathbb{R}^n} f(x+tu-\xi)\varphi(\xi)d\lambda(\xi) - \int_{\mathbb{R}^n} f(x-\xi)\varphi(\xi)d\lambda(\xi) \right) \\
&= \frac{1}{t}\left(\int_{\mathbb{R}^n} f(x-y)\varphi(y+tu)d\lambda(y) - \int_{\mathbb{R}^n} f(x-\xi)\varphi(\xi)d\lambda(\xi) \right) \text{ en posant } y=\xi-tu \\
&= \frac{1}{t}\left(\int_{\mathbb{R}^n} f(x-\xi)(\varphi(\xi+tu)-\varphi(\xi))d\lambda(\xi) \right) \\
&= \int_{\mathbb{R}^n} f(x-\xi)\frac{\varphi(tu+\xi)-\varphi(\xi)}{t}d\lambda(\xi)
\end{aligned}$$

On pose $M = \sup\{|d\varphi(y)(u)|; y \in \mathbb{R}^n\}$ (qui est bien défini et fini car $d\varphi$ est à support compact), et $K = supp(\varphi) + [-1,1]u$. On note que, comme $supp(\varphi)$ et $[-1,1]$ sont compacts, K l'est également.
Or φ est de classe $\mathcal{C}^\infty$ à support compact donc, en particulier, lipschitzienne. On note $M := Lip(\varphi)\|u\|$.
Pour tout $\xi \in \mathbb{R}^n$, pour tout $t \in [-1,1]\backslash\{0\}$, on a :

$$\left| f(x-\xi)\frac{\varphi(tu+\xi)-\varphi(\xi)}{t} \right| \leq g(\xi) := M|f(x-\xi)|\mathbb{1}_K(\xi)$$

Or $\int_{\mathbb{R}^n} g d\lambda \leq M \sup\{|f(x+\xi)|, \xi \in K\}\lambda(K) < \infty$ donc, d'après le théorème de convergence dominée, la dérivée directionnelle $\partial_u g(x)$ existe et :

$$\partial_u g(x) = \int_{\mathbb{R}^n} \lim_{t\to 0} f(x-\xi)\frac{\varphi(tu+\xi)-\varphi(\xi)}{t} = \int_{\mathbb{R}^n} f(x-\xi)d\varphi(\xi)(u)d\lambda(\xi) = \int_{\mathbb{R}^n} f(\xi)d\varphi(x-\xi)(u)d\lambda(\xi)$$

Or $\xi \to d\varphi(\xi)(u)$ est continue à support compact donc, d'après le théorème de continuité dans l'intégrale, $\partial_u g : \mathbb{R}^n \to \mathbb{R}$ est continue. □

Lemme 8. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une application lipschitzienne, $\varphi \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^+)$ une fonction positive à support compact vérifiant $\int_{\mathbb{R}^n} \varphi d\lambda = 1$. Pour tout $k \in \mathbb{N}$, on pose :*

$$g_k : \left\{ \begin{array}{lll} \mathbb{R}^n & \to & \mathbb{R} \\ x & \mapsto & 2^{kn} \int_{\mathbb{R}^n} f(x-u)\varphi(2^k u)d\lambda(u) \end{array} \right.$$

On note $\kappa = \int_{\mathbb{R}^n} \|\xi\|_2 \varphi(\xi)d\lambda(\xi)$. On a :

$$\forall x \in \mathbb{R}^n, \forall k \in \mathbb{N}, |g_k(x)-f(x)| \leq \kappa 2^{-k} Lip(f)$$

En particulier, (g_k) converge uniformément vers la fonction f.

Démonstration. Soit $k \in \mathbb{N}, x \in \mathbb{R}^n$.
On note, en faisant un changement de variable, que :

$$2^{kn} \int_{\mathbb{R}^n} f(x)\varphi(2^k\xi)d\lambda(\xi) = 2^{kn} f(x) \int_{\mathbb{R}^n} \varphi(u)\frac{d\lambda(u)}{2^{nk}} = f(x)$$

On a alors :

$$\begin{aligned}
|g_k(x) - f(x)| &= 2^{kn}\Big|\int_{\mathbb{R}^n} (f(x-\xi) - f(x))\varphi(2^k\xi)d\lambda(\xi)\Big| \\
&\leq 2^{kn} \int_{\mathbb{R}^n} |f(x-\xi) - f(x)|\varphi(2^k\xi)d\lambda(\xi) \\
&\leq 2^{kn} Lip(f) \int_{\mathbb{R}^n} \|\xi\|_2\varphi(2^k\xi)d\lambda(\xi) \\
&= 2^{kn} Lip(f) \int_{\mathbb{R}^n} \frac{1}{2^k}\|u\|_2\varphi(u)\frac{d\lambda(u)}{2^{kn}} \\
&= 2^{-k} Lip(f)\kappa
\end{aligned}$$

D'où : $\forall x \in \mathbb{R}^n, \forall k \in \mathbb{N}, |g_k(x) - f(x)| \leq \kappa 2^{-k} Lip(f)$. $\square$

On rappelle l'inégalité de Jensen :

Lemme 9 (Inégalité de Jensen)**.** *Soit $(\Omega, \mathscr{F}, \mu)$ un espace de probabilité, $f \in L^1(\mu)$ une fonction à valeurs réelles, $(a,b) \in \overline{\mathbb{R}}$ tels que pour tout $x \in \Omega$, $a < f(x) < b$, $\varphi :]a,b[\to \mathbb{R}$ une fonction convexe.*
On a l'inégalité suivante :

$$\varphi\left(\int_\Omega f d\mu\right) \leq \int_\Omega \varphi \circ f d\mu$$

en posant $\int_\Omega \varphi \circ f d\mu = +\infty$ si $\varphi \circ f \notin L^1(\mu)$.

Montrons désormais un dernier résultat intermédiaire.

Théorème 6. *Soit $f : \mathbb{R}^n \to \mathbb{R}$ une fonction 1-lipschitzienne. Pour tout $t > 0$, on a :*

$$\mathbb{P}(|f(G) - \mathbb{E}(f(G))| > t) \leq 2\exp\left(\frac{-2t^2}{\pi^2}\right)$$

Démonstration. Quitte à travailler avec $g = f - \mathbb{E}(f(G))$, on peut supposer que $\mathbb{E}(f(G)) = 0$. Supposons tout d'abord que f est de classe $\mathcal{C}^1$. Comme f est 1-lipschitzienne, d'après la remarque 4, on sait que pour tout $\xi \in \mathbb{R}^n$, $\|\nabla f(\xi)\|_2 \leq 1$.
Soit G' une copie indépendante de G (*i.e* un vecteur gaussien défini sur le même espace de probabilité que G tel que les g_i et g'_i soient mutuellement indépendantes et ayant les mêmes propriétés que G) tel que (G, G') soit un vecteur gaussien de dimension $2n$.
Pour tout $\alpha \in \mathbb{R}$, on pose :

$$\begin{cases} G_\alpha := G\sin(\alpha) + G'\cos(\alpha) \\ G'_\alpha := G\cos(\alpha) - G'\sin(\alpha) \end{cases}$$

On note que comme (G, G') est invariant par transformations orthogonales, pour tout $\alpha \in \mathbb{R}$, (G_α, G'_α) suit la même loi que (G, G'). En effet, on a :

$$\mathscr{L}(G_\alpha, G'_\alpha) = \mathscr{L}\begin{pmatrix} g_1 \sin(\alpha) + g'_1 \cos(\alpha) \\ \vdots \\ g_n \sin(\alpha) + g'_n \cos(\alpha) \\ g_1 \cos(\alpha) - g'_1 \sin(\alpha) \\ \vdots \\ g_n \cos(\alpha) - g'_n \sin(\alpha) \end{pmatrix} = \mathscr{L}\begin{pmatrix} g_1 \sin(\alpha) + g'_1 \cos(\alpha) \\ g_1 \cos(\alpha) - g'_1 \sin(\alpha) \\ \vdots \\ g_n \sin(\alpha) + g'_n \cos(\alpha) \\ g_n \cos(\alpha) - g'_n \sin(\alpha) \end{pmatrix}$$

Or :

$$\begin{pmatrix} g_1 \sin(\alpha) + g'_1 \cos(\alpha) \\ g_1 \cos(\alpha) - g'_1 \sin(\alpha) \\ \vdots \\ \vdots \\ g_n \sin(\alpha) + g'_n \cos(\alpha) \\ g_n \cos(\alpha) - g'_n \sin(\alpha) \end{pmatrix} = \begin{pmatrix} \sin(\alpha) & \cos(\alpha) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \cos(\alpha) & -\sin(\alpha) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \sin(\alpha) & \cos(\alpha) & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cos(\alpha) & -\sin(\alpha) & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sin(\alpha) & \cos(\alpha) \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \begin{pmatrix} g_1 \\ g'_1 \\ \vdots \\ \vdots \\ g_n \\ g'_n \end{pmatrix}$$

et $\begin{pmatrix} \sin(\alpha) & \cos(\alpha) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ \cos(\alpha) & -\sin(\alpha) & 0 & \cdots & \cdots & \cdots & \cdots & 0 \\ 0 & 0 & \sin(\alpha) & \cos(\alpha) & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \cos(\alpha) & -\sin(\alpha) & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sin(\alpha) & \cos(\alpha) \\ 0 & \cdots & \cdots & \cdots & \cdots & 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix} \in O_{2n}(\mathbb{R})$ donc on a bien $\mathscr{L}(G_\alpha, G'_\alpha) = \mathscr{L}(G, G')$.

Soit $\lambda > 0$. G' est une copie de G donc $\mathbb{E}(f(G')) = 0$ et exp est convexe d'où, d'après l'inégalité de Jensen :

$$1 = \exp(\mathbb{E}(-\lambda f(G'))) \leq \mathbb{E}(\exp(-\lambda f(G')))$$

Et comme G' est une copie indépendante de G, $\exp(\lambda f(G))$ et $\exp(-\lambda f(G'))$ sont indépendants donc on a :

$$\mathbb{E}(\exp(\lambda f(G))) \leq \mathbb{E}(\exp(\lambda f(G)))\mathbb{E}(\exp(-\lambda f(G'))) = \mathbb{E}(\exp(\lambda(f(G) - f(G'))))$$

Or, pour tout $\omega \in \Omega$:

$$\begin{aligned} f(G(\omega)) - f(G'(\omega)) &= f(G_{\frac{\pi}{2}}(\omega)) - f(G_0(\omega)) \\ &= \int_0^{\frac{\pi}{2}} \frac{d}{d\alpha} f(G_\alpha(\omega)) d\alpha \\ &= \int_0^{\frac{\pi}{2}} df(G_\alpha(\omega))(G'_\alpha(\omega)) d\alpha \\ &= \int_0^{\frac{\pi}{2}} \langle \nabla f(G_\alpha(\omega)), G'_\alpha(\omega) \rangle d\alpha \end{aligned}$$

Donc, en appliquant de nouveau l'inégalité de Jensen avec la fonction exp et grâce au théorème de Fubini-Tonelli, on a :

$$\begin{aligned}
\mathbb{E}(\exp(\lambda(f(G)-f(G')))) &= \mathbb{E}\left(\exp\left(\int_0^{\frac{\pi}{2}} \lambda\frac{\pi}{2}\langle\nabla f(G_\alpha), G'_\alpha\rangle \frac{d\alpha}{\frac{\pi}{2}}\right)\right) \\
&\leq \int_0^{\frac{\pi}{2}} \mathbb{E}\left(\exp\left(\lambda\frac{\pi}{2}\langle\nabla f(G_\alpha), G'_\alpha\rangle\right)\right)\frac{d\alpha}{\frac{\pi}{2}} \\
&= \frac{2}{\pi}\int_0^{\frac{\pi}{2}} \mathbb{E}\left(\exp\left(\lambda\frac{\pi}{2}\langle\nabla f(G), G'\rangle\right)\right) d\alpha \\
&= \mathbb{E}\left(\exp\left(\lambda\frac{\pi}{2}\langle\nabla f(G), G'\rangle\right)\right) \\
&= \mathbb{E}(F(G))
\end{aligned}$$

où l'on a posé $F : \left\{\begin{array}{lcl} \mathbb{R}^n & \to & \mathbb{R} \\ \xi & \mapsto & \mathbb{E}\left(\exp\left(\lambda\langle\frac{\pi}{2}\nabla f(\xi), G'\rangle\right)\right)\end{array}\right.$

Or, pour tout $\xi \in \mathbb{R}^n$, on a :

$$\begin{aligned}
F(\xi) &= \mathbb{E}\left(\exp\left(\lambda\frac{\pi}{2}\sum_{i=1}^n \frac{\partial f}{\partial x_i}(\xi)g'_i\right)\right) \\
&= \mathbb{E}\left(\prod_{i=1}^n \exp\left(\lambda\frac{\pi}{2}\frac{\partial f}{\partial x_i}(\xi)g'_i\right)\right) \\
&= \prod_{i=1}^n \mathbb{E}\left(\exp\left(\lambda\frac{\pi}{2}\frac{\partial f}{\partial x_i}(\xi)g'_i\right)\right) \\
&= \prod_{i=1}^n \int_{\mathbb{R}} e^{\lambda\frac{\pi}{2}\frac{\partial f}{\partial x_i}(\xi)x}\frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}}d\lambda(x) \\
&= \prod_{i=1}^n \int_{\mathbb{R}} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x^2-2\lambda\frac{\pi}{2}\frac{\partial f}{\partial x_i}(\xi)x+(\lambda\frac{\pi}{2})^2(\frac{\partial f}{\partial x_i}(\xi))^2)}e^{\frac{1}{2}(\lambda\frac{\pi}{2})^2(\frac{\partial f}{\partial x_i}(\xi))^2}d\lambda(x) \\
&= \prod_{i=1}^n \exp\left(\frac{1}{2}\left(\lambda\frac{\pi}{2}\right)^2\left(\frac{\partial f}{\partial x_i}(\xi)\right)^2\right) \times 1 \\
&= \exp\left(\frac{1}{2}\left(\lambda\frac{\pi}{2}\right)^2 \|\nabla f(\xi)\|_2^2\right)
\end{aligned}$$

Donc, pour tout $\xi \in \mathbb{R}^n$, puisque $\|\nabla f(\xi)\|_2 \leq 1$, $F(\xi) \leq \exp\left(\frac{\lambda^2\pi^2}{8}\right)$, d'où :

$$\mathbb{E}(\exp(\lambda f(G))) \leq \mathbb{E}(\exp(\lambda(f(G)-f(G')))) \leq \mathbb{E}(F(G)) \leq \exp\left(\frac{\lambda^2\pi^2}{8}\right)\mathbb{E}(1) = \exp\left(\frac{\lambda^2\pi^2}{8}\right)$$

De plus, pour tout $t \in \mathbb{R}$, on a $\exp(|t|) \leq \exp(t) + \exp(-t)$ et, puisque $-f$ est 1-lipschitzienne et vérifie également $\mathbb{E}(-f(G)) = -\mathbb{E}(f(G)) = 0$, on en déduit que :

$$\mathbb{E}(\exp(\lambda|f(G)|)) \leq \mathbb{E}(\exp(\lambda f(G))) + \mathbb{E}(\exp(-\lambda f(G))) \leq 2\exp\left(\frac{\lambda^2\pi^2}{8}\right)$$

Soit $t > 0$. En appliquant l'inégalité de Markov, on a alors :

$$\mathbb{P}(|f(G)| > t) = \mathbb{P}(\exp(\lambda|f(G)|) > \exp(\lambda t)) \leq \frac{2\exp\left(\frac{\lambda^2\pi^2}{8}\right)}{\exp(\lambda t)} = 2\exp\left(\frac{\lambda^2\pi^2 - 8\lambda t}{8}\right)$$

Ceci étant vrai pour tout λ strictement positif, on a en particulier avec $\lambda = \frac{4t}{\pi^2}$:

$$\mathbb{P}(|f(G)| > t) \leq 2\exp\left(-\frac{2t^2}{\pi^2}\right)$$

D'où, pour tout $t > 0$, on a :

$$\mathbb{P}(|f(G) - \mathbb{E}(f(G))| > t) \leq 2\exp\left(\frac{-2t^2}{\pi^2}\right)$$

On ne suppose désormais plus que f est de classe $\mathcal{C}^1$. Soit $t > 0$.
D'après les lemmes 7 et 8, pour tout $k \in \mathbb{N}$, il existe $g_k : \mathbb{R}^n \to \mathbb{R}$ de classe $\mathcal{C}^1$ et 1-lipschitzienne telle que $\|f - g_k\|_\infty \leq \frac{1}{2^k}$.
Puisque, pour tout $k \in \mathbb{N}$, $\{|f(G)| > t\} \subset \{|g_k(G)| > t - \frac{1}{2^k}\} \cup \{|(f - g_k)(G)| > \frac{1}{2^k}\} = \{|g_k(G)| > t - \frac{1}{2^k}\} \cup \varnothing$, on a, pour tout $k \in \mathbb{N}$:

$$\mathbb{P}(|f(G)| > t) \leq \mathbb{P}\left(|g_k(G)| > t - \frac{1}{2^k}\right) \leq 2\exp\left(-\frac{2}{\pi^2}\left(t - \frac{1}{2^k}\right)^2\right)$$

On obtient alors le résultat par passage à la limite. □

Nous avons désormais tous les outils pour démontrer le phénomène de concentration de la mesure, que l'on doit à Milman. Il est à noter que les constantes, qui ne sont sûrement pas optimales, importent peu. Le point-clé, en effet, se trouve dans la rapidité avec laquelle le deuxième membre de l'inégalité ci-dessous tend vers 0, qui nous permet d'en déduire qu'en grandes dimensions, les fonctions lipschitziennes sont quasiment égales à leur moyenne.

Théorème 7 (Le phénomène de concentration de la mesure)**.** *Soit $f : \mathcal{S}^{n-1} \to \mathbb{R}$ une fonction 1-lipschitzienne. On pose $\overline{f} = \int_{\mathcal{S}^{n-1}} f d\sigma_n$. Pour tout $t > 0$, on a :*

$$\sigma_n(|f - \overline{f}| > t) \leq 4\exp\left(\frac{-nt^2}{128\pi^2}\right)$$

Démonstration. Quitte à travailler avec $g = f - \overline{f}$, on peut supposer que $\overline{f} = 0$. Alors, pour tout $x \in \mathcal{S}^{n-1}$, $|f(x)| \leq 2$. Le résultat est donc clair pour les $t > 2$.
On prolonge f sur $\mathbb{R}^n$ en posant pour tout $x \in \mathbb{R}^n \backslash \{0\}$, $f(x) = \|x\|_2 f\left(\frac{x}{\|x\|_2}\right)$ et $f(0) = 0$.
Pour tout $(x, y) \in (\mathbb{R}^n \backslash \{0\})^2$, on a :

$$\begin{aligned}
\left\|\frac{x}{\|x\|_2} - \frac{y}{\|y\|_2}\right\|_2 &= \left\|\frac{x\|y\|_2 - y\|x\|_2}{\|x\|_2\|y\|_2}\right\|_2 \\
&= \frac{\|x(\|y\|_2 - \|x\|_2) + \|x\|_2(x - y)\|_2}{\|x\|_2\|y\|_2} \\
&\leq \frac{|\|y\|_2 - \|x\|_2|}{\|y\|_2} + \frac{\|x - y\|_2}{\|y\|_2} \\
&\leq 2\frac{\|x - y\|_2}{\|y\|_2}
\end{aligned}$$

On en déduit que, pour tout $(x,y) \in (\mathbb{R}^n\backslash\{0\})^2$:

$$\begin{aligned}
|f(x)-f(y)| &= \left| \|x\|_2 f\left(\frac{x}{\|x\|_2}\right) - \|y\|_2 f\left(\frac{y}{\|y\|_2}\right) \right| \\
&= \left| \|y\|_2 \left(f\left(\frac{x}{\|x\|_2}\right) - f\left(\frac{y}{\|y\|_2}\right)\right) + (\|x\|_2 - \|y\|_2) f\left(\frac{x}{\|x\|_2}\right) \right| \\
&\leq \|y\|_2 \left| f\left(\frac{x}{\|x\|_2}\right) - f\left(\frac{y}{\|y\|_2}\right) \right| + \left| \|x\|_2 - \|y\|_2 \right| \left| f\left(\frac{x}{\|x\|_2}\right) \right| \\
&\leq \|y\|_2 \times 2 \frac{\|x-y\|_2}{\|y\|_2} + \|x-y\|_2 \times 2 \\
&\leq 4\|x-y\|_2
\end{aligned}$$

Donc le prolongement de f à $\mathbb{R}^n$ est 4-lipschitzien. On note également que $\mathbb{E}(f(G)) = 0$ puisque, comme $\|G\|_2$ et $\frac{G}{\|G\|_2}$ sont indépendants, on a, d'après le théorème de transfert :

$$\mathbb{E}(f(G)) = \mathbb{E}\left(\|G\|_2 f\left(\frac{G}{\|G\|_2}\right)\right) = \mathbb{E}(\|G\|_2)\overline{f} = 0$$

car $\mathscr{L}\left(\frac{G}{\|G\|_2}\right) = \sigma_n$ (voir l'annexe B pour la démonstration).
Soit $t \in]0,2]$. Il suffit de montrer que $\mathbb{P}\left(\left|f\left(\frac{G}{\|G\|_2}\right)\right| > t\right) \leq 4\exp\left(\frac{-nt^2}{128\pi^2}\right)$ puisque :

$$\mathbb{P}\left(\left|f\left(\frac{G}{\|G\|_2}\right)\right| > t\right) = \int_\Omega \mathbb{1}_{\left\{\left|f\left(\frac{G}{\|G\|_2}\right)\right|>t\right\}} d\mathbb{P} = \int_{\mathbb{R}^n} \mathbb{1}_{\{|f(x)|>t\}} d\sigma_n(x) = \sigma_n(|f|>t)$$

Notons que :

$$\|G\|_1^2 = \sum_{i=1}^n |g_i|^2 + 2\sum_{i=1}^{n-1}\sum_{j=i+1}^n |g_i||g_j| \leq \sum_{i=1}^n |g_i|^2 + \sum_{i=1}^{n-1}\sum_{j=i+1}^n (|g_i|^2+|g_j|^2) \leq n\|G\|_2^2$$

donc :

$$\frac{1}{2}\sqrt{n} < \sqrt{\frac{2n}{\pi}} = \frac{1}{\sqrt{n}}\mathbb{E}\left(\sum_{j=1}^n |g_j|\right) \leq \mathbb{E}(\|G\|_2)$$

car :

$$\begin{aligned}
\mathbb{E}(|g_1|) &= \int_{\mathbb{R}} |x| \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} d\lambda(x) \\
&= 2\int_{[0,+\infty[} x \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} d\lambda(x) \\
&= \sqrt{\frac{2}{\pi}} \left[-e^{-\frac{x^2}{2}}\right]_0^{+\infty} \\
&= \sqrt{\frac{2}{\pi}}
\end{aligned}$$

On en déduit que si $\|G\|_2 < \frac{1}{4}\sqrt{n}$, alors $\mathbb{E}(\|G\|_2) - \|G\|_2 > \frac{1}{2}\sqrt{n} - \frac{1}{4}\sqrt{n} = \frac{1}{4}\sqrt{n}$. D'où, d'après le théorème 6, comme $\|.\|_2$ est 1-lipschitzienne, on a :

$$\mathbb{P}\left(\|G\|_2 < \frac{1}{4}\sqrt{n}\right) \leq \mathbb{P}\left(|\|G\|_2 - \mathbb{E}(\|G\|_2)| > \frac{1}{4}\sqrt{n}\right) \leq 2\exp\left(-\frac{n}{8\pi^2}\right)$$

De plus, $\frac{1}{4}f$ est 1-lipschitzienne donc, toujours d'après le théorème 6, on a :

$$\begin{aligned}
\mathbb{P}\left(|f(G)| > \frac{t\sqrt{n}}{4}\right) &= \mathbb{P}\left(\left|\frac{1}{4}f(G) - \mathbb{E}\left(\frac{1}{4}f(G)\right)\right| > \frac{t\sqrt{n}}{16}\right) \\
&\leq 2\exp\left(-2\frac{t^2 n}{256\pi^2}\right) \\
&= 2\exp\left(-\frac{nt^2}{128\pi^2}\right)
\end{aligned}$$

Or :

$$\left\{\left|f\left(\frac{G}{\|G\|_2}\right)\right| > t\right\} = \left\{\frac{f(G)}{\|G\|_2} > t\right\} \subset \left\{\|G\|_2 < \frac{1}{4}\sqrt{n}\right\} \cup \left\{|f(G)| > t\frac{\sqrt{n}}{4}\right\}$$

Donc, puisque $t \leq 2$:

$$\begin{aligned}
\mathbb{P}\left(\left|f\left(\frac{G}{\|G\|_2}\right)\right| > t\right) &\leq \mathbb{P}\left(\|G\|_2 < \frac{1}{4}\sqrt{n}\right) + \mathbb{P}\left(|f(G)| > t\frac{\sqrt{n}}{4}\right) \\
&\leq 2\exp\left(-\frac{n}{8\pi^2}\right) + 2\exp\left(-\frac{nt^2}{128\pi^2}\right) \\
&\leq 4\exp\left(\frac{-nt^2}{128\pi^2}\right)
\end{aligned}$$

$\square$

3 Le théorème de Dvoretzky

Démontrons désormais dans cette troisième partie le théorème de Dvoretzky.
On considère $\mathbb{R}^n$ muni de sa norme euclidienne canonique, $\|.\|_2 = \|.\|_{l_2^n}$, et $\|.\|_F$ une seconde norme sur $\mathbb{R}^n$.
On pose $\theta_F = \int_{\mathcal{S}^{n-1}} \|\xi\|_F d\sigma_n(\xi)$.
D'après le théorème de transfert, on a alors :

$$\theta_F = \mathbb{E}\left(\left\|\frac{G}{\|G\|_2}\right\|_F\right) = \mathbb{E}\left(\frac{\|G\|_F}{\|G\|_2}\right)$$

Commençons d'abord par démontrer les deux lemmes suivants :

Lemme 10. *Soit $(Y_0, \|.\|)$ un espace euclidien de dimension m. Soit $\varepsilon > 0$.
Il existe un ε-réseau $\{x_1, \cdots, x_N\}$ pour l'ensemble $\{x \in Y_0; \|x\| = 1\}$ avec $N \leq \left(1 + \frac{2}{\epsilon}\right)^m$.*

Démonstration. Il existe un sous-ensemble maximal $\{x_1, \cdots, x_N\}$ de $A := \{x \in Y_0; \|x\| = 1\}$ vérifiant :

$$\forall (i,j) \in [|1, N|], i \neq j \Rightarrow \|x_i - x_j\| \geq \varepsilon$$

En effet, on peut le construire de la manière suivante, par récurrence : soit $y_1 \in A$, si pour tout $y \in A, \|y - y_1\| < \varepsilon$, alors $\{y_1\}$ convient. Sinon il existe $y_2 \in A$ tel que $\|y_1 - y_2\| \geq \varepsilon$. Supposons ainsi construits $y_1, \cdots, y_n$. Si pour tout $y \in A$, il existe $i \in [|1, n|]$ tel que $\|y - y_i\| < \varepsilon$, alors la famille $\{y_1, \cdots, y_n\}$ convient. Sinon il existe $y_{n+1} \in A$ tel que pour tout $i \in [|1, n|]$, $\|y_{n+1} - y_i\| \geq \varepsilon$. De plus, comme A est compact, on sait que ce processus est fini d'où l'existence de ce sous-ensemble maximal.
$\{x_1, \cdots, x_N\}$ est un ε-réseau de A puisque sinon, il existe $y \in A$ tel que pour tout $i \in [|1, N|]$, $\|x_i - y\| \geq \varepsilon$, ce qui contredit la maximalité de $\{x_1, \cdots, x_N\}$.
On note que les boules ouvertes $\left\{x \in Y_0; \|x - x_j\| < \frac{1}{2}\varepsilon\right\}$, $j \in [|1, N|]$ sont deux à deux disjointes. Soit $j \in [|1, N|]$, $x \in \left\{x \in Y_0; \|x - x_j\| < \frac{1}{2}\varepsilon\right\}$. On a :

$$\|x\| \leq \|x_j\| + \|x - x_j\| \leq 1 + \frac{1}{2}\varepsilon$$

donc, $\left\{x \in Y_0; \|x - x_j\| < \frac{1}{2}\varepsilon\right\} \subset \left(1 + \frac{1}{2}\varepsilon\right) B_{Y_0}$, B_{Y_0} désignant la boule unité fermée de Y_0. D'où :

$$\bigsqcup_{j=1}^{N} \left\{x \in Y_0; \|x - x_j\| < \frac{1}{2}\varepsilon\right\} \subset \left(1 + \frac{1}{2}\varepsilon\right) B_{Y_0}.$$

On a alors :

$$\begin{aligned} N\left(\frac{\varepsilon}{2}\right)^m vol(B_{Y_0}) &= vol\left(\bigsqcup_{j=1}^{N} \left\{x \in Y_0; \|x - x_j\| < \frac{1}{2}\varepsilon\right\}\right) \\ &\leq vol\left(\left(1 + \frac{1}{2}\varepsilon\right) B_{Y_0}\right) \\ &= \left(1 + \frac{\varepsilon}{2}\right)^m vol(B_{Y_0}) \end{aligned}$$

On a donc bien $N \leq \left(\frac{2}{\varepsilon} + 1\right)^m$. □

Lemme 11. *Soit $(E, \|.\|)$ un espace vectoriel normé de dimension finie, $(X, \|.\|_X)$ un espace vectoriel normé et $\{x_1, \cdots, x_N\}$ un ε-réseau de la sphère unité $\mathcal{S}_E := \{e \in E; \|e\| = 1\}$, $0 < \varepsilon < 1$. Soit $T : E \to X$ une application linéaire telle que pour tout $j \in [|1, N|]$, $1 - \varepsilon \leq \|T(x_j)\|_X \leq 1 + \varepsilon$.*
Pour tout $e \in E$, on a :

$$\left(\frac{1-3\varepsilon}{1-\varepsilon}\right)\|e\| \leq \|T(e)\|_X \leq \left(\frac{1+\varepsilon}{1-\varepsilon}\right)\|e\|$$

Démonstration. Soit $e \in E$. Par homogénéité, on peut supposer $\|e\| = 1$.
Comme $\{x_1, \cdots, x_N\}$ est un ε-réseau de $\mathcal{S}_E$, il existe $j \in [|1, n|]$ tel que $\|e - x_j\| \leq \varepsilon$. On a alors :

$$\|T(e)\|_X \leq \|T(e) - T(x_j)\|_X + \|T(x_j)\|_X \leq \|T(e - x_j)\|_X + (1+\varepsilon) \leq \|T\|\varepsilon + (1+\varepsilon)$$

Donc $\|T\| \leq \|T\|\varepsilon + (1+\varepsilon)$ d'où $\|T\| \leq \frac{1+\varepsilon}{1-\varepsilon}$.
De plus, on a :

$$\|T(e)\|_X = \|T(x_j) - T(x_j - e)\|_X \geq \|T(x_j)\|_X - \|T(x_j - e)\|_X \geq (1-\varepsilon) - \|T\|\varepsilon$$

Or :

$$(1-\varepsilon) - \|T\|\varepsilon \geq \frac{(1-\varepsilon)^2 - \varepsilon - \varepsilon^2}{1-\varepsilon} = \frac{1-3\varepsilon}{1-\varepsilon}$$

D'où, pour tout $e \in \mathcal{S}_E$, on a :

$$\left(\frac{1-3\varepsilon}{1-\varepsilon}\right) \leq \|T(e)\|_X \leq \left(\frac{1+\varepsilon}{1-\varepsilon}\right)$$

On a donc bien le résultat recherché. $\square$

Nous aurons besoin, dans la démonstration du théorème ci-dessous, de la formule suivante :

$$\forall \xi_0 \in \mathcal{S}^{n-1}, \forall f \in \mathcal{C}(\mathcal{S}^{n-1}), \int_{\mathcal{S}^{n-1}} f d\sigma_n = \int_{O_n} f(U\xi_0) d\mu(U)$$

où μ est la mesure de Haar (normalisée) de O_n (voir l'annexe pour sa démonstration).

Théorème 8. *Soit $\|.\|_F$ une norme sur $\mathbb{R}^n$ telle que pour tout $x \in \mathbb{R}^n$, $\|x\|_F \leq \|x\|_2$.*
(a) Soit $0 < \varepsilon < \frac{1}{3}$. Il existe une sous-espace X_0 de $\mathbb{R}^n$ de dimension k tel que pour tout $x \in X_0$,

$$(1-\varepsilon)\theta_F\|x\|_2 \leq \|x\|_F \leq (1+\varepsilon)\theta_F\|x\|_2$$

si $k \leq C\theta_F^2 n \frac{\varepsilon^2}{|\ln(\varepsilon)|}$, où C est une constante universelle.
(b) Soit $0 < \varepsilon < \frac{1}{3}$. Il existe un sous-espace X_0 de $\mathbb{R}^n$ vérifiant $\dim(X_0) \geq k$ et $d_{X_0} \leq 1 + \varepsilon$ si $k \leq C_1\theta_F^2 n \frac{\varepsilon^2}{|\ln(\varepsilon)|}$, où C_1 est une constante universelle, et $d_{X_0} := d(X_0, \ell_2^n)$.

Démonstration. (a) Soit $Y_0 \subset \mathbb{R}^n$ un sous-espace vectoriel de dimension k. $(Y_0, \|.\|_2)$ est un espace euclidien de dimension k donc, d'après le lemme 10, il existe un $\frac{\varepsilon}{3}$-réseau $\{x_1, \cdots, x_N\}$ de $\{x \in Y_0; \|x\|_2 = 1\}$ avec $N \leq \left(1 + \frac{6}{\varepsilon}\right)^k$.
On note O_n le groupe orthogonal de degré n sur $\mathbb{R}$ et μ sa mesure de Haar normalisée. On pose :

$$A := \left\{U \in O_n; \forall j \in [|1, N|], \left(1 - \frac{\varepsilon}{3}\right)\theta_F \leq \|Ux_j\|_F \leq \left(1 + \frac{\varepsilon}{3}\right)\theta_F\right\}$$

Montrons que A est non vide.
On a :

$$A^c = \left\{ U \in O_n ; \exists j \in [|1, N|]; |\|Ux_j\|_F - \theta_F| > \frac{\varepsilon}{3}\theta_F \right\}$$
$$= \bigcup_{j=1}^{N} \left\{ U \in O_n ; |\|Ux_j\|_F - \theta_F| > \frac{\varepsilon}{3}\theta_F \right\}$$

Donc :

$$\mu(A^c) \leq \sum_{j=1}^{N} \mu \left(\left\{ U \in O_n ; |\|Ux_j\|_F - \theta_F| > \frac{\varepsilon}{3}\theta_F \right\} \right)$$
$$\leq \sum_{j=1}^{N} \int_{O_n} \mathbb{1}_{\left\{ |\|Ux_j\|_F - \theta_F| > \frac{\varepsilon}{3}\theta_F \right\}}(U) d\mu(U)$$
$$= \sum_{j=1}^{N} \int_{\mathcal{S}^{n-1}} \mathbb{1}_{\left\{ |\|\xi\|_F - \theta_F| > \frac{\varepsilon}{3}\theta_F \right\}}(\xi) d\sigma_n(\xi)$$
$$= N\sigma_n \left(\left\{ \xi \in \mathcal{S}^{n-1} ; |\|\xi\|_F - \theta_F| > \frac{\varepsilon}{3}\theta_F \right\} \right)$$
$$= N\sigma_n \left(|f - \overline{f}| > \frac{\varepsilon}{3} \right) \text{ avec } f = \|.\|_F$$
$$\leq 4N \exp \left(-\frac{n\varepsilon^2\theta_F^2}{9 \times 128\pi^2} \right)$$
$$= 4N \exp \left(-\frac{n\varepsilon^2\theta_F^2}{1152\pi^2} \right)$$

d'après le théorème 7 car $\|.\|_F$ est 1-lipschitzienne sur $\mathcal{S}^{n-1}$, que les mesures μ et σ_n sont régulières et que $\theta_F = \int_{\mathcal{S}^{n-1}} \|\xi\|_F d\sigma_n(\xi)$.
Or $k \geq 1$ donc on a :

$$4N \leq 4^k N \leq 4^k \left(\frac{\varepsilon + 6}{\varepsilon} \right)^k < 4^k \left(\frac{7}{\varepsilon} \right)^k = \exp(k(\ln(28) - \ln(\varepsilon)))$$

D'où :

$$\mu(A^c) < \exp \left(k(\ln(28) - \ln(\varepsilon)) - \frac{n\varepsilon^2\theta_F^2}{1152\pi^2} \right)$$

On en déduit que si $k \leq \frac{n\varepsilon^2\theta_F^2}{1152\pi^2(\ln(28)-\ln(\varepsilon))}$, alors $\mu(A^c) < 1$, donc $\mu(A) = \mu(O_n) - \mu(A^c) > 0$ d'où $A \neq \varnothing$. Comme $A \neq \varnothing$, il existe $U \in A$. U est tel que pour tout $j \in [|1, N|]$, on a :

$$\left(1 - \frac{\varepsilon}{3}\right) \theta_F \leq \|Ux_j\|_F \leq \left(1 + \frac{\varepsilon}{3}\right) \theta_F$$

donc, d'après le lemme 11, comme $U : (\mathbb{R}^n, \|.\|_2) \to (\mathbb{R}^n, \|.\|_F)$ est linéaire,pour tout $x \in Y_0$, on a :

$$\left(\frac{1 - \varepsilon}{1 - \frac{\varepsilon}{3}} \right) \theta_F \|x\|_2 \leq \|Ux\|_F \leq \left(\frac{1 + \frac{\varepsilon}{3}}{1 - \frac{\varepsilon}{3}} \right) \theta_F \|x\|_2$$

Or $\varepsilon^2 \leq \varepsilon$ donc $3 + \varepsilon \leq 3 + 3\varepsilon - \varepsilon - \varepsilon^2 = (3 - \varepsilon)(1 + \varepsilon)$, d'où, pour tout $x \in Y_0$:

$$(1 - \varepsilon)\theta_F \|x\|_2 \leq \left(\frac{1 - \varepsilon}{1 - \frac{\varepsilon}{3}} \right) \theta_F \|x\|_2 \leq \|Ux\|_F \leq (1 + \varepsilon)\theta_F \|x\|_2$$

On pose alors $X_0 = U(Y_0)$. On a $\dim(X_0) = k$, et pour tout $y = U(x) \in X_0$, on a :

$$(1-\varepsilon)\theta_F \|y\|_2 \leq \|y\|_F \leq (1+\varepsilon)\theta_F \|y\|_2$$

car $\|y\|_2 = \|Ux\|_2 = \|x\|_2$ puisque $U \in O_n$.
On a alors le résultat en posant $C = \frac{\ln(3)}{1152\pi^2(\ln(28)+\ln(3))}$, car on veut que $C \leq \frac{|\ln(\varepsilon)|}{1152(\ln(28)-\ln(\varepsilon))}$ en sachant que la fonction $f : \left\{ \begin{array}{lll}]0, \frac{1}{3}[& \to & \mathbb{R} \\ x & \mapsto & \frac{|\ln(x)|}{\ln(28)-\ln(x)} \end{array} \right.$ est strictement décroissante, et que $f\left(\frac{1}{3}\right) = \frac{\ln(3)}{\ln(28)+\ln(3)}$
(b) Soit $0 < \varepsilon < \frac{1}{3}$. On applique (a) avec $\frac{\varepsilon}{3}$ au lieu de ε. Soit $k \leq \frac{C}{9}\theta_F^2 n \frac{\varepsilon^2}{\ln(3)-\ln(\varepsilon)}$.
D'après (a), il existe un sous-espace X_0 de $\mathbb{R}^n$ de dimension k tel que pour tout $x \in X_0$, $\left(1-\frac{\varepsilon}{3}\right)\theta_F\|x\|_2 \leq \|x\|_F \leq \left(1+\frac{\varepsilon}{3}\right)\theta_F\|x\|_2$.
On pose $i : \left\{ \begin{array}{lll} (\mathbb{R}^n, \|.\|_F) & \to & \ell_2^n \\ x & \mapsto & x \end{array} \right.$.
On a alors :

$$d_{X_0} \leq \|i\|\|i^{-1}\| \leq \frac{1}{\left(1-\frac{\varepsilon}{3}\right)\theta_F} \times \left(1+\frac{\varepsilon}{3}\right)\theta_F = \frac{1+\frac{\varepsilon}{3}}{1-\frac{\varepsilon}{3}} \leq 1+\varepsilon$$

On a donc bien le résultat avec $C_1 = \frac{C}{18}$, car on veut que $C_1 \leq \frac{C|\ln(\varepsilon)|}{9(\ln(3)-\ln(\varepsilon))}$ en sachant que la fonction $g : \left\{ \begin{array}{lll}]0, \frac{1}{3}[& \to & \mathbb{R} \\ x & \mapsto & \frac{|\ln(x)|}{\ln(3)-\ln(x)} \end{array} \right.$ est strictement décroissante, et que $g\left(\frac{1}{3}\right) = \frac{\ln(3)}{2\ln(3)} = \frac{1}{2}$. □

Dans le théorème précédent, la seule information que nous avons sur θ_F est qu'il est compris entre 0 et 1. Cependant, pour pouvoir appliquer ce résultat dans la démonstration du théorème de Dvoretzky, il faut que θ_F soit d'ordre beaucoup plus grand que $n^{-\frac{1}{2}}$.
Le lemme suivant nous donne une première minoration de θ_F en fonction de $n^{-\frac{1}{2}}$:

Lemme 12. *Soit $\|.\|_F$ une norme sur $\mathbb{R}^n$ telle que pour tout $x \in \mathbb{R}^n$, $\|x\|_F \leq \|x\|_2$.*
(a) $\frac{1}{\|I\|_{F\to\ell_2^n}} \leq \theta_F$;
(b) Si G est un vecteur gaussien centré réduit de dimension n, alors $\theta_F \geq n^{-\frac{1}{2}}\mathbb{E}(\|G\|_F)$.

Démonstration. (a) On a :

$$1 = \sigma_n(\mathcal{S}^{n-1}) = \int_{\mathcal{S}^{n-1}} \|\xi\|_2 \times \frac{\|\xi\|_F}{\|\xi\|_F} d\sigma_n(\xi) \leq \int_{\mathcal{S}^{n-1}} \|I\|_{F\to\ell_2^n}\|\xi\|_F d\sigma_n(\xi) = \|I\|_{F\to\ell_2^n}\theta_F$$

D'où le résultat.
(b) Comme les variables aléatoires $\frac{G}{\|G\|_2}$ et $\|G\|_2$ sont indépendantes, on a :

$$\mathbb{E}(\|G\|_2)\theta_F = \mathbb{E}(\|G\|_2)\mathbb{E}\left(\left\|\frac{G}{\|G\|_2}\right\|_F\right) = \mathbb{E}\left(\|G\|_2\frac{\|G\|_F}{\|G\|_2}\right) = \mathbb{E}(\|G\|_F)$$

De plus, la fonction $x \to x^2$ est convexe donc $\mathbb{E}(\|G\|_2)^2 \leq \mathbb{E}(\|G\|_2^2)$ d'où :

$$\mathbb{E}(\|G\|_2) \leq (\mathbb{E}(\|G\|_2^2))^{\frac{1}{2}} = \mathbb{E}(g_1^2 + \cdots + g_n^2)^{\frac{1}{2}} = \left(\sum_{i=1}^{n} \mathbb{E}(g_i^2)\right)^{\frac{1}{2}} = n^{\frac{1}{2}}$$

car, pour tout $i \in [|1, n|]$, $\mathbb{E}(g_i^2) = Var(g_i) + \mathbb{E}(g_i)^2 = 1 + 0^2 = 1$.
Donc $\mathbb{E}(\|G\|_F) \leq n^{\frac{1}{2}}\theta_F$, ce qui donne le résultat. □

Pour pouvoir démontrer le théorème de Dvoretzky, qui s'applique à tout espace vectoriel normé $(X, \|.\|_X)$ de dimension n, à l'aide du théorème 8, il faut munir X d'une norme euclidienne. Or la norme induite sur X par son ellipsoïde de John est euclidienne, on pourrait donc utiliser cette dernière. Cependant, la meilleure minoration de θ_X dont on dispose, grâce au théorème 4 et au lemme 12, est :

$$\theta_X \geq \frac{1}{\|I\|_{X \to \mathcal{E}_X}} \geq n^{-\frac{1}{2}}$$

ce qui ne suffit pas d'après ce que l'on a dit un peu plus haut.
C'est pourquoi on commence par utiliser l'ellipsoïde de John pour ensuite se ramener à un espace plus petit, à l'aide de la proposition suivante :

Proposition 4 (Lemme de Dvoretzky-Rogers)**.** *Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n. Notons $\|.\|_{\mathcal{E}_X}$ la norme induite sur X par l'ellipsoïde de John.*
Il existe une base orthonormale $(e_1, \cdots, e_n)$ de $(X, \|.\|_{\mathcal{E}_X})$ vérifiant :

$$\forall j \in [|1, n|], \|e_j\|_X \geq 2^{-\frac{n}{n-j+1}}$$

En particulier, pour tout $j \leq \frac{n}{2} + 1$, on a : $\|e_j\|_X \geq \frac{1}{4}$.

Démonstration. On rappelle que l'ellipsoïde de John de X est l'ellipsoïde de volume maximal contenu dans B_X, la boule unité fermée de X.
Construisons la base $(e_1, \cdots, e_n)$ par récurrence.
D'après le lemme 4, il existe $e_1 \in X$ tel que $\|e_1\|_{\mathcal{E}_X} = \|e_1\|_X = 1$. Supposons ensuite construits $(e_1, \cdots, e_{j-1})$, $2 \leq j \leq n$ orthonormés vérifiant la condition de l'énoncé.
On pose :

$$t_j := \max\{\|x\|_X ; \|x\|_{\mathcal{E}_X} = 1, \forall 1 \leq i \leq j-1, \ \langle x, e_i \rangle = 0\}$$

t_j est bien défini car l'ensemble $A_j := \{x \in X; \|x\|_{\mathcal{E}_X} = 1, \forall 1 \leq i \leq j-1, \ \langle x, e_i \rangle = 0\}$ est compact (fermé et borné dans X de dimension finie) et que $\|.\|_X$ est continue. Par conséquent, il existe $e_j \in A_j$ tel que $\|e_j\|_X = t_j$.
Montrons que la famille $(e_1, \cdots, e_n)$ ainsi construite convient.
On note que c'est une famille orthonormale par construction et que c'est une base de X puisque c'est une famille libre de n vecteurs de X, supposé de dimension n. De plus, pour tout $j \in [|1, n|]$, pour tout $x \in Vect\{e_j, \cdots, e_n\}$, $\|x\|_X \leq t_j \|x\|_{\mathcal{E}_X}$ (car s'il existait $x \in Vect\{e_j, \cdots, e_n\}$ tel que $\|x\|_X > t_j \|x\|_{\mathcal{E}_X}$, $\frac{x}{\|x\|_{\mathcal{E}_X}}$ contredirait la définition de t_j).
Soit $j \in [|1, n|]$, $a, b > 0$. L'application définie par :

$$\left\{ \begin{array}{lll} X \times X & \to & \mathbb{R} \\ (x, y) & \mapsto & a^{-2} \sum_{i=1}^{j-1} \langle x, e_i \rangle \langle y, e_i \rangle + b^{-2} \sum_{i=j}^{n} \langle x, e_i \rangle \langle y, e_i \rangle \end{array} \right.$$

est un produit scalaire sur X dont on note $\mathcal{E}_{a,b}$ l'ellipsoïde associée :

$$\mathcal{E}_{a,b} = \left\{ x \in X; a^{-2} \sum_{i=1}^{j-1} |\langle x, e_i \rangle|^2 + + b^{-2} \sum_{i=j}^{n} |\langle x, e_i \rangle|^2 \leq 1 \right\}$$

Montrons que $\mathcal{E}_{a,b} \subset B_X$ si $a + bt_j \leq 1$.
Soit $x = \sum_{i=1}^{n} \langle x, e_i \rangle e_i \in \mathcal{E}_{a,b}$. On note en particulier que :

$$\left\| \sum_{i=1}^{j-1} \langle x, e_i \rangle e_i \right\|_X \leq \left\| \sum_{i=1}^{j-1} \langle x, e_i \rangle e_i \right\|_{\mathcal{E}_X} = \left(\sum_{i=1}^{j-1} |\langle x, e_i \rangle|^2 \right)^{\frac{1}{2}} \leq (a^2)^{\frac{1}{2}} = a$$

et que :

$$\left\|\sum_{i=j}^{n}\langle x,e_i\rangle e_i\right\|_X \leq t_j\left\|\sum_{i=j}^{n}\langle x,e_i\rangle e_i\right\|_{\mathcal{E}_X} \leq t_j\left(\sum_{i=j}^{n}|\langle x,e_i\rangle|^2\right)^{\frac{1}{2}} \leq t_j(b^2)^{\frac{1}{2}} = t_j b$$

donc on a :

$$\|x\|_X \leq \left\|\sum_{i=1}^{j-1}\langle x,e_i\rangle e_i\right\|_X + \left\|\sum_{i=j}^{n}\langle x,e_i\rangle e_i\right\|_X \leq a + bt_j$$

Ce qui assure que $\mathcal{E}_{a,b} \subset B_X$ si $a + bt_j \leq 1$.
Montrons que le volume de $\mathcal{E}_{a,b}$ vaut $a^{j-1}b^{n-j+1}$ fois le volume de $\mathcal{E}_X$.
D'après la proposition 1, il existe une application linéaire $S : \ell_2^n \to X$ telle que $\mathcal{E}_X = S(B_{\ell_2^n})$, et on note que l'application $T \in \mathscr{L}(X)$ définie par :

$$T : \left\{\begin{array}{lll} X & \to & X \\ (x_1,\cdots,x_n) & \mapsto & (ax_1,\cdots,ax_{j-1},bx_j,\cdots,bx_n) \end{array}\right.$$

est un isomorphisme vérifiant $T(\mathcal{E}_X) = \mathcal{E}_{a,b}$. On a alors :

$$\frac{vol_S(\mathcal{E}_{a,b})}{vol_S(\mathcal{E}_X)} = |\det(T)| = a^{j-1}b^{n-j+1}$$

d'après ce qui suit la proposition 1 et la démonstration du théorème 1.
On pose alors $b = (2t_j)^{-1} > 0$. On note que $(1 - bt_j) > 0$ et que $\mathcal{E}_{1-bt_j,b} \subset B_X$ donc :

$$2^{-n}t_j^{-(n-j+1)} = \frac{1}{2^{j-1}}\frac{1}{2^{n-j+1}t_j^{n-j+1}} = (1-bt_j)^{j-1}b^{n-j-1} \leq 1.$$

On en déduit qu'on a bien trouvé une base orthonormale $(e_1,\cdots,e_n)$ de $(X,\|.\|_{\mathcal{E}_X})$ telle que, pour tout $j \in [|1,n|]$, on a :

$$\|e_j\|_X = t_j \geq 2^{\frac{-n}{n-j+1}}$$

□

Nous aurons besoin d'un dernier lemme, donnant une minoration de la moyenne du maximum de m gaussiennes :

Lemme 13. *Il existe une constante universelle $C_3 > 0$ telle que pour tout vecteur gaussien $G = (g_1,\cdots,g_m)$ de dimension $m \geq 2$ dont les composantes sont indépendantes, centrées, réduites, on ait :*

$$\mathbb{E}(\|G\|_\infty) \geq C_3(\ln(m))^{\frac{1}{2}}$$

Démonstration. Soit g une gaussienne centrée réduite, $s > 0$. On a :

$$\begin{aligned}
\mathbb{P}(|g| > s) &= \mathbb{P}(g < -s) + \mathbb{P}(g > s) \\
&= \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{-s} e^{-\frac{1}{2}\xi^2}d\xi + \frac{1}{\sqrt{2\pi}}\int_{s}^{+\infty} e^{-\frac{1}{2}\xi^2}d\xi \\
&= \frac{2}{\sqrt{2\pi}}\int_{s}^{+\infty} e^{-\frac{1}{2}\xi^2}d\xi \\
&= \sqrt{\frac{2}{\pi}}\int_{s}^{+\infty} e^{-\frac{1}{2}\xi^2}d\xi \\
&\geq \sqrt{\frac{2}{\pi}}se^{-2s^2}
\end{aligned}$$

car la fonction $\varphi : s \mapsto \int_s^{+\infty} e^{-\frac{s^2}{2}} - se^{-2s^2}$ vérifie $\varphi(0) \geq 0$ et, pour tout $s > 0$, $\varphi'(s) = e^{-\frac{s^2}{2}}((4s^2-1)e^{-\frac{3s^2}{2}} + 1) \geq 0$.
On en déduit que, pour tout $s > 0$:

$$\begin{aligned}
\mathbb{P}(\|G\|_\infty > s) &= 1 - \mathbb{P}(\|G\|_\infty \leq s) \\
&= 1 - \mathbb{P}\left(\bigcap_{i=1}^{m}\{|g_i| \leq s\}\right) \\
&= 1 - \prod_{i=1}^{m} \mathbb{P}(|g_i| \leq s) \text{ car les variables sont indépendantes} \\
&= 1 - \prod_{i=1}^{m}(1 - \mathbb{P}(|g_i| > s)) \\
&= 1 - (1 - \mathbb{P}(|g| > s))^m \\
&\geq 1 - \left(1 - \sqrt{\frac{2}{\pi}} se^{-2s^2}\right)^m
\end{aligned}$$

Soit $t > 0$. Comme $m \geq 2$, $t(\ln(m))^{\frac{1}{2}} > 0$, donc :

$$\mathbb{P}\left(\|G\|_\infty > t(\ln(m))^{\frac{1}{2}}\right) \geq 1 - \left(1 - \sqrt{\frac{2}{\pi}} t(\ln(m))^{\frac{1}{2}} e^{-2t^2\ln(m)}\right)^m \geq 1 - d(t)$$

où l'on a posé :

$$d(t) := \sup_{m \geq 2}\left(1 - \sqrt{\frac{2}{\pi}} t(\ln(m))^{\frac{1}{2}} m^{-2t^2}\right)^m$$

Avec $t = \frac{1}{2\sqrt{2}}$, comme $m \mapsto \left(1 - \sqrt{\frac{2}{\pi}} t(\ln(m))^{\frac{1}{2}} m^{-2t^2}\right)^m$ décroît vers 0 quand m tend vers l'infini, on voit que $0 < d(t) < 1$.
On en déduit, d'après l'inégalité de Markov, qu'on a :

$$\mathbb{E}(\|G\|_\infty) \geq \frac{1}{2\sqrt{2}}(\ln(m))^{\frac{1}{2}}\mathbb{P}\left(\|G\|_\infty > \frac{1}{2\sqrt{2}}(\ln(m))^{\frac{1}{2}}\right) \geq \left(1 - d\left(\frac{1}{2\sqrt{2}}\right)\right)\frac{1}{2\sqrt{2}}(\ln(m))^{\frac{1}{2}}$$

On a alors le résultat avec $C_3 = \frac{1}{2\sqrt{2}}\left(1 - d\left(\frac{1}{2\sqrt{2}}\right)\right)$. □

Nous avons désormais tous les outils pour démontrer le théorème de Dvoretzky.

Théorème 9 (Théorème de Dvoretzky)**.** *Il existe une constante universelle $C_4 > 0$ vérifiant la propriété suivante : Si $(X, \|.\|_X)$ est un espace vectoriel normé de dimension n et $0 < \varepsilon < \frac{1}{3}$, alors X admet un sous-espace vectoriel X_0 de dimension k tel que $d_{X_0} \leq 1 + \varepsilon$ dès que*

$$k \leq C_4 \ln(n)\frac{\varepsilon^2}{|\ln(\varepsilon)|}.$$

Démonstration. On pose $C_4 = \min\left\{7, \frac{C_1C_3^2}{16}\right\} > 0$. Soit $(X, \|.\|_X)$ un espace vectoriel normé de dimension n, $0 < \varepsilon < \frac{1}{3}$.

Notons $\|.\|_{\mathcal{E}_X}$ la norme induite sur X par son ellipsoïde de John.
Pour tout $n \in [|1,4|]$, on a :

$$C_4 \ln(n) \frac{\varepsilon^2}{|\ln(\varepsilon)|} \leq 7\ln(4) \frac{1}{9\ln(3)} < 1$$

donc le théorème est vrai si $n \in [|1,4|]$.
Supposons alors $n \geq 5$. D'après le lemme de Dvoretzky-Rogers, il existe un sous-espace Y de X tel que $m = \dim Y \geq \frac{n}{2}$ (en particulier $m \geq 3$), et tel que $(Y, \|.\|_{\mathcal{E}_X})$ admet une base orthonormée $(e_1, \cdots, e_m)$ vérifiant : pour tout $j \in [|1,m|]$, $\|e_j\|_X \geq \frac{1}{4}$.
Soit $G = (g_1, \cdots, g_m)$ un vecteur gaussien centré de dimension m de matrice de corrélation I_m défini sur un espace probabilisé $(\Omega, \mathscr{F}, \mathbb{P}_G)$. Soit $(\varepsilon_1, \cdots, \varepsilon_m)$ un vecteur aléatoire de Rademacher indépendant de G défini sur un espace probabilisé $(\Omega', \mathscr{F}', \mathbb{P}_\varepsilon)$.
On pose $G_0 = \sum_{j=1}^m g_j e_j$.
Montrons que pour tout $j \in [|1,m|]$, pour tout $x = \sum_{i=1}^m x_i e_i \in Y$, on a :

$$\|x_j e_j\|_X \leq \mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i x_i e_j\right\|_X\right) = \frac{1}{2^m} \sum_{\alpha \in \{-1,1\}^m} \left\|\sum_{i=1}^m \alpha_i x_i e_i\right\|_X .$$

Par symétrie, il suffit de le montrer pour $j = m$. Soit $x = \sum_{i=1}^m x_i e_i \in Y$. On a :

$$\begin{aligned}
\|x_m e_m\| &= \frac{1}{2^{m-1}} \left\| \sum_{\alpha \in \{-1,1\}^{m-1}} \left(\sum_{j=1}^{m-1} \alpha_j x_j e_j + x_m e_m \right) \right\|_X \\
&= \frac{1}{2^m} \left\| \sum_{\alpha \in \{-1,1\}^{m-1}} \left(\sum_{j=1}^{m-1} \alpha_j x_j e_j + x_m e_m \right) \right\|_X + \frac{1}{2^m} \left\| \sum_{\alpha \in \{-1,1\}^{m-1}} \left(\sum_{j=1}^{m-1} -\alpha_j x_j e_j - x_m e_m \right) \right\|_X \\
&= \frac{1}{2^m} \left\| \sum_{\alpha \in \{-1,1\}^{m-1}} \left(\sum_{j=1}^{m-1} \alpha_j x_j e_j + x_m e_m \right) \right\|_X + \frac{1}{2^m} \left\| \sum_{\alpha \in \{-1,1\}^{m-1}} \left(\sum_{j=1}^{m-1} \alpha_j x_j e_j - x_m e_m \right) \right\|_X \\
&\leq \frac{1}{2^m} \sum_{\alpha \in \{-1,1\}^{m-1}} \left(\left\| \sum_{j=1}^{m-1} \alpha_j x_j e_j + x_m e_m \right\|_X + \left\| \sum_{j=1}^{m-1} \alpha_j x_j e_j - x_m e_m \right\|_X \right) \\
&= \frac{1}{2^m} \sum_{\alpha \in \{-1,1\}^m} \left\| \sum_{j=1}^m \alpha_j x_j e_j \right\|_X
\end{aligned}$$

On a donc bien : pour tout $j \in [|1,m|]$, pour tout $x = \sum_{i=1}^m x_i e_i \in Y$,

$$\|x_j e_j\|_X \leq \mathbb{E}_\varepsilon\left(\left\|\sum_{i=1}^m \varepsilon_i x_i e_j\right\|_X\right).$$

On en déduit que, comme pour tout $\omega \in \Omega$, $G_0(\omega) \in Y$, on a, d'après le théorème de Fubini-

Tonelli :

$$\begin{aligned}
\mathbb{E}(\|G_0\|_X) &= \mathbb{E}\left(\left\|\sum_{j=1}^{m}\varepsilon_j g_j e_j\right\|_X\right) \\
&= \mathbb{E}_G\left(\mathbb{E}_\varepsilon\left(\left\|\sum_{j=1}^{m}\varepsilon_j g_j e_j\right\|_X\right)\right) \\
&\geq \mathbb{E}_G(\max\{\|g_j e_j\|_X, j\in[|1,m|]\}) \\
&\geq \mathbb{E}_G\left(\max\left\{\frac{1}{4}|g_j|, j\in[|1,m|]\right\}\right) \\
&= \frac{1}{4}\mathbb{E}_G(\|G\|_\infty) \\
&\geq \frac{C_3}{4}(\ln(m))^{\frac{1}{2}} \text{ d'après le lemme 13.}
\end{aligned}$$

En identifiant $(Y, \|.\|_{\mathcal{E}_X})$ et $(\mathbb{R}^m, \|.\|_2)$ grâce à la bijection donnée dans la démonstration du théorème 1, on peut définir $\theta_Y = \mathbb{E}\left(\frac{\|G_0\|_X}{\|G_0\|_{\mathcal{E}_X}}\right)$. On a alors, d'après le lemme 12 :

$$\theta_Y \geq m^{-\frac{1}{2}}\mathbb{E}(\|G_0\|_X) \geq \frac{C_3}{4}m^{-\frac{1}{2}}(\ln(m))^{\frac{1}{2}}$$

Or, la suite $\left(\frac{\ln(n)}{n}\right)_{n\geq 3}$ est décroissante donc :

$$\frac{\ln(n)m}{\ln(m)n} \leq 1 \leq \frac{C_1C_3^2}{16C_4}$$

d'où :

$$k \leq C_4\ln(n)\frac{\varepsilon^2}{|\ln(\varepsilon)|} \leq \frac{C_1C_3^2}{16}m^{-1}\ln(m)n\frac{\varepsilon^2}{|\ln(\varepsilon)|} \leq C_1\theta_Y^2 n\frac{\varepsilon^2}{|\ln(\varepsilon)|}$$

L'identification (isométrique) entre $(X, \|.\|_{\mathcal{E}_X})$ et $(\mathbb{R}^n, \|.\|_2)$ permet alors de conclure grâce au théorème 8. □

Dans le cas où $(X, \|.\|_X) = (\mathbb{R}^n, \|.\|_p)$, $p \in [1, +\infty[$, le théorème de Dvoretzky n'est pas optimal, comme le confirme le théorème suivant.

Théorème 10. *Soit* $1 \leq p < \infty$, $n \in \mathbb{N}^*$.
Pour tout $0 < \varepsilon < \frac{1}{3}$, *l'espace* ℓ_p^n *contient un sous-espace* X_0 *de dimension* k *tel que* $d_{X_0} \leq 1+\varepsilon$, *dès que :*
- $k \leq C_2 n^{\frac{2}{p}}\varepsilon^2|\ln(\varepsilon)|^{-1}$ *si* $p \geq 2$ *;*
- $k \leq C_2 n\varepsilon^2|\ln(\varepsilon)|^{-1}$ *si* $1 \leq p \leq 2$ *;*

où $C_2 > 0$ *est une constante universelle.*

Démonstration. On considère $\mathbb{R}^n$ muni de la norme $\|.\|_F = \|.\|_p$ et on note désormais θ_p au lieu de θ_F. On pose $C_2 = \frac{2}{\pi}C_1$.
Supposons d'abord $p \geq 2$. L'inégalité de Hölder assure que pour tout $x \in \mathbb{R}^n$:

$$\|x\|_p \leq \|x\|_2 \leq n^{\frac{1}{2}-\frac{1}{p}}\|x\|_p$$

$i.e$ $\|I\|_{\ell_2^n \to \ell_p^n} \leq 1$ et $\|I\|_{\ell_p^n \to \ell_2^n} \leq n^{\frac{1}{2}-\frac{1}{p}}$.
D'après le lemme 12, on a alors $\theta_p^2 \geq n^{\frac{2}{p}-1}$ d'où :

$$k \leq C_2 n^{\frac{2}{p}} \varepsilon^2 |\ln(\varepsilon)|^{-1} \leq C_2 n \theta_p^2 \leq C_1 n \theta_p^2 \varepsilon^2 |\ln(\varepsilon)|^{-1}$$

Ceci assure le résultat pour $p \geq 2$ grâce au théorème 8.
Supposons désormais $1 \leq p \leq 2$. Toujours d'après l'inégalité de Hölder, pour tout $x \in \mathbb{R}^n$, on a :

$$\frac{1}{\sqrt{n}} \|x\|_1 \leq \frac{1}{n^{\frac{1}{p}-\frac{1}{2}}} \|x\|_p \leq \|x\|_2$$

On pose $\|.\|_{X_1} = \frac{1}{\sqrt{n}} \|.\|_1$. On note que pour tout $x \in \mathbb{R}^n$, $\|x\|_{X_1} \leq \|x\|_2$. Soit G un vecteur gaussien centré réduit de dimension n, g une gaussienne centrée réduite.
D'après le lemme 12, on a :

$$\theta_{X_1} \geq n^{-\frac{1}{2}} \mathbb{E}(\|G\|_{X_1}) = n^{-1} \mathbb{E}(\|G\|_1) = \mathbb{E}(|g|) = \sqrt{\frac{2}{\pi}}$$

Donc, si l'on note $\|.\|_{X_p} = n^{\frac{1}{2}-\frac{1}{p}} \|.\|_p$, on a $\theta_{X_p} \geq \theta_{X_1} \geq \sqrt{\frac{2}{\pi}}$. On en déduit que :

$$k \leq C_2 n \varepsilon^2 |\ln(\varepsilon)|^{-1} \leq C_2 n \frac{\pi}{2} \theta_{X_p}^2 \varepsilon^2 |\ln(\varepsilon)|^{-1} = C_1 n \theta_{X_p}^2 \varepsilon^2 |\ln(\varepsilon)|^{-1}$$

Le théorème 8 assure à nouveau le résultat, puisque les espaces $(\mathbb{R}^n, \|.\|_p)$ et $(\mathbb{R}^n, \|.\|_{X_p})$ sont isométriques grâce à l'application linéaire $T : \left\{ \begin{array}{lll} (\mathbb{R}^n, \|.\|_p) & \to & (\mathbb{R}^n, \|.\|_{X_p}) \\ x & \mapsto & n^{\frac{1}{p}-\frac{1}{2}} \times x \end{array} \right.$. $\square$

Remarque 5. *D'après un résultat de Gordon prouvé en 1985, on peut supprimer le terme* $|\ln(\varepsilon)|$ *dans le théorème de Dvoretzky. Cependant, la dépendance en* n *ne peut être optimisée.*

Pour conclure, nous pouvons dire que la démonstration du théorème de Dvoretzky témoigne de l'utilité des méthodes probabilistes dans la géométrie des Banach, où elles ne sont apparues que dans les années 1960-1970.
Ce théorème, qui est un résultat de structure important, peut se reformuler de la façon suivante : un espace de Hilbert est finiment représentable dans tout espace de Banach de dimension infinie, ce qui garantit la minimalité des espaces de Hilbert pour la relation de représentabilité finie (si E et F sont deux espaces vectoriels normés, on dit que F est finiment représentable dans E si et seulement si pour tout sous-espace $U \subset F$ de dimension finie et $\varepsilon > 0$, il existe un sous-espace $V \subset E$ et un isomorphisme $T : U \to V$ tels que $\|T\|\|T^{-1}\| < 1 + \varepsilon$, $i.e$ si tout sous-espace de dimension finie de F est à distance de Banach-Mazur arbitrairement proche de 1 d'un sous-espace de dimension finie de E).
Pour compléter ce projet, des résultats d'existence et unicité concernant la mesure de Haar et la mesure invariante par rotation sur la sphère euclidienne sont démontrés en annexe.

A Existence et unicité de la mesure de Haar

A.1 Théorème de représentation de Riesz

Pour démontrer l'existence de la mesure de Haar, nous allons avoir besoin du théorème de représentation de Riesz, qui nécessite lui-même les préliminaires topologiques ci-dessous.

Définition 6. *Soit $(X, \mathscr{B}(X))$ un espace topologique mesurable muni de sa tribu borélienne, μ une mesure sur $(X, \mathscr{B}(X))$.*
μ est une mesure de Borel sur $(X, \mathscr{B}(X))$ si pour tout compact $K \subset X$, $\mu(K)$ est finie.

Définition 7. *Soit $(X, \mathscr{B}(X), \mu)$ un espace topologique mesuré muni de sa tribu borélienne.*
(a) μ est dite extérieurement régulière si et seulement si pour tout $A \in \mathscr{B}(X)$, on a :

$$\mu(A) = \inf\{\mu(O), O \text{ ouvert}, A \subset O\}$$

(b) μ est dite intérieurement régulière si et seulement si pour tout $A \in \mathscr{B}(X)$, on a :

$$\mu(A) = \sup\{\mu(K), K \text{ compact}, K \subset A\}$$

(c) μ est dite régulière si et seulement si elle est à la fois extérieurement et intérieurement régulière.

Soit K un espace topologique compact. On note $C(K) = (C(K, \mathbb{R}), \|.\|_\infty)$ l'espace vectoriel normé des fonctions continues sur K muni de la norme uniforme.

Lemme 14. *Soit F un fermé de K, V un ouvert de K tels que $F \subset V$.*
Il existe W ouvert de K tel que $F \subset W \subset \overline{W} \subset V$.

Démonstration. On pose $G = K \backslash V$. G est fermé et $F \cap G = \varnothing$ car $F \subset V$.
Soit $x \in F$. Comme K est compact donc en particulier séparé, pour tout $y \in G$, comme $y \neq x$, il existe U_y un voisinage ouvert de x, V_y un voisinage ouvert de y tels que $U_y \cap V_y = \varnothing$.
On note que G est compact (car c'est un fermé de K compact) et que $G \subset \bigcup_{y \in G} V_y$ donc il existe $(y_1, \cdots, y_n) \in G^n$ tel que $G \subset \bigcup_{i=1}^n V_{y_i}$.
On pose $O_x = \bigcup_{i=1}^n V_{y_i}$, $W_x = \bigcap_{i=1}^n U_{y_i}$. Ce sont des ouverts, $x \in W_x$ et $W_x \cap O_x = \varnothing$, puisque $W_x \subset \bigcap_{i=1}^n (K \backslash V_{y_i}) = K \backslash O_x$.
F est fermé dans K compact, donc compact, et $F \subset \bigcup_{x \in F} W_x$ avec pour tout $x \in F$, W_x ouvert donc il existe $(x_1, \cdots, x_k) \in F^k$ tel que $F \subset \bigcup_{i=1}^k W_{x_i}$.
On pose $W = \bigcup_{i=1}^k W_{x_i}$, $O = \bigcap_{i=1}^k O_{x_i}$. Ce sont des ouverts et ils vérifient : $F \subset W$, $G \subset O$, $O \cap W = \varnothing$.
On a alors $W \subset K \backslash O$ avec $K \backslash O$ fermé donc $\overline{W} \subset K \backslash O \subset K \backslash G = V$.
D'où : $F \subset W \subset \overline{W} \subset V$. □

Théorème 11 (Lemme d'Urysohn)**.** *Soit A et B deux fermés disjoints non vide de K.*
Il existe $f \in C(K)$ telle que $0 \leq f \leq 1$, $f_{|A} = 0$ et $f_{|B} = 1$.

Démonstration. On pose $A_0 = A$, $A_1 = K \backslash B$. A_0 est fermé, A_1 est ouvert et $A_0 \subset A_1$ donc d'après le lemme 14, il existe $A_{\frac{1}{2}}$ ouvert tel que $A_0 = \overline{A_0} \subset A_{\frac{1}{2}} \subset \overline{A_{\frac{1}{2}}} \subset A_1$.
Désormais, A_0 et $\overline{A_{\frac{1}{2}}}$ sont fermés, $A_{\frac{1}{2}}$ et A_1 sont ouverts et $A_0 \subset A_{\frac{1}{2}} \subset \overline{A_{\frac{1}{2}}} \subset A_1$ donc, d'après le lemme 14, il existe $A_{\frac{1}{4}}$ et $A_{\frac{3}{4}}$ ouverts tels que :

$$A_0 \subset A_{\frac{1}{4}} \subset \overline{A_{\frac{1}{4}}} \subset A_{\frac{1}{2}} \subset A_{\frac{3}{4}} \subset \overline{A_{\frac{3}{4}}} \subset A_1$$

On pose $D = \left\{ \frac{k}{2^n}, n \in \mathbb{N}, k \in [|0, 2^n|] \right\}$. On construit ainsi par récurrence sur n une famille d'ensembles $(A_t)_{t\in D}$ vérifiant :
(i) $A_0 = A$, $A_1 = K\backslash B$;
(ii) $\forall t \in D\backslash\{0\}$, A_t est ouvert ;
(iii) $\forall (s,t) \in D^2$, $s < t \implies \overline{A_s} \subset A_t$.
Posons alors :

$$\alpha : \begin{cases} K & \to & [0,1] \\ x & \mapsto & \sup\{s \in D, x \notin \overline{A_s}\} & \text{si } x \notin A \\ x & \mapsto & 0 & \text{si } x \in A \end{cases}$$

et :

$$\beta : \begin{cases} K & \to & [0,1] \\ x & \mapsto & \inf\{t \in D, x \in A_t\} & \text{si } x \notin B \\ x & \mapsto & 1 & \text{si } x \in B \end{cases}$$

Soit $x \in K$. Montrons que $\alpha(x) \leq \beta(x)$.
Si $x \in A$ ou $x \in B$, le résultat est clair. Supposons alors que $x \notin A$ et $x \notin B$.
Soit $t \in [0,1]$ tel que $t < \alpha(x)$. Il suffit de montrer que $t \leq \beta(x)$. Or par définition de $\alpha(x)$, $x \notin \overline{A_t}$ puisque $(\overline{A_s})_{s\in D}$ est croissante. Donc $x \notin A_t$ et pour tout $u \leq t$, $x \notin A_u$ puisque $(A_s)_{s\in D}$ est croissante. D'où pour tout $u \in \{s \in D; x \in A_s\}$, $t \leq u$, d'où $t \leq \beta(x)$, et finalement, on a bien $\alpha(x) \leq \beta(x)$.
Montrons désormais que $\alpha(x) = \beta(x)$. Supposons que $\alpha(x) < \beta(x)$.
Il existe alors $(s,t) \in D^2$ tel que $\alpha(x) < s < t < \beta(x)$. Donc par définition de $\alpha(x)$ et $\beta(x)$, $x \in \overline{A_s}$ et $x \notin A_t$. Or $s < t$ donc $\overline{A_s} \subset A_t$, ce qui est absurde. Donc pour tout $x \in K$, $\alpha(x) = \beta(x)$. On pose $f = \alpha (= \beta)$.
On a bien $0 \leq f \leq 1$, $f_{|A} = 0$ et $f_{|B} = 1$. Montrons alors que $f \in C(K)$. Il suffit de montrer que pour tout $\gamma \in [0,1]$, $f^{-1}(]-\infty, \gamma[)$ et $f^{-1}(]\gamma, +\infty[)$ sont ouverts. Soit $\gamma \in [0,1]$.
• Si $\gamma = 0$, $f^{-1}(]-\infty, \gamma[) = \varnothing$ est ouvert. Supposons alors $\gamma > 0$. Comme D est dense dans $[0,1]$, on a :

$$\begin{aligned} \{f < \gamma\} &= \bigcup_{0<\beta<\gamma, \beta\in D} \{f < \beta\} \\ &= \bigcup_{0<\beta<\gamma, \beta\in D} \{x \in K; \beta(x) < \beta\} \\ &= \bigcup_{0<\beta<\gamma, \beta\in D} \{x \in K; \exists t < \beta, t \in D\backslash\{0\}, x \in A_t\} \\ &= \bigcup_{0<\beta<\gamma, \beta\in D} \bigcup_{t<\beta, t\in D\backslash\{0\}} A_t \end{aligned}$$

Or pour tout $t \in D\backslash\{0\}$, A_t est ouvert donc $\{f < \gamma\}$ est ouvert.
• Si $\gamma = 1$, $f^{-1}(]\gamma, +\infty[) = \varnothing$ est ouvert. Supposons alors $\gamma < 1$. Comme D est dense dans

$[0,1]$, on a :

$$\begin{aligned}
\{f > \gamma\} &= \bigcup_{\gamma<\beta<1, \beta\in D} \{f > \beta\} \\
&= \bigcup_{\gamma<\beta<1, \beta\in D} \{x \in K; \alpha(x) > \beta\} \\
&= \bigcup_{\gamma<\beta<1, \beta\in D} \{x \in K; \exists t > \beta, t \in D\backslash\{0\}, x \notin \overline{A_t}\} \\
&= \bigcup_{\gamma<\beta<1, \beta\in D} \bigcup_{t>\beta, t\in D\backslash\{0\}} (\overline{A_t})^c
\end{aligned}$$

Donc $\{f > \gamma\}$ est ouvert.
Donc f est bien continue. □

Notations 1. *Soit O un ouvert de K, F un fermé de K, $f \in C(K)$.*
Nous noterons :
- *$f \prec O$ si $0 \leq f \leq 1$ et $Supp(f) \subset O$;*
- *$F \prec f$ si $0 \leq f \leq 1$ et $f_{|F} = 1$.*

Remarque 6. *Le lemme d'Urysohn affirme alors que si F est un fermé, O est un ouvert, et $F \subset O$, alors il existe $f \in C(K)$ telle que $F \prec f \prec O$. En effet, en posant $B = F$ et $A = K\backslash O$ dans le théorème 11, on obtient le résultat.*

Corollaire 1 (Partitions de l'unité)**.** *Soit $F \subset K$ fermé, $O_1, \cdots, O_n$ ouverts tels que $F \subset \bigcup_{i=1}^n O_i$.*
Il existe $(h_1, \cdots, h_n) \in C(K)^n$ tel que, pour tout $i \in [|1,n|]$, on ait $h_i \prec O_i$ et $F \prec \sum_{i=1}^n h_i$.

Démonstration. Pour tout $x \in F$, il existe $i_x \in [|1,n|]$ tel que $x \in O_{i_x}$ ouvert donc il existe W_x voisinage ouvert de x tel que $\overline{W_x} \subset O_{i_x}$, d'après le lemme 14.
F est fermé, donc compact et $F \subset \bigcup_{x\in F} W_x$ donc il existe $(x_1, \cdots, x_p) \in F^p$ tel que :

$$F \subset \bigcup_{k=1}^{p} W_{x_k} \subset \bigcup_{k=1}^{p} \overline{W_{x_k}}$$

Pour tout $i \in [|1,n|]$, on pose :

$$F_i := \bigcup_{k=1}^{p} \{\overline{W_{x_k}}; \overline{W_{x_k}} \subset O_i\}$$

Pour tout $i \in [|1,n|]$, F_i est fermé (union finie de fermés) et $F \subset \bigcup_{i=1}^n F_i$.
D'après le lemme d'Urysohn, pour tout $i \in [|1,n|]$, comme $F_i \subset O_i$, il existe $g_i \in C(K)$ telle que $F_i \prec g_i \prec O_i$.
Pour tout $i \in [|1,n|]$, on pose :

$$h_i = (\prod_{k=1}^{i-1}(1-g_k))g_i$$

On note qu'on a toujours : pour tout $i \in [|1,n|]$, $h_i \prec O_i$, et on montre par récurrence que pour tout $m \in [|1,m|]$, on a :

$$h_1 + \cdots + h_m = 1 - (1-g_1)\cdots(1-g_m)$$

Or, pour tout $i \in [|1,n|]$, $F_i \prec g_i$, donc, comme $0 \leq h_1 + \cdots + h_n \leq 1$, on a bien :

$$F \subset \bigcup_{i=1}^{n} F_i \prec h_1 + \cdots + h_n$$

$\square$

Définition 8. *Une application linéaire $\phi : C(K) \to \mathbb{R}$ est dite positive si et seulement si :*

$$\forall f \in C(K), f \geq 0 \implies \phi(f) \geq 0.$$

Nous avons désormais tous les outils pour démontrer le théorème de représentation de Riesz :

Théorème 12 (Riesz)**.** *Soit ϕ une forme linéaire positive sur $C(K)$.*
Il existe une unique mesure régulière finie μ sur $\mathscr{B} := \mathscr{B}(K)$ telle que :

$$\forall f \in C(K), \phi(f) = \int_K f d\mu$$

Démonstration. Montrons tout d'abord l'existence.
Pour tout ouvert O, on pose :

$$\mu(O) := \sup\{\phi(f), f \in C(K), f \prec O\}$$

On note que pour tout $O_1 \subset O_2$ ouverts, on a $\mu(O_1) \leq \mu(O_2)$ $(*)$ car $\{f \in C(K); f \prec O_1\} \subset \{f \in C(K); f \prec O_2\}$.
Pour tout $E \subset K$ quelconque, on pose alors :

$$\mu(E) := \inf\{\mu(O); E \subset O, O \text{ ouvert}\}$$

On note que cette nouvelle application μ est bien définie car pour tout O ouvert, $\mu(O) = \inf\{\mu(O'); O \subset O', O' \text{ ouvert}\}$ par $(*)$, que $\mu(K) = \phi(\mathbb{1}_K)$ et que pour tout $E_1 \subset E_2 \subset K$, $\mu(E_1) \leq \mu(E_2)$.

Etape 1 : μ est σ-sous-additive
Commençons par montrer que pour tout O_1, O_2 ouverts, $\mu(O_1 \cup O_2) \leq \mu(O_1) + \mu(O_2)$. Soit O_1, O_2 ouverts.
Soit $g \in C(K)$ telle que $g \prec (O_1 \cup O_2)$. Alors : $F := Supp(g) \subset (O_1 \cup O_2)$.
D'après le corollaire 1, il existe $(h_1, h_2) \in C(K)^2$ tel que $h_1 \prec O_1$, $h_2 \prec O_2$ et $F \prec h_1 + h_2$.
D'où $g = h_1 g + h_2 g$ car $h_1 + h_2 = 1$ sur $F = Supp(g)$.
De plus, comme $h_1 g \prec O_1$ et $h_2 g \prec O_2$, on a $\phi(h_1 g) \leq \mu(O_1)$ et $\phi(h_2 g) \leq \mu(O_2)$ d'où, pour toute $g \in C(K)$ telle que $g \prec (O_1 \cup O_2)$, on a :

$$\phi(g) = \phi(h_1 g) + \phi(h_2 g) \leq \mu(O_1) + \mu(O_2)$$

On en déduit que, pour tout O_1, O_2 ouverts :

$$\mu(O_1 \cup O_2) = \sup\{\phi(g), g \in C(K), g \prec (O_1 \cup O_2)\} \leq \mu(O_1) + \mu(O_2) (*)$$

Montrons alors la σ-sous additivité. Soit $(E_i)_{i \in \mathbb{N}^*} \in \mathcal{P}(K)^{\mathbb{N}^*}$. Soit $\varepsilon > 0$.
Par définition de μ, pour tout $i \in \mathbb{N}^*$, il existe O_i ouvert tel que $E_i \subset O_i$ et :

$$\mu(O_i) \leq \mu(E_i) + \frac{\varepsilon}{2^i}$$

On pose $O := \bigcup_{i=1}^{\infty} O_i$. Soit $f \in C(K)$ telle que $f \prec O$. On a : $F := Supp(g) \subset O$ donc, comme F est compact, il existe $n \in \mathbb{N}^*$ tel que $F \subset \bigcup_{i=1}^{n} O_i$. D'où : $f \prec \bigcup_{i=1}^{n} O_i$.
On en déduit, grâce à $(*)$ et à la définition de μ que, pour toute $f \in C(K)$ telle que $f \prec O$, il existe $n \in \mathbb{N}^*$ tel que :

$$\phi(f) \leq \mu\left(\bigcup_{i=1}^{n} O_i\right) \leq \sum_{i=1}^{n} \mu(O_i) \leq \sum_{i=1}^{n} \left(\mu(E_i) + \frac{\varepsilon}{2^i}\right) \leq \sum_{i=1}^{\infty} \mu(E_i) + \varepsilon$$

D'où, pour tout $\varepsilon > 0$, on a :

$$\begin{aligned} \mu\left(\bigcup_{i=1}^{\infty} E_i\right) &\leq \mu(O) \\ &= \sup\{\phi(f), f \in C(K), f \prec O\} \\ &\leq \sum_{i=1}^{\infty} \mu(E_i) + \varepsilon \end{aligned}$$

Finalement, par passage à la limite, on a :

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) \leq \sum_{i=1}^{\infty} \mu(E_i)$$

Donc μ est bien σ-sous-additive.

Etape 2 : Pour tout F fermé, $\mu(F) = \inf\{\phi(f), F \prec f\}$
Soit F fermé, $\varepsilon > 0$. Par définition de $\mu(F)$, il existe O ouvert tel que $F \subset O$ et $\mu(O) \leq \mu(F)+\varepsilon$; et par définition de $\mu(O)$ et d'après le lemme d'Urysohn, il existe $f \in C(K)$ telle que $F \prec f \prec O$ et $\phi(f) \leq \mu(O) \leq \mu(F) + \varepsilon$.
Donc, pour tout $\varepsilon > 0$, il existe $f \in C(K)$ telle que $F \prec f$ et $\phi(f) \leq \mu(F) + \varepsilon$. D'où :

$$\mu(F) \geq \inf\{\phi(f), f \in C(K), F \prec f\}$$

Montrons maintenant l'égalité. Pour cela, montrons que pour toute $f \in C(K)$ telle que $F \prec f$, pour tout $\alpha \in]0, 1[$, $\mu(F) \leq \frac{1}{\alpha}\phi(f)$.
Soit $f \in C(K)$ telle que $F \prec f$, $\alpha \in]0, 1[$.
On pose $O_\alpha := \{f > \alpha\}$. On a $F \subset O_\alpha$ puisque $\alpha < 1$.
Soit $g \in C(K)$ telle que $g \prec O_\alpha$. Alors $g \leq \mathbb{1}_{O_\alpha} = \mathbb{1}_{\left\{\frac{1}{\alpha}f > 1\right\}}$ donc $g \leq \frac{1}{\alpha}f$, d'où $\phi\left(g - \frac{1}{\alpha}f\right) \leq 0$ puisque ϕ est positive. On a alors, pour toute $g \in C(K)$ telle que $g \prec O_\alpha$:

$$\phi(g) \leq \frac{1}{\alpha}\phi(f)$$

Donc : $\mu(O_\alpha) \leq \frac{1}{\alpha}\phi(f)$. Or $F \subset O_\alpha$ d'où, pour toute $f \in C(K)$ telle que $F \prec f$, pour tout $\alpha \in]0, 1[$:

$$\mu(F) \leq \mu(O_\alpha) \leq \frac{1}{\alpha}\phi(f)$$

On a donc bien :

$$\mu(F) \leq \inf\{\phi(f), f \in C(K), F \prec f\}$$

Et finalement, pour tout F fermé :

$$\mu(F) = \inf\{\phi(f), f \in C(K), F \prec f\}$$

On pose $\mathscr{T}_\phi = \{E \subset K, \mu(E) = \sup\{\mu(F); F \subset E, F \text{ fermé}\}\}$.

Etape 3 : Montrons que :
(i) F fermé $\implies$ $F \in \mathscr{T}_\phi$;
(ii) O ouvert $\implies$ $O \in \mathscr{T}_\phi$;
(iii) μ est σ-additive sur $\mathscr{T}_\phi$.

(i) Comme pour tout $A, B \in \mathcal{P}(K)$, on a :

$$A \subset B \implies \mu(A) \leq \mu(B)$$

on a bien que pour tout F fermé, $F \in \mathscr{T}_\phi$.
(ii) Soit O ouvert, $\varepsilon > 0$. Montrons qu'il existe F fermé tel que $\mu(O) - \varepsilon \leq \mu(F)$.
Par définition de $\mu(O)$, il existe $f \in C(K)$ telle que $f \prec O$ et $\mu(O) - \varepsilon \leq \phi(f)$. On pose $F := \mathit{Supp}(f) \subset O$. D'après le lemme 14, il existe un ouvert W tel que :

$$F \subset W \subset \overline{W} \subset O$$

Alors, comme $\mathit{Supp}(f) = F \subset W$ ouvert, $f \prec W$ donc :

$$\mu(O) - \varepsilon \leq \phi(f) \leq \mu(W) \leq \mu(\overline{W})$$

D'où, pour tout $\varepsilon > 0$, il existe F fermé tel que $\mu(O) - \varepsilon \leq \mu(F)$. Donc $O \in \mathscr{T}_\phi$.
(iii) Commençons par montrer l'additivité sur un nombre fini de fermés.
Soit F_1, F_2 deux fermés disjoints, $\varepsilon > 0$. D'après le lemme d'Urysohn, il existe $f \in C(K)$ telle que $0 \leq f \leq 1$, $f_{|F_1} = 1$, $f_{|F_2} = 0$, *i.e* $F_1 \prec f$ et $F_2 \prec 1 - f$. D'après l'étape 2, comme $F_1 \cup F_2$ est fermé, il existe $g \in C(K)$ telle que $F_1 \cup F_2 \prec g$ et :

$$\mu(F_1 \cup F_2) \leq \phi(g) \leq \mu(F_1 \cup F_2) + \varepsilon$$

On note que $F_1 \prec fg$ et $F_2 \prec (1 - f)g$, d'où, toujours d'après l'étape 2, $\mu(F_1) \leq \phi(fg)$ et $\mu(F_2) \leq \phi((1 - f)g)$.
On a alors, pour tout $\varepsilon > 0$:

$$\mu(F_1) + \mu(F_2) \leq \phi(fg) + \phi((1 - f)g) = \phi(g) \leq \mu(F_1 \cup F_2) + \varepsilon$$

D'où, pour tous F_1, F_2 fermés disjoints, d'après l'étape 1, on a bien :

$$\mu(F_1 \cup F_2) = \mu(F_1) + \mu(F_2)$$

Montrons désormais la σ-additivité sur $\mathscr{T}_\phi$. Soit $(E_i)_{i \in \mathbb{N}^*} \in \mathscr{T}_\phi^{\mathbb{N}^*}$ des ensembles deux à deux disjoints, $\varepsilon > 0$.
Pour tout $i \in \mathbb{N}^*$, $E_i \in \mathscr{T}_\phi$ donc il existe F_i fermé tel que $F_i \subset E_i$ et :

$$\mu(E_i) - \frac{\varepsilon}{2^i} \leq \mu(F_i)$$

On a alors, pour tout $n \in \mathbb{N}^*$, pour tout $\varepsilon > 0$:

$$\sum_{i=1}^{n} \mu(E_i) - \varepsilon \leq \sum_{i=1}^{n} \left(\mu(E_i) - \frac{\varepsilon}{2^i}\right) \leq \sum_{i=1}^{n} \mu(F_i) = \mu\left(\bigcup_{i=1}^{n} F_i\right) \leq \mu\left(\bigcup_{i=1}^{\infty} E_i\right)$$

car $F_1, \cdots, F_n$ sont deux-à-deux disjoints. Et finalement, d'après l'étape 1 :

$$\mu\left(\bigcup_{i=1}^{\infty} E_i\right) = \sum_{i=1}^{\infty} \mu(E_i)$$

Etape 4 : $\mathscr{T}_\phi$ est une tribu
• D'après l'étape 3, $K \in \mathscr{T}_\phi$.
• Soit $A, B \in \mathscr{T}_\phi$, $\varepsilon > 0$. Montrons que $A \backslash B \in \mathscr{T}_\phi$. Pour cela, il suffit de montrer qu'il existe $F \subset A \backslash B$ fermé tel que $\mu(A \backslash B) - \varepsilon \leq \mu(F)$.
$A \in \mathscr{T}_\phi$ donc il existe F_1 fermé tel que $F_1 \subset A$ et $\mu(A) - \frac{\varepsilon}{2} \leq \mu(F_1)$. De plus, il existe O_1 ouvert tel que $A \subset O_1$ et $\mu(O_1) \leq \mu(A) + \frac{\varepsilon}{2}$. D'où, comme $O_1 \backslash F_1, O_1, F_1 \in \mathscr{T}_\phi$:

$$\mu(O_1 \backslash F_1) = \mu(O_1) - \mu(F_1) \leq \mu(A) + \frac{\varepsilon}{2} - \mu(A) + \frac{\varepsilon}{2} = \varepsilon$$

De même, il existe F_2 fermé, O_2 ouvert tels que $F_2 \subset B \subset O_2$ et $\mu(O_2 \backslash F_2) \leq \varepsilon$.
On a alors :

$$F_1 \backslash O_2 \subset A \backslash B \subset O_1 \backslash F_2$$

et :

$$\mu(F_1 \backslash O_2) \leq \mu(A \backslash B) \leq \mu(O_1) - \mu(F_2) - \mu(F_1) + \mu(F_1) + \mu(O_2) - \mu(O_2) \leq \mu(F_1 \backslash O_2) + 2\varepsilon$$

avec $F_1 \backslash O_2$ fermé.
Donc $A \backslash B \in \mathscr{T}_\phi$. En particulier, pour tout $A \in \mathscr{T}_\phi$, $K \backslash A \in \mathscr{T}_\phi$.
• Soit $(E_i)_{i \in \mathbb{N}^*} \in \mathscr{T}_\phi^{\mathbb{N}^*}$. On a :

$$E := \bigcup_{i \in \mathbb{N}^*} E_i = E_1 \cup \left(\bigcup_{i=2}^{\infty} E_i \backslash E_{i-1}\right)$$

avec, pour tout $i \geq 2$, $E_i \backslash E_{i-1} \in \mathscr{T}_\phi$ d'après ce qui précède.
On peut donc supposer les E_i deux à deux disjoints.
Soit $\varepsilon > 0$. D'après l'étape 3, $\mu(E) = \sum_{i=1}^{\infty} \mu(E_i)$, donc il existe $n \in \mathbb{N}^*$ tel que :

$$\mu(E) - \varepsilon \leq \sum_{i=1}^{n} \mu(E_i)$$

Or pour tout $i \in [|1, n|]$, $E_i \in \mathscr{T}_\phi$ donc il existe F_i fermé tel que $F_i \subset E_i$ et :

$$\mu(E_i) - \frac{\varepsilon}{2^i} \leq \mu(F_i)$$

On a alors :

$$\mu(E) - 3\varepsilon \leq \sum_{i=1}^{n} \mu(E_i) - 2\varepsilon \leq \sum_{i=1}^{n} \left(\mu(E_i) - \frac{\varepsilon}{2^i}\right) \leq \sum_{i=1}^{n} \mu(F_i) = \mu(F)$$

où $F = \bigcup_{i=1}^n F_i$ est fermé.
D'où $E \in \mathscr{T}_\phi$.

Finalement, $\mathscr{T}_\phi$ est bien une tribu, contenant les ouverts de K, ce qui implique que $\mathscr{B} \subset \mathscr{T}_\phi$.

Etape 5 : Pour toute $f \in C(K)$, $\phi(f) = \int_K f d\mu$
Montrons que pour toute $f \in C(K)$, $\phi(f) \leq \int_K f d\mu$, ce qui concluera puisqu'alors, on aura :

$$-\phi(f) = \phi(-f) \leq \int_K -f d\mu = -\int_K f d\mu$$

i.e $\int_K f d\mu \leq \phi(f)$.
Soit $f \in C(K)$, $\varepsilon > 0$. Puisque f est continue sur K compact, il existe $(a,b) \in \mathbb{R}^2$ tel que $f(K) = [a,b]$ et il existe $(y_0, \cdots, y_n) \in \mathbb{R}^n$ tel que $y_0 = a < y_1 < \cdots < y_n = b$ soit une subdivision de $[a,b]$ de pas inférieur ou égal à ε.
Pour tout $i \in [|1,n|]$, $E_i := f^{-1}(]y_{i-1}, y_i]) \in \mathscr{B} \subset \mathscr{T}_\phi$. De plus, pour tout $i \in [|1,n|]$, il existe O_i ouvert tel que $E_i \subset O_i$ et :

$$\mu(O_i) \leq \mu(E_i) + \frac{\varepsilon}{n}$$

et $f < y_i + \varepsilon$ sur O_i.
On a $K = \bigcup_{i=1}^n$ fermé donc d'après le corollaire 1, il existe $(h_1, \cdots, h_n) \in C(K)^n$ tel que pour tout $i \in [|1,n|]$, $h_i \prec O_i$ et $K \prec \sum_{i=1}^n h_i$ *i.e* $\sum_{i=1}^n h_i \equiv 1$.
D'où $f = \sum_{i=1}^n h_i f$ et, pour tout $i \in [|1,n|]$, comme $h_i \geq 0$, on a :

$$h_i f \leq (y_i + \varepsilon) h_i$$

On a alors, pour tout $n \in \mathbb{N}$:

$$\begin{aligned}
\phi(f) &= \sum_{i=1}^n \phi(h_i f) \\
&\leq \sum_{i=1}^n (y_i + \varepsilon)\phi(h_i) \\
&\leq \sum_{i=1}^n ((y_{i-1} + \varepsilon) + \varepsilon) \left(\mu(E_i) + \frac{\varepsilon}{n}\right) \\
&= \sum_{i=1}^n y_{i-1}\mu(E_i) + 2\varepsilon\mu\left(\bigcup_{i=1}^n E_i\right) + \frac{\varepsilon}{n}\sum_{i=1}^n y_{i-1} + 2\varepsilon^2 \frac{n}{n} \\
&\leq \sum_{i=1}^n y_{i-1}\mu(E_i) + 2\varepsilon\mu(K) + \frac{\varepsilon}{n}\sum_{i=1}^n b + 2\varepsilon^2 \\
&= \sum_{i=1}^n y_{i-1}\mu(E_i) + 2\varepsilon\mu(K) + \varepsilon(b + 2\varepsilon)
\end{aligned}$$

car les E_i sont deux à deux disjoints.
Donc, pour tout $\varepsilon > 0$, on a :

$$\phi(f) \leq \int_K \left(\sum_{i=1}^n y_{i-1} \mathbb{1}_{E_i}\right) d\mu + \varepsilon(2\mu(K) + b + 2\varepsilon) \leq \int_K f d\mu + \varepsilon(2\mu(K) + b + 2\varepsilon)$$

D'où :

$$\phi(f) \leq \int_K f d\mu.$$

On a donc construit une mesure μ finie sur $\mathscr{B}$ et extérieurement régulière par construction, qui est également intérieurement régulière d'après l'étape 4 et qui vérifie :

$$\forall f \in C(K), \phi(f) = \int_K f d\mu.$$

Montrons alors l'unicité. Soit λ et μ deux mesures régulières qui conviennent.
Soit F un fermé, $\varepsilon > 0$.
$F \in \mathscr{B}$ et μ est extérieurement régulière donc il existe un ouvert O vérifiant :

$$\mu(F) \leq \mu(O) \leq \mu(F) + \varepsilon$$

D'après le lemme d'Urysohn, on a alors l'existence de $f \in C(K)$ telle que $F \prec f \prec O$. On en déduit que, pour tout $\varepsilon > 0$:

$$\lambda(F) \leq \int_K f d\lambda = \phi(f) = \int_K f d\mu \leq \mu(O) \leq \mu(F) + \varepsilon$$

Donc, pour tout F fermé, $\lambda(F) \leq \mu(F)$. Par symétrie, on en déduit également que pour tout fermé F, $\mu(F) \leq \lambda(F)$.
Soit $A \in \mathscr{B}$. Montrons que $\mu(A) = \lambda(A)$.
Comme λ et μ sont intérieurement régulières, on a :

$$\mu(A) = \sup\{\mu(F), F \text{ compact}, F \subset A\} = \sup\{\lambda(F), F \text{ compact}, F \subset A\} = \lambda(A)$$

Finalement, on a bien $\lambda = \mu$ sur $\mathscr{B}$. □

A.2 Unicité de la mesure de Haar

Commençons tout d'abord par montrer l'unicité de la mesure de Haar, qui est une conséquence du théorème de Radon-Nikodym.

Théorème 13 (Radon-Nikodým, cas des mesures finies)**.** *Soit μ et ν deux mesures finies sur un espace mesurable $(X, \mathscr{A})$. Les assertions suivantes sont équivalentes :*
(i) $\forall A \in \mathscr{A}$, $\mu(A) = 0 \implies \nu(A) = 0$,
(ii) il existe $f \in \mathscr{L}^1_{\mathbb{R}^+}(\mu)$ telle que, pour tout $A \in \mathscr{A}$, $\nu(A) = \int_A f d\mu$.
En outre, la fonction f est unique (dans $L^1_{\mathbb{R}^+}(\mu)$ i.e à une égalité μ-p.p. près).

Démonstration. Le sens (ii) implique (i) est clair. Montrons alors que (i) implique (ii).
Etape 1 : $\nu \leq \mu$
Supposons que $\nu \leq \mu$ (au sens où pour tout $A \in \mathscr{A}$, $\nu(A) \leq \mu(A)$ ou encore pour toute fonction mesurable positive g, $\int_X g d\nu \leq \int_X g d\mu$). On note que μ et ν vérifient la condition (i). On pose :

$$\phi : \left\{ \begin{array}{lll} L^2_{\mathbb{R}}(\mu) & \to & \mathbb{R} \\ f & \mapsto & \int_X f d\nu \end{array} \right.$$

Cette application ϕ est continue puisque, d'après l'inégalité de Hölder, comme ν est finie, on a, pour toute $g \in L^2_{\mathbb{R}}(\mu) \subset L^2_{\mathbb{R}}(\nu)$:

$$\left|\int_X g d\nu\right| = \left|\int_X g 1 d\nu\right| \leq \left(\int_X g^2 d\nu\right)^{\frac{1}{2}} \nu(X)^{\frac{1}{2}} \leq \left(\int_X g^2 d\mu\right)^{\frac{1}{2}} \nu(X)^{\frac{1}{2}}$$

D'après le théorème de représentation du dual de $L^2_{\mathbb{R}}(\mu)$, il existe alors $f \in L^2_{\mathbb{R}}(\mu)$ telle que :

$$\forall g \in L^2_{\mathbb{R}}(\mu), \int_X g d\nu = \phi(g) = \int_X g f d\mu$$

Comme la mesure μ est finie, $f \in L^2_{\mathbb{R}}(\mu)$, donc $f \in L^1_{\mathbb{R}}(\mu)$ d'après l'inégalité de Hölder. On peut évidemment assimiler f à l'un de ses représentants dans $\mathscr{L}^1_{\mathbb{R}}(\mu)$. Et plus généralement comme $\mathscr{L}^\infty_{\mathbb{R}}(\mu) \subset \mathscr{L}^2_{\mathbb{R}}(\mu)$, on a que pour tout $A \in \mathscr{A}$:

$$\nu(A) = \int_A f d\mu$$

Montrons enfin que f est μ-p.p. à valeurs dans l'intervalle $[0,1]$.
Supposons $\mu(\{f < 0\}) > 0$. Comme $\{f < 0\} = \bigcup_{n \geq 1} \left\{f \leq -\frac{1}{n}\right\}$ (union croissante), il existe $n_0 \in \mathbb{N}^*$ tel que $\mu\left(\left\{f \leq -\frac{1}{n_0}\right\}\right) > 0$. On a alors :

$$0 \leq \nu\left(\left\{f \leq -\frac{1}{n_0}\right\}\right) = \int_{\left\{f \leq -\frac{1}{n_0}\right\}} f d\mu \leq -\frac{1}{n_0}\mu\left(\left\{f \leq -\frac{1}{n_0}\right\}\right) < 0$$

Ce qui est absurde. Donc $\mu(\{f < 0\}) = 0$ et, quitte à remplacer f par $f\mathbb{1}_{f \geq 0}$, on peut supposer f positive. On montre de la même façon que $\mu(\{f > 1\}) = 0$.
Donc f est μ-p.p. à valeurs dans l'intervalle $[0,1]$.
Etape 2 : Cas général
Supposons que μ et ν vérifient (i). D'après l'étape 1 appliquée aux deux mesures finies μ et $\mu + \nu$, il existe une fonction $f \in \mathscr{L}^1_{\mathbb{R}^+}(\mu + \nu)$ telle que $0 \leq f \leq 1$ μ-p.p. et :

$$\forall A \in \mathscr{A}, \nu(A) = \int_A f d(\mu + \nu)$$

On en déduit alors l'égalité entre mesures finies : $(1 - f).\nu = f.\mu$.
On pose $N := \{f = 1\}$. On a : $\mu(N) = \int_N f d\mu = \int_N (1 - f) d\nu = 0$. Or μ et ν vérifient (i) donc $\nu(N) = 0$. On a alors, pour tout $A \in \mathscr{A}$:

$$\begin{aligned}\nu(A) &= \nu(A \cap N) + \nu(A \cap N^c) \\ &= 0 + \int_{N^c} \frac{\mathbb{1}_A}{1 - f}(1 - f) d\nu \\ &= \int_{N^c} \frac{\mathbb{1}_A}{1 - f} f d\mu = \int_A \mathbb{1}_{N^c} \frac{f}{1 - f} d\mu\end{aligned}$$

Or $\int_X \mathbb{1}_{N^c} \frac{f}{1-f} d\mu = \nu(N^c) = \nu(X) < +\infty$ donc $\mathbb{1}_{N^c} \frac{f}{1-f} \in \mathscr{L}^1_{\mathbb{R}^+}(\mu)$ convient.
Etape 3 : Unicité
Si f et $\tilde{f}$ vérifient (ii), alors :

$$\nu(\{f > \tilde{f}\}) = \int_{\{f > \tilde{f}\}} f d\mu = \int_{\{f > \tilde{f}\}} \tilde{f} d\mu$$

Donc $\int_{\{f>\tilde{f}\}}(f-\tilde{f})d\mu = 0$ d'où $\mu(\{f > \tilde{f}\}) = 0$ et par symétrie $\mu(\{f < \tilde{f}\}) = 0$.
On en déduit que $\mu(\{f \neq \tilde{f}\}) = 0$ *i.e* $f = \tilde{f}$ μ-p.p.

□

Bien que nous puissions nous contenter du théorème précédent puisque les mesures considérées dans le cadre de ce projet sont finies, démontrons le théorème de Radon-Nikodým dans le cas général.

Théorème 14 (Radon-Nikodým). *Soit μ et ν deux mesures σ-finies sur un espace mesurable $(X, \mathscr{A})$. Les assertions suivantes sont équivalentes :*
(i) $\forall A \in \mathscr{A}$, $\mu(A) = 0 \implies \nu(A) = 0$,
(ii) il existe $f : (X, \mathscr{A}) \to \mathbb{R}^+$ mesurable telle que, pour tout $A \in \mathscr{A}$, $\nu(A) = \int_A f d\mu$.
En outre, la fonction f est unique (à une égalité μ-p.p. près).

Démonstration. Le sens (ii) implique (i) est clair. Montrons alors que (i) implique (ii).
Les mesures μ et ν sont σ-finies donc il existe deux partitions $\mathscr{A}$-mesurables de X, $(F_n)_{n\in\mathbb{N}}$ et $(G_n)_{n\in\mathbb{N}}$ vérifiant : pour tout $n \in \mathbb{N}$, $\mu(F_n) + \nu(G_n) < +\infty$. On note que les $E_{k,l} := F_k \cap G_l$, $k, l \in \mathbb{N}$, forment une partition $\mathscr{A}$-mesurable de X vérifiant : pour tout $k, l \in \mathbb{N}$, $\mu(E_{k,l}) \leq \mu(F_k) < +\infty$ et $\nu(E_{k,l}) \leq \nu(G_l) < +\infty$.
Or $\mathbb{N}^2$ est dénombrable infini donc on peut supposer que les ensembles $(E_{k,l})$ sont indexés par $\mathbb{N}$. On les notera désormais $(E_n)_{n\in\mathbb{N}}$.
Pour tout $n \in \mathbb{N}$, on pose $\mu_n := \mu(.\cap E_n)$ et $\nu_n := \nu(.\cap E_n)$. Ces mesures sont finies et vérifient les hypothèses du théorème 13 donc, d'après ce même théorème, pour tout $n \in \mathbb{N}$, il existe une fonction $f_n \in \mathscr{L}^1_{\mathbb{R}^+}(\mu_n)$ telle que $\nu_n = f_n.\mu_n = (f_n \mathbb{1}_{E_n}).\mu$. On pose alors $f := \sum_{n\in\mathbb{N}} f_n \mathbb{1}_{E_n}$.
D'après le théorème de Beppo-Levi, pour tout $A \in \mathscr{A}$, on a :

$$\begin{aligned}
\int_A f d\mu &= \int_X \mathbb{1}_A \sum_{n\in\mathbb{N}} f_n \mathbb{1}_{E_n} d\mu \\
&= \sum_{n\in\mathbb{N}} \int_X \mathbb{1}_A f_n d\mu_n \\
&= \sum_{n\in\mathbb{N}} \nu(A \cap E_n) = \nu(A)
\end{aligned}$$

D'où $\nu = f.\mu$.
L'unicité se montre comme dans le cas fini du théorème 13. □

Lemme 15. *Soit $(K, \mathscr{B}(K), \mu)$ un espace mesuré, avec K un espace topologique compact, μ une mesure régulière.*
L'ensemble $C(K)$ des fonctions continues est dense dans $L^1(\mu)$.

Démonstration. L'ensemble des fonctions étagées intégrables est dense dans $L^1(\mu)$ donc il suffit de montrer que pour tout $E \in \mathscr{B}(K)$ de mesure finie, pour tout $\varepsilon > 0$, il existe $\phi \in C(K)$ telle que $\|\phi - \mathbb{1}_E\|_1 \leq \varepsilon$.
Soit $\varepsilon > 0$, $E \in \mathscr{B}(K)$ tel que $\mu(E) < +\infty$.
μ est régulière donc il existe O ouvert, F fermé tels que $F \subset E \subset O$ et $\mu(O \backslash F) \leq \varepsilon$. D'après le lemme d'Urysohn, il existe alors $\phi \in C(K)$ telle que $\mathbb{1}_F \leq \phi \leq \mathbb{1}_O$.
On a alors :

$$\mathbb{1}_F - \mathbb{1}_O \leq \mathbb{1}_F - \mathbb{1}_E \leq \phi - \mathbb{1}_E \leq \mathbb{1}_O - \mathbb{1}_F$$

donc $|\phi - \mathbb{1}_E| \leq \mathbb{1}_{O\backslash F}$.
D'où $\|\phi - \mathbb{1}_E\|_1 \leq \mu(O\backslash F) \leq \varepsilon$, ce qui assure le résultat. □

Définition 9. *Un groupe topologique G est un espace topologique qui vérifie :*
(i) G est un groupe,
(ii) $\left\{\begin{array}{lll} G\times G & \to & G \\ (x,y) & \mapsto & xy \end{array}\right.$ *et* $\left\{\begin{array}{lll} G & \to & G \\ x & \mapsto & x^{-1} \end{array}\right.$ *sont continues (où l'on a muni $G\times G$ de la topologie produit)*

Définition 10. *Soit G un groupe topologique, $\mathscr{B}(G)$ la tribu des boréliens de G. Une mesure de Haar μ sur G est une mesure de Borel régulière, non nulle sur $\mathscr{F}$ qui est invariante par translation à gauche, i.e qui vérifie :*

$$\forall A\in\mathscr{F}, \forall g\in G, \mu(gA)=\mu(A)$$

Quand G est compact, on exige parfois de plus que μ soit normalisée, i.e que μ soit une mesure de probabilité.

Théorème 15. *Soit μ et ν deux mesures de Haar non nulles sur un groupe topologique compact G, muni de sa tribu borélienne $\mathscr{B}(G)$.*
Il existe $c\in]0,+\infty[$ tel que $d\mu=cd\nu$.

Démonstration. On pose $\eta=\mu+\nu$. On note que η est une mesure de Haar et comme $\mu\leq\eta$, d'après le théorème de Radon-Nikodym (théorème 14), il existe $h\in L^1_{\mathbb{R}^+}(\eta)$ telle que $d\mu=hd\eta$. On montre comme dans la démonstration du théorème 13 que $0\leq h\leq 1$.
Soit $f\in C(G)$. Pour tout $y\in G$, on a :

$$\begin{aligned}\int_G f(x)h(x)d\eta(x) &= \int_G f(x)d\mu(x)\\ &= \int_G f(yx)d\mu(x)\\ &= \int_G f(yx)h(x)d\eta(x)\\ &= \int_G f(x)h(y^{-1}x)d\eta(x)\end{aligned}$$

car η et μ sont des mesures de Haar. On en déduit que, pour tout $y\in G$:

$$\int_G f(x)(h(x)-h(y^{-1}x))d\eta(x)=0$$

Or $h\in L^\infty(\eta)$ et $C(G)$ est dense dans $L^1(\eta)$ d'après le lemme 15 donc, pour tout $f\in L^1(\eta)$, pour tout $y\in G$:

$$\int_G f(x)(h(x)-h(y^{-1}x))d\eta(x)=0$$

Le dual de $L^1(\eta)$ étant $L^\infty(\eta)$, on a alors : pour tout $y\in G$, pour η-presque tout $x\in G$, $h(x)-h(y^{-1}x)=0$. D'après le théorème de Fubini-Tonelli, on a :

$$\int_G\left(\int_G |h(x)-h((x^{-1}y)^{-1})|d\eta(y)\right)d\eta(x)=\int_G\left(\int_G |h(x)-h(y^{-1}x)|d\eta(x)\right)d\eta(y)=0$$

Or η est une mesure de Haar donc, toujours d'après le théorème de Fubini-Tonelli :

$$\int_G\left(\int_G |h(x)-h(y^{-1})|d\eta(x)\right)d\eta(y)=\int_G\left(\int_G |h(x)-h(y^{-1})|d\eta(y)\right)d\eta(x)=0$$

Donc pour η-presque tout $y \in G$, on a : pour η-presque tout $x \in G$, $h(y^{-1}) = h(x)$.
Soit $y_0 \in G$ tel que pour η-presque tout $x \in G$, $h(x) = h(y_0^{-1})$. On pose $k = h(y_0^{-1})$. Pour η-presque tout $x \in G$, on a donc $h(x) = k$, d'où $d\mu = kd\eta$.
On a alors $d\nu = (1-k)d\eta$ donc $0 < k < 1$, d'où $d\mu = k(1-k)^{-1}d\nu$, ce qui assure le résultat recherché. □

A.3 Existence de la mesure de Haar

Montrons désormais l'existence d'une mesure de Haar sur tout groupe métrique compact. Celle-ci est une conséquence du théorème de Markov-Kakutani, qui nécessite plusieurs résultats topologiques.

Lemme 16. *Tout espace compact métrisable possède une suite dense, i.e tout espace métrique compact est séparable.*

Démonstration. Soit (K, d) un espace métrique compact.
Pour tout $m \in \mathbb{N}^*$, $\left\{B\left(x, \frac{1}{m}\right), x \in K\right\}$ est un recouvrement de K par des ouverts donc on peut en extraire un sous-recouvrement fini $\left\{B\left(x_{1,m}, \frac{1}{m}\right), \cdots, B\left(x_{n_m,m}, \frac{1}{m}\right)\right\}$.
On pose $A := \{x_{i,m}; i \in [|1, n_m|], m \in \mathbb{N}^*\}$.
A est une partie dénombrable de K puisque c'est une réunion dénombrable d'ensembles finis.
Montrons que A est dense dans K. Soit $x \in K$, $\varepsilon > 0$.
Il existe $m \in \mathbb{N}^*$ tel que $\frac{1}{m} < \varepsilon$, donc il existe $i \in [|1, n_m|]$ tel que $x \in B\left(x_{i,m}, \frac{1}{m}\right)$ avec $x_{i,m} \in A$.
On note que $d(x, x_{i,m}) < \frac{1}{m} < \varepsilon$ donc A est bien dense dans K. □

Proposition 5. *Soit Y un ensemble non vide tel que pour tout $y \in Y$, il existe $\mathscr{B}_y \subset \mathscr{P}(Y)$ vérifiant :*
(i) $\forall V \in \mathscr{B}_y$, $y \in V$;
(ii) $Y \in \mathscr{B}_y$;
(iii) $\forall U, V \in \mathscr{B}_y$, $\exists W \in \mathscr{B}_y; W \subset U \cap V$.
Alors il existe une unique topologie τ sur Y telle que pour tout $y \in Y$, $\mathscr{B}_y$ soit une base de voisinages pour τ (i.e : $\forall y \in Y, \forall U \in \tau, y \in U \implies \exists V \in \mathscr{B}_y; y \in V \subset U$).

Démonstration. Commençons par montrer l'existence. Pour cela, on pose :

$$\tau = \{\varnothing\} \cup \{O \in \mathscr{P}(Y); \forall y \in O, \exists U \in \mathscr{B}_y; U \subset O\}$$

Montrons que τ définit une topologie sur Y.
• $\varnothing \in \tau$ par définition de τ et $Y \in \tau$ par (ii).
• Soit $U, V \in \tau$. Si $U = \varnothing$ ou $V = \varnothing$, on a bien $U \cap V \in \tau$. Sinon, $U \neq \varnothing$ et $V \neq \varnothing$. Soit $y \in U \cap V$. Comme $U \in \tau$ et $V \in \tau$, il existe $O_1, O_2 \in \mathscr{B}_y$ tels que $O_1 \subset U$ et $O_2 \subset V$. Or, d'après (iii), il existe $O \in \mathscr{B}_y$ tel que $O \subset O_1 \cap O_2 \subset U \cap V$, d'où $U \cap V \in \tau$.
• Soit $(O_i)_{i \in I} \in \tau^I$. Si, pour tout $i \in I$, $O_i = \varnothing$, alors $\bigcup_{i \in I} O_i = \varnothing \in \tau$. Sinon, soit $y \in \bigcup_{i \in I} O_i$. Il existe $i \in I$ tel que $y \in O_i$. Or $O_i \in \tau$ donc il existe $U \in \mathscr{B}_y$ tel que $U \subset O_i \subset \bigcup_{i \in I} O_i$. D'où $\bigcup_{i \in I} O_i \in \tau$.
Donc τ définit bien une topologie sur Y. Montrons que pour tout $y \in Y$, $\mathscr{B}_y$ est une base de voisinages pour τ. Soit $y \in Y$, $U \in \tau$ tel que $y \in U$. Par définition de τ, il existe $O \in \mathscr{B}_y$ tel que $O \subset U$. De plus, d'après (i), $y \in O$ donc, pour tout $y \in Y$, $\mathscr{B}_y$ est en effet une base de voisinages pour τ.
Montrons désormais l'unicité.

Soit τ, τ' deux topologies sur Y admettant, pour tout $y \in Y$, $\mathscr{B}_y$ pour base de voisinages. Par symétrie, il suffit de montrer que $\tau \subset \tau'$. Soit donc $U \in \tau$.
Pour tout $y \in U$, comme $U \in \tau$, il existe $V \in \mathscr{B}_y$ tel que $y \in V \subset U$ d'où, comme B_y est une base de voisinage pour τ', U est voisinage de chacun de ses points. Donc $U \in \tau'$ et on a bien l'unicité recherchée. $\square$

Dans la suite, on considère un espace de Banach $(E, \|.\|)$.
Pour tout $x^* \in E^*$, on pose :

$$\mathcal{V}_{x^*} = \{V_{x^*,x_1,\cdots,x_n,\varepsilon} = \{y^* \in E^*; \forall i \in [|1,n|], |(x^*-y^*)(x_i)| < \varepsilon\}, n \in \mathbb{N}^*, \varepsilon > 0, (x_1, \cdots, x_n) \in E^n\}$$

D'après la proposition 5, il existe une unique topologie, appelée topologie préfaible et notée ω^* ou $\sigma(E^*, E)$ sur E^* telle que pour tout $x^* \in E^*$, $\mathcal{V}_{x^*}$ soit une base de voisinages pour $\sigma(E^*, E)$.
En effet, soit $x^* \in E^*$. Montrons que $\mathcal{V}_{x^*}$ vérifie (i), (ii) et (iii) de la proposition 5.
(i) est clair. De plus, (ii) est vérifié car :

$$E^* = \{y \in E^*; |(x^* - y^*)(0)| < 1\} \in \mathcal{V}_{x^*}$$

Finalement, montrons que (iii) est également vérifié. Soit $U_1, U_2 \in \mathcal{V}_{x^*}$. Il existe $(n_1, n_2) \in (\mathbb{N}^*)^2$, $(\varepsilon_1, \varepsilon_2) \in (\mathbb{R}^{+*})^2$, $((x_1, \cdots, x_{n_1}), (y_1, \cdots, y_{n_2})) \in E^{n_1} \times E^{n_2}$ tels que $U_1 = V_{x^*,x_1,\cdots,x_{n_1},\varepsilon_1}$ et $U_2 = V_{x^*,y_1,\cdots,y_{n_2},\varepsilon_2}$. On a alors :

$$V_{x^*,x_1,\cdots,x_{n_1},y_1,\cdots,y_{n_2},\min(\varepsilon_1,\varepsilon_2)} \subset U_1 \cap U_2$$

Donc (iii) est bien vérifiée.

Remarque 7. *Soit $(x_n^*)_{n\in\mathbb{N}} \subset E^*$, $x^* \in E^*$. On vérifie que :*

$$x_n^* \underset{n\to\infty}{\overset{\omega^*}{\longrightarrow}} x^* \Leftrightarrow \forall x \in E, x_n^*(x) \underset{n\to\infty}{\longrightarrow} x^*(x)$$

Cette topologie n'est jamais métrisable lorsque E est de dimension infinie. On a toutefois le théorème suivant :

Théorème 16 (Alaoglu-Bourbaki)**.** *Si E est séparable, alors la boule unité fermée de E^* munie de la topologie préfaible, (B_{E^*}, ω^*), est compacte métrisable.*

Démonstration. On suppose E séparable.
Montrons que (B_{E^*}, ω^*) est métrisable. Comme E est séparable, la sphère unité de E, $\mathcal{S}_E$, l'est également. On note alors $(x_n)_{n\in\mathbb{N}}$ une suite dense dans $\mathcal{S}_E$. On pose :

$$d : \left\{ \begin{array}{lll} B_{E^*} \times B_{E^*} & \to & \mathbb{R}^+ \\ (x^*, y^*) & \mapsto & \sum_{n\in\mathbb{N}} \frac{1}{2^n}|(x^* - y^*)(x_n)| \end{array} \right.$$

Montrons que d est bien définie et est une distance sur B_{E^*}.
On note que pour tout $(x^*, y^*) \in (B_{E^*})^2$, $d(x^*, y^*) \leq 4 < +\infty$, donc d est bien définie.
• Il est clair que, pour tout $(x^*, y^*) \in (B_{E^*})^2$, $d(x^*, y^*) = d(y^*, x^*)$.

• Soit $(x^*, y^*, z^*) \in (B_{E^*})^3$. On a :

$$\begin{aligned}
d(x^*, z^*) &= \sum_{n\in\mathbb{N}} \frac{1}{2^n}|(x^* - y^* + y^* - z^*)(x_n)| \\
&\leq \sum_{n\in\mathbb{N}} \frac{1}{2^n}(|(x^* - y^*)(x_n)| + |(y^* - z^*)(x_n)|) \\
&\leq \sum_{n\in\mathbb{N}} \frac{1}{2^n}|(x^* - y^*)(x_n)| + \sum_{n\in\mathbb{N}} \frac{1}{2^n}|(y^* - z^*)(x_n)| \\
&= d(x^*, y^*) + d(y^*, z^*)
\end{aligned}$$

Donc l'inégalité triangulaire est vérifiée.
• Soit $(x^*, y^*) \in (B_{E^*})^2$ tel que $d(x^*, y^*) = 0$.
Alors, pour tout $n \in \mathbb{N}$, $x^*(x_n) = y^*(x_n)$. Montrons que $x^* = y^*$. Soit $x \in E\backslash\{0\}$.
Comme $(x_n)_{n\in\mathbb{N}}$ une suite dense dans $\mathcal{S}_E$, il existe une sous-suite $(x_{n_k})_{k\in\mathbb{N}}$ de $(x_n)_{n\in\mathbb{N}}$ telle que :

$$\lim_{l\to+\infty} \left\| x_{n_l} - \frac{x}{\|x\|} \right\| = 0$$

On a alors :

$$x^*(x) = \|x\|x^*\left(\frac{x}{\|x\|}\right) = \lim_{l\to+\infty} \|x\|x^*(x_{n_l}) = \lim_{l\to+\infty} \|x\|y^*(x_{n_l}) = \|x\|y^*\left(\frac{x}{\|x\|}\right) = y^*(x)$$

D'où, comme $x^*(0) = 0 = y^*(0)$, on a bien $x^* = y^*$.
Finalement, d définit bien une distance sur B_{E^*}.
Il reste à montrer que la topologie que définit cette distance, notée ici τ_d est ω^* sur B_{E^*}. Pour cela, d'après la proposition 5, il suffit de montrer que pour tout $x^* \in B_{E^*}$, $\{A \cap B_{E^*}, A \in \mathcal{V}_{x^*}\}$ est une base de voisinage pour τ_d.
Soit $x^* \in B_{E^*}$, $\delta > 0$. Il suffit de montrer qu'il existe $V \in \mathcal{V}_{x^*}$ tel que $V \cap B_{E^*} \subset B(x^*, \delta)$.
On pose $\varepsilon = \frac{\delta}{8}$, $n \in \mathbb{N}$ tel que $\frac{1}{2^{n-1}} < \frac{\delta}{2}$. Montrons que $V_{x^*,x_0,\cdots,x_n,\varepsilon} \cap B_{E^*} \subset B(x^*, \delta)$.
Soit $z^* \in V_{x^*,x_0,\cdots,x_n,\varepsilon} \cap B_{E^*}$. On a :

$$\begin{aligned}
d(x^*, z^*) &= \sum_{i\in\mathbb{N}} \frac{1}{2^i}|(x^* - z^*)(x_i)| \\
&= \sum_{i=0}^{n} \frac{1}{2^i}|(x^* - z^*)(x_i)| + \sum_{i=n+1}^{\infty} \frac{1}{2^i}|(x^* - z^*)(x_i)| \\
&\leq \varepsilon \times 2 + \sum_{i=n+1}^{\infty} \frac{1}{2^i} \times 2 \text{ car } x^*, z^* \in B_{E^*} \\
&= \frac{\delta}{4} + \frac{1}{2^{n-1}} \\
&< \frac{\delta}{2} + \frac{\delta}{2} \leq \delta
\end{aligned}$$

D'où $V_{x^*,x_0,\cdots,x_n,\varepsilon} \cap B_{E^*} \subset B(x^*, \delta)$.
On en déduit que (B_{E^*}, ω^*) est métrisable. Montrons désormais que c'est un compact. Soit $(y_n)_{n\in\mathbb{N}} \subset B_{E^*}$.
Pour tout $k \in \mathbb{N}$, $(y_n(x_k))_{n\in\mathbb{N}}$ est une suite bornée de $\mathbb{R}$ (car pour tout $n \in \mathbb{N}$, $|y_n(x_k)| \leq \|y_n\|\|x_k\| \leq 1$) donc, à l'aide du procédé d'extraction diagonale, on construit $\varphi : \mathbb{N} \to \mathbb{N}$

strictement croissante, $(\lambda_k)_{k\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$ telles que, pour tout $k \in \mathbb{N}$, $\lim_{l\to+\infty} y_{\varphi(l)}(x_k) = \lambda_k$.
Soit $x \in \mathcal{S}_E$. Montrons que $(y_{\varphi(n)}(x))_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{R}$ complet.
Soit $\varepsilon > 0$. La suite $(x_k)_{k\in\mathbb{N}}$ est dense dans $\mathcal{S}_E$ donc il existe $k \in \mathbb{N}$ tel que $\|x - x_k\| \leq \varepsilon$. Or $(y_{\varphi(n)}(x_k))_{n\in\mathbb{N}}$ converge donc est de Cauchy donc il existe $N \in \mathbb{N}$ tel que pour tout $(p,q) \in \mathbb{N}^2$ tel que $p \geq N$, $q \geq N$, $|y_{\varphi(p)}(x_k) - y_{\varphi(q)}(x_k)| \leq \varepsilon$. On a alors, pour tout $(p,q) \in \mathbb{N}^2$ tel que $p \geq N$, $q \geq N$:

$$\begin{aligned}
|y_{\varphi(p)}(x) - y_{\varphi(q)}(x)| &= |y_{\varphi(p)}(x) - y_{\varphi(p)}(x_k) + y_{\varphi(p)}(x_k) - y_{\varphi(q)}(x_k) + y_{\varphi(q)}(x_k) - y_{\varphi(q)}(x)| \\
&\leq |y_{\varphi(p)}(x - x_k)| + |y_{\varphi(p)}(x_k) - y_{\varphi(q)}(x_k)| + |y_{\varphi(q)}(x_k - x)| \\
&\leq 1 \times \|x - x_k\| + |y_{\varphi(p)}(x_k) - y_{\varphi(q)}(x_k)| + 1 \times \|x - x_k\| \\
&\leq 3\varepsilon
\end{aligned}$$

Donc la suite $(y_{\varphi(n)}(x))_{n\in\mathbb{N}}$ est de Cauchy dans $\mathbb{R}$ complet donc il existe $\lambda(x) \in \mathbb{R}$ tel que $(y_{\varphi(n)}(x))_{n\in\mathbb{N}}$ converge vers $\lambda(x)$. En posant $\lambda(0) = 0$, comme pour tout $n \in \mathbb{N}$, $y_{\varphi(n)} \in B_{E^*}$, ceci définit une application linéaire $\lambda \in B_{E^*}$ telle que pour tout $x \in E$, $\lim_{l\to+\infty} y_{\varphi(l)}(x) = \lambda(x)$.
D'où $y_{\varphi(n)} \xrightarrow[n\to+\infty]{\omega^*} \lambda \in B_{E^*}$. On en déduit, puisque (B_{E^*}, ω^*) est métrisable, que (B_{E^*}, ω^*) est un compact. □

Lemme 17. *Soit (G, d) un groupe métrique compact.*
L'espace $E := (C(G,\mathbb{R}), \|.\|_\infty)$ des fonctions continues de G dans $\mathbb{R}$, muni de la norme uniforme, est séparable.

Démonstration. D'après le lemme 16, G est séparable donc il existe une suite $(g_n)_{n\in\mathbb{N}^*} \subset G$ dense dans G.
On pose alors $f_0 \equiv 1$ et, pour tout $n \in \mathbb{N}^*$:

$$f_n : \left\{ \begin{array}{lcl} G & \to & \mathbb{R} \\ x & \mapsto & d(x, g_n) \end{array} \right.$$

On pose également $A := \mathbb{R}[(f_n)_{n\in\mathbb{N}}]$ que l'on assimile à la sous-algèbre de E des fonctions polynômiales à coefficients réels en $(f_n)_{n\in\mathbb{N}}$. Puisque $\mathbb{Q}[(f_n)_{n\in\mathbb{N}}]$ est dénombrable et dense dans A par densité de $\mathbb{Q}$ dans $\mathbb{R}$, il suffit de montrer que A est dense dans E. On a :
(i) A sépare les points de G. En effet, soit $(x,y) \in G^2$ tel que $x \neq y$. Alors il existe $n \in \mathbb{N}^*$ tel que $f_n(x) \neq f_n(y)$ puisque sinon, comme $(g_n)_{n\in\mathbb{N}^*}$ est dense dans G, il existe $\varphi : \mathbb{N} \to \mathbb{N}$ telle que :

$$0 = d(x,x) = \lim_{l\to+\infty} d(x, g_{\varphi(l)}) = f_{\varphi(l)}(x) = f_{\varphi(l)}(y) = \lim_{l\to+\infty} d(y, g_{\varphi(l)}) = d(y,x) \neq 0$$

(ii) Pour tout $x \in G$, il existe $f \in A$ telle que $f(x) \neq 0$ puisque $f_0 \in A$.
Le théorème de Stone-Weierstrass assure alors, puisque G est compact, que A est dense dans E. Donc E est bien séparable. □

Théorème 17 (Markov-Kakutani)**.** *Soit E un espace de Banach séparable, $K \subset E^*$ une partie convexe, bornée et ω^*-fermée, $T : K \to K$ une application affine (i.e pour tout $(x,y) \in K^2$, pour tout $t \in [0,1]$, $T(tx + (1-t)y) = tT(x) + (1-t)T(y)$), et $\omega^* - \omega^*$-continue.*
T admet un point fixe dans K.

Démonstration. Dans la preuve, K sera muni exclusivement de la topologie préfaible.
Soit $x_0 \in K$. Pour tout $n \in \mathbb{N}$, on pose :

$$x_n := \frac{1}{n+1} \sum_{i=0}^{n} T^i(x_0)$$

Comme K est convexe, pour tout $n \in \mathbb{N}$, $x_n \in K$. Or K est borné donc il existe $\lambda > 0$ tel que $K \subset \lambda B_{E^*}$, d'où, comme K est ω^*-fermée, (K, ω^*) est compact métrisable d'après le théorème 16 puisque E est séparable. On en déduit qu'il existe une sous-suite $(x_{n_k})_{k\in\mathbb{N}}$ de $(x_n)_{n\in\mathbb{N}}$, $x \in K$ tels que :

$$x_{n_k} \xrightarrow[k\to\infty]{\omega^*} x$$

Montrons que $T(x) = x$. Soit $v \in E$.
Comme $x_{n_k} \xrightarrow[k\to\infty]{\omega^*} x$, $\lim\limits_{l\to+\infty} x_{n_l}(v) = x(v)$.
De plus, comme K est borné, $M := \sup_{y\in K} |y(v)|$ est fini. On a alors, pour tout $k \in \mathbb{N}$:

$$\begin{aligned}
|(T(x_{n_k}) - x_{n_k})(v)| &= \frac{1}{n_k+1} \left| \left(\sum_{i=0}^{n_k} T^{i+1}(x_0) - \sum_{i=0}^{n_k} T^i(x_0) \right)(v) \right| \\
&= \frac{1}{n_k+1} |T^{n_k+1}(x_0)(v) - x_0(v)| \\
&\leq \frac{2M}{n_k+1}
\end{aligned}$$

D'où : $\lim\limits_{l\to+\infty} (T(x_{n_l}) - x_{n_l})(v) = 0$. Or T est continue (pour la topologie préfaible) donc, pour tout $v \in E$:

$$(T(x) - x)(v) = \lim_{l\to+\infty} (T(x_{n_l}) - x_{n_l})(v) = 0$$

Finalement, on a bien $T(x) = x$, *i.e* T possède un point fixe dans K. □

Théorème 18. *Soit G un groupe métrique compact.*
Il existe une unique mesure de Haar normalisée μ sur G. De plus, μ est invariante par translation à droite.

Démonstration. C'est une conséquence du théorème de Markov-Kakutani.
On pose $E := (C(G, \mathbb{R}), \|.\|_\infty)$, $K := Prob(G)$ l'ensemble des mesures de probabilité sur $(G, \mathscr{B}(G))$.
Soit $\mu \in K$. On peut voir μ comme un élément du dual en voyant que μ définit une forme linéaire positive, encore notée $\mu \in E^*$:

$$\left\{ \begin{array}{lll} E & \to & \mathbb{R} \\ f & \mapsto & \int_G f d\mu \end{array} \right.$$

qui vérifie également $\mu(\mathbb{1}_G) = 1$, ce qui assure qu'avec cette identification, $K \subset B_{E^*}$.
On note que K est convexe. Montrons que c'est un fermé de (E^*, ω^*).
Soit $(\mu_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$, $\phi \in E^*$ telles que $\mu_n \xrightarrow[n\to+\infty]{\omega^*} \phi$.
Pour toute $f \in E$ positive, pour tout $n \in \mathbb{N}$, $\mu_n(f) \geq 0$ donc, par passage à la limite, $\phi(f) \geq 0$.
Donc ϕ est une forme linéaire positive, vérifiant $\phi(\mathbb{1}_G) = 1$ puisque pour tout $n \in \mathbb{N}$, $\mu_n(\mathbb{1}_G) = 1$.
Le théorème de Riesz assure alors l'existence d'une unique mesure μ régulière finie sur $\mathscr{B}(G)$

telle que pour toute $f \in E$, $\phi(f) = \int_G f d\mu$. D'après l'identification faite plus haut, on a alors $\phi = \mu$, et puisque $\phi(\mathbb{1}_G) = 1$, ϕ est une mesure de probabilité, *i.e* $\phi \in K$.
Donc K est bien fermé, puisque, comme E est séparable d'après le lemme 17, B_{E^*} est métrisable d'après le théorème 16.
D'où, comme B_{E^*} est compacte d'après le théorème 16, K est une partie convexe, ω^*-compacte de E^*.
Pour tout $g \in G$, on pose :

$$T_g : \left\{ \begin{array}{lll} K & \to & K \\ \mu & \mapsto & \left\{ \begin{array}{lll} E & \to & \mathbb{R} \\ f & \mapsto & \int_G \tau_g(f) d\mu = \int_G f(gx) d\mu(x) \end{array} \right. \end{array} \right.$$

Pour tout $g \in G$, T_g est bien définie car, pour tout $\mu \in K$, $T_g(\mu)$ est une forme linéaire positive et que, pour toute $\mu \in K$, $T_g(\mu)(\mathbb{1}_G) = 1$, donc en faisant le même raisonnement que précédemment, le théorème de Riesz assure que $T_g(\mu) \in K$.
Soit $g \in G$. Montrons que T_g est ω^*-continue. Pour cela, on utilise le critère séquentiel de continuité puisque, d'après le théorème 16 et le lemme 17, (K, ω^*) est métrisable.
Soit $(\mu_n)_{n\in\mathbb{N}} \in K^{\mathbb{N}}$, $\mu \in K$ telles que $\mu_n \underset{n\to+\infty}{\overset{\omega^*}{\longrightarrow}} \mu$. Soit $(f_1, \cdots, f_k) \in E^k$, $\varepsilon > 0$. Montrons qu'il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$:

$$T_g(\mu_n) \in V_{T_g(\mu), f_1, \cdots, f_k, \varepsilon} \cap K = \{\nu \in K; \forall i \in [|1, k|], |(T_g(\mu) - \nu)(f_i)| < \varepsilon\}$$

Pour tout $i \in [|1, k|]$, on pose $F_i : \left\{ \begin{array}{lll} G & \to & \mathbb{R} \\ x & \mapsto & f_i(gx) \end{array} \right. \in E$.
Comme $\mu_n \underset{n\to+\infty}{\overset{\omega^*}{\longrightarrow}} \mu$, il existe $n_0 \in \mathbb{N}$ tel que pour tout $n \geq n_0$, $\mu_n \in V_{\mu, F_1, \cdots, F_k, \varepsilon} \cap K$. On en déduit que, pour tout $n \geq n_0$, pour tout $i \in [|1, k|]$, on a :

$$|(T_g(\mu) - T_g(\mu_n))(f_i)| = \left| \int_G f_i(gx) d\mu(x) - \int_G f_i(gx) d\mu_n(x) \right| = |(\mu - \mu_n)(F_i)| < \varepsilon$$

D'où, pour tout $n \geq n_0$, $T_g(\mu_n) \in V_{T_g(\mu), f_1, \cdots, f_k, \varepsilon} \cap K$.
On en déduit donc que, pour tout $g \in G$, T_g est ω^*-continue.
On note que $\mu \in K$ est une mesure de Haar si et seulement si pour tout $g \in G$, $T_g(\mu) = \mu$.
D'après le lemme 16, il existe $(g_n)_{n\in\mathbb{N}^*}$ une suite dense d'éléments de G. On pose :

$$S : \left\{ \begin{array}{lll} K & \to & K \\ \mu & \mapsto & \sum_{n\in\mathbb{N}^*} \frac{1}{2^n} T_{g_n}(\mu) \end{array} \right.$$

On note que S est bien définie puisque $\sum_{n\in\mathbb{N}^*} \frac{1}{2^n} = 1$, qu'elle est affine (car pour tout $n \in \mathbb{N}^*$, T_{g_n} l'est) et continue pour la topologie préfaible sur K, puisque, comme pour tout $n \in \mathbb{N}$, $\|T_{g_n}\| \leq 1$, S est la limite uniforme de la suite de fonctions continues $\left(\sum_{n=0}^N \frac{1}{2^n} T_{g_n}(\mu)\right)_{N\in\mathbb{N}^*}$.
D'après le théorème de Markov-Kakutani, on a alors l'existence d'un point fixe pour S, *i.e* il existe $\mu \in K$ telle que :

$$\mu = S(\mu) = \sum_{n\in\mathbb{N}^*} \frac{1}{2^n} T_{g_n}(\mu)$$

ou encore pour toute $f \in E$:

$$\int_G f d\mu = \sum_{n\in\mathbb{N}^*} \frac{1}{2^n} \int_G f(g_n x) d\mu(x) (*)$$

Montrons que μ est une mesure de Haar. Soit $f \in E$. On pose :

$$F : \left\{ \begin{array}{lll} G & \to & \mathbb{R} \\ g & \mapsto & T_g(\mu)(f) = \int_G f(gx)d\mu(x) \end{array} \right.$$

Comme f est uniformément continue, $F \in E$. De plus, par $(*)$, pour tout $g \in G$, on a :

$$\begin{aligned} \sum_{n\in\mathbb{N}^*} \frac{1}{2^n} F(gg_n) &= \sum_{n\in\mathbb{N}^*} \frac{1}{2^n} \int_G f(gg_nx)d\mu(x) \\ &= \sum_{n\in\mathbb{N}^*} \frac{1}{2^n} \int_G (\tau_g(f))(g_nx)d\mu(x) \\ &= \int_G \tau_g(f)d\mu \text{ car } \tau_g(f) \in E \\ &= \int_G f(gx)d\mu(x) \\ &= F(g) \end{aligned}$$

Or F est une fonction continue sur le compact G donc admet un maximum en un point $g_0 \in G$, en lequel on a donc :

$$\sum_{n\in\mathbb{N}^*} \frac{1}{2^n} F(g_0g_n) = F(g_0)$$

D'où, puisque $\sum_{n\in\mathbb{N}^*} \dfrac{1}{2^n} = 1$ et que pour tout $n \in \mathbb{N}^*$, $F(g_0g_n) \leq F(g_0)$, on a que pour tout $n \in \mathbb{N}^*$, $F(g_0g_n) = F(g_0)$.
Or $(g_n)_{n\in\mathbb{N}^*}$ est dense dans G et F est continue donc pour tout $g \in G$, $F(g_0g) = F(g_0)$, *i.e* , comme G est un groupe, F est constante sur G. Ce qui signifie que, pour tout $g \in G$, on a :

$$\int_G f(gx)d\mu(x) = F(g) = F(e) = \int_G f(x)d\mu(x)$$

i.e μ est invariante par translation à gauche.
Posons K' l'ensemble des mesures de probabilités sur G qui sont invariantes à gauche. K' est un sous ensemble convexe et compact de K, non vide d'après ce qui précède. On montre par le même raisonnement qu'il existe $\tilde{\mu} \in K'$ invariante par translation à droite.
Montrons désormais l'unicité de $\tilde{\mu}$ comme mesure invariante par translation à gauche.
Soit $\nu \in K$ une autre mesure de probabilité invariante par translation à gauche.
Pour toute $f \in E$, on a, puisque $\tilde{\mu}$ est invariante à droite :

$$\int_G f(x)d\tilde{\mu}(x) = \int_G \int_G f(x)d\tilde{\mu}(x)d\nu(y) = \int_G \int_G f(xy)d\tilde{\mu}(x)d\nu(y)$$

et, puisque ν est invariante par translation à gauche, d'après le théorème de Fubini, on a :

$$\int_G f(x)d\tilde{\mu}(x) = \int_G \int_G f(xy)d\nu(y)d\tilde{\mu}(x) = \int_G \int_G f(y)d\nu(y)d\tilde{\mu}(x) = \int_G f(y)d\nu(y)$$

D'où $\nu = \tilde{\mu}$. $\square$

B Autour de la mesure sur la sphère

B.1 Existence d'une mesure invariante par rotation sur la sphère

On note $\mathcal{S}^{n-1} := \{u \in \mathbb{R}^n; \|u\|_2 = 1\}$ la sphère unité de $\mathbb{R}^n$ muni de sa norme euclidienne usuelle $\|.\|_2$, et pour tout borélien A de $\mathcal{S}^{n-1}$, on pose :

$$\sigma^{n-1}(A) := n.\lambda^{(n)}(C_1(A))$$

où $C_a(A) := \{t.u; 0 < t < a, u \in A\}$ est le cône sur A de rayon $a > 0$.

Proposition 6. *σ^{n-1} est une mesure borélienne sur $\mathcal{S}^{n-1}$.*
De plus, cette mesure est invariante par rotation, i.e pour tout $Q \in O_n(\mathbb{R})$, pour tout borélien $A \subset \mathcal{S}^{n-1}$, $\sigma^{n-1}(Q(A)) = \sigma^{n-1}(A)$.

Démonstration. Montrons que σ^{n-1} est une mesure borélienne sur $\mathcal{S}^{n-1}$.
• pour tout borélien $A \subset \mathcal{S}^{n-1}$, $C_1(A)$ est un borélien de $\mathbb{R}$ et $\lambda^{(n)}(A) \in \mathbb{R}^+$ donc $\sigma^{n-1}(A) \in \mathbb{R}^+$;
• $\sigma^{n-1}(\varnothing) = n\lambda^{(n)}(\varnothing) = 0$;
• Soit $(A_k)_{k\in\mathbb{N}} \in \mathscr{B}(\mathcal{S}^{n-1})$ des boréliens deux à deux disjoints. Les $C_1(A_n)$, $n \in \mathbb{N}$, sont alors également deux à deux disjoints, d'où :

$$\sigma^{n-1}\left(\bigcup_{k\in\mathbb{N}} A_k\right) = n.\lambda^{(n)}\left(C_1\left(\bigcup_{k\in\mathbb{N}} A_k\right)\right) = n.\lambda^{(n)}\left(\bigcup_{k\in\mathbb{N}} C_1(A_k)\right) = \sum_{k\in\mathbb{N}} n.\lambda^{(n)}(C_1(A_k)) = \sum_{k\in\mathbb{N}} \sigma^{n-1}(A_k)$$

donc σ^{n-1} est bien une mesure borélienne sur $\mathcal{S}^{n-1}$.
Soit $Q \in O_n(\mathbb{R})$, $A \in \mathscr{B}(\mathcal{S}^{n-1})$. On a :

$$\sigma^{n-1}(Q(A)) = n.\lambda^{(n)}(C_1(Q(A))) = n|\det(Q)|\lambda^{(n)}(C_1(A)) = \sigma^{n-1}(A)$$

donc σ^{n-1} est bien invariante par rotation. □

Notons ϕ l'homéomorphisme suivant :

$$\phi : \begin{cases} \mathbb{R}^{+*} \times \mathcal{S}^{n-1} & \to & \mathbb{R}^n \backslash \{0\} \\ (r, u) & \mapsto & r.u \end{cases}$$

Le mesure de Lebesgue et la mesure sphérique sont reliées par la formule suivante :

Théorème 19.
$$\forall A \in \mathscr{B}(\mathbb{R}^n), \lambda^{(n)}(A) = (r^{n-1}dr \otimes d\sigma^{n-1})(\phi^{-1}(A))$$

Démonstration. On rappelle qu'un π-système de $\mathbb{R}^n$ est une famille de parties de $\mathbb{R}^n$ contenant $\mathbb{R}^n$ et stable par intersection finie. D'après un résultat de théorie de la mesure, il suffit de montrer que cette formule est vraie pour tous les éléments du π-système $\Pi = \{C_a(U) := \{t \cdot u; 0 < t < a, u \in U\}, a \in \mathbb{R}^+, U \subset \mathcal{S}^{n-1}$ ouvert$\}$ qui engendre $\mathscr{B}(\mathbb{R}^n)$.
Soit $a \in \mathbb{R}^+$, $U \subset \mathcal{S}^{n-1}$ ouvert. On a :

$$\lambda^{(n)}(C_a(U)) = \lambda^{(n)}(aC_1(U)) = a^n\lambda^{(n)}(C_1(U)) = \frac{a^n}{n}\sigma^{n-1}(U)$$

Or $\frac{a^n}{n} = \int_0^a r^{n-1}dr$ donc, d'après le théorème de Fubini-Tonelli, on a :

$$\begin{aligned}
\lambda^{(n)}(C_a(U)) &= \int_0^a r^{n-1}dr \int_U d\sigma^{n-1} \\
&= \int_{]0,a[\times U} r^{n-1}dr \otimes d\sigma^{n-1} \\
&= (r^{n-1}dr \otimes d\sigma^{n-1})(]0,a[\times U) \\
&= (r^{n-1}dr \otimes d\sigma^{n-1})(\phi^{-1}(C_a(U)))
\end{aligned}$$

Ce qui assure le résultat. □

Du théorème précédent et du théorème de Fubini, on déduit le corollaire suivant :

Corollaire 2. *Pour toute $f \in \mathscr{L}^1(\mathbb{R}^n)$, on a :*

$$\begin{aligned}
\int_{\mathbb{R}^n} f d\lambda^{(n)} &= \int_0^{+\infty} \int_{\mathcal{S}^{n-1}} f(r.u) r^{n-1} d\sigma^{n-1}(u) dr \\
&= \int_{\mathcal{S}^{n-1}} \int_0^{+\infty} f(r.u) r^{n-1} dr d\sigma^{n-1}(u)
\end{aligned}$$

B.2 Indépendance de $\|G\|_2$ et $\frac{G}{\|G\|_2}$

Pour démontrer le théorème de Dvoretzky, nous avons eu besoin de l'indépendance des vecteurs aléatoires $\|G\|_2$ et $\frac{G}{\|G\|_2}$ où $G = (g_i)_{i=1}^n$ avec les g_i, $i \in [|1, n|]$, i.i.d de loi normale centrée réduite, que l'on démontre ici tout d'abord en identifiant les lois respectives de $\|G\|_2$ et $\frac{G}{\|G\|_2}$, puis en revenant à la définition de vecteurs aléatoires indépendants.

Lemme 18. $\|G\|_2$ *admet pour densité la fonction :*

$$x \mapsto x^{n-1} \sigma^{n-1}(\mathcal{S}^{n-1}) \frac{e^{-\frac{x^2}{2}}}{(2\pi)^{\frac{n}{2}}} \mathbb{1}_{[0,+\infty[}(x)$$

Démonstration. Pour toute $f : \mathbb{R} \to \mathbb{R}$ borélienne bornée, on a :

$$\begin{aligned}
\mathbb{E}(f(\|G\|_2)) &= \int_{\mathbb{R}^n} f(\|x\|_2) \frac{e^{-\frac{\|x\|_2^2}{2}}}{(2\pi)^{\frac{n}{2}}} d\lambda^{(n)}(x) \\
&= \int_0^{+\infty} \int_{\mathcal{S}^{n-1}} f(\alpha) \frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}} \alpha^{n-1} d\sigma^{n-1}(u) d\alpha \\
&= \int_0^{+\infty} f(\alpha) \frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}} \alpha^{n-1} \sigma^{n-1}(\mathcal{S}^{n-1}) d\alpha
\end{aligned}$$

d'où le résultat. □

Lemme 19. $\frac{G}{\|G\|_2}$ *admet pour loi* $\frac{1}{\sigma^{n-1}(\mathcal{S}^{n-1})}\sigma^{n-1}$.

Démonstration. Pour toute $f : \mathbb{R}^n \to \mathbb{R}$ borélienne bornée, on a :

$$\begin{aligned}\mathbb{E}\left(f\left(\frac{G}{\|G\|_2}\right)\right) &= \int_{\mathbb{R}^n} f\left(\frac{x}{\|x\|_2}\right)\frac{e^{-\frac{\|x\|_2^2}{2}}}{(2\pi)^{\frac{n}{2}}}d\lambda^{(n)}(x) = \int_0^{+\infty}\int_{\mathcal{S}^{n-1}} f(u)\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}d\sigma^{n-1}(u)d\alpha \\ &= \int_{\mathcal{S}^{n-1}} f(u)\left(\int_0^{+\infty}\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}d\alpha\right)d\sigma^{n-1}(u)\end{aligned}$$

Or, on a :

$$\begin{aligned}\int_0^{+\infty}\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}d\alpha &= \int_0^{+\infty}\int_{\mathcal{S}^{n-1}}\frac{e^{-\frac{\|\alpha u\|_2^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}\frac{d\sigma^{n-1}(u)}{\sigma^{n-1}(\mathcal{S}^{n-1})}d\alpha \\ &= \int_{\mathbb{R}^n}\frac{e^{-\frac{\|x\|_2^2}{2}}}{(2\pi)^{\frac{n}{2}}}\frac{d\lambda^{(n)}(x)}{\sigma^{n-1}(\mathcal{S}^{n-1})} \\ &= \frac{1}{\sigma^{n-1}(\mathcal{S}^{n-1})}\end{aligned}$$

D'où :

$$\mathscr{L}\left(\frac{G}{\|G\|_2}\right) = \frac{1}{\sigma^{n-1}(\mathcal{S}^{n-1})}\sigma^{n-1}$$

□

Théorème 20. *Les variables aléatoires $\|G\|_2$ et $\frac{G}{\|G\|_2}$ sont indépendantes.*

Démonstration. Soit $r \in \mathbb{R}^+$, $A \in \mathscr{B}(\mathcal{S}^{n-1})$. Il suffit de montrer que :

$$\mathbb{P}\left(\|G\|_2 \leq r, \frac{G}{\|G\|_2} \in A\right) = \mathbb{P}(\|G\|_2 \leq r)\mathbb{P}\left(\frac{G}{\|G\|_2} \in A\right)$$

Faisons-le :

$$\begin{aligned}\mathbb{P}\left(\|G\|_2 \leq r, \frac{G}{\|G\|_2} \in A\right) &= \mathbb{P}(G \in C_r(A)) \\ &= \int_{\mathbb{R}^n}\mathbb{1}_{C_r(A)}(x)\frac{e^{-\frac{\|x\|_2^2}{2}}}{(2\pi)^{\frac{n}{2}}}d\lambda^{(n)}(x) \\ &= \int_0^{+\infty}\int_{\mathcal{S}^{n-1}}\mathbb{1}_{C_r(A)}(\alpha.u)\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}d\sigma^{n-1}(u)d\alpha \\ &= \int_0^{+\infty}\int_{\mathcal{S}^{n-1}}\mathbb{1}_{[0,r]}(\alpha)\mathbb{1}_A(u)\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}d\sigma^{n-1}(u)d\alpha \\ &= \left(\int_0^{+\infty}\mathbb{1}_{[0,r]}(\alpha)\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}d\alpha\right)\left(\int_{\mathcal{S}^{n-1}}\mathbb{1}_A(u)d\sigma^{n-1}(u)\right) \\ &= \left(\int_0^{r}\frac{e^{-\frac{\alpha^2}{2}}}{(2\pi)^{\frac{n}{2}}}\alpha^{n-1}\sigma^{n-1}(\mathcal{S}^{n-1})d\alpha\right)\left(\int_A\frac{d\sigma^{n-1}(u)}{\sigma^{n-1}(\mathcal{S}^{n-1})}\right) \\ &= \mathbb{P}(\|G\|_2 \leq r)\mathbb{P}\left(\frac{G}{\|G\|_2} \in A\right)\end{aligned}$$

D'où le résultat. □

B.3 Unicité de cette mesure sur la sphère

Commençons par établir un résultat d'unicité :

Théorème 21. *Soit μ et σ deux mesures de probabilités sur la sphère de $\mathbb{R}^n$ munie de sa tribu borélienne $\mathscr{B}(\mathcal{S}^{n-1})$, invariantes par les rotations.*
Alors $\mu = \sigma$.

Démonstration. On note $\|.\|$ la norme euclidienne usuelle sur $\mathbb{R}^n$. D'après le cours de probabilités, il suffit de montrer que les fonctions caractéristiques de ces deux probabilités, notées respectivement $\tilde{\mu}$ et $\tilde{\sigma}$ sont égales.
Par hypothèse, pour toute rotation ϕ, pour tout $\xi \in \mathbb{R}^n$, on a :

$$\begin{aligned}\tilde{\mu}(\xi) &= \int_{\mathbb{R}^n} e^{i\xi\cdot x} d\mu(x) = \int_{\mathbb{R}^n} e^{i\xi\cdot\phi^{-1}(x)} d\mu(x) \\ &= \int_{\mathbb{R}^n} e^{i\phi(\xi)\cdot x} d\mu(x) = \tilde{\mu}(\phi(\xi))\end{aligned}$$

Or, pour tout $\xi \in \mathbb{R}^n$, il existe une rotation R_ξ telle que $R_\xi(\xi) = (\|\xi\|, 0, \cdots, 0)$. Donc, pour tout $\xi \in \mathbb{R}^n$, on a :

$$\tilde{\mu}(\xi) = \tilde{\mu}(R_\xi(\xi)) = \tilde{\mu}((\|\xi\|, 0, \cdots, 0))$$

De même, pour tout $\xi \in \mathbb{R}^n$, $\tilde{\sigma}(\xi) = \tilde{\sigma}((\|\xi\|, 0, \cdots, 0))$.
Or, pour tout $r > 0$, on a :

$$\begin{aligned}\tilde{\sigma}((r, 0, \cdots, 0)) &= \int_{\mathbb{R}^n} \tilde{\sigma}((r\|x\|, 0, \cdots, 0)) d\mu(x) \text{ car } Supp(\mu) \subset \mathcal{S}^{n-1} \\ &= \int_{\mathbb{R}^n} \tilde{\sigma}(rx) d\mu(x) = \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{i(rx)\cdot\xi} d\sigma(\xi) d\mu(x) \\ &= \int_{\mathbb{R}^n}\int_{\mathbb{R}^n} e^{ix\cdot(r\xi)} d\mu(x) d\sigma(\xi) \text{ par Fubini} \\ &= \int_{\mathbb{R}^n} \tilde{\mu}(r\xi) d\sigma(\xi) = \int_{\mathbb{R}^n} \tilde{\mu}((r\|\xi\|, 0, \cdots, 0)) d\sigma(\xi) \\ &= \int_{\mathbb{R}^n} \tilde{\mu}((r, 0, \cdots, 0)) d\sigma(\xi) \text{ car } Supp(\sigma) \subset \mathcal{S}^{n-1} \\ &= \tilde{\mu}((r, 0, \cdots, 0))\end{aligned}$$

D'où, comme $\tilde{\mu}(0) = 1 = \tilde{\sigma}(0)$, pour tout $\xi \in \mathbb{R}^n$:

$$\tilde{\mu}(\xi) = \tilde{\mu}((\|\xi\|, 0, \cdots, 0)) = \tilde{\sigma}((\|\xi\|, 0, \cdots, 0)) = \tilde{\sigma}(\xi)$$

On a donc bien $\mu = \sigma$. □

Notons $\sigma_n = \frac{\sigma^{n-1}}{\sigma^{n-1}(\mathcal{S}^{n-1})}$. C'est, d'après le théorème 21, l'unique mesure de probabilité sur $\mathcal{S}^{n-1}$ invariante par rotation. On a la formule suivante :

Proposition 7. *Soit μ la mesure de Haar (normalisée) du groupe orthogonal O_n.*
Pour toute $f \in \mathcal{C}(\mathcal{S}^{n-1})$, pour tout $\xi_0 \in \mathcal{S}^{n-1}$, on a :

$$\int_{\mathcal{S}^{n-1}} f(\xi) d\sigma_n(\xi) = \int_{O_n} f(U\xi_0) d\mu(U)$$

Démonstration. Soit $\xi_0 \in \mathcal{S}^{n-1}$. On pose :

$$\phi : \left\{ \begin{array}{lcl} \mathcal{C}(\mathcal{S}^{n-1}) & \to & \mathbb{R} \\ f & \mapsto & \int_{O_n} f(U\xi_0) d\mu(U) \end{array} \right.$$

C'est une forme linéaire positive sur $\mathcal{C}(\mathcal{S}^{n-1})$ avec $\mathcal{S}^{n-1}$ compact donc, d'après le théorème de Riesz, il existe une unique mesure régulière finie ν sur $\mathscr{B}(\mathcal{S}^{n-1})$ telle que :

$$\forall f \in \mathcal{C}(\mathcal{S}^{n-1}), \int_{O_n} f(U\xi_0) d\mu(U) = \phi(f) = \int_{\mathcal{S}^{n-1}} f(\xi) d\nu(\xi)$$

Or, d'après le lemme 15, pour toute rotation ϕ de $\mathbb{R}^n$, pour tout $A \in \mathscr{B}(\mathcal{S}^{n-1})$, on a :

$$\nu(\phi(A)) = \int_{O_n} \mathbb{1}_{\phi(A)}(U\xi_0) d\mu(U) = \int_{O_n} \mathbb{1}_A(M^{-1}U\xi_0) d\mu(U) = \int_{O_n} \mathbb{1}_A(V\xi_0) d\mu(V) = \nu(A)$$

où M est la matrice de ϕ entre les bases canoniques de $\mathbb{R}^n$, car μ est une mesure de Haar ; et $\nu(\mathcal{S}^{n-1}) = \mu(O_n) = 1$ donc, d'après le théorème 21, $\nu = \sigma_n$.
On a donc bien le résultat. $\square$

Bibliographie :

— Albiac F., Kalton N. (2016). Topics in Banach Space Theory, Second Edition.

— Brezis H. (1999). Analyse fonctionnelle : Théorie et applications.

— Rudin W. (1998). Analyse réelle et complexe, Troisième édition.

— Li D., Queffélec H. (2004). Introduction à l'étude des espaces de Banach, Analyse et probabilités.

— Milman V., Schechtman G. (2001). Asymptotic Theory of Finite Dimensional Normed Spaces.

— Simon B. (2015). Real Analysis, A Comprehensive Course in Analysis, Part 1.

— Daniel J., Puchol M. (2009). Mesures sur la sphère et harmoniques sphériques.

— Troyanov M. (2008). Mesures et Intégration.

— Briane M., Pagès G. (2006). Théorie de l'intégration : cours et exercices : licence et master de mathématiques