Autour du théorème de Szemerédi : Le théorème de Sárközy-Furstenberg et le théorème de Roth

Claire Burrin

EA Systèmes dynamiques (MAT571)
Sous la direction de Jérôme Buzzi

École Polytechnique
Décembre 2010

Table des matières

Le principe de correspondance de Furstenberg marque le début de la théorie de Ramsey ergodique. Jusqu'alors, la théorie de Ramsey étudiait par la combinatoire des structures hautement organisées[1] en prêtant une attention particulière aux sous-structures préservées par une partition finie de l'espace ambiant. Un de ses résultats fondateurs fut la preuve qu'apporta van der Waerden en 1927 à la conjecture de Baudet. Une progression arithmétique de raison r et de terme initial u_0 est une suite $(u_n)_{n\in\mathbb{N}_0}$ définie récursivement par $u_{n+1} = u_n + r$. Une progression arithmétique de longueur l et de terme initial a est donc de la forme $a, a+r, a+2r, ..., a+(l-1)r$. Un ensemble A est k-colorié si il existe une application $\varphi : A \to [1,k]$, où $[1,k]$ est l'intervalle discret $\{1,2,...,k\}$.

Théorème 0.1 (van der Waerden, 1927)**.** *Soit $k \in \mathbb{N}$. Si $\mathbb{N}$ est k-colorié, il y existe une progression arithmétique monochromatique de longueur arbitraire.*

Définition 0.2 (densité de Banach supérieure)**.** *La densité de Banach supérieure d'un ensemble $E \subseteq \mathbb{N}$ ou $\mathbb{Z}$ est*

$$d^*(E) = \limsup_{b-a\to\infty} \frac{\left|E \cap [a,b]\right|}{b-a}.$$

En d'autres termes, il existe une suite d'intervalles (a_k, b_k) dans $\mathbb{N}$ ou $\mathbb{Z}$ avec $b_k - a_k \to \infty$ quand $k \to \infty$ et

$$\lim_{k\to\infty} \frac{\left|E \cap [a_k,b_k]\right|}{b_k - a_k} \longrightarrow d^*(E)$$

et pour tout autre suite d'intervalles avec $b_k - a_k \to \infty$

$$\limsup_{k\to\infty} \frac{\left|E \cap [a_k,b_k]\right|}{b_k - a_k} \leq d^*(E).$$

En 1936, Erdös et Turàn firent la conjecture la suivante : tout sous-ensemble de $\mathbb{N}$ de densité de Banach supérieure positive contient une progression arithmétique de longueur l arbitraire. Le cas $l = 3$ fût démontré en 1952 par Roth puis, après avoir prouvé le cas $l = 4$, Szemerédi démontra la généralisation suivante de cette conjecture.

Théorème 0.3 (Szemerédi, 1975)**.** *Tout ensemble $E \subseteq \mathbb{Z}$ de densité de Banach supérieure positive contient une progression arithmétique de longueur arbitraire.*

Deux ans plus tard, Furstenberg donne une preuve ergodique du théorème de Szemerédi à travers une extension de la récurrence de Poincaré[2].

Théorème 0.4 (Récurrence multiple de Furstenberg)**.** *Soit $(X, \mathcal{B}, \mu, T)$ un système probabiliste et soit $A \in \mathcal{B}$ un mesurable tel que $\mu(A) > 0$. Pour tout $k \geq 1$, il existe $n \in \mathbb{N}$ tel que*

$$\mu(A \cap T^{-n}A \cap T^{-2n}A \cap ... \cap T^{-kn}A) > 0.$$

[1] P.ex., les (semi-)groupes $(\mathbb{N}, \mathbb{Z}, \mathbb{Z}^d, ...)$, les espaces de vecteurs ou encore les graphes complet K_n, i.e., les graphes dont les n sommets sont tous reliés deux à deux par une arête.

[2] L'affirmation de Furstenberg est en fait plus précise et l'extension de la récurrence de Poincaré son corollaire direct, cf. [4].

Le théorème de Szemerédi résulte de la récurrence multiple par le principe de correspondance de Furstenberg. L'idée est la suivante, il s'agit de lier au sous-ensemble $E \subseteq \mathbb{Z}$ de densité de Banach supérieure positive un système dynamique probabiliste de mesure préservée auquel on peut appliquer un énoncé ergodique.

Dans la première partie de ce travail, ce procédé est illustré dans deux situations : le théorème de Szemerédi (0.3) et le théorème de Sárközy-Furstenberg. Dans un deuxième temps, on donne la démonstration du théorème de récurrence polynomiale, due à Furstenberg, et du théorème de Roth qui stipule l'existence de progressions arithmétiques de longueur 3 dans $\mathbb{N}$ (c'est-à-dire, par le principe de correspondance, le cas $k = 2$ du théorème de récurrence multiple 0.4). La démonstration du théorème de récurrence multiple dans le cas général, énoncé ci-dessus, dépasse le cadre de ce travail. On pourra se référer pour celle-ci à [3] ou, pour une excellente présentation de la méthodologie de cette démonstration, à [6].

1 Le principe de correspondance de Furstenberg

Un *système dynamique probabiliste* $(X, \mathcal{B}, \mu, T)$ est la donnée d'un espace de probabilité borélien $(X, \mathcal{B}, \mu)$ et d'une application $T : X \to X$ bijective, mesurable et préservant la mesure μ, c'est-à-dire $\mu(T^{-1}A) = \mu(A)$ pour tout $A \in \mathcal{B}$ (ou μ est *T-invariante*). Le système probabiliste donné par μ est appelé *ergodique* si tout ensemble mesurable strictement invariant est de mesure triviale, i.e. $T^{-1}A = A$ implique $\mu(A) \in \{0, 1\}$. Un système probabiliste est *uniquement ergodique* si il n'existe qu'une seule mesure de probabilité T-invariante En substance, l'ergodicité signifie que l'espace X ne peut pas être décomposé en sous-systèmes de mesures positives sur lesquels T agirait séparément. De cette façon, on peut imaginer que tout système probabiliste se décompose en parties ergodiques. C'est la donnée du théorème suivant.

Théorème 1.1 (Décomposition ergodique)**.** *Soit $(X, \mu, \mathcal{B}, T)$ un système probabiliste. Alors il existe une application mesurable $x \mapsto \mu_x$ telle que $\mu_x : X \to [0, 1]$ est une mesure de probabilité T-invariante et ergodique pour presque tout $x \in X$ et*

$$\mu = \int_X \mu_x d\mu(x).$$

En particulier, dans de nombreux cas, un système probabiliste peut être supposé ergodique puisque les propriétés vérifiées par le sous-système ergodique sont relevés au système probabiliste. C'est notamment le cas pour la récurrence multiple. Soit $(X, \mathcal{B}, \mu, T)$ un système probabiliste avec $A \in \mathcal{B}$ de mesure $\mu(A) > 0$, alors

$$\mu(\{x \in X : \mu_x(A) > 0\}) > 0,$$

où μ_x, $x \in X$, sont les mesures données par la décomposition ergodique (théorème 1.1). Si le théorème de récurrence multiple de Furstenberg est vérifié par tout système ergodique, il s'applique alors aussi aux systèmes probabilistes. En effet, si pour tout sous-système ergodique donné par μ_x et tout $k \in \mathbb{N}$, il existe $n \in \mathbb{N}$ tel que $\mu_x(A \cap ... \cap T^{-kn}A) > 0$, alors

$$\mu(A \cap ... \cap T^{-kn}A) = \int_X \mu_x(A \cap ... \cap T^{-kn}A)\mathrm{d}\mu(x) > 0.$$

Pour la démonstration du théorème de décomposition ergodique (théorème 1.1) ainsi que des résultats de théorie ergodique utilisés tout au long de ce travail, on se référera à [3] ou [2].

1.1 Décalage de Bernoulli

A titre de rappel, le paragraphe suivant expose la construction du système probabiliste du décalage de Bernoulli. Supposons qu'on fait un jeu à n débouchés compris dans $\{1, ..., n\}$ (imaginons un dé à n faces). Chaque débouché i a une probabilité attribuée p_i. Ces probabilités doivent satisfaire $\sum_{i=1}^{n} p_i = 1$ et $p_i \geq 0$ pour chaque $1 \leq i \leq n$. Alors $\mathbf{p} = (p_1, p_2, ..., p_n)$ définit une mesure de probabilité, appelée mesure de Bernoulli, sur l'espace $\Lambda = \{1, ..., n\}$ muni de la topologie discrète. On répète l'expérience une infinité de fois. L'espace considéré est alors

$$\Omega = \Lambda^{\mathbb{Z}} = \{\ \mathbf{x} = (x(n))_{n\in\mathbb{Z}} : x(n) \in \Lambda \text{ pour chaque } i \in \mathbb{Z}\}.$$

L'espace produit Ω est équipé de la topologie produit et de la métrique induite par

$$\mathrm{d}(\mathbf{x}, \mathbf{y}) := 2^{-\min\{|n|:x(n)\neq y(n)\}}.$$

Alors le décalage (vers la droite)

$$T : (x(n))_{n\in\mathbb{Z}} \mapsto (x(n+1))_{n\in\mathbb{Z}}$$

est une application continue sur l'espace métrique compact Ω. Soit $I \subset \mathbb{Z}$ un ensemble fini. La suite finie $\mathbf{a} = (a(i))_{i\in I}$ d'éléments à valeur dans Λ définit le *cylindre*

$$I(\mathbf{a}) = \{\mathbf{x} \in X : x(i) = a(i) \text{ pour tout } i \in I\}.$$

L'ensemble des cylindres est une σ-algèbre $\mathcal{B}$ sur laquelle est définie une unique mesure de probabilité

$$\mu(I(\mathbf{a})) = \prod_{i\in I} p_{a(i)}.$$

Ainsi, $(X, \mathcal{B}, \mu)$ est un espace de probabilité et le décalage T est $\mathcal{B}$-mesurable et préserve la mesure μ.

1.2 Le principe de correspondance

Lemme 1.2 (Principe de correspondance)**.** *Soit $E \subseteq \mathbb{Z}$. Le point $\mathbf{1}_E \in \Omega = \{0,1\}^{\mathbb{Z}}$ définit par*

$$1_E(n) = \begin{cases} 1 & \text{si } n \in E \\ 0 & \text{si } n \notin E. \end{cases}$$

pour tous $n \in \mathbb{Z}$ engendre $X = \overline{\{T^n \mathbf{1}_E : n \in \mathbb{Z}\}}$ pour le décalage de Bernoulli T. Soit $A = \{\omega \in X : \omega(0) = 1\}$. Alors $d^*(E) > 0$ si et seulement si il existe un système dynamique probabiliste $(X, \mathcal{B}, \mu, T)$ où $\mu(A) > 0$.

Démonstration. Soit $E \subseteq \mathbb{Z}$ avec $\mathrm{d}^*(E) > 0$. Rappelons que tout point x d'un espace induit la mesure de Dirac δ_x définie par

$$\delta_x(S) = \begin{cases} 1 & \text{si } x \in S \\ 0 & \text{si } x \notin S, \end{cases},$$

pour tout ensemble S mesurable. Par définition de la densité de Banach supérieure, il existe une suite d'intervalles $(a_k, b_k)_{k \in \mathbb{N}} \subset \mathbb{Z}$ pour laquelle $b_k - a_k \to \infty$ quand $k \to \infty$ et

$$\mathrm{d}^*(E) = \lim_{k \to \infty} \frac{\left|E \cap [a_k, b_k]\right|}{b_k - a_k} = \lim_{k \to \infty} \frac{1}{b_k - a_k} \sum_{n=a_k}^{b_k} \delta_n(E).$$

Définissons une suite de mesures de probabilité (μ_k) par

$$\mu_k = \frac{1}{b_k - a_k} \sum_{n=a_k}^{b_k} \delta_{T^n \mathbf{1}_E}.$$

D'après le théorème de Kryloff-Bogoliouboff (cf. [2]), la suite (μ_k) converge dans la topologie faible* le long d'une sous-suite vers une mesure T-invariante μ. On obtient ainsi un système probabiliste $(X, \mathcal{B}, \mu, T)$. D'autre part, l'énoncé implique directement l'équivalence

$$T^n \mathbf{1}_E \in A \text{ si et seulement si } n \in E \tag{1.1}$$

pour chaque $n \in \mathbb{Z}$. Ainsi,

$$\begin{aligned} \mu(A) &= \lim_{k \to \infty} \frac{1}{b_k - a_k} \sum_{n=a_k}^{b_k} \delta_{T^n \mathbf{1}_E}(A) \\ &= \lim_{k \to \infty} \frac{1}{b_k - a_k} \sum_{n=a_k}^{b_k} \delta_n(E) \\ &= \mathrm{d}^*(E), \end{aligned}$$

qui est strictement positive par hypothèse. Supposons maintenant qu'il existe une mesure T-invariante telle que l'ensemble mesurable A soit de mesure strictement positive. Par le théorème de décomposition ergodique 1.1, supposons de plus que le système $(X, \mathcal{B}, \mu, T)$ soit ergodique. Alors, d'après le théorème de Birkhoff (cf. [2]), pour presque tout $x \in X$,

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=0}^{N-1} \delta_{T^n x} = \mu$$

dans la topologie faible*. Donc, il existe $m \in \mathbb{Z}$ tel que le point $T^m \mathbf{1}_E$ vérifie l'équation ci-dessus et

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \delta_{T^{n+m}\mathbf{1}_E}(A) = \mu(T^{-m}A) = \mu(A) > 0.$$

D'autre part, selon l'équivalence (1.1), cette limite est exactement

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \delta_{n+m}(E) = \lim_{N\to\infty} \frac{|E \cap [m, N-1+m]|}{N} \leq \mathrm{d}^*(E).$$

□

Dans la suite de cette section, on présente et discute quelques applications du principe de correspondance. En commençant par le théorème de Szemerédi.

1.3 Szemerédi et le théorème de récurrence multiple

Le théorème de Szemerédi 0.3 affirme que tout ensemble de densité de Banach supérieure strictement positive contient une progression arithmétique de longueur arbitraire. Supposons la validité du théorème de récurrence multiple 0.4.

Démonstration du théorème de Szemerédi. Soit $k \in \mathbb{N}$ et $E \subseteq \mathbb{Z}$ de densité de Banach supérieure. L'existence d'un entier $n \in \mathbb{N}$ tel que

$$\mu(A \cap T^{-n}A \cap ... \cap T^{-kn}A) > 0$$

découle alors directement du principe de correspondance 1.2 et du théorème de récurrence multiple 0.4. Si $\mathbf{x} \in X$ appartient à l'intersection $A \cap T^{-n}A \cap ... \cap T^{-kn}A$ alors l'ensemble $\{\mathbf{x}, T^n\mathbf{x}, ..., T^{kn}\mathbf{x}\}$ est contenu dans A. De plus, comme $\mathbf{x} \in X$, il existe $m \in \mathbb{Z}$ tel que $x(j) = T^m(1_E(j))$ pour $0 \leq j \leq kn$ et donc

$$1_E(m) = 1_E(m+n) = ... = 1_E(m+kn) = 1.$$

L'ensemble E contient la progression arithmétique $m, m+n, ..., m+kn$ pour $k \in \mathbb{N}$ arbitraire. □

1.4 Théorème de Sárközy-Furstenberg

En 1978, Sárközy [8] et Furstenberg [4] démontrèrent de façon indépendante le résultat suivant à partir d'une conjecture de Lovász.

Théorème 1.3 (Sárközy-Furstenberg). *Soit $E \subseteq \mathbb{Z}$ un ensemble de densité supérieure positive et $p(n) \in \mathbb{Z}[n]$ un polynôme à coefficients entiers vérifiant $p(0) = 0$. Alors il existe $x, y \in E$ et $n \in \mathbb{N}$ tels que $x - y = p(n)$.*

En d'autres termes, l'ensemble des polynômes à coefficients entiers qui disparaissent en 0 et l'ensemble $E - E = \{x - y : x, y \in E\}$ pour tout E de densité supérieure positive sont d'intersection non nulle. Par le principe de correspondance, il suffit d'admettre la version polynomiale du théorème de Poincaré dont la démonstration fera l'objet de la section 2.

Théorème 1.4 (Récurrence polynomiale)**.** *Soit $(X, \mathcal{B}, \mu, T)$ un système probabiliste et soit $p(t) \in \mathbb{Z}[t]$ un polynôme de coefficient constant nul. Si $A \in \mathcal{B}$ est de mesure $\mu(A) > 0$, alors il existe $n \in \mathbb{N}$ tel que*

$$\mu(A \cap T^{p(n)} A) > 0.$$

Remarquons que si le polynôme p est de coefficient constant non nul, ou, en particulier, si $0 \notin p(\mathbb{N})$, alors il existe un entier non nul $m \in \mathbb{Z} \setminus \{0\}$ tel que l'intersection $p(\mathbb{N}) \cap m\mathbb{Z}$ est vide. Alors, X ne contient aucune récurrence polynomiale. En effet, par l'absurde, si il existe $n \in \mathbb{N}$ tel que $\mu(A \cap T^{-p(n)} A) > 0$, alors en appliquant le théorème de Poincaré (cf. [6]) au système probabiliste $(X, \mathcal{B}, T^m, \mu)$, il existe également un $n' \in \mathbb{N}$ tel que $\mu(A \cap T^{-p(n)} A) > 0$. En conclusion, $p(n) = mn'$ pour deux éléments $n, n' \in \mathbb{N}$, ce qui est une contradiction.

Démonstration du théorème de Sárközy-Furstenberg. Soit $E \subset \mathbb{Z}$ un ensemble de densité de Banach supérieure, alors, dans la notation du lemme 1.2 (principe de correspondance), il existe $n \in \mathbb{N}$ tel que $\mu(A \cap T^{-p(n)} A) > 0$. Tout $\mathbf{x} \in A \cap T^{-p(n)} A$ vérifie $x(0) = x(p(n)) = 1$ et puisque $\mathbf{x} \in X$, il existe $m \in \mathbb{Z}$ tel que $\mathbf{x}$ et $T^m \mathbf{1}_E$ coincident sur les indices $0, 1, ..., p(n)$. Donc $m, m + p(n) \in E$ et par conséquent $p(n) \in E - E$. $\square$

2 Théorème de récurrence polynomiale

Soit U un opérateur unitaire sur un espace de Hilbert $\mathcal{H}$. Pour démontrer le théorème de récurrence polynomiale 2.6, nous décomposerons (dans l'esprit de la démonstration du théorème de von Neumann (cf. [2])) l'espace de Hilbert en $\mathcal{H} = \mathcal{H}_{\text{rat}} \oplus \mathcal{H}_{\text{rat}}^{\perp}$, où

$$\mathcal{H}_{\text{rat}} = \overline{\{x \in \mathcal{H} : U^m x = x \text{ pour un } m \in \mathbb{N}\}} \tag{2.1}$$

est l'union des sous-espaces fermés et U-invariants

$$\mathcal{H}_m = \{x \in \mathcal{H} : U^m x = x\}.$$

2.1 Ergodicité et théorie spectrale

Tout système probabiliste $(X, \mathcal{B}, \mu, T)$ induit un opérateur associé (ou *opérateur de Koopman*)

$$U_T : L^2(X, \mathcal{B}, \mu) \to L^2(X, \mathcal{B}, \mu), f \mapsto f \circ T.$$

Dorénavant, nous écrirons $L^2(\mu) = L^2(X, \mathcal{B}, \mu)$. L'application T étant inversible, il est facile de vérifier que U_T est unitaire. Cet opérateur permet d'étudier l'ergodicité par l'analyse harmonique plutôt que par les des considérations géométriques. En particulier, l'ergodicité est caractérisée par le fait que les seules fonctions U_T-invariantes sont les fonctions constantes (cf. [2]). L'ergodicité est donc une propriété spectrale. Dans le cadre plus général des espaces de Hilbert, nous rappelons l'énoncé du théorème spectral pour les opérateurs unitaires (cf. [7]). On écrit $e_n(x) = e^{2\pi in\cdot x}$.

Théorème 2.1 (Théorème spectral pour les opérateurs unitaires)**.** *Soit $U : \mathcal{H} \to \mathcal{H}$ un opérateur unitaire sur un espace de Hilbert $\mathcal{H}$. Alors $\mathcal{H} = \oplus_{n\in\mathbb{N}}\mathcal{H}_n$, où*

- *Pour chaque $n \in \mathbb{N}$, $\mathcal{H}_n = \overline{\{U^n x : n \in \mathbb{Z}\}}$ pour un élément $x \in \mathcal{H}_n$,*
- *Il existe un isomorphisme unitaire $\phi : \mathcal{H}_n \to L^2(\mathbb{T}, \mu_n)$, où μ_n est une mesure finie définie sur le tore $\mathbb{T}$, tel que $\phi \circ U|_{\mathcal{H}_n} = e_1 \cdot \phi$, i.e. le diagramme*

$$\begin{tikzcd} \mathcal{H}_n \arrow[r, "U|_{\mathcal{H}_n}"] \arrow[d, "\phi"'] & \mathcal{H}_n \arrow[d, "\phi"] \\ L^2(\mathbb{T}, \mu_n) \arrow[r, "x\mapsto e_1 x"'] & L^2(\mathbb{T}, \mu_n) \end{tikzcd}$$

commute.

Nous ne retiendrerons de la démonstration que l'expression des mesures finies μ_n, $n \in \mathbb{N}$ et de l'isomorphisme unitaire ϕ. Remarquons que chaque sous-espace U-invariant $\mathcal{H}_n$ est identifié à l'adhérence de l'orbite d'un de ses points. Soit x celui-ci. Alors, il existe d'après le théorème de Bochner (cf. [7]), une mesure finie sur le tore définie par x :

$$\langle U^n x, x\rangle = \int_{\mathbb{T}} e_n \mathrm{d}\mu_x.$$

L'isomorphisme unitaire est

$$\phi : \overline{\{U^n x : n \in \mathbb{Z}\}} \to L^2(\mathbb{T}, \mu_x), \sum c_n U^n x \mapsto \sum c_n e_n,$$

où les sommes ci-dessus sont finies et l'application ϕ est une isométrie linéaire qu'on démontre surjective par des arguments de densité. Notons que l'élément x est envoyé par l'isomorphisme ϕ sur la fonction constante $1 \in L^2(\mathbb{T}, \mu_x)$.

Enfin, le sous-espace $\mathcal{H}_{\mathrm{rat}}$ défini par (2.1) est identifié à l'ensemble des éléments de l'espace de Hilbert $\mathcal{H}$ tels que $Ux = e^{2\pi i\alpha}x$ pour un α rationnel, ce qui explique la notation.

2.2 Décomposition de l'espace de Hilbert

Lemme 2.2. *Le supplément orthogonal de $\mathcal{H}_{rat}$, noté $\mathcal{H}_{rat}^{\perp}$, est l'adhérence de l'ensemble*

$$\left\{x \in \mathcal{H} : \lim_{N\to\infty} \left\| \frac{1}{N}\sum_{n=0}^{N-1} U^{p(n)}x \right\| = 0 \textit{ pour tout polynôme } p(t) \in \mathbb{Z}[t] \right\}.$$

La démonstration de ce lemme nécessite quelques préparatifs sur l'équidistribution des polynômes. Une suite réelle $(y_n)_{n\in\mathbb{N}}$ est *équidistribuée* pour la mesure de Lebesgue λ si

$$\lim_{n\to\infty}\frac{1}{n}\sum_{n=1}^{N}\delta_{y_n}=\lambda$$

dans la topologie faible*. Une suite réelle $(x_n)_{n\in\mathbb{N}}$ est *équidistribuée modulo 1* si la suite associée $(y_n)_{n\in\mathbb{N}}$ définie par les parties fractionnaires $y_n=\{x_n\}$ est équidistribuée.

Lemme 2.3 (Critère d'équidistribution de Weyl)**.** *Une suite réelle $(x_n)_{n\in\mathbb{N}}$ d'éléments à valeur dans $[0,1]$ est équidistribuée si et seulement si, pour tout entier $k\neq 0$,*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}e^{2\pi ikx_n}=0.$$

Théorème 2.4 (Théorème d'équidistribution polynomial de Weyl)**.** *Soit $p(t)=a_kt^k+\ldots+a_0$ un polynôme réel possédant au moins un coefficient irrationnel. Alors la suite $(p(n))_{n\in\mathbb{N}}$ est équidistribuée modulo 1 pour la mesure de Lebesgue.*

Remarquons que le lemme 2.3 est facilement démontré à l'aide du théorème d'approximation de Weierstrass. Il existe une démonstration ergodique très élégante du théorème 2.4 due à Furstenberg dans son monographe [4].

Démonstration du lemme 2.2. Par l'absurde, supposons que $x\in\mathcal{H}_m$ pour un $m\in\mathbb{N}$ arbitrairement choisi. Alors,

$$\frac{1}{N}\sum_{n=0}^{N-1}U^{mn}x=x$$

pour tout $N\in\mathbb{N}$. Réciproquement, soit $x\in\mathcal{H}_{\mathrm{rat}}^{\perp}$. Il existe par le théorème spectral une unique mesure finie borélienne μ_x sur le tore vérifiant

$$\langle U^n x,x\rangle=\int_{\mathbb{T}}e_n\mathrm{d}\mu_x$$

pour tout $n\in\mathbb{Z}$.

$$\begin{aligned}
\left\|\frac{1}{N}\sum_{n=0}^{N-1}U^{p(n)}x\right\|^2 &= \frac{1}{N^2}\sum_{n_1,n_2=0}^{N-1}\left\langle U^{p(n_1)-p(n_2)}x,x\right\rangle\\
&= \int_0^1\frac{1}{N^2}\sum_{n_1,n_2=0}^{N-1}e^{2\pi it(p(n_1)-p(n_2))}\mathrm{d}\mu_x(t)\\
&= \int_0^1\left|\frac{1}{N}\sum_{n=0}^{N-1}e^{2\pi itp(n)}\right|^2\mathrm{d}\mu_x(t)\\
&= \int_{\mathbb{Q}\cap[0,1)}\left|\frac{1}{N}\sum_{n=0}^{N-1}e^{2\pi itp(n)}\right|^2\mathrm{d}\mu_x(t)+\int_{\mathbb{R}\setminus\mathbb{Q}\cap[0,1)}\left|\frac{1}{N}\sum_{n=0}^{N-1}e^{2\pi itp(n)}\right|^2\mathrm{d}\mu_x(t)
\end{aligned}$$

Le second terme converge par le théorème de convergence dominée et le critère d'équidistribution de Weyl vers 0 puisque la suite polynomiale $tp(n)$ est équidistribuée modulo 1 pour t irrationnel (d'après le théorème 2.4). Dans le cas où l'argument t est rationnel, supposons que $\mu_x(\{t\}) > 0$. Il existe un entier positif m tel que $mt \in \mathbb{Z}$. Ainsi, la fonction caractéristique $1_{\{t\}} \in L^2(\mathbb{T}, \mu_x)$ est invariante pour la multiplication avec le caractère e_m, i.e.

$$e_m(s)1_{\{t\}}(s) = e_m(t)1_{\{t\}}(s) = 1_{\{t\}}(s).$$

Puisque $L^2(\mathbb{T}, \mu_x)$ et le sous-espace $\overline{\{U^n x : n \in \mathbb{Z}\}}$ sont isomorphes d'après le théorème spectral, la fonction $1_{\{t\}}$ est envoyée de façon unique sur un élément $y \in \mathcal{H}_m \subset \mathcal{H}_{\mathrm{rat}}$ impliquant la contradiction. Donc $\mu_x(\{t\}) = 0$, pour chaque t rationnel.

□

2.3 Théorème de récurrence polynomiale

Proposition 2.5. *Soit $U : \mathcal{H} \to \mathcal{H}$ un opérateur unitaire dans un espace de Hilbert $\mathcal{H}$ et $p(t) \in \mathbb{Z}[t]$ un polynôme de coefficient constant 0 (i.e. $p(0) = 0$). Si $x \in \mathcal{H}$ vérifie*

$$\langle U^{p(n)}x, x\rangle = 0$$

pour tous $n \in \mathbb{N}$, alors $x \in \mathcal{H}_{rat}^{\perp}$.

Démonstration. Décomposons $x \in \mathcal{H}$ en $x = x_1 + x_2$ où $x_1 \in \mathcal{H}_{\mathrm{rat}}$ et $x_2 \in \mathcal{H}_{\mathrm{rat}}^{\perp}$. Pour tout $\varepsilon > 0$, il existe $m \in \mathbb{N}$ tel qu'un élément $x_1' \in \mathcal{H}_m$ vérifie $\|x_1 - x_1'\| < \varepsilon$ car $\bigcup_{m\geq 1} \mathcal{H}_m$ est dense dans $\mathcal{H}_{\mathrm{rat}}$. Puisque le polynôme $p(t) \in \mathbb{Z}[t]$ est de coefficient constant nul, $m|p(mn)$ pour tous $n \in \mathbb{N}$. En particulier, il existe une suite d'entiers non-nuls (k_n) définis par $mk_n = p(mn)$. Ainsi, l'inégalité pour tout $k \in \mathbb{Z}$,

$$\begin{aligned}\left\|U^{mk}x_1 - x_1\right\| &\leq \left\|U^{mk}x_1 - U^{mk}x_1'\right\| + \|x_1' - x_1\| \\ &\leq \left(\left\|U^{mk}\right\| + 1\right)\|x_1 - x_1'\| \\ &< 2\varepsilon,\end{aligned}$$

implique

$$\begin{aligned}\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^{p(mn)}x - x_1\right\| &\leq \lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^{mk_n}x_1 - x_1\right\| + \lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} U^{p(mn)}x_2\right\| \\ &< 2\varepsilon.\end{aligned}$$

Finalement, on montre avec Cauchy-Schwarz que $x_1 = 0$. En effet,

$$|\langle x_1, x\rangle| = \left|\left\langle \frac{1}{N}\sum_{n=0}^{N-1} U^{p(mn)}x - x_1, x\right\rangle\right| < 2\varepsilon\,\|x\|\,.$$

Si on laisse $\varepsilon \to 0$, alors $|\langle x_1, x\rangle| = \|x_1\|^2 = 0$. □

Enfin, rappellons l'énoncé du théorème de récurrence poylnomiale.

Théorème 2.6 (Récurrence polynomiale)**.** *Soit* $(X,\mathcal{B},\mu,T)$ *un système probabiliste et soit* $p(t)\in\mathbb{Z}[t]$ *un polynôme de coefficient constant nul. Si* $A\in\mathcal{B}$ *est de mesure* $\mu(A)>0$*, alors il existe* $n\in\mathbb{N}$ *tel que*

$$\mu(A\cap T^{p(n)}A)>0.$$

Démonstration. Par l'absurde, on suppose, pour tout $n\in\mathbb{N}$, que $\mu(T^{-p(n)}A\cap A)=0$. Les hypothèses de la proposition 2.5 sont alors satisfaites. En effet,

$$\left\langle U_T^{p(n)}1_A,1_A\right\rangle=\int_X U_T^{p(n)}(1_A)1_A\mathrm{d}\mu=\int_X 1_{T^{-p(n)}A\cap A}\mathrm{d}\mu=0$$

et ainsi la fonction $1_A\in L^2(\mu)$ vérifie

$$\frac{1}{N}\sum_{n=0}^{N-1}U_T^{p(n)}1_A\longrightarrow 0 \tag{2.2}$$

quand N tend vers l'infini. Toutefois, pour chaque $N\in\mathbb{N}$, la somme

$$\frac{1}{N}\sum_{n=0}^{N-1}\mu(T^{-p(n)}A)\quad=\quad\int_X\frac{1}{N}\sum_{n=0}^{N-1}U_T^{p(n)}1_A\mathrm{d}\mu$$

doit être strictement positive de par l'invariance de la mesure, i.e. $\mu(T^{-p(n)}A)=\mu(A)>0$ pour tous $n\geq 0$. On obtient alors une contradiction avec le théorème de convergence dominée et (2.2). □

3 Théorème de Roth

Pour démontrer le théorème de Roth, les systèmes probabilistes considérés sont supposés ergodiques.

3.1 Facteur de Kronecker

Un système $(Y,\mathcal{A},\nu,S)$ est appelé un *facteur* de $(X,\mathcal{B},\mu,T)$ s'il existe une application surjective $\pi:X\to Y$ telle que le diagramme

$$\begin{array}{ccc} X & \xrightarrow{T} & X\\ \pi\downarrow & & \downarrow\pi\\ Y & \xrightarrow[S]{} & Y\end{array}$$

commute et que $\mu=\nu\circ\pi$ presque partout. Lorsque, de plus, l'application π est injective, les deux systèmes sont dits isomorphes. Le système entier est un exemple de facteur trivial. En particulier, le système $(X,\mathcal{A},\mu,T)$ où $\mathcal{A}$ est le sous-σ-algèbre T-invariant de $\mathcal{B}$

est un facteur de $(X, \mathcal{B}, \mu, T)$. Le *facteur de Kronecker* du système ergodique $(X, \mathcal{B}, \mu, T)$ est le plus petit sous-σ-algèbre de $\mathcal{B}$ pour lequel toutes les fonctions propres de l'opérateur unitaire U_T sont mesurables. (Exemple de facteur de Kronecker). Nous verrons que c'est le plus grand facteur sur lequel un système ergodique se comporte comme une rotation sur un groupe compact et abélien. Pour cela, nous rappelons dans la section suivante les propriétés de la rotation.

3.2 Rotation sur un groupe compact abélien

Théorème 3.1. *Soit G un groupe compact et métrisable et R_g la rotation sur G par un élément fixe $g \in G$, $R_g(x) = gx$. Les énoncés suivants sont équivalents.*

(1) Le groupe G est abélien et la rotation R_g est uniquement ergodique (avec comme unique mesure invariante m_G, la mesure de Haar sur G,

(2) R_g est ergodique pour la mesure de Haar m_G,

(3) le sous-groupe $H = \{g^n : n \in \mathbb{Z}\}$ est dense dans G,

En fait, il existe une propriété équivalente supplémentaire qui veut que $\chi(g) \neq 1$ pour tout caractère non trivial $\chi \in \widehat{G}$. La démonstration de cette dernière équivalence est purement calculatoire mais est laissée de côté car elle ne sera d'aucune utilité pour la suite de ce travail. Une rotation ergodique R_g sur un groupe compact et abélien vérifie alors le théorème de récurrence multiple. En effet, pour tout $A \in \mathcal{B}$ avec $m_G(A) > 0$ et pour tout $k \in \mathbb{N}$,

$$m_G(A \cap R_g^{-n}A \cap ... \cap R_g^{-kn}A) = \int_G 1_A(x)1_A(g^n x) \cdots 1_A(g^{kn}x)\mathrm{d}m_G(x)$$

doit être positive pour un $n \in \mathbb{N}$ car $\{g^n : n \in \mathbb{Z}\}$ est dense dans G.

Démonstration du théorème 3.1. Supposons que le sous-groupe H ne soit pas dense dans G. Alors, l'application quotient $G \to G/H$ est non triviale et donc, il existe une fonction non constante $f \in L^2(G/H, \mu|_{G/H})$. Or, $U_{R_g}f(x) = f(gx) = f(x)$ et donc U_{R_g} n'est pas ergodique. Supposons à présent que H est dense dans G. Donc, pour tout $\delta > 0$ et tout $x \in G$, il existe $n \in \mathbb{Z}$ tel que $\mathrm{d}(g^n, x) < \delta$, où $\mathrm{d}(\cdot,\cdot)$ est la métrique sur G. Soit f une fonction continue sur G et $\varepsilon > 0$. Par continuité, il suffit de choisir un δ adéquat pour que $|f(g^n y) - f(xy)| < \varepsilon$. Soit μ une mesure de probabilité R_g-invariante, alors celle-ci est également invariante par translations à gauche. En effet,

$$\left| \int_G f(xy)\mathrm{d}\mu(y) - \int_G f(y)\mathrm{d}\mu(y) \right| \leq \int_G |f(xy) - f(g^n y)|\, \mathrm{d}\mu(y) < \varepsilon.$$

L'unique mesure invariante par translations est la mesure de Haar, c'est-à-dire que μ est l'unique mesure de probabilité R_g-invariante et donc uniquement ergodique. Enfin, il est clair que l'unique ergodicité implique l'ergodicité. □

Théorème 3.2 (Halmos, von Neumann). *Soit $(X, \mathcal{B}, \mu, T)$ un système ergodique. Son facteur de Kronecker est isomorphe à la rotation ergodique $R_g(x) = gx$ sur un groupe compact abélien G.*

Démonstration. L'isomorphisme se construit de la manière suivante. Soit $\{f_i : i \in \mathbb{N}\}$ l'ensemble des fonctions propres de $U_T : L^2(\mu) \to L^2(\mu)$ associés aux valeurs propres $\{\lambda_i : i \in \mathbb{N}\}$. (Cet ensemble est dénombrable car $L^2(\mu)$ est un espace séparable.) Posons $g = (\lambda_1, \lambda_2, ...)$. Alors le sous-groupe abélien $G = \overline{\{g^n : n \in \mathbb{N}\}} \leqslant S^{\mathbb{N}}$ est compact et métrisable et donc la rotation $R_g : G \to G, x \mapsto gx$ est ergodique pour la mesure de Haar m_G par le théorème 3.1. Définissons l'application

$$\pi : X \to S^{\mathbb{N}}, x \mapsto (\frac{f_1}{|f_1|}(x), \frac{f_2}{|f_2|}(x), ...).$$

Celle-ci est bien définie, injective et $\pi(T(x)) = (U_T \frac{f_1(x)}{|f_1|}, ...) = (\lambda_1 \frac{f_1(x)}{|f_1|}, ...) = g\pi(x) = R_g(\pi(x))$. Par définition, la mesure $\nu = \mu \circ \pi^{-1}$ est R_g-invariante. Or, l'unique mesure R_g invariante est m_G. Donc $\nu = m_G$. Ainsi la rotation sur G est isomorphe au facteur de Kronecker de $(X, \mathcal{B}, \mu, T)$. □

En particulier, le facteur de Kronecker d'un système ergodique vérifie également le théorème de récurrence multiple.

Théorème 3.3. *Tout facteur de Kronecker $(X, \mathcal{A}, \mu, T)$ d'un système ergodique $(X, \mathcal{B}, \mu, T)$ vérifie le théorème de récurrence multiple. De plus, pour tout $A \in \mathcal{A}$ tel que $\mu(A) > 0$ et tout $k \in \mathbb{N}$,*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap ... \cap T^{-kn}A) > 0.$$

Démonstration. D'après le théorème 3.2 de Halmos et von Neumann, le facteur de Kronecker est isomorphe à la rotation uniquement ergodique R_g définie sur un groupe compact et abélien G muni de la mesure de Haar m_G. Soit 1_A la fonction indicatrice de l'ensemble A. Posons

$$\phi : G \to \mathbb{R}, h \mapsto \int_G 1_A(x) 1_A(hx) \cdots 1_A(h^k x) \mathrm{d}m_G.$$

Cette fonction est non négative (i.e. $\phi \geq 0$) avec $\phi(e) = 1$ pour e l'élément neutre de G et $\phi(g^n) = m_G(A \cap R_g^{-1}A ... \cap R_g^{-kn}A) > 0$ pour au moins un $n \in \mathbb{N}$. De plus, ϕ est continue. Rappelons que l'unique ergodicité de R_g se caractérise par la convergence suivante (cf. [2]),

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \phi(R_g^n e) = \int_G \phi \mathrm{d}m_G.$$

Par continuité, cette limite doit être positive. □

3.3 Facteur caractéristique

Un facteur $(Y, \mathcal{A}, \nu, S)$ d'un système ergodique $(X, \mathcal{B}, \mu, T)$ est dit *caractéristique* pour la moyenne

$$\frac{1}{N} \sum_{n=0}^{N-1} U_T^n f_1 \cdot U_T^{2n} f_2 \cdots U_T^{kn} f_k,$$

où $k \in \mathbb{N}$ et $f_1, ..., f_k \in L^\infty(\mu)$, si le comportement asymptotique de cette moyenne coïncide avec celui de

$$\frac{1}{N}\sum_{n=0}^{N-1} U_T^n \mathbb{E}(f_1|\mathcal{A}) \cdots U_T^{kn} \mathbb{E}(f_k|\mathcal{A})$$

dans $L^2(\mu)$, i.e. si

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_1 \cdot U_T^{2n} f_2 \cdots U_T^{kn} f_k - \frac{1}{N}\sum_{n=0}^{N-1} U_T^n \mathbb{E}(f_1|\mathcal{A}) \cdots U_T^{kn} \mathbb{E}(f_k|\mathcal{A})\right\|_{L^2} \to 0$$

lorsque N tend vers l'infini. Dans le cas de la moyenne simple ($k = 1$), le facteur caractéristique est donné par l'ensemble des fonctions constantes dans $L^2(\mu)$ suivant le théorème de von Neumann (cf. [2]). L'objet de cette sous-section est de montrer que le facteur de Kronecker est caractéristique pour la moyenne double. Pour cela, nous commençons par introduire le lemme de van der Corput.

Lemme 3.4 (van der Corput). *Soit $(u_n)_{n\geq 0}$ une suite bornée dans un espace de Hilbert $\mathcal{H}$. On définit une suite réelle*

$$\gamma_h = \limsup_{N\to\infty}\left|\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h}, u_n\rangle\right|$$

pour tout $h \geq 0$. Alors

$$\limsup_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} u_n\right\|^2 \leq \limsup_{H\to\infty}\frac{1}{H}\sum_{h=0}^{H-1}\gamma_h.$$

Démonstration. Soit $M \in \mathbb{R}$ tel que $\|u_n\| \leq M$ pour tout $n \geq 0$. Alors, par Cauchy-Schwarz, la suite $(\gamma_h)_{h\geq 0}$ est bornée par M^2. Donc, pour un $\varepsilon > 0$ fixe, il existe $h_0 = h_0(\varepsilon) \geq 0$ tel que $\gamma_h < \varepsilon$ pour tout $h \geq h_0$. De la même façon, il existe $H_0 = H_0(\varepsilon)$ tel que pour tout $H \geq H_0$

$$\frac{1}{H}\sum_{h=0}^{H-1}\gamma_h < \varepsilon. \tag{3.1}$$

Soit $N > H$, en développant $\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1} u_{n+h}$, on obtient

$$\begin{aligned}\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1} u_{n+h} &= \frac{1}{NH}\left(u_0 + ... + (H-1)u_{H-2} + (H-1)u_N + ...u_{N+H-2}\right)\\ &\quad + \frac{1}{N}\left(u_H + ... + u_{N-1}\right).\end{aligned}$$

Ainsi, seuls les premiers et derniers termes de cette série et de $\frac{1}{N}\sum_{n=0}^{N-1} u_n$ diffèrent et il existe N suffisamment grand tel que pour $H \geq H_0$

$$\left|\frac{1}{N}\sum_{n=0}^{N-1} u_n - \frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1} u_{n+h}\right| < \varepsilon. \tag{3.2}$$

On peut donc considérer la moyenne double en lieu de la moyenne simple. Par convexité

$$\left\|\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}\right\|^2 \leq \frac{1}{N}\sum_{n=0}^{N-1}\left\|\frac{1}{H}\sum_{h=0}^{H-1}u_{n+h}\right\|^2$$

où la norme dans le terme de droite peut être réarranger à travers le produit scalaire hilbertien en

$$\left\|\frac{1}{H}\sum_{h=0}^{H-1}u_{n+h}\right\|^2 = \frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\langle u_{n+h_1}, u_{n+h_2}\rangle.$$

Cette somme étant donc positive, on obtient avec l'inégalité de Minkowski,

$$\begin{aligned}
\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\langle u_{n+h_1}, u_{n+h_2}\rangle &= \left|\frac{1}{N}\sum_{n=0}^{N-1}\frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\langle u_{n+h_1}, u_{n+h_2}\rangle\right| \\
&= \left|\frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h_1}, u_{n+h_2}\rangle\right| \\
&\leq \frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\left|\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h_1}, u_{n+h_2}\rangle\right|.
\end{aligned}$$

En résumé, on obtient des étapes précédentes la borne

$$\begin{aligned}
\limsup_{N\to\infty}\left\|\frac{1}{NH}\sum_{n=0}^{N-1}\sum_{h=0}^{H-1}u_{n+h}\right\|^2 &\leq \limsup_{N\to\infty}\frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\left|\frac{1}{N}\sum_{n=0}^{N-1}\langle u_{n+h_1}, u_{n+h_2}\rangle\right| \\
&= \frac{1}{H^2}\sum_{h_1,h_2=0}^{H-1}\gamma_{|h_1-h_2|}.
\end{aligned}$$

On remarque qu'en choisissant initialement H très grand (en particulier, beaucoup plus grand que H_0), le terme ci-dessus peut être majoré par ε d'après (3.1) $\square$

Théorème 3.5 (Furstenberg)**.** *Soit $(X, \mathcal{B}, \mu, T)$ un système ergodique, $\mathcal{A}$ le σ-algèbre correspondant au facteur de Kronecker de X et $f_1, f_2 \in L^\infty(\mu)$. Alors*

$$\left\|\frac{1}{N}\sum_{n=0}^{N-1}U_T^n f_1 U_T^{2n} f_2 - \frac{1}{N}\sum_{n=0}^{N-1}U_T^n \mathbb{E}(f_1|\mathcal{A}) U_T^{2n}\mathbb{E}(f_2|\mathcal{A})\right\|_{L^2(\mu)} \longrightarrow 0$$

lorsque $N \to \infty$.

Démonstration. L'expression

$$\frac{1}{N}\sum_{n=0}^{N-1}U_T^n f_1 U_T^{2n} f_2 - \frac{1}{N}\sum_{n=0}^{N-1}U_T^n \mathbb{E}(f_1|\mathcal{A}) U_T^{2n}\mathbb{E}(f_2|\mathcal{A})$$

réécrite

$$\frac{1}{N}\sum_{n=0}^{N-1} U_T^n(f_1-\mathbb{E}(f_1|\mathcal{A}))U_T^{2n}f_2+\frac{1}{N}\sum_{n=0}^{N-1} U_T^n\mathbb{E}(f_1|\mathcal{A})U_T^{2n}(f_2-\mathbb{E}(f_2|\mathcal{A}))$$

suggère que l'énoncé de Furstenberg peut être reformulé de la façon suivante : Si $\mathbb{E}(f_1|\mathcal{A})=0$ ou $\mathbb{E}(f_2|\mathcal{A})=0$, alors

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1} U_T^n f_1 U_T^{2n} f_2\right\|_{L^2(\mu)}=0.$$

Admettons que $\mathbb{E}(f_1|\mathcal{A})=0$ et dérivons ce résultat grâce au lemme de van der Corput. Pour cela, on vérifie que $\frac{1}{H}\sum_{h=0}^{H-1}\gamma_h\to 0$ quand $H\to\infty$ dans la notation du lemme 3.4, où la suite $(u_n)\subset L^2(\mu)$ est définie par $u_n=U_T^n f_1 U_T^{2n} f_2$. En calculant

$$\begin{aligned}\langle u_{n+h},u_n\rangle &= \left\langle U_T^{n+h}f_1\cdot U_T^{2(n+h)}f_2, U_T^n f_1\cdot U_T^{2n}f_2\right\rangle\\ &= \left\langle U_T^{h}f_1\cdot U_T^{n+2h}f_2, f_1\cdot U_T^{n}f_2\right\rangle\\ &= \int_X f_1U_T^hf_1\cdot U_T^n\left(f_2U_T^{2h}f_2\right)\mathrm{d}\mu,\end{aligned}$$

on obtient

$$\gamma_h=\limsup_{N\to\infty}\left|\int_X f_1U_T^hf_1\frac{1}{N}\sum_{n=0}^{N-1}U_T^n\left(f_2U_T^{2h}f_2\right)\mathrm{d}\mu\right|.$$

Par le théorème ergodique moyen de von Neumann (cf. [2]), la somme à l'intérieur de l'intégrale ci-dessus converge vers la projection orthogonale $\mathbb{P}(f_2U_T^{2h}f_2)$ sur le sous-espace fermé des fonctions T-invariantes de $L^2(\mu)$. Puisque T est ergodique, ce sous-espace ne contient que des fonctions constantes. Or, f_1 est orthogonale à l'ensemble des fonctions constantes. En effet, soit $g\in L^2(\mu)$ une fonction constante, alors $\langle f_1,g\rangle=\int_X f_1g\mathrm{d}\mu=\int_X\mathbb{E}(f_1g|\mathcal{A})\mathrm{d}\mu=\int_X g\mathbb{E}(f_1|\mathcal{A})\mathrm{d}\mu=0$. Alors $\gamma_h=0$ pour tout $h\geq 0$ et le lemme de van der Corput permet de conclure

$$\lim_{N\to\infty}\left\|\frac{1}{N}\sum_{n=0}^{N-1}U_T^nf_1\cdot U_T^{2n}f_2\right\|^2=0,$$

i.e., le facteur de Kronecker est caractéristique pour la moyenne double. □

3.4 Théorème de Roth

Théorème 3.6 (Roth)**.**

Par le principe de correspondance de Furstenberg (et en particulier, son application au théorème de Szémeredi), il nous suffit de démontrer le cas $k=2$ su théorème de récurrence multiple.

Théorème 3.7 (Théorème de récurrence multiple pour $k = 2$)**.** *Soit* $(X, \mathcal{B}, \mu, T)$ *un système probabiliste et* $A \in \mathcal{B}$ *un ensemble mesurable de mesure positive. Alors il existe* $n \in \mathbb{N}$ *tel que* $\mu(A \cap T^{-n}A \cap T^{-2n}A) > 0$.

Démonstration. L'affirmation est claire par les propositions précédentes. Etant donné un système ergodique, tout facteur de Kronecker de celui-ci est isomorphe à une rotation sur un groupe compact abélien (théorème 3.2 de Halmos et von Neumann). Or, ce dernier système vérifie le théorème de récurrence multiple (théorème 3.3). Dès lors que le facteur de Kronecker est caractéristique pour la moyenne double, le système ergodique de départ satisfait alors le cas de récurrence double. On vérifie ici que les conditions de ces propositions sont bien... vérifiées. Dénotons par $\mathcal{A}$ le sous-σ-algèbre associé au facteur de Kronecker et posons $f = \mathbb{E}(1_A|\mathcal{A})$. De par sa définition, f est $\mathcal{A}$-mesurable, positive p.p. et une fonction de $L^\infty(X, \mathcal{B}, \mu)$. De plus, $\int_X f \mathrm{d}\mu = \mu(A) > 0$ et donc, par le théorème 3.3, la limite de

$$\frac{1}{N} \sum_{n=1}^{N} f U_T^n f U_T^{2n} f \mathrm{d}\mu$$

existe et est strictement positive. Il reste à vérifier que cette moyenne est bien celle que nous cherchons. D'après le théorème 3.5 de Furstenberg, la limite de

$$\frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap T^{-2n}A) = \frac{1}{N} \sum_{n=1}^{N} \int_X 1_A U_T^n 1_A U_T^{2n} 1_A \mathrm{d}\mu$$

coïncide avec celle de

$$\frac{1}{N} \sum_{n=1}^{N} \int_X 1_A U_T^n f U_T^{2n} f \mathrm{d}\mu.$$

Dans cette expression, les termes 1_A et $U_T^n f U_T^{2n} f$ appartiennent tous deux à $L^\infty(X, \mathcal{B}, \mu)$. En appliquant une seconde fois le théorème 3.5 et la propriété XX de l'espérance conditionnelle, nous obtenons

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int_X 1_A U_T^n f U_T^{2n} f \mathrm{d}\mu = \lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int_X f U_T^n f U_T^{2n} f \mathrm{d}\mu.$$

□

3.5 Le théorème de récurrence multiple pour $k \geq 3$?

Pourquoi les arguments précédents sont obsolètes quand $k \geq 3$. Se traduit dans le language de la combinatoire additive également.

Références

[1] V. Bergelson, *Ergodic Ramsey theory : a dynamical approach to static theorems*

[2] J. Buzzi, *Systèmes dynamiques*, MAT551, Ecole Polytechnique.

[3] M. Eisiedler et T. Ward, *Ergodic Theory with a View towards Number Theory*, Springer London, 2010.

[4] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, 1981.

[5] P.R. Halmos, What does the spectral theorem say ?, *The American Mathematical Monthly*, **70** (1963).

[6] B. Kra, Ergodic methods in additive combinatorics

[7] P.D. Lax, *Functional Analysis*, John Wiley and Sons, 2002.

[8] A. Sàrközy, On difference sets of integers III. *Acta Math. Acad. Sci. Hungar.* **31** (1978) 125–149.