Solution du sujet de Mathématiques Générales, Cachan 3A 2009

Avertissement: Je tiens à préciser que je ne suis pas lié à l'École Normale Supérieure de Cachan; par suite les affirmations vraies ou fausses contenues dans ces pages ne sauraient engager l'École.

Olivier Garet, le 22 avril 2009

I Théorème ergodique en moyenne de Von Neumann

1. On va montrer que quels que soient $A, B, C \in \mathcal{M}(\mu)$, on a

$$\mu(A\Delta C) \le \mu(A\Delta B) + \mu(B\Delta C).$$

En effet, d'après l'inégalité triangulaire,

$$|1\!\!1_A - 1\!\!1_C| \le |1\!\!1_A - 1\!\!1_B| + |1\!\!1_B - 1\!\!1_C|,$$

soit

$$1\!\!1_{A\Delta C} \le 1\!\!1_{A\Delta B} + 1\!\!1_{B\Delta C}.$$

Il suffit alors d'intégrer par rapport à μ pour conclure.

2. $$\mu(T^{-2}(A)\Delta A) \le \mu(T^{-2}(A)\Delta T^{-1}(A)) + \mu(T^{-1}(A)\Delta A).$$

Cependant $T^{-2}(A)\Delta T^{-1}(A) = T^{-1}(T^{-1}(A)\Delta A))$. Comme T préserve μ, $\mu(T^{-2}(A)\Delta T^{-1}(A)) = \mu(T^{-1}(A)\Delta A))$. Ainsi

$$\mu(T^{-2}(A)\Delta A) \le 2\mu(T^{-1}(A)\Delta A) = 0,$$

d'où $\mu(T^{-2}(A)\Delta A) = 0$.

3. L'égalité est établie pour $n = 0, 1$. Maintenant

$$\mu(T^{-(n+1)}(A)\Delta A) \le \mu(T^{-(n+1)}(A)\Delta T^{-n}(A)) + \mu(T^{-n}(A)\Delta A).$$

On a $T^{-(n+1)}(A)\Delta T^{-n}(A) = T^{-n}(T^{-1}(A)\Delta A)$. Comme T préserve μ, ses itérées aussi, d'où $\mu(T^{-(n+1)}(A)\Delta T^{-n}(A)) = \mu(T^{-1}(A)\Delta A) = 0$. Ainsi, $\mu(T^{-(n+1)}(A)\Delta A) \le \mu(T^{-n}(A)\Delta A)$ et la récurrence est alors évidente.

4. Remarquons que quelque soient les ensembles $A, (B_i)_{i\in\mathbb{N}}$, on a l'inclusion

$$A\Delta(\cup_{i\in\mathbb{N}}B_i) \subset \cup_{i\in\mathbb{N}}A\Delta B_i.$$

Cela se voit aisément en passant au complémentaire: si pour tout entier i, $1\!\!1_A(\omega) = 1\!\!1_{B_i}(\omega)$, alors $1\!\!1_A(\omega) = \max_{i\in\mathbb{N}}(1\!\!1_{B_i}(\omega)) = 1\!\!1_{\cup_i A_i}(\omega)$. Par sous-additivité, on aura donc

$$\mu(A\Delta(\cup_{i\in\mathbb{N}}B_i)) \le \sum_{i\in\mathbb{N}} \mu(A\Delta B_i).$$

En particulier, avec $B_i = T^{-(n+i+1)}(A)$, cela donne

$$\mu(A\Delta(\cup_{n\geq k+1}T^{-n}(A))) \leq \sum_{n\geq k+1} \mu(A\Delta T^{-n}(A)) = 0.$$

Remarquons encore que quelque soient les ensembles $A, (B_i)_{i\in\mathbb{N}}$, on a

$$A\Delta(\cap_{i\in\mathbb{N}}B_i) \subset \cup_{i\in\mathbb{N}}A\Delta B_i.$$

Là encore, cela se voit aisément en passant au complémentaire: si pour i, $1\!\!1_A(\omega) = 1\!\!1_{B_i}(\omega)$, alors $1\!\!1_A(\omega) = \inf_{i\in\mathbb{N}}(1\!\!1_{B_i}(\omega)) = 1\!\!1_{\cap_i A_i}(\omega)$. Par sous-additivité, on aura donc

$$\mu(A\Delta(\cap_{i\in\mathbb{N}}B_i)) \leq \sum_{i\in\mathbb{N}} \mu(A\Delta B_i).$$

Cela donne le résultat voulu en prenant $B_i = \underset{n=i+2}{\overset{+\infty}{\cup}} T^{-n}(A)$.

5.

$$\begin{aligned}
\hat{B} &= \underset{k=1}{\overset{+\infty}{\cap}} T^{-k}(\underset{n=1}{\overset{+\infty}{\cup}} T^{-n}(B)) \\
&= \underset{k=1}{\overset{+\infty}{\cap}} (\underset{n=1}{\overset{+\infty}{\cup}} T^{-k}(T^{-n}(B))) \\
&= \underset{k=1}{\overset{+\infty}{\cap}} (\underset{n=1}{\overset{+\infty}{\cup}} T^{-(n+k)}(B)) \\
&= \underset{k=1}{\overset{+\infty}{\cap}} (\underset{i\geq k+1}{\cup} T^{-i}(B))
\end{aligned}$$

Ainsi $\hat{B}$ apparait comme l'intersection d'une famille décroissante d'ensembles: on peut sans dommage enlever le premier terme de l'intersection:

$$\begin{aligned}
\hat{B} &= \underset{k=2}{\overset{+\infty}{\cap}} (\underset{i\geq k+1}{\overset{+\infty}{\cup}} T^{-i}(B)) \\
&= \underset{k=2}{\overset{+\infty}{\cap}} T^{-k}(\underset{n=1}{\overset{+\infty}{\cup}} T^{-n}(B)) \\
&= \underset{k=1}{\overset{+\infty}{\cap}} T^{-(k+1)}(\underset{n=1}{\overset{+\infty}{\cup}} T^{-n}(B)) \\
&= \underset{k=1}{\overset{+\infty}{\cap}} T^{-1}(T^{-k}(\underset{n=1}{\overset{+\infty}{\cup}} T^{-n}(B))) \\
&= T^{-1}(\underset{k=1}{\overset{+\infty}{\cap}} T^{-k}(\underset{n=1}{\overset{+\infty}{\cup}} T^{-n}(B))) \\
&= T^{-1}(\hat{B})
\end{aligned}$$

6. D'après la question précédente $\hat{A}$ est invariant par T et on a montré au début de cette même question que

$$\hat{A} = \underset{k=1}{\overset{+\infty}{\cap}} (\underset{i\geq k+1}{\cup} T^{-i}(A)).$$

Vu cette écriture, $\mu(A\Delta\hat{A}) = 0$ est exactement le résultat de I.4.

7. Soit f un point fixe de U_T et f_1 un représentant de la classe de f. On a $f_1 = f_1 \circ T$, μ-presque partout, d'où

$$\mathbb{1}_{(-\infty,a)} \circ f_1 = \mathbb{1}_{(-\infty,a)} \circ f_1 \circ T \quad \mu \text{ p.p.},$$

soit

$$\mathbb{1}_{f_1^{-1}((-\infty,a))} = \mathbb{1}_{T^{-1}(f_1^{-1}((-\infty,a)))} \quad \mu \text{ p.p.}$$

ce qui veut dire que $\mathbb{1}_{f_1^{-1}((-\infty,a))\Delta T^{-1}(f_1^{-1}((-\infty,a)))}$ est μ presque sûrement nulle, soit $\mu(f_1^{-1}((-\infty,a))\Delta T^{-1}(f_1^{-1}((-\infty,a)))) = 0$: $\mu(f_1^{-1}((-\infty,a))$ est presque invariant pour T.

Remarquons enfin que si f_2 est un autre représentant de la classe de f, $(f_1^{-1}((-\infty,a)))\Delta(f_2^{-1}((-\infty,a))) \subset \{f_1 \neq f_2\}$, donc est de mesure nulle: la classe de $f_1^{-1}((-\infty,a))$ modulo μ ne dépend pas du représentant choisi, ce qui légitime l'écriture $f^{-1}((-\infty,a))$.

8. Remarque préliminaire: quels que soient A et B mesurables, on a

$$\begin{aligned} |\mu(A) - \mu(B)| &= |\int \mathbb{1}_A - \mathbb{1}_B \, d\mu| \\ &\leq \int |\mathbb{1}_A - \mathbb{1}_B| \, d\mu \\ &= \int \mathbb{1}_{A\Delta B} \, d\mu = \mu(A\Delta B) \end{aligned}$$

En particulier si $\mu(A\Delta B) = 0$, $\mu(A) = \mu(B)$.

Soit f un point fixe de U_T et f_1 un représentant de la classe de f. Soit a un réel queconque. Comme $\mu(f_1^{-1}((-\infty,a))$ est presque invariant pour T, il existe B_a invariant par T, avec $\mu(f_1^{-1}((-\infty,a)\Delta B_a) = 0$. D'après la remarque ci-dessus $\mu(f_1^{-1}((-\infty,a)) = \mu(B_a)$; comme B_a est invariant par T et T ergodique, on a donc $\mu(f_1^{-1}((-\infty,a)) \in \{0;1\}$. La fonction $a \mapsto \mu(f_1^{-1}((-\infty,a))$ est croissante. D'après le théorème de convergence monotone, sa limite en $-\infty$ est 0, sa limite en $+\infty$ est 1. Soit a_0 la borne inférieure des a tels $\mu(f_1^{-1}((-\infty,a)) > 0$. Pour tout n, on a $\mu(f_1^{-1}((-\infty,a+1/n)) > 0$, donc $\mu(f_1^{-1}((-\infty,a_0+1/n)) > 0$, donc $\mu(f_1^{-1}((-\infty,a_0+1/n)) = 1$, et $\mu(f_1^{-1}((-\infty,a_0-1/n)) = 0$. Finalement, pour tout n, $\mu(f_1^{-1}([a_0-1/n,a_0+1/n)) = 1$, d'où, d'après le théorème de continuité séquentielle décroissante, $\mu(f_1^{-1}(\{a_0\})) = 1$, ce qui signifie que f_1 est μ-presque sûrement égale à a_0: la fonction constante égale à a_0 est bien un représentant de la classe de f. Réciproquement, il est bien évident que les classes de fonctions constantes sont invariantes par U_T.

9. Soient $f \in R$, $g \in K$: il existe $h \in L^2$ avec $f = h - U_T h$ et on sait que $g = U_T g$. Il vient donc

$$\begin{aligned}
\langle f, g \rangle &= \langle h - U_T h, g \rangle \\
&= \langle h, g \rangle - \langle U_T h, g \rangle \\
&= \langle h, g \rangle - \langle U_T h, U_T g \rangle \\
&= \langle h, g \rangle - \langle h, g \rangle \\
&= 0
\end{aligned}$$

En effet, comme noté dans l'énoncé, U_T étant unitaire préserve le produit scalaire. R et K sont donc bien orthogonaux.

10. Soit $f \in R^\perp$.

$$\begin{aligned}
\|f - U_T f\|_2^2 &= \|f\|^2 + \|U_T f\|^2 - 2\langle f, U_T f \rangle \\
&= \|f\|^2 + \|f\|^2 - 2\langle f, U_T f \rangle \\
&= 2\langle f, f \rangle - 2\langle f, U_T f \rangle \\
&= 2\langle f, f - U_T f \rangle \\
&= 0
\end{aligned}$$

car $f - U_T f \in R$ et $f \in R^\perp$. Ainsi, $f - U_T f = 0$, soit $f \in K$.

11. $(R \oplus K)^\perp \subset R^\perp \cap K^\perp \subset K \cap K^\perp$ d'après la question précédente. Mais un élément de $K \cap K^\perp$ est orthogonal à lui même, donc nul, d'où $(R \oplus K)^\perp \subset \{0\}$, soit $(R \oplus K)^\perp = \{0\}$. On en déduit que $((R \oplus K)^\perp)^\perp = \{0\}^\perp = L^2$. Mais on sait que dans un espace de Hilbert, le double orthogonal d'un espace vectoriel est son adhérence: ainsi $\overline{R \oplus K} = L^2$.

12. $$S_n f = (\frac{1}{n} \sum_{k=0}^{n-1} U_T^k) f.$$

Ainsi si $f \in K$, on a par une récurrence évidente $U_T^k f = f$ pour tout $k \geq 1$, d'où $S_n f = f$ pour tout n: la convergence est alors évidente.

13. Soit $f \in R$: écrivons $f = (I - U_T) g$ avec $g \in L^2$.

$$S_n f = \frac{1}{n} \sum_{k=0}^{n-1} U_T^k (I - U_T) g = \frac{1}{n} \sum_{k=0}^{n-1} (U_T^k g - U_T^{k+1} g) = \frac{1}{n}(f - U_T^n g).$$

U_T préserve la norme donc ses itérées aussi:

$$\|S_n f\|_2 \leq \frac{1}{n}(\|g\|_2 + \|U_T^n g\|_2) = \frac{2}{n}\|g\|_2,$$

ce qui entraîne clairement que $S_n(f)$ tend vers 0 dans L^2.

14. Pour $f \in K$, $P_K(f) = f$, donc avec I.12, $S_n(f)$ tend vers $P_K(f)$. Pour $f \in R$, $P_K(f) = 0$ d'après I.9, donc avec I.13, $S_n(f)$ tend vers $0 = P_K(f)$. Si $f = g+h \in R \oplus K$, alors, $S_n(f) = S_n(g)+S_n(h)$ tend vers $g+0 = P_K(f)$.

    Soit maintenant f quelconque et $\varepsilon > 0$: par densité, on peut trouver f_ε, avec $f_\varepsilon \in R \oplus K$ et $\|f - f_\varepsilon\|_2 \leq \varepsilon$. Un projecteur orthogonal est contractant (au sens large), donc

$$\|P_K(f - f_\varepsilon)\|_2 \leq \|f - f_\varepsilon\|_2 \leq \varepsilon.$$

    Pour g quelconque, on a

$$\|S_n(g)\|_2 \leq \frac{1}{n}\sum_{k=0}^{n-1}\|U_T^k g\|_2 = \frac{1}{n}\sum_{k=0}^{n-1}\|g\|_2 = \|g\|_2.$$

    Comme

$$S_n(f) - P_K(f) = S_n(f) - S_n(f_\varepsilon) + S_n(f\varepsilon) - P_K(f\varepsilon) + P_K(f_\varepsilon) - P_K(f)$$

$$\begin{aligned}\|S_n(f) - P_K(f)\|_2 &\leq \|S_n(f - f_\varepsilon)\|_2 + \|S_n(f\varepsilon) - P_K(f_\varepsilon)\|_2 + \|P_K(f - f\varepsilon)\|_2 \\ &\leq \|f - f_\varepsilon\|_2 + \|S_n(f_\varepsilon) - P_K(f_\varepsilon)\|_2 + \|f - f\varepsilon\|_2 \\ &\leq 2\varepsilon + \|S_n(f_\varepsilon) - P_K(f_\varepsilon)\|_2\end{aligned}$$

    Comme $S_n(f_\varepsilon)$ tend vers $P_K(f_\varepsilon)$, on en déduit que

$$\limsup_{n\to+\infty} \|S_n(f) - P_K(f)\|_2 \leq 2\varepsilon.$$

    Comme ε est quelconque, $\limsup\limits_{n\to+\infty} \|S_n(f) - P_K(f)\|_2 \leq 0$, ce qui montre que $S_n(f)$ tend dans L^2 vers $P_K(f)$.

15. Vu la question précédente, il suffit de montrer que si T est ergodique $P_K(f) = (\int_X f\ d\mu)1$. Or P_K est le projecteur orthogonal sur K, espace vectoriel des fonctions invariantes. Mais nous avons vu au I.6 que si T est ergodique, les seules fonctions invariantes sont les constantes, donc P_K est le projecteur orthogonal sur l'espace des fonctions constantes: cet espace est de dimension 1, la fonction constante égale à 1 en forme une base orthonormée et l'on a donc

$$P_K(f) = \langle f, 1\rangle 1 = (\int_X f\ d\mu)1.$$

16. Supposons T ergodique. $f = 1\!\!1_A \in L^2$, donc $S_n(1\!\!1_A)$ tend dans L^2 vers $(\int_X 1\!\!1_A\ d\mu)1 = \mu(A)1$. Par continuité du produit scalaire, $\langle S_n(1\!\!1_A), 1\!\!1_B\rangle$ tend

vers $\langle \mu(A)1, 1\!\!1_B\rangle = \mu(A)\mu(B)$. Il reste donc à remarquer que

$$\begin{aligned}
\langle S_n(1\!\!1_A), 1\!\!1_B\rangle &= \int_X S_n(1\!\!1_A)1\!\!1_B \; d\mu \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\int_X (1\!\!1_A \circ T^k)1\!\!1_B \; d\mu \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\int_X 1\!\!1_{T^{-k}(A)\cap B} \; d\mu \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}(A)\cap B)
\end{aligned}$$

Inversement, supposons que la relation est réalisée pour tous A et B mesurables. Soit A invariant par T et posons $B = A$: $T^{-k}(A) \cap B = A \cap B = A \cap A = A$, donc

$$\frac{1}{n}\sum_{k=0}^{n-1}\mu(T^{-k}(A)\cap B) = \mu(A),$$

qui converge vers $\mu(A)\mu(B) = \mu(A)^2$: on a donc $\mu(A) = \mu(A)^2$, soit $\mu(A)(1-\mu(A)) = 0$: $\mu(A) \in \{0,1\}$ et donc T est ergodique.

II Théorème d'équirépartition de Weyl

1.

$$\begin{aligned}
S_n(e_j) &= \frac{1}{n}\sum_{k=0}^{n-1} e_j \circ T^k \\
S_n(e_j)(x) &= \frac{1}{n}\sum_{k=0}^{n-1} e_j(T^k x) \\
&= \frac{1}{n}\sum_{k=0}^{n-1} e_j(x + k\alpha) \\
&= \frac{1}{n}\sum_{k=0}^{n-1} \exp(2i\pi j(x+k\alpha)) \\
&= \exp(2i\pi j(x))\frac{1}{n}\sum_{k=0}^{n-1}\exp(2i\pi j\alpha)^k \\
&= \alpha_n e_j(x)
\end{aligned}$$

Avec $\alpha_n = 1$ si $\exp(2i\pi j\alpha) = 1$ et

$$\alpha_n = \frac{1}{n}\frac{1-\exp(2i\pi jn\alpha)}{1-\exp(2i\pi j\alpha)}$$

sinon. Ainsi, e_j est une fonction propre pour S_n, avec $S_n e_j = \alpha_n e_j$. En particulier $S_n e_0 = e_0$.

2.
$$\|S_n e_j\|_\infty = \|\alpha_n e_j\|_\infty = |\alpha_n| \|e_j\|_\infty = |\alpha_n|.$$
Si α est irrationnel et $j \neq 0$, $ij\alpha$ ne peut être entier, donc $\exp(2i\pi j\alpha) \neq 1$ et on a la majoration
$$|\alpha_n| \leq \frac{1}{n} \frac{2}{|1 - \exp(2i\pi j\alpha)|}$$
qui donne la convergence uniforme de $S_n e_j$ vers 0.

3.
$$\int_{\mathbb{T}^1} e_j \ d\mu = \frac{1}{2\pi} \int_0^1 \exp(2i\pi jx) \ dx = \delta_{0,j}$$
par un calcul classique. Si f est dans le sous espace vectoriel engendré par les e_j, il existe un entier N et des complexes $\alpha_{-N}, \dots, \alpha_N$, avec $f = \sum_{j=-N}^{N} \alpha_j e_j$.

Pour $j \neq 0$, $S_n e_j$ tend en norme infinie vers $0 = \delta_{0,j} = \int_{\mathbb{T}^1} e_j \ d\mu$. Pour $j = 0$, $S_n e_j = S_n e_0$ tend en norme infinie vers $e_0 = 1 = \delta_{0,j} = \int_{\mathbb{T}^1} e_j \ d\mu$. Par linéarité,
$$S_n f = \sum_{j=-N}^{N} \alpha_j S_n e_j$$
converge donc en norme infinie vers
$$\sum_{j=-N}^{N} \alpha_j \int_{\mathbb{T}^1} e_j \ d\mu = \int_{\mathbb{T}^1} \sum_{j=-N}^{N} \alpha_j e_j \ d\mu = \int_{\mathbb{T}^1} f \ d\mu$$

4. Puisque $\mathcal{C}(\mathbb{T}^1, \mathbb{C})$, muni de la norme infinie, est un espace complet, la série $\sum_{j \in \mathbb{Z}} c_j e_j$, qui est normalement convergente (au sens ou $\sum_{j \in \mathbb{Z}} \|c_j e_j\|_\infty = \sum_{j \in \mathbb{Z}} |c_j| < +\infty$) est également convergente dans $\mathcal{C}(\mathbb{T}^1, \mathbb{C})$.

5. C'est essentiellement la même preuve qu'au I.14. Soit $\varepsilon > 0$ on peut trouver N tel que $\sum_{|j|>N} |c_j| \leq \varepsilon$, ce qui entraîne que si on pose $f_\varepsilon = \sum_{|j| \leq N} c_j e_j$, on a
$$\|f - f_\varepsilon\|_\infty = \| \sum_{|j|>N} c_j e_j \|_\infty \leq \sum_{|j|>N} \|c_j e_j\|_\infty = \sum_{|j|>N} |c_j| \leq \varepsilon.$$
Par ailleurs, on a pour tout h et tout n: $\|S_n h\|_\infty \leq \|h\|_\infty$.
$$\begin{aligned} S_n(f) - \langle f, e_0 \rangle &= S_n(f) - S_n(f_\varepsilon) + S_n(f_\varepsilon) - \langle f_\varepsilon, e_0 \rangle + \langle f_\varepsilon - f, e_0 \rangle \\ &= S_n(f - f_\varepsilon) + S_n(f_\varepsilon) - \langle f_\varepsilon, e_0 \rangle + \langle f_\varepsilon - f, e_0 \rangle \end{aligned}$$

Cela nous donne

$$\begin{aligned}
\|S_n(f) - \langle f, e_0\rangle\|_\infty &\leq \|S_n(f - f_\varepsilon)\|_\infty + \|S_n(f_\varepsilon) - \langle f_\varepsilon, e_0\rangle\|_\infty + \|f_\varepsilon - f\|_\infty \\
&\leq \|f - f_\varepsilon\|_\infty + \|S_n(f_\varepsilon) - \langle f_\varepsilon, e_0\rangle\|_\infty + \|f_\varepsilon - f\|_\infty \\
&\leq 2\varepsilon + \|S_n(f_\varepsilon) - \langle f_\varepsilon, e_0\rangle\|_\infty
\end{aligned}$$

D'après la question II.5, $S_n(f_\varepsilon)$ tend uniformément vers $\langle f_\varepsilon, e_0\rangle$, donc

$$\limsup_{n\to+\infty} \|S_n(f) - \langle f, e_0\rangle\|_\infty \leq 2\varepsilon.$$

Comme ε est quelconque,

$$\limsup_{n\to+\infty} \|S_n(f) - \langle f, e_0\rangle\|_\infty \leq 0$$

et $S_n(f)$ converge uniformément vers $\langle f, e_0\rangle = \int_{\mathbb{T}^1} f \, d\mu$.

6. Soit f une fonction de classe C^1, 1 périodique. Une intégration par parties donne

$$\begin{aligned}
\frac{1}{2\pi}\int_0^{2\pi} f'(x)e^{-2i\pi jx}\, dx &= \frac{1}{2\pi}[f(.)e^{-2i\pi j.}]_0^1 + \frac{1}{2\pi}\int_0^{2\pi} f(x)2i\pi j e^{-2i\pi jx}\, dx \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(x)2i\pi j e^{-2i\pi jx}\, dx
\end{aligned}$$

ce qui entraîne que pour tout $f \in \mathcal{C}^1(\mathbb{T}^1, \mathbb{C})$, on a

$$\langle f', e_j\rangle = ij\langle f, e_j\rangle,$$

d'où pour $j \neq 0$ $|c_j| = |\langle f, e_j\rangle| = \frac{1}{j}\langle f', e_j\rangle$. D'après l'inégalité de Cauchy-Schwarz, on a donc

$$\sum_{j\neq 0} |c_j| \leq (\sum_{j\neq 0} \frac{1}{j^2})^{1/2}(\sum_{j\neq 0} |\langle f', e_j\rangle|^2)^{1/2} \leq \frac{\pi}{\sqrt{3}}\|f'\|_2,$$

avec l'inégalité de Parseval-Bessel. Comme la famille (e_j) est une base orthonormée de L^2, on a bien l'identité $f = \sum_{j\in\mathbb{Z}} c_j e_j$ au sens de la convergence dans L^2.

Par ailleurs, d'après la question II.4, la série $\sum_{j\in\mathbb{Z}} c_j e_j$ converge pour la norme uniforme vers une fonction h. La convergence pour la norme infinie entraînant la convergence en moyenne quadratique et l'unicité de la limite en moyenne quadratique donne $\|f - h\|_2 = 0$. Comme f et h sont continues, on a $f = h$.

7. Si $b - a = 1$, alors $\mathbb{1}_{[a,b]} = 1 = e_0$ et on peut prendre $f_b = f_h = e_0$. Sinon, quitte à composer par une translation, on peut supposer que $0 < a < b < 1$. Posons $P(x) = 6x^5 - 15x^4 + 10x^3$. Il est facile de voir que $P(0) = P'(0) = P''(0) = P'(1) = P''(1) = 0$ et que $P(1) = 1$. On a

$P'(x) = 30x^2(x-1)^2 \geq 0$ donc pour $x \in [0,1]$, $0 = P(0) \leq P(x) \leq P(1) = 1$. En conséquence, la fonction Ψ définie sur $\mathbb{R}$ par $\Psi(x) = 0$ pour $x < 0$, $\Psi(x) = 1$ pour $x > 1$ et $\Psi(x) = P(x)$ pour $x \in [0,1]$ est une fonction de classe C^2 à valeurs dans $[0,1]$. Soit n tel que $2/n \leq \varepsilon$. Posons $f_h(x) = \Psi(n(x-(a-1/n))$ pour $x \in [0, \frac{a+b}{2}]$ et $f_h(x) = 1 - \Psi(n(x-b))$ pour $x \in (\frac{a+b}{2}, 1]$. Pour $x \in [a, \frac{a+b}{2}]$, on a $n(n(x-(a-1/n)) \geq 1$ et donc $f_h(x) = 1$. Pour $x \in (\frac{a+b}{2}, b]$, on a $n(x-b) < 0$, et donc $f_h(x) = 1 - 0 = 1$. Ainsi f_h est constante égale à 1 entre a et b et donc f_h est C^2 sur $[0,1]$. Pour $x \in [0, a-1/n]$, on a $n(x-(a-1/n) \leq 0$ et donc $f_h(x) = 0$ Pour $x \in [b+1/n, 1]$, on a $n(x-b) \geq 1$, et donc $f_h(x) = 1-1 = 0$. Ainsi, f_h se prolonge en une fonction continue 1-périodique, C^2 sur $[0,1]$. Comme f_h est nulle sur un voisinage de 0, elle est en fait C^2 sur $\mathbb{R}$. En passant au quotient, on obtient donc une fonction $f_h \in \mathcal{C}^2(\mathbb{T}^1, \mathbb{C})$. Par construction f_h est à valeurs dans $[0,1]$. On a les inégalités

$$\mathbb{1}_{[a,b]} \leq f_h \leq \mathbb{1}_{[a-1/n, b+1/n]}$$

De la deuxième inégalité, on déduit que

$$\int_{\mathbb{T}^1} f_h \, d\mu \leq \int_{\mathbb{T}^1} \mathbb{1}_{[a-1/n,b+1/n]} \, d\mu = (b-a) + \frac{2}{n} \leq \int_{\mathbb{T}^1} \mathbb{1}_{[a,b]} + \varepsilon,$$

ce qui montre que la fonction f_h vérifie bien les propriétés voulues.

Comme $1 - \mathbb{1}_{[a,b]}$ périodisée est aussi l'indicatrice d'un segment, on peut lui appliquer les mêmes techniques, ce qui permet de construire de manière analogue la fonction f_b.

8. Soit $\varepsilon > 0$. Choisissons f_b et f_h comme dans la question précédente. Il existe N tel que $n \geq N$ entraîne $\|S_n f_b - \int f_b \, d\mu\|_\infty \leq \varepsilon$ et $\|S_n f_h - \int f_h \, d\mu\|_\infty \leq \varepsilon$. $f \mapsto S_n f$ est linéaire et envoie toute fonction positive sur une fonction positive. Les inégalités $f_b \leq \mathbb{1}_{[a,b]} \leq f_h$ entraînent donc

   $$S_n f_b \leq S_n \mathbb{1}_{[a,b]} \leq S_n f_h$$

   ce qui fait que pour tout $x \in \mathbb{T}^1$, on a

   $$\int f_b \; - \varepsilon \leq S_n f_b(x) \leq S_n \mathbb{1}_{[a,b]}(x) \leq S_n f_h(x) \leq \int f_h \; + \varepsilon,$$

   ce qui, vus les contrôles pour les intégrales de f_b et f_b entraîne

   $$(b-a) - 2\varepsilon \leq S_n \mathbb{1}_{[a,b]}(x) \leq (b-a) + 2\varepsilon.$$

   Ainsi $n \geq N$ entraîne

   $$\|S_n \mathbb{1}_{[a,b]} - (b-a)\|_\infty \leq 2\varepsilon.$$

   Comme ε est arbitraire, cela montre bien que $S_n \mathbb{1}_{[a,b]}$ converge uniformément vers $b-a$. Comme

   $$S_n \mathbb{1}_{[a,b]}(x) = \frac{1}{n} \sum_{k=0}^{n-1} \mathbb{1}_{[a,b]}(T^k(x)),$$

on obtient bien le résultat souhaité.

9. D'après I.14, pour tout $f \in L^2$, $S_n f$ tend dans L^2 vers $P_K(f)$. D'après la question précédente, pour $f = 1\!\!1_{[a,b]}$, $S_n f$ tend en norme infinie vers $\int_{\mathbb{T}^1} f\ d\mu = \langle f, e_0 \rangle$. Comme la convergence en norme infinie entraîne la convergence dans L^2, on en déduit que pour $f = 1\!\!1_{[a,b]}$, $P_K(f) = \int_{\mathbb{T}^1} f\ d\mu = \langle f, e_0 \rangle$. $f \mapsto P_K(f)$ et $f \mapsto \langle f, e_0 \rangle$ sont deux formes linéaires continues sur L^2 qui coïncident sur l'espace vectoriel engendré par les indicatrices d'intervalles. Comme cet espace vectoriel est dense dans L^2, elles sont égales, ce qui fait que pour tout f dans L^2, d'après I.14, $S_n f$ tend dans L^2 vers $P_K(f) = \langle f, e_0 \rangle = \int_{\mathbb{T}^1} f\ d\mu$.

10. Soit A un événement invariant. Posons $f = 1\!\!1_A$. On a, pour tout entier naturel k, $1\!\!1_A \circ T^k = 1\!\!1_A$, donc pour tout n on a $S_n f = 1\!\!1_A$. Comme $S_n f$ tend dans L^2 vers $\int_{\mathbb{T}^1} f\ d\mu$, on a donc $1\!\!1_A = \int_{\mathbb{T}^1} f\ d\mu$. Comme $1\!\!1_A$ est presque sûrement égale à une constante, $\mu(A) = 0$ ou 1 suivant que cette constante est 0 ou 1 (ça ne peut pas être autre chose). Ainsi, T est ergodique pour μ.

III Existence et singularité mutuelle des mesures invariantes

1. On a

$$\begin{aligned}\int_X \phi\ d\mu_{x_0,n} &= \frac{1}{n}\sum_{k=0}^{n-1}\int \phi\ d\delta_{T^k x_0} \\ &= \frac{1}{n}\sum_{k=0}^{n-1}\int \phi(T^k x_0)\end{aligned}$$

En appliquant cette identité à la fonction $\phi \circ T$, on obtient

$$\int_X \phi \circ T\ d\mu_{x_0,n} = \frac{1}{n}\sum_{k=0}^{n-1}\int \phi(T^{k+1}x_0)$$

Ainsi, en faisant la différence, les termes se télescopent et on obtient

$$\int_X \phi\ d\mu_{x_0,n} - \int_X \phi \circ T\ d\mu_{x_0,n} = \frac{1}{n}(\phi(x) - \phi(T^n x)),$$

d'où, avec l'inégalité triangulaire

$$|\int_X \phi\ d\mu_{x_0,n} - \int_X \phi \circ T\ d\mu_{x_0,n}| \le \frac{2}{n}\|\phi\|_\infty.$$

2. En particulier, pour tout k,

$$|\int_X \phi\ d\mu_{x_0,n_k} - \int_X \phi \circ T\ d\mu_{x_0,n_k}| \le \frac{2}{n_k}\|\phi\|_\infty.$$

On fait l'hypothèse supplémentaire que T est continue. Alors $\phi \circ T$ est continue, donc, par définition de la convergence faible, on obtient

$$\int_X \phi \, d\mu_{x_0,*} = \int_X \phi \circ T \, d\mu_{x_0,*},$$

quelque soit la fonction continue ϕ: comme tester sur les fonctions continues suffit à caractériser la loi, on peut dire que $\mu_{x_0,*}$ est invariante par T.

3. D'après le théorème ergodique, $S_n 1\!\!1_A$ tend dans $L^2(\mu_1)$ vers $\int_X 1\!\!1_A \, d\mu_1 = \mu_1(A)$. On peut donc en extraire une sous-suite $\phi_1(n)$ telle que $S_{\phi_1(n)}$ converge μ_1-presque partout vers $\mu_1(A)$, donc partout sur un ensemble B_1 tel que $\mu_1(B_1) = 1$

4. D'après le théorème ergodique, $S_n 1\!\!1_A$ tend dans $L^2(\mu_2)$ vers $\int_X 1\!\!1_A \, d\mu_2 = \mu_2(A)$. La suite extraite $S_{\phi_1(n)} 1\!\!1_A$ tend aussi dans $L^2(\mu_2)$ vers $\int_X 1\!\!1_A \, d\mu_2 = \mu_2(A)$. On peut donc en extraire une sous-suite $\phi_1(\phi_2(n))$ telle que $S_{\phi_1(\phi_2(n))}$ converge μ_2-presque partout vers $\mu_2(A)$, donc partout sur un ensemble B_2 tel que $\mu_2(B_2) = 1$.

5. Sur B_1, $S_{\phi_1(\phi_2(n))}$ vers $\mu_1(A)$ (suite extraite d'une suite qui converge vers $\mu_1(A)$). Sur B_2, $S_{\phi_1(\phi_2(n))}$ vers $\mu_2(A)$. Comme $\mu_1(A) \neq \mu_2(A)$, aucun point ne peut être dans les deux à la fois.

6. Comme $\mu_i(B_i) = 1$, $\mu_i(B_i^c) = 0$, donc on a pour tout A mesurable et tout i: $\mu_i(A) = \mu_i(A \cap B_i) + \mu_i(A \cap B_i^c) = \mu_i(A \cap B_i) + 0 = \mu_i(A \cap B_i)$. $\mu_1(B_1) = 1$, $\mu_2(B_1) = \mu_2(B_1 \cap B_2) = \mu_2(\varnothing) = 0$. $\mu_2(B_2) = 1$, $\mu_1(B_2) = \mu_1(B_1 \cap B_2) = \mu_1(\varnothing) = 0$. Ainsi μ_1 et μ_2 sont mutuellement singulières.

7. Il est bien connu que l'ensemble des mesures positives est un cône. On va donc se contenter de montrer qu'une combinaison convexe de mesures de probabilités sur $\mathcal{B}$ invariantes par T est de masse totale 1 et est invariante par T. Soient $\mu_1, \ldots \mu_n$ de telles mesures, et $\alpha_1, \ldots, \alpha_n$ des réels positifs de somme 1. On pose $\mu = \sum_{i=1}^n \alpha_i \mu_i$. Pour tout A mesurable

$$\mu(A) = \sum_{i=1}^n \alpha_i \mu_i(A).$$

Si $A = X$, $\mu(X) = \sum_{i=1}^n \alpha_i \mu_i(X) = \sum_{i=1}^n \alpha_i = 1$. Donc μ est une mesure de probabilités. De même

$$\mu(T^{-1}(A)) = \sum_{i=1}^n \alpha_i \mu_i(T^{-1}(A)).$$

Comme μ_i est préservée par T, on a pour tout i: $\mu_i(A) = \mu_i(T^{-1}(A))$, d'où en faisant la somme $\mu(A) = \mu(T^{-1}(A))$.

8. Si μ n'est pas ergodique alors il existe A invariant par T avec $0 < \mu(A) < 1$. A^c est aussi invariant par T car $T^{-1}(A^c) = (T^{-1}(A))^c$. Posons $\alpha_1 = \mu(A), \alpha_2 = \mu(A^c)$. On définit μ_1 et μ_2 par $\mu_1(X) = \mu(X \cap A)/\mu(A)$ et $\mu_2(X) = \mu(X \cap A^c)/\mu(A^c)$. Il est facile de voir que μ_1 et μ_2 sont des mesures de probabilité. On a la combinaison convexe $\mu = \alpha_1\mu_1 + \alpha_2\mu_2$. $\mu_1(T^{-1}(X)) = \mu(T^{-1}(X \cap A))/\mu(A) = \mu(T^{-1}(X) \cap T^{-1}(A))/\mu(A) = \mu(X \cap A)/\mu(A) = \mu_1(A)$. Ainsi $\mu_1 \in \mathcal{C}$. De même $\mu_2 \in \mathcal{C}$: μ est une combinaison convexe stricte de deux éléments de $\mathcal{C}$: ce n'est pas un point extrémal. On a donc montré que les points extrémaux de $\mathcal{C}$ sont des probabilités ergodiques.

   Réciproquement, montrons que tout élément de $\mathcal{C}$ ergodique est extrémal. Soit μ un élément de $\mathcal{C}$, combinaison convexe de $\mu = \lambda\mu_g + (1-\lambda)\mu_d$. On suppose que μ est ergodique et que $0 < \lambda < 1$. Montrons que μ_d et μ_g sont ergodiques. Soit A invariant par translation. $\mu(A) \in \{0,1\}$. Si $\mu(A) = 0$, comme $\mu(A) \geq \lambda\mu_g(A)$, on a $\mu_g(A) = 0$. Si $\mu(A) = 1$, comme $0 = \mu(A^C) \geq \lambda\mu_g(A^c)$, on a $\mu_g(A^c) = 0$, d'où $\mu_g(A) = 1$. Dans les deux cas $\mu_g(A) \in \{0,1\}$ Ainsi μ_g est ergodique. On fait de même pour μ_d.

   On va maintenant montrer que $\mu_d = \mu_g$. Soit A mesurable quelconque. D'après le théorème ergodique de Von Neumann, $S_n1\!\!1_A$ tend dans $L^2(\mu)$ vers $\mu(A)1$. Par continuité de la norme $\|S_n1\!\!1_A\|_2$ tend vers $\mu(A)$, soit $\int_X S_n^2\ d\mu \to \mu(A)^2$. De même, en appliquant le théorème ergodique dans $L^2(\mu_d)$ et $L^2(\mu_g)$, $\int_X S_n^2\ d\mu_d \to \mu_d(A)^2$ et $\int_X S_n^2\ d\mu_g \to \mu_g(A)^2$. En faisant une combinaison affine des deux dernières égalités, on a

$$\int_X S_n^2\ d(\lambda\mu_g + (1-\lambda)\mu_d) \to \lambda\mu_g(A)^2 + (1-\lambda)\mu_d(A)^2.$$

   soit

$$\int_X S_n^2\ d\mu \to \lambda\mu_g(A)^2 + (1-\lambda)\mu_d(A)^2.$$

   Par unicité de la limite, on en déduit que

$$\mu(A)^2 = \lambda\mu_g(A)^2 + (1-\lambda)\mu_d(A)^2,$$

   soit

$$(\lambda\mu_g(A) + (1-\lambda)\mu_d(A))^2 = \lambda\mu_g(A)^2 + (1-\lambda)\mu_d(A)^2.$$

   Comme la fonction $x \to x^2$ est strictement convexe, on en déduit que $\mu_d(A) = \mu_g(A)$. Ainsi $\mu_d = \mu_g$, ce qui montre que μ est extrémal.

IV Ergodicité du décalage de Bernoulli

1. $T^{-1}(A) = T^{-1}(A) \cap [0,1/2) \cup T^{-1}(A) \cap [1/2,1)$

$$\begin{aligned} T^{-1}(A) \cap [0,1/2) &= \{x \in [0,1/2); T(x) \in A\} \\ &= \{x \in [0,1/2); 2x \in A\} \\ &= [0,1/2) \cap \{x \in \mathbb{T}^1; 2x \in A\} \\ &= \{x \in \mathbb{T}^1; 2x \in A\} \end{aligned}$$

Car $2x \in [0,1)$ entraîne $x \in [0,1/2)$. De même,

$$\begin{aligned} T^{-1}(A) \cap [1/2) &= \{x \in [1/2,); T(x) \in A\} \\ &= \{x \in [0,1/2); 2x-1 \in A\} \\ &= [1/2,) \cap \{x \in \mathbb{T}^1; 2x \in A\} \\ &= \{x \in \mathbb{T}^1; 2x-1 \in A\} \end{aligned}$$

Car $2x \in [0,1)$ entraîne $x \in [1/2,1)$. Ainsi

$$T^{-1}(A) = \{x \in \mathbb{T}^1; 2x \in A\} \cup \{x \in \mathbb{T}^1; 2x-1 \in A\}$$

Les deux ensembles considérés sont mesurables (images réciproques de A mesurable par une application continue, donc mesurable), donc $T^{-1}(A)$ est mesurable. Ainsi T est mesurable.

$$T^{-1}(A) = A/2 \cup (A+1)/2.$$

On sait que les applications affines envoient un ensemble mesurable sur un ensemble dont la mesure est celle de l'ensemble de départ multiplié par le coefficient de l'affinité:

$$\begin{aligned} \mu(T^{-1}(A)) &= \lambda(A/2 \cap [0,1)) + \lambda((A+1)/2 \cap [0,1)) \\ &= \lambda(A/2) + \lambda((A+1)/2\cap) \\ &= \lambda(A)/2 + \lambda(A)/2 = \lambda(A) \\ &= \mu(A) \end{aligned}$$

donc T préserve μ.

2. On vient de faire la preuve pour $k = 1$. On va montrer la propriété par récurrence. Supposons la propriété acquise jusqu'au rang k. Soit j entre 0 et $2^k - 1$. Posons $B = T^{-k}(A) \cap [\frac{j}{2^k}, \frac{j+1}{2^k})$. D'après la propriété au rang 1

$$\mu(T^{-1}(B) \cap [0,1/2)) = \mu(T^{-1}(B) \cap [1/2,1)) = \mu(B)/2,$$

soit

$$\mu(T^{-1}(B) \cap [0,1/2)) = \mu(T^{-1}(B) \cap [1/2,1)) = 2^{-(k+1)}\mu(A)$$

d'après l'hypothèse de récurrence. Or

$$\begin{aligned} T^{-1}(B) \cap [0,1/2) &= T^{-(k+1)}(A) \cap T^{-1}([\frac{j}{2^k}, \frac{j+1}{2^k})) \cap [0,1/2) \\ &= T^{-(k+1)}(A) \cap [\frac{j}{2^{k+1}}, \frac{j+1}{2^{k+1}}) \end{aligned}$$

et

$$\begin{aligned} T^{-1}(B) \cap [1/2,1) &= T^{-(k+1)}(A) \cap T^{-1}([\frac{j}{2^k}, \frac{j+1}{2^k})) \cap [1/2,1) \\ &= T^{-(k+1)}(A) \cap [\frac{2^k+j}{2^{k+1}}, \frac{2^k+j+1}{2^{k+1}}) \end{aligned}$$

Tout entier entre 0 et $2^{k+1}-1$ est soit entre 0 et 2^k-1, soit la fomme de 2^k est d'un entier entre 0 et 2^k-1: on a ainsi couvert tous les entiers entre 0 et $2^{k+1}-1$, ce qui montre bien que la proposition est héréditaire.

3. Supposons d'abord que les extrémités de intervalles soient des nombres dyadiques: il existe x, y, n entiers avec $a = \frac{x}{2^n} \leq \frac{y}{2^n} = b$ Alors

$$\begin{aligned}
\mu(A \cap I) &= \sum_{j=x}^{y-1} \mu(A \cap [\frac{j}{2^n}, \frac{j+1}{2^n}) \\
&= \sum_{j=x}^{y-1} \mu(T^{-k}(A) \cap [\frac{j}{2^n}, \frac{j+1}{2^n}) \\
&= \sum_{j=x}^{y-1} 2^{-k}\mu(A) \\
&= (y-x)/2^k \mu(A) \\
&= \mu(I)\mu(A)
\end{aligned}$$

Les intervalles à extrémités dyadiques forment un π-système qui engendre la tribu borélienne de $[0,1)$. Les mesures $I \mapsto \mu(A \cap I)$ et $I \mapsto \mu(A)\mu(I)$ coïncident sur un π-système qui engendre la tribu borélienne de $[0,1)$: elles sont donc égales sur toute la tribu borélienne de $[0,1)$, donc en particulier sur les intervalles.

4. On a montré que pour A invariant et I borélien $\mu(A \cap I) = \mu(A)\mu(I)$. En prenant $I = A$, avec A invariant, on a $\mu(A) = \mu(A)^2$, donc $\mu(A) \in \{0,1\}$: T est ergodique pour μ.

5. Traitons d'abord le cas $0 < p < 1$. On va montrer que pour tout n, il existe au moins 2^n réels ayant la propriété cherchée. Cela donnera le résultat voulu.

Soit $(\alpha_j)_{j\geq 1}$ une suite de variables aléatoires indépendantes suivant la loi de Bernoulli de paramètre $1-p$ sur l'espace $(\Omega, \mathcal{F}, \mathbb{P})$. La loi forte des grands nombres dit que $\mathbb{P}$ presque sûrement

$$\frac{1}{k}\sum_{j=0}^{k-1}(1-\alpha_i) \to \mathbb{E}[1-\alpha_1] = p,$$

c'est à dire que $\mathbb{P}(F) = 1$, avec

$$F = \{\frac{1}{k}\sum_{j=0}^{k-1} 1\!\!1_{\{\alpha_i=0\}} \to p\}.$$

Soit $y \in \{0,1\}^n$. On pose $A_y = \cap_{i=1}^n \{\alpha_i = y_i\}$. Comme $\mathbb{P}(F) = 1$,

$$\mathbb{P}(F \cap A_y) = \mathbb{P}(\cap_{i=1}^n \{\alpha_i = y_i\}) = \prod_{i=1}^n \mathbb{P}(\alpha_i = y_i) > 0;$$

Soit $\mathbb{P}(F \cap A_y) > 0$ donc $F \cap A_y \neq \varnothing$. Soit $\omega \in F \cap A_y$ Posons $x(y) = \sum_{j=0}^{+\infty} \alpha_j(\omega) 2^{-(j+1)}$. Les $x(y)$ sont tous distincts (les combinaisons des n premiers chiffres sont différentes et les développements sont tous propres puisqu'ils ont une infinité de 0) et ont la propriété voulue: on a donc trouvé 2^n nombres avec la propriété voulue.

$p = 0$: prenons les n premiers chiffres comme on veut, puis pour $k > n$, $\alpha_k = 0$ si k est un carré parfait, 1 sinon. On fabrique ainsi autant de nombre qu'il y a de suites finies de 0 et de 1, donc une infinité. Pour $p = 1$, il suffit de considérer $1 - x$, pour tous les x trouvés pour $p = 0$.

6. Prenons u_1, v_1 quelconques, puis $u_{n+1} = 2(v_1 + v_2 + \cdots + v_n)$, $v_{n+1} = 2(u_1 + u_2 + \cdots + u_n)$. On alterne u_1 "1", v_1 " "0", u_2 "1", v_2 " "0", etc. À la fin d'un bloc de 0 la proportion de 0 dépasse $2/3$, à la fin d'un bloc de 1 la proportion de 1 dépasse $2/3$, ce qui montre qu'il ne peut y avoir de limite pour la proportion.

FIN