

# THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS-EST

présentée par  
Qing CHU

pour obtenir le grade de  
DOCTEUR DE L'UNIVERSITÉ PARIS-EST  
*Spécialité : Mathématiques*

---

## Quelques problèmes de convergence et de récurrence multiple en théorie ergodique

---

Soutenue le 6 Juillet 2010 devant le jury composé des professeurs :

M. Bernard HOST	Directeur de thèse
Mme Bryna KRA	Examineur
M. Emmanuel LESIGNE	Rapporteur
M. Alejandro MAASS	Examineur
M. Mathieu MEYER	Examineur
M. Jean-Paul THOUVENOT	Rapporteur

# Remerciements

---

Je tiens tout d'abord à exprimer ma très profonde reconnaissance à Bernard Host, qui m'a fait l'honneur d'accepter de diriger cette thèse.

Sa disponibilité et sa ténacité ont permis à ces travaux de se dérouler dans les meilleures conditions possibles. Sa rigueur mathématique et ses conseils avisés m'ont été d'une grande aide, de même que son investissement scientifique et humain.

Je remercie également sa femme Bati Chétanian pour son hospitalité ainsi que son soutien.

Pour l'attention minutieuse qu'ils ont portée à mon travail, pour leurs remarques pertinentes, et pour m'avoir fait l'honneur de rapporter cette thèse, je tiens à remercier Emmanuel Lesigne ainsi que Jean-Paul Thouvenot, qui a suivi l'évolution de mes recherches avec attention.

Je souhaiterais également remercier Bryna Kra; ses conseils judicieux et ses commentaires constructifs m'ont été d'un grand secours. Je suis très honorée par sa présence dans le jury, ainsi que par celle d'Alejandro Maass et Mathieu Meyer.

Ma gratitude va également à Nikos Frantzikinakis, pour sa générosité en temps et en idées, et dont le dynamisme et la perspicacité m'ont constamment stimulée.

Je remercie Aihua Fan, mon ancien professeur, qui m'a aidée à venir en France et m'a permis de rencontrer Bernard Host.

Un grand merci à l'ensemble du Laboratoire d'Analyse et de Mathématiques Appliquées de l'Université Paris-Est - Marne-la-Vallée pour leur accueil chaleureux, ainsi qu'à Anne Raskine et Pascal Romon pour m'avoir permis d'enseigner. Pour leurs encouragements et leur aide, je remercie également François Blanchard, Marco Cannone et Damien Lambertson. Merci enfin à Christiane Lafargue pour sa gentillesse et son efficacité.

Merci à Barbera Goossen et à l'ensemble du Bureau international des chercheurs invités de l'Université Paris-Est, qui m'ont bien aidée alors que j'étais un peu perdue à mon arrivée.

Merci également à Sylvie Cach, responsable de l'École Doctorale de l'Université Paris-Est, pour m'avoir guidée dans mes démarches administratives.

Ce travail n'a été possible que grâce au soutien financier de la bourse du gouvernement Chinois.

Je voudrais tout particulièrement exprimer ma reconnaissance à l'ensemble des doctorants du LAMA, à mes amis et à toutes les personnes qui m'ont soutenue par leur gentillesse et leur dévouement.

Enfin, un grand merci à mes parents, qui n'ont eu de cesse de m'encourager durant ces trois années.

# Avant-propos

---

Dans cette thèse, nous nous intéressons à certaines questions de convergence et de récurrence multiple en théorie ergodique. Elle est composée de cinq chapitres.

Le premier chapitre présente l'historique. Nous rassemblons des résultats de convergence et de récurrence multiple dans un système muni d'une ou de plusieurs transformations qui commutent, et en même temps, nous introduisons le contexte dans lequel nous travaillons.

Le deuxième chapitre donne une présentation de la méthode et des outils que nous utilisons à plusieurs endroits dans cette thèse. En particulier, nous détaillons le mécanisme de facteurs caractéristiques et le rôle des nilsystèmes que nous employons souvent pour les systèmes munis d'une transformation, ainsi que la machinerie des systèmes magiques que nous employons pour les systèmes munis de plusieurs transformations qui commutent.

Les trois chapitres suivants sont consacrés respectivement aux études de trois questions indépendantes, chacune de ces questions donne lieu à un article.

Dans le troisième chapitre, nous considérons des moyennes ergodiques multiples pondérées le long de polynômes, dont la convergence pour le cas linéaire a été démontrée par Host et Kra. Nous établissons la convergence pour ces moyennes et comme une conséquence, nous montrons que pour toute fonction  $f$  mesurable bornée sur un système ergodique, la suite  $(f(T^n x))$  est universellement bonne pour la convergence de moyennes ergodiques multiples le long de polynômes pour presque tout  $x$ .

Dans les deux derniers chapitres, nous nous intéressons tout particulièrement aux systèmes munis de plusieurs transformations qui commutent. Nous considérons une question de convergence et une question de récurrence, mais pour les preuves, nous utilisons toute la machinerie des systèmes magiques introduite récemment par Host.

Le quatrième chapitre concerne des moyennes sur les cubes, dont la convergence dans le cas où toutes les transformations sont égales a été démontrée par Host et Kra. Nous proposons une preuve pour la convergence générale.

Le cinquième chapitre est consacré à une étude de la récurrence multiple d'aspect quantitatif pour deux transformations qui commutent. Nous obtenons un résultat qui est similaire mais pas identique à celui en cas d'une transformation établi par Bergelson, Host et Kra.

Les références bibliographiques renvoient à la bibliographie générale en fin de thèse.



## CHAPITRE I

### Introduction

Dans ce chapitre, nous présenterons un bref historique de résultats de récurrence multiple et de convergence en théorie ergodique et le contexte dans lequel nous travaillons.

Nous commençons par une présentation de principes de correspondance que nous allons utiliser pour déduire des résultats combinatoires correspondant aux résultats de récurrence que nous allons présenter dans la deuxième section. La troisième section sera consacrée à la présentation des résultats de convergence qui sont étroitement liés à ces résultats de récurrence.

#### 1. Principes de correspondance

Le premier principe de correspondance a été établi par Furstenberg ([29],[30]) dans sa fameuse démonstration ergodique du théorème de Szemerédi [60]. D'autres généralisations ont été données subséquemment. Ces résultats établissent un lien entre les deux domaines : la théorie ergodique et la théorie combinatoire, et permettent de montrer, à partir de résultats de récurrence de systèmes dynamiques mesurés, des résultats combinatoires. Ils constituent un nouveau domaine appelé la théorie ergodique de Ramsey où des problèmes issus de la combinatoire sont démontrés en utilisant la théorie ergodique. Beaucoup de résultats combinatoires issus de ce domaine n'ont pas jusqu'à présent d'autres preuves (par exemple la version polynomiale du théorème de Szemerédi qu'on donne plus bas) :

Rappelons que la *densité* d'un sous-ensemble d'entiers  $E \subset \mathbb{N}$  est définie par :

$$d^*(E) = \limsup_{N \rightarrow \infty} \frac{1}{N} \text{Card}(E \cap [0, N - 1]) .$$

**THÉORÈME I.1** (Furstenberg ([29], [30])). *Soit  $E \subset \mathbb{N}$  un ensemble de densité positive. Alors il existe un espace probabilisé  $(X, \mathcal{X}, \mu)$  muni d'une transformation  $T$  qui est inversible mesurable et préservant  $\mu$  et un ensemble  $A \in \mathcal{X}$  avec  $\mu(A) = d^*(E)$ , tels que*

$$\mu(T^{m_1} A \cap \dots \cap T^{m_k} A) \leq d^*((E + m_1) \cap \dots \cap (E + m_k))$$

pour tout entier  $k \geq 1$  et tous  $m_1, \dots, m_k \in \mathbb{Z}$ .

La densité d'un sous-ensemble de  $\mathbb{N}^k$  (ou de  $\mathbb{Z}^k$ ) est définie de façon parallèle à celle d'un sous-ensemble de  $\mathbb{N}$ , nous avons la généralisation multidimensionnelle suivante du principe de correspondance :

**THÉORÈME I.2** (Furstenberg et Katznelson [31]). *Soit  $E \subset \mathbb{N}^k$  un ensemble de densité positive. Alors il existe  $k$  transformations inversibles mesurables  $T_1, T_2, \dots, T_k$  qui commutent et préservent la mesure de l'espace probabilisé  $(X, \mathcal{X}, \mu)$  et un ensemble  $A \in \mathcal{X}$  avec  $\mu(A) = d^*(E)$  tels que*

$$\mu\left(\bigcap_{(m_1, \dots, m_k) \in F} T_1^{m_1} \dots T_k^{m_k} A\right) \leq d^*\left(\bigcap_{(m_1, \dots, m_k) \in F} (E + (m_1, \dots, m_k))\right)$$

pour tout sous-ensemble fini  $F$  de  $\mathbb{Z}^k$ .

Il existe aussi le principe de correspondance pour des actions de  $\mathbb{R}^k$ , voir [33].

## 2. Résultats de récurrence et combinatoires correspondant

En dynamique, la notion de récurrence signifie que partant d'un ensemble, on y revient d'une manière significative sous l'action d'une transformation ou de plusieurs transformations.

**2.1. Récurrence simple.** Dans la suite, aucun résultat n'utilise l'hypothèse d'inversibilité des transformations, mais on l'utilise en vue de simplifier la notation. Commençons par deux résultats classiques de récurrence dans la théorie ergodique. Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Le théorème de récurrence de Poincaré énonce que

$$\mu(A \cap T^n A) > 0 \text{ pour une infinité de valeurs de } n.$$

Khintchine [45] a renforcé ce résultat en montrant que sous les mêmes hypothèses, l'ensemble

$$\{n \in \mathbb{Z} : \mu(A \cap T^n A) > \mu(A)^2 - \epsilon\}$$

est syndétique pour tout  $\epsilon > 0$ . Rappelons qu'un sous-ensemble  $E$  de  $\mathbb{Z}$  est dit syndétique si  $\mathbb{Z}$  peut être couvert par un nombre fini de translatés de  $E$ .

Nous remarquons que la borne inférieure dans ce théorème est optimale en considérant les schémas de Bernoulli.

Le théorème de Poincaré peut être considéré comme un résultat de récurrence qualitatif alors que le théorème de Khintchine peut être considéré comme un résultat de récurrence quantitatif.

Furstenberg [30] a démontré l'extension polynomiale suivante du théorème de Poincaré :

Un *polynôme à valeurs entières*  $p(n)$  est un polynôme qui prend une valeur entière pour chaque entier  $n$ .

**THÉORÈME I.3** (Furstenberg [30]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $p(n)$  un polynôme à valeurs entières avec  $p(0) = 0$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Alors*

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

*pour une infinité de valeurs de  $n$ .*

En appliquant le principe de correspondance de Furstenberg (le théorème I.1), on obtient le théorème suivant qui a été démontré indépendamment par Sárközy [59] lorsque  $p(n) = n^2$  en utilisant la méthode du cercle :

**THÉORÈME I.4** (Sárközy [59], Furstenberg [30]). *Si  $E \subset \mathbb{N}$  un ensemble de densité positive,  $p(n)$  un polynôme à valeurs entières avec  $p(0) = 0$ . Alors il existe  $x, y \in E$ , et  $n \in \mathbb{N}$ ,  $n > 0$ , tels que  $x - y = p(n)$ .*

**2.2. Récurrence multiple.** Les trois théorèmes présentés ci-dessus sont des résultats de récurrence simple. Furstenberg [29] a démontré un théorème de récurrence multiple :

**THÉORÈME I.5** (Furstenberg [29]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Alors pour tout entier  $k \geq 1$ ,*

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

*pour une infinité de valeurs de  $n$ .*

En vertu du principe de correspondance de Furstenberg (le théorème I.1), ce théorème entraîne le théorème de Szemerédi [60] :

**THÉORÈME I.6** (Szemerédi [60]). *Tout sous-ensemble d'entiers de densité positive contient des progressions arithmétiques de longueur arbitraire.*

Le théorème de Szemerédi a été conjecturé par Erdős et Turán [24], et a été démontré pour la première fois par Szemerédi [60]. Outre la preuve ergodique de Furstenberg [29] résumée ci-dessus, de nombreuses autres preuves aussi importantes ont été données ensuite dans les articles [34], [35], [57], [58], [64], et [36] via différentes approches.

Le théorème de récurrence multiple de Furstenberg peut être considéré comme une généralisation d'une grande portée du théorème de récurrence de Poincaré qui correspond au cas où  $k = 1$ . Le théorème de Furstenberg a suscité de nombreuses études et généralisations que nous allons décrire dans la suite.

D'abord, il est naturel de se demander s'il existe un théorème quantitatif de récurrence qui a la même relation avec le théorème de Khintchine que le théorème de récurrence de Furstenberg avec celui de Poincaré. Plus précisément, on peut se demander si la mesure  $\mu(A \cap T^n A \cap \dots \cap T^{kn} A)$  admet une borne inférieure sous la forme  $\mu(A)^c$  pour un certain entier  $c > 0$ . Cette question a été résolue par Bergelson, Host et Kra [14] :

**THÉORÈME I.7** (Bergelson, Host et Kra [14]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Si  $(X, \mathcal{X}, \mu, T)$  est ergodique, alors pour tout  $\epsilon > 0$ , les ensembles*

$$\{n \in \mathbb{Z} : \mu(A \cap T^n A \cap T^{2n} A) > \mu(A)^3 - \epsilon\}$$

et

$$\{n \in \mathbb{Z} : \mu(A \cap T^n A \cap T^{2n} A \cap T^{3n} A) > \mu(A)^4 - \epsilon\}$$

sont syndétiques.

De même, en considérant les schémas de Bernoulli, on voit que la borne inférieure dans ce théorème est optimale.

Nous remarquons qu'un résultat analogue n'est pas vrai si on enlève l'hypothèse d'ergodicité ou si la longueur des progressions arithmétiques  $k$  est supérieure à 4 (voir des contre-exemples construits dans [14]). De ce fait, pour la récurrence multiple, nous avons des résultats quantitatifs analogues au théorème de Khintchine dans certains cas, mais pas dans tous les cas.

**2.3. Récurrence pour plusieurs transformations.** Furstenberg et Katznelson [31] ont généralisé le théorème de Furstenberg (le théorème I.5) au cas de plusieurs transformations qui commutent :

**THÉORÈME I.8** (Furstenberg et Katznelson [31]). *Soient  $k \geq 1$  un entier et  $T_1, T_2, \dots, T_k$  des transformations inversibles mesurables qui*

commutent et préservent la mesure de l'espace probabilisé  $(X, \mathcal{X}, \mu)$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Alors

$$\mu(A \cap T_1^n A \cap T_2^n A \cap \dots \cap T_k^n A) > 0$$

pour une infinité de valeurs de  $n$ .

Le théorème de Furstenberg I.5 correspond au cas où  $T_j = T^j$ ,  $j = 1, \dots, k$ . Nous remarquons que le théorème n'est plus valable sans l'hypothèse de commutativité des transformations (voir par exemple [13]).

En vertu du principe de correspondance multidimensionnel (le théorème I.2), ce théorème entraîne la version multidimensionnelle suivante du théorème de Szemerédi :

**THÉORÈME I.9** (Furstenberg et Katznelson [31]). *Soit  $E \subset \mathbb{N}^k$  un ensemble de densité positive. Alors pour tout sous-ensemble fini  $F$  de  $\mathbb{N}^k$ , il existe  $a \in \mathbb{N}^k$  et un entier  $d > 0$ , tels que  $a + d \cdot F \subset E$ .*

Ici,  $a + d \cdot F = \{a + dx : x \in F\}$ .

Aucune preuve combinatoire de ce résultat n'a été connue jusqu'à récemment. Gowers [35] en a donné une pour la première fois, des résultats similaires avec des conséquences similaires ont été obtenus indépendamment dans [57], [58], et [65].

Il est naturel de se demander s'il existe un résultat de récurrence quantitatif par rapport au théorème de récurrence multiple de Furstenberg et Katznelson comme le théorème de Bergelson, Host et Kra par rapport au théorème de récurrence multiple de Furstenberg. Cette thèse comporte une étude sur le cas de deux transformations qui commutent pour cette question. Cette étude est faite dans le Chapitre V et a fait l'objet d'un article [19].

**2.4. Récurrence polynomiale.** Bergelson et Leibman [11] ont démontré le théorème suivant qui généralise simultanément le théorème de récurrence simple polynomiale de Furstenberg (le théorème I.3) et le théorème de récurrence multiple de Furstenberg et Katznelson I.8 :

**THÉORÈME I.10** (Bergelson et Leibman [11]). *Soient  $k \geq 1$  un entier et  $T_1, T_2, \dots, T_k$  des transformations inversibles mesurables qui commutent et préservent la mesure de l'espace probabilisé  $(X, \mathcal{X}, \mu)$ . Soient  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes à valeurs entières avec  $p_j(0) = 0$ ,  $j = 1, 2, \dots, k$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Alors*

$$\mu(A \cap T_1^{p_1(n)} A \cap T_2^{p_2(n)} A \cap \dots \cap T_k^{p_k(n)} A) > 0 .$$

pour une infinité de valeurs de  $n$ .

En vertu du principe de correspondance multidimensionnel (le théorème I.2), la contrepartie combinatoire de ce théorème est :

**THÉORÈME I.11** (Bergelson et Leibman [11]). *Soient  $E \in \mathbb{N}^k$  un ensemble de densité positive et  $p_1(n), p_2(n), \dots, p_l(n)$  des polynômes à valeurs entières avec  $p_j(0) = 0$ ,  $j = 1, 2, \dots, l$ . Alors pour tous  $v_1, \dots, v_l \in \mathbb{Z}^k$ , il existe  $u \in \mathbb{Z}^k$  et un entier  $n$ , tels que  $u + p_i(n)v_i \in E$  pour tous  $i = 1, \dots, l$ .*

En particulier, quand  $T_1 = T_2 = \dots = T_k = T$ , on applique le principe de correspondance de Furstenberg (le théorème I.1), on en déduit la version polynomiale suivante du théorème de Szemerédi :

**THÉORÈME I.12** (Bergelson et Leibman [11]). *Soient  $E \subset \mathbb{N}$  un ensemble de densité positive et  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes à valeurs entières avec  $p_j(0) = 0$ ,  $j = 1, 2, \dots, k$ . Alors il existe  $a \in \mathbb{N}$  et un entier  $d > 0$ , tels que*

$$\{a, a + p_1(d), a + p_2(d), \dots, a + p_k(d)\} \subset E .$$

Frantzikinakis et Kra [28] ont considéré un cas particulier où les polynômes sont linéairement indépendants, ils ont obtenu un résultat de récurrence quantitative du même type que le théorème de Khintchine :

Rappelons qu'une famille de polynômes  $\{p_1(n), \dots, p_k(n)\}$  est dite *linéairement indépendante* si pour tous entiers  $m_1, \dots, m_k$  avec au moins un  $m_j \neq 0$ ,  $j \in \{1, \dots, k\}$ , le polynôme  $\sum_{j=1}^k m_j p_j(n)$  n'est pas une constante.

**THÉORÈME I.13** (Frantzikinakis et Kra [28]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soient  $k \geq 1$  un entier et  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes linéairement indépendants à valeurs entières avec  $p_j(0) = 0$ ,  $j = 1, 2, \dots, k$ . Soit  $A \in \mathcal{X}$  un ensemble de mesure positive. Alors pour tout  $\epsilon > 0$ , l'ensemble*

$$\{n \in \mathbb{Z} : \mu(A \cap T^{p_1(n)} A \cap T^{p_2(n)} A \cap \dots \cap T^{p_k(n)} A) \geq \mu(A)^{k+1} - \epsilon\}$$

*est syndétique.*

Nous remarquons qu'ici le résultat est valable sans hypothèse d'ergodicité et pour tout  $k$ , ce qui contraste avec le cas linéaire.

**2.5. Récurrence cubique quantitative.** D'autres types de récurrence ont été également étudiés. Par exemple la récurrence cubique quantitative suivante a été démontrée par Bergelson pour  $k = 2$  et par Host et Kra [41] pour le cas général :

Soit  $k \geq 1$  un entier, les éléments de  $\{0, 1\}^k$  s'écrivent comme  $\epsilon = \epsilon_1 \epsilon_2 \dots \epsilon_k$ , sans parenthèse ni virgule, avec  $\epsilon_i \in \{0, 1\}$ ,  $i = 1, \dots, k$ .

**THÉORÈME I.14** (Host et Kra [41]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $A \in \mathcal{X}$  un sous-ensemble de  $X$ . Alors pour tout  $\delta > 0$ , l'ensemble*

$$\{n = (n_1, \dots, n_k) \in \mathbb{Z}^k : \mu\left(\bigcap_{\epsilon \in \{0,1\}^k} T^{n_1 \epsilon_1 + \dots + n_k \epsilon_k} A\right) \geq \mu(A)^{2^k} - \delta\}$$

*est syndétique.*

Nous remarquons que la preuve pour ce résultat quantitatif provient de l'identification de la limite des moyennes associées. Cette thèse comporte une étude sur la généralisation de ce résultat au cas de plusieurs transformations qui commutent. Cette étude est faite dans le Chapitre IV et a fait l'objet d'un article [18].

En vertu du principe de correspondance de Furstenberg (le théorème I.1), la contrepartie combinatoire de ce théorème est :

**THÉORÈME I.15** (Host et Kra [41]). *Soit  $E \subset \mathbb{N}$  un ensemble de densité  $d^*(E) > \delta > 0$ . Soit  $k \geq 1$  un entier. Alors le sous-ensemble de  $\mathbb{Z}^k$*

$$\{n = (n_1, \dots, n_k) \in \mathbb{Z}^k : d^*\left(\bigcap_{\epsilon \in \{0,1\}^k} (E + n_1 \epsilon_1 + \dots + n_k \epsilon_k)\right) \geq \delta^{2^k}\}$$

*est syndétique.*

### 3. Résultats de convergence

**3.1. Convergence de moyennes simples.** Rappelons le théorème ergodique de von Neumann :

**THÉORÈME I.16** (von Neumann). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $f \in L^2(\mu)$ . Alors les moyennes*

$$\frac{1}{N} \sum_{n=1}^N f(T^n x)$$

*convergent dans  $L^2(\mu)$  vers une fonction  $\tilde{f}$  qui est invariante sous  $T$ . Si  $(X, \mathcal{X}, \mu, T)$  est ergodique, alors la limite  $\tilde{f}$  est la fonction constante  $\int f d\mu$ .*

Soient  $A, B \in \mathcal{X}$ . En appliquant le théorème ergodique de von Neumann à la fonction indicatrice  $f = \mathbf{1}_B$ , on voit que les moyennes

$$\frac{1}{N} \sum_{n=1}^N \mu(A \cap T^n B) = \frac{1}{N} \sum_{n=1}^N \int \mathbf{1}_A(x) f(T^n x) d\mu(x)$$

convergent.

**3.2. Convergence de moyennes multiples.** En fait, pour montrer le théorème I.5, Furstenberg a montré un énoncé plus fort : il a considéré les moyennes des

$$\mu(A \cap T^n A \cap T^{2n} A \cap \dots \cap T^{kn} A)$$

et il a montré que la limite inférieure de ces moyennes est positive :

**THÉORÈME I.17** (Furstenberg [29]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soit  $A$  un ensemble de  $X$  de mesure positive. Alors pour tout entier  $k \geq 1$ ,*

$$\liminf_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \mu(A \cap T^n A \cap T^{2n} A \cap \dots \cap T^{kn} A) > 0 .$$

D'ailleurs, les preuves pour la récurrence multiple de Furstenberg et Katznelson (le théorème I.8) et celle de Bergelson et Leibman (le théorème I.10) prennent le même chemin et se déduisent du fait que les limites inférieures des moyennes associées sont positives.

Il est naturel de se demander si les limites inférieures dans le théorème de récurrence multiple de Furstenberg ci-dessus et aussi dans ses généralisations sont en fait des limites. De nombreux travaux ont été réalisés pour établir la convergence de différents types de moyennes ergodiques multiples :

Dans le cas d'une transformation  $T$ , le premier théorème montre la convergence des moyennes ergodiques multiples le long des progressions arithmétiques  $\{n, 2n, \dots, kn\}$  :

**THÉORÈME I.18** (Host et Kra [41]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soient  $k \geq 1$  un entier et  $f_1, f_2, \dots, f_k$  des fonctions mesurables et bornées sur  $X$ . Alors les moyennes*

$$(1) \quad \frac{1}{N} \sum_{n=1}^N f_1(T^n x) f_2(T^{2n} x) \cdots f_k(T^{kn} x) .$$

*convergent dans  $L^2(\mu)$ .*

Ce théorème a été démontré par Furtenberg [29] pour  $k = 2$ . De nombreux travaux ([21], [22], [23], [32], [40], [56]) ont été consacrés au cas où  $k = 3$ . Le cas général a été démontré par Host et Kra [41], et plus tard par Ziegler [70].

**3.3. Convergence de moyennes polynomiales.** Un résultat de convergence est aussi valable pour les moyennes ergodiques multiples le long de polynômes :

**THÉORÈME I.19** (Host et Kra [42], Leibman [52]). *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé et  $T$  une transformation de  $X$  inversible mesurable et préservant  $\mu$ . Soient  $k \geq 1$  un entier,  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes à valeurs entières, et  $f_1, f_2, \dots, f_k$  des fonctions mesurables et bornées sur  $X$ . Alors les moyennes*

$$(2) \quad \frac{1}{N} \sum_{n=1}^N f_1(T^{p_1(n)}x) f_2(T^{p_2(n)}x) \cdots f_k(T^{p_k(n)}x) .$$

*convergent dans  $L^2(\mu)$ .*

Ce théorème a été d'abord démontré par Bergelson [8] sous l'hypothèse que le système est faiblement mélangeant. Furstenberg et Weiss [32] l'ont démontré dans deux cas particuliers pour  $k = 2$  : pour  $p_1(n) = n$ ,  $p_2(n) = n^2$ , et pour  $p_1(n) = n^2$ ,  $p_2(n) = n^2 + n$ . Le cas général a été démontré par Host et Kra [42] sous de petites restrictions. Ces restrictions ont été surmontées par Leibman [52]. Frantzikinakis et Kra ont montré que si le système est totalement ergodique et les polynômes sont linéairement indépendants, alors la limite de ces moyennes est en fait une constante, égale au produit des intégrales de chaque fonction  $f_i$ .

**3.4. D'autres types de convergence.** D'autres types de moyennes ergodiques multiples comme les moyennes pondérées de la forme

$$\frac{1}{N} \sum_{n=0}^{N-1} c_n f_1(T^n x) f_2(T^{2n} x) \cdots f_k(T^{kn} x) .$$

avec une suite bornée  $\{c_n : n \in \mathbb{Z}\}$  ont été étudiées par Host et Kra [43]. Cette thèse comporte une étude consacrée à généraliser leurs résultats au cas polynomial. Cette étude est faite dans le Chapitre III et a fait l'objet d'un article [17].

Le résultat suivant de convergence sur les cubes est associé à la récurrence cubique, et il est utilisé pour démontrer cette dernière :

Rappelons que nous écrivons  $\epsilon = \epsilon_1 \epsilon_2 \dots \epsilon_k$  avec  $\epsilon_i \in \{0, 1\}$  pour un élément dans  $\{0, 1\}^k$ .

THÉORÈME I.20 (Host et Kra [41]). Soit  $T$  une transformation mesurable inversible préservant la mesure de l'espace probabilisé  $(X, \mathcal{X}, \mu)$ . Soient  $k \geq 1$  un entier et  $f_\epsilon, \epsilon \in \{0, 1\}^k \setminus \{00 \dots 0\}$ ,  $2^k - 1$  fonctions mesurables et bornées sur  $X$ . Alors les moyennes

$$\frac{1}{N^k} \sum_{n_1, n_2, \dots, n_k=0}^{N-1} \prod_{\substack{\epsilon \in \{0,1\}^k \\ \epsilon \neq 00 \dots 0}} f_\epsilon(T^{n_1 \epsilon_1 + \dots + n_k \epsilon_k} x) .$$

convergent dans  $L^2(\mu)$ .

Par exemple, lorsque  $k = 2$  et toutes les fonctions sont égales, l'expression ci-dessus est

$$\frac{1}{N^2} \sum_{n=1}^N \sum_{m=1}^N f(T^n x) f(T^m x) f(T^{n+m} x) .$$

**3.5. Convergence pour plusieurs transformations qui commutent.** Pour plusieurs transformations qui commutent  $T_1, T_2, \dots, T_k$ , les résultats de convergence sont récents et les méthodes sont différentes.

THÉORÈME I.21 (Tao [65]). Soient  $k \geq 1$  un entier et  $T_1, T_2, \dots, T_k$  des transformations mesurables inversibles qui commutent et préservent la mesure de l'espace probabilisé  $(X, \mathcal{X}, \mu)$ . Soient  $f_1, f_2, \dots, f_k$  des fonctions mesurables et bornées sur  $X$ . Alors les moyennes

$$(3) \quad \frac{1}{N} \sum_{n=1}^N f_1(T_1^n x) f_2(T_2^n x) \cdots f_k(T_k^n x) .$$

convergent dans  $L^2(\mu)$ .

Ce théorème a été d'abord démontré par Conze et Lesigne [21] pour  $k = 2$ . Sous l'hypothèse très particulière que toutes les transformations  $T_i$  et toutes les transformations  $T_i T_j^{-1}$ ,  $i \neq j$  sont ergodiques, il a été démontré par Zhang [69] pour  $k = 3$ , et par Frantzikinakis et Kra [27] pour  $k$  arbitraire. Le cas général a été initialement démontré par Tao [65] en utilisant une méthode finitaire. D'autres démonstrations ont été données par Austin [3], Host [39] et Towsner [67]. Parmi celles-ci, les démonstrations d'Austin et de Host ont procédé dans le cadre de la théorie ergodique, celle de Towsner utilise l'analyse nonstandard.

La convergence de moyennes ergodiques multiples pour plusieurs transformations qui commutent le long de polynômes reste une conjecture (explicitement donnée dans [9]) :

CONJECTURE (Bergelson et Leibman [9]). Soient  $k \geq 1$  un entier et  $T_1, T_2, \dots, T_k$  des transformations mesurables inversibles qui commutent et préservent la mesure de l'espace probabilisé  $(X, \mathcal{X}, \mu)$ . Soient  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes à valeurs entières et  $f_1, f_2, \dots, f_k$  des fonctions mesurables et bornées sur  $X$ . Alors les moyennes

$$\frac{1}{N} \sum_{n=1}^N f_1(T_1^{p_1(n)} x) f_2(T_2^{p_2(n)} x) \cdots f_k(T_k^{p_k(n)} x) .$$

convergent dans  $L^2(\mu)$ .

Jusqu'à présent, quelques résultats partiels se trouvent dans une série d'articles ([4], [5], [6]), [20] et [44].

**3.6. Convergence ponctuelle.** Parallèlement à la convergence en norme, il est naturel d'espérer l'existence des résultats de convergence ponctuelle de ces moyennes. Le théorème ergodique de Birkhoff est la contrepartie de convergence ponctuelle du théorème ergodique de von Neumann. Néanmoins, la convergence ponctuelle de différentes moyennes ergodiques multiples est un problème plus difficile. Bourgain [16] a démontré la convergence presque partout des moyennes

$$\frac{1}{N} \sum_{n=1}^N f(T^{an} x) g(T^{bn} x) ,$$

où  $a, b$  sont des entiers. Assani ([1], [2]) a montré la convergence presque partout de certaines moyennes ergodiques multiples sur les cubes.

Dans cette thèse, nous nous restreignons à la convergence en norme.

## CHAPITRE II

### Outils

Dans ce chapitre, nous présentons les outils et les méthodes, souvent récents, de la théorie ergodique que nous allons utiliser à plusieurs endroits dans cette thèse.

#### 1. Notation générale

Généralement, nous écrivons  $(X, \mu)$  pour un espace probabilisé, omettant la  $\sigma$ -algèbre. En cas de besoin, la  $\sigma$ -algèbre de l'espace probabilisé  $(X, \mu)$  s'écrit comme  $\mathcal{X}$ . Tout sous-ensemble de  $X$  est supposé implicitement être mesurable.

Nous supposons implicitement que tous les espaces probabilisés que nous considérons sont standards : c'est-à-dire que  $(X, \mathcal{X}, \mu)$  est isomorphe au segment  $[0, 1]$  muni de sa  $\sigma$ -algèbre de Borel et de la mesure de Lebesgue.

Étant donné un espace probabilisé  $(X, \mathcal{X}, \mu)$ , si  $\mathcal{A}$  et  $\mathcal{B}$  sont deux sous- $\sigma$ -algèbres de  $\mathcal{X}$ , on désigne par  $\mathcal{A} \vee \mathcal{B}$  la plus petite sous- $\sigma$ -algèbre de  $\mathcal{X}$  contenant  $\mathcal{A}$  et  $\mathcal{B}$ .

Si  $(X, \mathcal{X}, \mu)$  est un espace probabilisé et  $S$  est une transformation de  $X$  mesurable et préservant la mesure, nous écrivons  $\mathcal{I}(S)$  pour la sous- $\sigma$ -algèbre de  $\mathcal{X}$  formée des ensembles qui sont invariants sous  $S$ .

Un *système*  $(X, \mu, T_1, \dots, T_k)$  est un espace probabilisé  $X$  muni d'une ou de plusieurs transformations inversibles mesurables qui commutent et préservent la mesure  $\mu$ .

Toutes les fonctions que nous considérons sont supposées être à valeurs réelles.

Une fonction  $f$  est dite *bornée* si elle est mesurable et bornée.

#### 2. Facteurs

**2.1. Facteur.** Un *facteur* d'un système  $(X, \mathcal{X}, \mu, T_1, \dots, T_k)$  est défini par les deux façons suivantes qui sont différentes mais équivalentes :

D'abord, un facteur est une sous- $\sigma$ -algèbre de  $\mathcal{X}$  invariante sous chaque transformation  $T_i$ ,  $i = 1, \dots, k$ .

L'autre caractérisation est qu'un facteur est la donnée d'un système  $(Y, \mathcal{Y}, \nu, S_1, \dots, S_k)$  et d'une application mesurable  $\pi: X \rightarrow Y$  telle que l'image  $\pi(\mu)$  de  $\mu$  sous  $\pi$  soit égale à  $\nu$  et que

$$S_i \circ \pi = \pi \circ T_i ,$$

$i = 1, \dots, k$ ,  $\mu$ -presque partout.

On identifie toujours la  $\sigma$ -algèbre  $\mathcal{Y}$  avec son image réciproque  $\pi^{-1}(\mathcal{Y})$ , qui est une sous- $\sigma$ -algèbre invariante de  $\mathcal{X}$ , donc un facteur au sens de la première définition. Par abus de langage, on dit brièvement que  $Y$  est un facteur de  $X$  et l'application  $\pi$  est appelée l'*application facteur* dans ce cas.

Par exemple, si  $(X, \mathcal{X}, \mu, T)$  et  $(Y, \mathcal{Y}, \nu, S)$  sont deux systèmes, alors chacun est un facteur du système produit  $(X \times Y, \mathcal{X} \times \mathcal{Y}, \mu \times \nu, T \times S)$  et l'application facteur associée est la projection sur la coordonnée appropriée.

Voici un autre exemple plus intéressant :

**EXEMPLE.** Soit  $X = \mathbb{T} \times \mathbb{T}$  muni de la  $\sigma$ -algèbre de Borel  $\mathcal{X}$  et de la mesure de Haar  $\mu$ ,  $\alpha \in \mathbb{T}$ , et l'application  $T: X \rightarrow X$  est définie par

$$T(x, y) = (x + \alpha, y + x) ,$$

alors la projection sur la première coordonnée  $\pi: X \rightarrow \mathbb{T}, (x, y) \mapsto x$  est continue, elle est donc mesurable. Posons  $Y = \mathbb{T}$ , on munit  $Y$  de la  $\sigma$ -algèbre de Borel  $\mathcal{Y}$  et de la mesure de Haar  $\nu$ , et de la transformation  $S$  définie par  $x \mapsto x + \alpha$ , alors on vérifie facilement que  $\nu = \mu \circ \pi^{-1}$  et  $S \circ \pi = \pi \circ T$ . Il vient donc que  $(Y, \mathcal{Y}, \nu, S)$  est un facteur de  $(X, \mathcal{X}, \mu, T)$ .

Si  $Y$  est un facteur de  $X$  avec l'application facteur  $\pi$  et si  $f$  est une fonction intégrable sur  $X$ , nous considérons  $\mathbb{E}(f | \mathcal{Y})$  comme une fonction définie sur  $X$ . Nous appelons l'*espérance conditionnelle de  $f$  par rapport à  $Y$*  et nous notons  $\mathbb{E}(f | Y)$  la fonction sur  $Y$  définie par

$$\mathbb{E}(f | Y) \circ \pi = \mathbb{E}(f | \mathcal{Y}) .$$

Elle est caractérisée par

$$\text{pour toute fonction } g \in L^\infty(\nu), \quad \int_X f \cdot g \circ \pi d\mu = \int_Y \mathbb{E}(f|Y) \cdot g d\nu .$$

**2.2. Facteur de Kronecker.** Un facteur important est le facteur de Kronecker. Nous rappelons ici la définition et quelques propriétés classiques. Dans cette section,  $(X, \mathcal{X}, \mu, T)$  est un système ergodique avec une seule transformation. *Le facteur de Kronecker* de  $(X, \mathcal{X}, \mu, T)$  est la sous- $\sigma$ -algèbre  $\mathcal{Z}_1$  de  $\mathcal{X}$  engendrée par les fonctions propres de

l'opérateur unitaire associé  $U$  au système  $X$ , où  $U: L^2(\mu) \rightarrow L^2(\mu)$  est défini par

$$(Uf)(x) = f(Tx) .$$

Cette  $\sigma$ -algèbre est aussi la plus petite sous- $\sigma$ -algèbre de  $\mathcal{X}$  telle que toute fonction invariante du système  $(X \times X, \mu \times \mu, T \times T)$  soit mesurable par rapport à  $\mathcal{Z}_1 \times \mathcal{Z}_1$ . Un résultat classique est que le facteur de Kronecker peut être muni d'une structure de rotation sur un groupe :

**THÉORÈME II.1.** *Le facteur de Kronecker d'un système ergodique est isomorphe à un système  $(Z_1, \mathcal{Z}_1, m, R)$ , où  $Z_1$  est un groupe compact abélien,  $\mathcal{Z}_1$  est sa  $\sigma$ -algèbre de Borel,  $m$  est sa mesure de Haar, et  $Rz = z + \alpha$  pour un  $\alpha \in Z_1$  fixé.*

**EXEMPLE.** Soit  $X = \mathbb{T} \times \mathbb{T}$  muni de la  $\sigma$ -algèbre de Borel  $\mathcal{X}$  et de la mesure de Haar  $m$ ,  $\alpha \in \mathbb{T}$  un nombre irrationnel, et l'application  $T: X \rightarrow X$  est définie par

$$T(x, y) = (x + \alpha, y + x) ,$$

alors on peut montrer que le système  $(X, \mathcal{X}, m, T)$  est ergodique. La sous- $\sigma$ -algèbre  $\mathcal{Z}_1$  est celle engendrée par la projection sur la première coordonnée  $\pi: X \rightarrow \mathbb{T}, (x, y) \mapsto x$ . Il vient que la rotation  $x \mapsto x + \alpha$  sur  $\mathbb{T}$  est le facteur de Kronecker de  $X$  et la projection  $\pi$  est l'application facteur.

### 3. Facteur caractéristique

La notion de facteur caractéristique a été utilisée depuis longtemps de façon implicite dans des ouvrages (par exemple dans [29]) et cette notion a été introduite explicitement pour la première fois dans [32]. La définition précise dépend du contexte. Pour nous, étant donné un système  $(X, \mu, T_1, \dots, T_k)$ , un *facteur caractéristique* pour une moyenne ergodique  $A_N(f_1, \dots, f_l)$  (qui est une fonction multilinéaire de  $(f_1, \dots, f_l)$ ) dont nous étudions la convergence est un facteur  $Y$  de  $X$  tel qu'on peut remplacer une des fonctions  $f_i$  par son espérance conditionnelle  $\mathbb{E}(f_i \mid Y)$  sans changer le comportement asymptotique des moyennes  $A_N(f_1, \dots, f_l)$ . Autrement dit, s'il existe un entier  $1 \leq i \leq l$  tel que  $\mathbb{E}(f_i \mid Y) = 0$ , alors les moyennes  $A_N(f_1, \dots, f_l)$  convergent vers 0 dans  $L^2(\mu)$ .

Le système lui-même est toujours un facteur caractéristique. Un exemple simple est que le facteur caractéristique pour le système  $(X, \mu, T)$  et la moyenne  $\frac{1}{N} \sum_{n=1}^N f(T^n x)$  est la sous- $\sigma$ -algèbre  $T$ -invariante  $\mathcal{I}(T)$  d'après le théorème ergodique de von Neumann.

Un autre exemple est que le facteur caractéristique pour le système ergodique  $(X, \mu, T)$  et la moyenne double

$$\frac{1}{N} \sum_{n=1}^N f_1(T^n x) f_2(T^{2n} x)$$

est le facteur de Kronecker :

THÉORÈME II.2 (Furstenberg [29]). *Soient  $(X, \mu, T)$  un système ergodique,  $(Z, \mathcal{Z}_1, m, T)$  son facteur de Kronecker, et  $f_1, f_2$  des fonctions bornées. Alors la suite*

$$\left\| \frac{1}{N} \sum_{n=1}^N f_1(T^n x) f_2(T^{2n} x) - \frac{1}{N} \sum_{n=1}^N \mathbb{E}(f_1 | \mathcal{Z}_1)(T^n x) \mathbb{E}(f_2 | \mathcal{Z}_1)(T^{2n} x) \right\|_{L^2(\mu)}$$

tend vers 0 lorsque  $N \rightarrow \infty$ .

La notion devient utile seulement quand on peut trouver un facteur caractéristique qui a de bonnes propriétés, par exemple qui a de riches structures géométriques ou algébriques. Pour illustrer comment on se sert d'un facteur caractéristique, nous montrons comment on obtient la convergence des moyennes

$$\frac{1}{N} \sum_{n=1}^N f_1(T^n x) f_2(T^{2n} x),$$

en admettant que leur facteur caractéristique est le facteur de Kronecker. D'après le théorème ci-dessus, il suffit de montrer que les moyennes

$$\frac{1}{N} \sum_{n=1}^N \mathbb{E}(f_1 | \mathcal{Z}_1)(T^n x) \mathbb{E}(f_2 | \mathcal{Z}_1)(T^{2n} x)$$

convergent. On peut donc se restreindre au cas où les fonctions  $f_1$  et  $f_2$  sont mesurables par rapport à  $\mathcal{Z}_1$ . Comme les combinaisons linéaires de fonctions propres de  $X$  sont denses dans  $L^2(\mathcal{Z}_1, \mu)$ , il suffit de montrer la convergence quand  $f_1, f_2$  sont des fonctions propres de  $X$ . Supposons  $f_1(Tx) = \lambda_1 f_1(x)$ ,  $f_2(Tx) = \lambda_2 f_2(x)$ , pour  $\lambda_1, \lambda_2 \in \mathbb{C}$  quelconques. Alors les moyennes

$$\frac{1}{N} \sum_{n=1}^N f_1(T^n x) f_2(T^{2n} x) = f_1(x) f_2(x) \frac{1}{N} \sum_{n=1}^N \lambda_1^n \lambda_2^{2n}$$

convergent, ce qui conclut la preuve.

## 4. Semi-normes

**4.1. Semi-normes.** Les facteurs dans cette thèse sont construits à partir de semi-normes. Dans cette section, nous présenterons une définition générale de semi-normes : les *semi-normes boîtes*  $\|\cdot\|_{T_1, \dots, T_k}$  introduites dans [39] associées à plusieurs transformations qui commutent  $T_1, T_2, \dots, T_k$ .

Dans le cas où toutes les transformations sont égales à  $T$ , soit  $T_1 = T_2 = \dots = T_k = T$ , les semi-normes obtenues sont en fait les *semi-normes de Host-Kra* introduites dans [41]. Dans ce cas, on écrit  $\|\cdot\|_k$  pour les semi-normes.

**La construction :** Les semi-normes sont définies à partir de certaines mesures qui se définissent par récurrence. Soient  $k \geq 1$  un entier et  $(X, \mu, T_1, T_2, \dots, T_k)$  un système où  $T_1, T_2, \dots, T_k$  commutent. On pose  $X^* = X^{2^k}$ , alors les points de  $X^*$  s'écrivent comme  $\mathbf{x} = (\mathbf{x}_\epsilon : \epsilon \in \{0, 1\}^k)$ .

Pour  $1 \leq i \leq k$ ,  $T_i^\Delta$  désigne la *transformation diagonale*  $T_i \times \dots \times T_i$  sur  $X^*$  :

$$(T_i^\Delta x)_\epsilon = T_i x_\epsilon, \quad \text{pour tout } \epsilon \in \{0, 1\}^k,$$

et la *transformation latérale*  $T_i^*$  sur  $X^*$  est définie par

$$(4) \quad (T_i^* x)_\epsilon = \begin{cases} T_i x_\epsilon & \text{si } \epsilon_i = 0, \\ x_\epsilon & \text{si } \epsilon_i = 1, \end{cases} \quad \text{pour tout } \epsilon \in \{0, 1\}^k.$$

D'abord nous définissons la mesure  $\mu_{T_1}$  sur  $X^2$  comme la *mesure produit relativement indépendant de  $\mu$  au-dessus de  $\mathcal{I}(T_1)$* , c'est-à-dire pour tous  $f_0, f_1 \in L^\infty(\mu)$ , nous avons

$$\int (f_0 \otimes f_1)(x_0, x_1) d\mu_{T_1}(x_0, x_1) = \int \mathbb{E}(f_0 | \mathcal{I}(T_1)) \cdot \mathbb{E}(f_1 | \mathcal{I}(T_1)) d\mu.$$

Alors  $\mu_{T_1}$  est invariante sous toute transformation  $T_i \times T_i$ ,  $i \in \{1, \dots, k\}$  et aussi sous  $T_1 \times \text{Id}$ .

Ensuite nous définissons la mesure  $\mu_{T_1, T_2}$  sur  $X^4 = X^2 \times X^2$  comme la mesure produit relativement indépendant de  $\mu_{T_1}$  au-dessus de  $\mathcal{I}(T_2 \times T_2)$  :

$$\mu_{T_1, T_2} = \mu_{T_1} \times_{\mathcal{I}(T_2 \times T_2)} \mu_{T_1}.$$

Alors  $\mu_{T_1, T_2}$  est invariante sous toute transformation diagonale  $T_i \times T_i \times T_i \times T_i$ ,  $i \in \{1, \dots, k\}$ , et aussi sous toutes les transformations latérales  $T_1 \times \text{Id} \times T_1 \times \text{Id}$  et  $T_2 \times T_2 \times \text{Id} \times \text{Id}$  sur  $X^4$ .

De la même manière, pour  $j \leq k$ , nous obtenons la mesure  $\mu_{T_1, T_2, \dots, T_j}$  sur  $X^{2^j}$  qui est invariante sous toute transformation diagonale  $T_i \times T_i \times \dots \times T_i$ ,  $i \in \{1, \dots, k\}$  et aussi sous toutes les transformations

latérales associées à  $T_1, T_2, \dots, T_j$  comme dans (4) avec  $j$  substitué à  $k$ . Nous définissons

$$\mu_{T_1, \dots, T_{j+1}} = \mu_{T_1, \dots, T_j} \times_{\mathcal{I}(T_{j+1} \times \dots \times T_{j+1})} \mu_{T_1, \dots, T_j}.$$

En itérant cette construction, après  $k$  étapes, nous obtenons une mesure  $\mu^* = \mu_{T_1, \dots, T_k}$  sur  $X^* = X^{2^k}$ .

Dans le cas d'une transformation  $T$ , autrement dit si  $T_1 = T_2 = \dots = T_k = T$ , on note  $\mu^{[k]}$  ([41]) pour la mesure  $\mu^* = \mu_{T, T, \dots, T}$ .

Une propriété importante pour la mesure  $\mu^*$  est le lemme suivant :

**THÉORÈME II.3** (Lemme 1, [39]). *Soient  $f_\epsilon \in L^\infty(\mu)$ ,  $\epsilon \in \{0, 1\}^k$ , alors*

$$(5) \quad \int \prod_{\epsilon \in \{0, 1\}^k} f_\epsilon(x_\epsilon) d\mu^* = \lim_{N_k \rightarrow \infty} \frac{1}{N_k} \sum_{n_k=0}^{N_k-1} \dots$$

$$\lim_{N_2 \rightarrow \infty} \frac{1}{N_2} \sum_{n_2=0}^{N_2-1} \lim_{N_1 \rightarrow \infty} \frac{1}{N_1} \sum_{n_1=0}^{N_1-1} \int \prod_{\epsilon \in \{0, 1\}^k} T_1^{(1-\epsilon_1)n_1} \dots T_k^{(1-\epsilon_k)n_k} f_\epsilon d\mu.$$

De plus, la relation (5) est aussi valable pour les moyennes sur toutes suites d'intervalles  $[M_1, N_1), \dots, [M_k, N_k)$  dont la longueur  $N_i - M_i$  ( $1 \leq i \leq k$ ) tend vers  $\infty$ .

Pour  $f \in L^\infty(\mu)$ , nous définissons

$$(6) \quad \|f\|_{T_1, \dots, T_k} = \left( \int_{X^*} \prod_{\epsilon \in \{0, 1\}^k} f(x_\epsilon) d\mu^*(x) \right)^{1/2^k},$$

Il est montré dans la proposition 2 de [41] que  $\|\cdot\|_{T_1, \dots, T_k}$  est une semi-norme sur  $L^\infty(\mu)$ .

Dans le cas d'une transformation  $T_1 = T_2 = \dots = T_k = T$ ,  $\|\cdot\|_k = \|\cdot\|_{T, T, \dots, T}$  est la *semi-norme de Host-Kra* ([41]).

**4.2. Les semi-normes contrôlent les moyennes.** Les semi-normes servent à contrôler le comportement asymptotique de moyennes ergodiques multiples, par exemple les semi-normes de Host-Kra  $\|\cdot\|_k$  contrôlent celui des moyennes ergodiques multiples le long des progressions arithmétiques  $\{n, 2n, \dots, kn\}$  :

**THÉORÈME II.4** (Host et Kra [41]). *Soient  $(X, \mu, T)$  un système et  $f_1, \dots, f_k$  des fonctions bornées avec  $|f_i| \leq 1, i = 1, \dots, k$ . Alors*

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N f_1(T^n x) f_2(T^{2n} x) \dots f_k(T^{kn} x) \right\|_{L^2(\mu)} \leq \min_{1 \leq i \leq k} (i \|f_i\|_k).$$

Un résultat similaire est aussi valable pour les moyennes ergodiques multiples le long de polynômes :

**THÉORÈME II.5** (Host et Kra [42], Leibman [52]). *Soient  $(X, \mu, T)$  un système et  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes non constants à valeurs entières tels que pour tous  $1 \leq i \neq j \leq k$ , le polynôme  $p_i - p_j$  n'est pas une constante. Alors il existe un entier  $l \geq 1$ , tel que pour toutes fonctions bornées  $f_1, f_2, \dots, f_k$  sur  $X$ ,*

$$\lim_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N f_1(T^{p_1(n)}x) f_2(T^{p_2(n)}x) \cdots f_k(T^{p_k(n)}x) \right\|_{L^2(\mu)} = 0 .$$

lorsque  $\|f_i\|_l = 0$  pour au moins une valeur des  $i = 1, \dots, k$ .

Les semi-normes boîtes  $\|\cdot\|_{T_1, \dots, T_k}$  contrôlent le comportement asymptotique des moyennes ergodiques multiples pour plusieurs transformations qui commutent :

**THÉORÈME II.6** (Host [39]). *Soit  $(X, \mu, T_1, \dots, T_k)$  un système. Alors pour toutes fonctions  $f_1, \dots, f_k \in L^\infty(\mu)$  avec  $\|f_i\|_{L^\infty(\mu)} \leq 1$  pour  $2 \leq i \leq k$ ,*

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N f_1(T_1^n x) f_2(T_2^n x) \cdots f_k(T_k^n x) \right\|_{L^2(\mu)} \leq \|f_1\|_{T_1, T_2 T_1^{-1}, \dots, T_k T_1^{-1}} .$$

## 5. Une $\sigma$ -algèbre caractéristique sur $X$

Pour chaque  $k \geq 1$ , il existe une sous- $\sigma$ -algèbre sur  $X$  qui correspond à la semi-norme  $\|\cdot\|_{T_1, \dots, T_k}$ . Nous décrivons dans cette section cette  $\sigma$ -algèbre.

Il convient ici d'utiliser une autre notation : nous écrivons  $[k] = \{1, \dots, k\}$  et identifions  $\{0, 1\}^k$  avec la famille de sous-ensembles de  $[k]$ . Par conséquent, pour  $\epsilon \in \{0, 1\}^k$  et  $1 \leq i \leq k$ , l'assertion " $\epsilon_i = 1$ " est équivalente à " $i \in \epsilon$ ". En particulier, nous écrivons  $\emptyset$  pour  $00 \dots 0 \in \{0, 1\}^k$ .

Nous écrivons  $X^\sharp = X^{2^k - 1}$  et identifions  $X^*$  avec  $X \times X^\sharp$  en isolant la coordonnée  $\emptyset$  de chaque point : chaque point  $x \in X^*$  est écrit comme

$$x = (x_\emptyset, x^\sharp), \text{ où } x^\sharp = (x_\epsilon : \epsilon \in \{0, 1\}^k, \epsilon \neq \emptyset) \in X^\sharp.$$

Nous écrivons  $\mu^\sharp$  pour l'image de  $\mu^*$  sous la projection  $x \mapsto x^\sharp$ .

**LEMME II.7** (Lemma 4, [39]). *Soit  $\mathcal{Z}$  la sous- $\sigma$ -algèbre sur  $X$  formée des ensembles  $B$  tels qu'il existe un sous-ensemble  $A$  de  $X^\sharp$  vérifiant la relation*

$$(7) \quad \mathbf{1}_B(x_\emptyset) = \mathbf{1}_A(x^\sharp), \quad \text{pour } \mu^* \text{-presque partout } x = (x_\emptyset, x^\sharp) \in X^*.$$

Alors pour toute  $f \in L^\infty(\mu)$ , nous avons

$$(8) \quad \|f\|_{T_1, \dots, T_k} = 0 \text{ si et seulement si } \mathbb{E}_\mu(f \mid \mathcal{Z}) = 0 .$$

En particulier, si le système  $X$  est muni d'une transformation  $T$ , nous désignons par  $\mathcal{Z}_{k-1}$  la sous- $\sigma$ -algèbre associée à la semi-norme  $\|\cdot\|_k$  ( $k \geq 1$ ) et le facteur correspondant  $Z_{k-1}$  est appelé le *facteur de Host-Kra*.

DÉFINITION II.8 (Host et Kra [41], voir aussi [20]). Le *facteur de Host-Kra*  $Z_{k-1}$  est le facteur correspondant à la sous- $\sigma$ -algèbre caractérisée par

$$(9) \quad \|f\|_k = 0 \text{ si et seulement si } \mathbb{E}_\mu(f \mid \mathcal{Z}_{k-1}) = 0 .$$

Si le système  $(X, \mu, T)$  est ergodique, il est montré dans [41] que le facteur  $Z_{k-1}$  possède une structure "algébrique" que nous allons détailler dans la section suivante. Nous avons vu que  $Z_0$  est le facteur trivial et  $Z_1$  est le facteur de Kronecker.

Revenons au théorème I.18, en appliquant le théorème II.4, on en déduit que les moyennes (1) convergent vers 0 dans  $L^2(\mu)$  lorsqu'il existe au moins une valeur de  $i \in \{1, \dots, k\}$  telle que  $\mathbb{E}(f_i \mid \mathcal{Z}_{k-1}) = 0$ . Il s'ensuit donc que le facteur  $Z_{k-1}$  est un facteur caractéristique pour la moyenne (1). Par conséquent, pour montrer le théorème I.18, il suffit de le montrer quand toutes les fonctions  $f_i$ ,  $i = 1, \dots, k$  sont mesurables par rapport à  $Z_{k-1}$ .

De même, revenons au théorème I.19, en appliquant le théorème II.5, on en déduit qu'il existe un entier  $l \geq 1$ , tel que les moyennes (2) convergent vers 0 dans  $L^2(\mu)$  lorsqu'il existe au moins une valeur de  $i \in \{1, \dots, k\}$  telle que  $\mathbb{E}(f_i \mid \mathcal{Z}_{l-1}) = 0$ . Il s'ensuit donc qu'il existe un entier  $l \geq 1$  tel que le facteur  $Z_{l-1}$  est un facteur caractéristique pour la moyenne (2). Par conséquent, pour montrer le théorème I.19, il suffit de le montrer quand toutes les fonctions  $f_i$ ,  $i = 1, \dots, k$  sont mesurables par rapport à  $Z_{l-1}$ .

Revenons au théorème I.21, désignons par  $\mathcal{Z}$  la sous- $\sigma$ -algèbre associée à la semi-norme  $\|\cdot\|_{T_1, T_2 T_1^{-1}, \dots, T_k T_1^{-1}}$ . On en déduit, en appliquant le théorème II.6, que les moyennes (3) convergent vers 0 dans  $L^2(\mu)$  lorsque  $\mathbb{E}(f_1 \mid \mathcal{Z}) = 0$ . Il s'ensuit donc que le facteur  $Z$  correspondant à  $\mathcal{Z}$  est un facteur caractéristique pour la moyenne (3).

Contrairement au cas d'une seule transformation, le facteur  $Z$  ne possède généralement qu'une structure faible et nous ne pouvons pas l'utiliser directement pour démontrer le théorème I.21.

## 6. Nilsystèmes et nilsuites

Dans cette section, nous traitons des systèmes munis d'une transformation  $(X, \mu, T)$ .

**6.1. Nilsystèmes.** Supposons que  $(X, \mu, T)$  est un système ergodique. Un des résultats principaux de [41] est le théorème de structure suivant qui donne une description "algébrique" de  $Z_k$  :

**THÉORÈME II.9** (Host et Kra [41]). *Pour chaque  $k \geq 1$ ,  $Z_k$  est une limite projective de nilsystèmes d'ordre  $k$ .*

**DÉFINITION II.10.** Soient  $k \geq 1$  un entier,  $G$  un groupe de Lie nilpotent d'ordre  $k$  et  $\Gamma$  un sous-groupe discret cocompact de  $G$ .

La variété  $X = G/\Gamma$  est appelée *une nilvariété* d'ordre  $k$ .

Le groupe  $G$  agit naturellement sur  $X$  par translation à gauche : si  $g \in G$  et  $x \in X$ , la translation  $T_g$  par  $g$  est donnée par  $T_g(x\Gamma) = (gx)\Gamma$ . Il existe une mesure de probabilité unique  $\mu$  (la mesure de Haar) sur  $X$  qui est invariante sous cette action. Fixons un élément  $g \in G$  ; le système  $(X, \mu, T_g)$  est appelé un *nilsystème d'ordre  $k$*  et l'action  $T_g$  est appelée une *nilrotation*.

Le résultat de convergence suivant pour les nilsystèmes a été montré par Lesigne [54] au cas où le groupe  $G$  est connexe, et par Leibman pour le cas général [50].

**THÉORÈME II.11** (Lesigne [54], Leibman [50]). *Soient  $(X = G/\Gamma, \mu, T)$  un nilsystème et  $f$  une fonction continue sur  $X$ . Alors les moyennes*

$$\frac{1}{N} \sum_{n=1}^N f(T^n x)$$

*convergent pour tout  $x \in X$ .*

En appliquant ce théorème au système

$$(X^k, S: = T \times T^2 \times \cdots \times T^k)$$

avec la fonction continue

$$F(x_1, x_2, \cdots, x_k) = f_1(x_1)f_2(x_2) \cdots f_k(x_k) ,$$

on en déduit que le théorème I.18 est valable dans un nilsystème. Ainsi, il est valable dans une limite projective de nilsystèmes par un argument de densité. D'après le théorème de structure II.9 donné plus haut, les moyennes (1) convergent si toutes les fonctions  $f_i, i = 1, \dots, k$  sont mesurables par rapport à  $Z_{k-1}$ , ce qui montre le théorème I.18 d'après les raisonnements donnés dans la Section 5.

Le théorème précédent a été généralisé par Leibman [51] :

THÉORÈME II.12 (Leibman [51]). Soient  $X = G/\Gamma$  une nilvariété,  $T_1, \dots, T_k \in G$ , et  $p_1, \dots, p_k$  des polynômes  $\mathbb{Z}^d \rightarrow \mathbb{Z}$ . Alors les moyennes

$$\frac{1}{|\Phi_N|} \sum_{u \in \Phi_N} f(T_1^{p_1(u)} \dots T_k^{p_k(u)} x)$$

convergent lorsque  $N \rightarrow \infty$  pour toute  $f \in C(X)$  et toute suite de Følner  $\{\Phi_N\}_{N=1}^\infty$  dans  $\mathbb{Z}^d$ .

En appliquant ce théorème au système  $(X^k, S_1, S_2, \dots, S_k)$  où

$$S_1: = T_1 \times \text{Id} \times \dots \times \text{Id},$$

$$S_2: = \text{Id} \times T_2 \times \dots \times \text{Id},$$

...

$$S_k: = \text{Id} \times \dots \times \text{Id} \times T_k,$$

avec la fonction continue

$$F(x_1, x_2, \dots, x_k) = f_1(x_1) f_2(x_2) \dots f_k(x_k),$$

on en déduit que le théorème I.19 est valable dans un nilsystème. Par le même argument comme ci-dessus, le théorème I.19 s'en déduit.

**6.2. Nilsuites.** La notion de *nilsuite* est une généralisation de celle de suite presque-périodique, obtenue en remplaçant les groupes abéliens par des groupes nilpotents.

DÉFINITION II.13 (Bergelson, Host et Kra [14]). Soient  $k \geq 1$  un entier et  $(X, \mu, T)$  un nilsystème d'ordre  $k$ . Soient  $f: X \rightarrow \mathbb{C}$  une fonction continue et  $x_0 \in X$ . La suite  $(f(T^n x_0): n \in \mathbb{Z})$  est appelée une *nilsuite d'ordre  $k$  de base*. Une *suite d'ordre  $k$*  est une limite uniforme de nilsuites d'ordre  $k$  de base.

Les nilsuites ont été introduites pour la première fois en théorie ergodique dans l'étude des corrélations multiples :

Avant de donner le théorème, nous rappelons une définition :

DÉFINITION II.14. Soit  $\{a_n: n \in \mathbb{Z}\}$  une suite bornée de nombres réels, on dit que  $a_n$  tend vers 0 en densité uniforme, si

$$\lim_{N \rightarrow \infty} \sup_{M \in \mathbb{Z}} \frac{1}{N} \sum_{n=M}^{M+N-1} |a_n| = 0.$$

THÉORÈME II.15 (Bergelson, Host et Kra [14]). Soient  $(X, \mu, T)$  un système et  $f \in L^\infty(\mu)$ . Pour chaque  $k \geq 1$  un entier, on définit

$$I_f(k, n): = \int f(x) f(T^n x) \dots f(T^{kn} x) d\mu(x).$$

Alors la suite  $I_f(k, n)$  est une somme d'une suite tendant vers zéro en densité uniforme et une nilsuite d'ordre  $k$ .

Ce résultat a été utilisé dans [14] pour montrer le théorème de récurrence multiple quantitatif (le théorème I.7) concernant l'intersection  $A \cap T^n A \cap \dots \cap T^{kn} A$ ,  $k = 2, 3$ .

## 7. Systèmes magiques

Dans cette section, nous traitons des systèmes munis de plusieurs transformations  $(X, \mu, T_1, \dots, T_k)$ . Les systèmes magiques sont introduits par Host [39] dans sa démonstration du théorème I.21. Désignons par  $Z$  le facteur associé à la semi-norme  $\|\cdot\|_{T_1, T_2 T_1^{-1}, \dots, T_k T_1^{-1}}$ . Nous avons vu dans la section 5 que  $Z$  est un facteur caractéristique pour la moyenne (3), mais nous ne pouvons pas utiliser ce facteur puisqu'il ne possède qu'une structure faible. Les approches d'Austin [3] et de Host [39] pour le théorème I.21 procèdent dans la direction opposée en construisant une extension du système original, appelée *pleasant* (suivant la terminologie dans [3]) et *magique* (suivant la terminologie dans [39]), dans laquelle nous pouvons trouver un facteur caractéristique qui possède de bonnes propriétés pour montrer la convergence. Dans l'extension introduite par Host, ce facteur caractéristique est le facteur  $Z$  par construction.

Dans cette section, nous présenterons les systèmes magiques et leurs propriétés.

DÉFINITION II.16 (Host [39]). Un système  $(X, \mu, T_1, \dots, T_k)$  est dit *magique* si pour toute fonction  $f \in L^\infty(\mu)$  vérifiant

$$\mathbb{E}(f \mid \bigvee_{i=1}^k \mathcal{I}(T_i)) = 0 ,$$

nous avons  $\|f\|_{T_1, \dots, T_k} = 0$ .

Un des résultats principaux de [39] est le théorème suivant :

THÉORÈME II.17 (Théorème 2, [39]). *Le système  $(X^*, \mu^*, T_1^*, \dots, T_k^*)$  (voir Section 4.1) est magique.*

COROLLAIRE II.18. *Chaque système admet une extension magique.*

Nous montrons dans cette thèse que ce résultat est aussi valable si nous nous restreignons aux systèmes ergodiques :

THÉORÈME II.19 (Théorème 3.12, [19]). *Chaque système ergodique admet une extension magique ergodique.*

## 8. Lemme de van der Corput

Pour montrer les inégalités dans la section 4.2, nous allons utiliser plusieurs extensions de la classique inégalité de van der Corput, voir par exemple [48]. Ces extensions traitent des suites dans un espace de Hilbert  $\mathcal{H}$ , avec la norme  $\|\cdot\|$  et le produit hermitien  $\langle \cdot | \cdot \rangle$ . On désigne par  $Re(z)$  la partie réelle du nombre complexe  $z$ .

THÉORÈME II.20 (voir par exemple [8]). *Soit  $\{x_n\}$  une suite dans  $\mathcal{H}$ . Alors pour tous entiers  $N$  et  $H$  tels que  $1 \leq H \leq N$ ,*

$$H^2 \left\| \sum_{n=1}^N x_n \right\|^2 \leq H(N+H-1) \sum_{n=1}^N \|x_n\|^2 + 2(N+H-1) \sum_{h=1}^{H-1} (H-h) \sum_{n=1}^{H-h} Re \langle x_n | x_{n+h} \rangle .$$

En passant à la limite, nous obtenons :

THÉORÈME II.21. *Soit  $\{x_n\}$  une suite bornée dans  $\mathcal{H}$ . Alors*

$$\limsup_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N x_n \right\|^2 \leq \limsup_{H \rightarrow \infty} \frac{1}{H} \sum_{h=1}^H \limsup_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n=1}^N \langle x_n | x_{n+h} \rangle \right| .$$

Le lemme suivant est une version multidimensionnelle du lemme de van der Corput qui peut être retrouvée dans la preuve du lemme A6 de [12]. Écrivons  $n = (n_1, \dots, n_k)$  pour un point dans  $\mathbb{Z}^k$ .

THÉORÈME II.22. *Soit  $\{x_n : n \in \mathbb{Z}^k\}$  une suite bornée dans  $\mathcal{H}$ . Supposons que pour tout  $h = (h_1, \dots, h_k) \in \mathbb{Z}^k$*

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1 \leq n_1 < N_1 \\ \dots \\ M_k \leq n_k < N_k}} Re \langle x_{n+h} | x_n \rangle \longrightarrow \gamma_h$$

*lorsque  $N_1 - M_1, \dots, N_k - M_k \rightarrow +\infty$  et que*

$$\sum_{\substack{-H \leq h_1 \leq H \\ \dots \\ -H \leq h_k \leq H}} \prod_{i=1}^k \frac{H - |h_i|}{H^2} \cdot \gamma_h \longrightarrow 0$$

*lorsque  $H \rightarrow +\infty$ . Alors*

$$\left\| \prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{M_1 \leq n_1 < N_1 \\ \dots \\ M_k \leq n_k < N_k}} x_n \right\| \rightarrow 0$$

*lorsque*  $N_1 - M_1, \dots, N_k - M_k \rightarrow +\infty$ .

## CHAPITRE III

### Convergence de moyennes ergodiques multiples pondérées le long de polynômes

L'étude présentée dans ce chapitre fait l'objet de l'article [17].

#### 1. Introduction

Commençons par un résultat classique dans la théorie ergodique : Pour  $t \in \mathbb{T}$ , on utilise la notation standard  $e(t) = e^{2\pi it}$ .

**THÉORÈME III.1** (Théorème ergodique de Wiener-Wintner). *Soient  $(X, \mu, T)$  un système ergodique et  $\phi$  une fonction bornée sur  $X$ . Alors il existe un sous-ensemble  $X_0 \subset X$  avec  $\mu(X_0) = 1$  tel que les moyennes*

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(T^n x) e(nt)$$

*convergent lorsque  $N \rightarrow \infty$  pour tout  $x \in X_0$  et tout  $t \in \mathbb{T}$ .*

Le point important ici est que l'ensemble  $X_0$  ne dépend pas du choix de  $t$ .

Le résultat suivant est un corollaire immédiat du théorème spectral :

**COROLLAIRE III.2** (du théorème spectral). *Soit  $a = (a_n : n \in \mathbb{Z})$  une suite bornée telle que les moyennes*

$$\frac{1}{N} \sum_{n=0}^{N-1} a_n e(nt)$$

*convergent lorsque  $N \rightarrow \infty$  pour tout  $t \in \mathbb{T}$ . Alors pour tout système  $(Y, \nu, S)$  et toute fonction  $f \in L^\infty(\nu)$ , les moyennes*

$$\frac{1}{N} \sum_{n=0}^{N-1} a_n f(S^n x)$$

*convergent dans  $L^2(\nu)$ .*

En intégrant ces deux résultats, on obtient :

THÉORÈME III.3. Soient  $(Y, \nu, S)$  un système ergodique et  $\phi \in L^\infty(\nu)$ . Alors il existe  $Y_0 \subset Y$  avec  $\nu(Y_0) = 1$  tel que, pour tout  $y \in Y_0$ , tout système  $(X, \mu, T)$ , et toute fonction  $f \in L^\infty(\mu)$ , les moyennes

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(S^n y) f(T^n x) .$$

convergent dans  $L^2(\mu)$ .

On dit que pour tout  $y \in Y_0$ , la suite  $(\phi(S^n y))$  est *universellement bonne pour la convergence en moyenne de moyennes ergodiques*.

Host et Kra [43] ont généralisé ces résultats pour des moyennes ergodiques multiples. Ils ont montré une généralisation suivante du théorème de Wiener-Wintner, où la suite exponentielle est remplacée par une nilsuite :

THÉORÈME III.4 (Host et Kra [43]). Soient  $(X, \mu, T)$  un système ergodique et  $\phi$  une fonction bornée sur  $X$ . Alors il existe un sous-ensemble  $X_0 \subset X$  avec  $\mu(X_0) = 1$  tel que les moyennes

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(T^n x) b_n$$

convergent lorsque  $N \rightarrow \infty$  pour tout  $x \in X_0$  et toute nilsuite  $b = (b_n : n \in \mathbb{Z})$ .

Parallèlement, ils ont montré la version nilpotente suivante du corollaire du théorème spectral :

THÉORÈME III.5 (Host et Kra [43]). Soient  $k \geq 1$  un entier et  $a = (a_n : n \in \mathbb{Z})$  une suite bornée telle que les moyennes

$$\frac{1}{N} \sum_{n=0}^{N-1} a_n b_n$$

convergent lorsque  $N \rightarrow \infty$  pour toute nilsuite  $b = (b_n : n \in \mathbb{Z})$  d'ordre  $k$ . Alors pour tout système  $(X, \mu, T)$  et toutes fonctions  $f_1, \dots, f_k \in L^\infty(\nu)$ , les moyennes

$$\frac{1}{N} \sum_{n=0}^{N-1} a_n f_1(T^n x) f_2(T^{2n} x) \cdots f_k(T^{kn} x)$$

convergent dans  $L^2(\mu)$ .

En intégrant ces deux résultats, ils ont obtenu :

**THÉORÈME III.6** (Host et Kra [43]). *Soient  $(Y, \nu, S)$  un système ergodique et  $\phi \in L^\infty(\nu)$ . Alors il existe  $Y_0 \subset Y$  avec  $\nu(Y_0) = 1$  tel que, pour tout  $y \in Y_0$ , tout système  $(X, \mu, T)$ , tout entier  $k \geq 1$  et toutes fonctions  $f_1, \dots, f_k \in L^\infty(\mu)$  les moyennes*

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(S^n y) f_1(T^n x) f_2(T^{2n} x) \cdots f_k(T^{kn} x) .$$

*convergent dans  $L^2(\mu)$ .*

Par conséquent, pour tout  $y \in Y_0$ , la suite  $(\phi(S^n y))$  est *universellement bonne pour la convergence de moyennes ergodiques multiples le long de progressions arithmétiques*.

On dit qu'une suite  $(a_n)$  bornée vérifie la *propriété  $\mathbf{P}$* , si pour tout entier  $k \geq 2$ , pour toute suite  $b = (b_n : n \in \mathbb{Z})$  d'ordre  $k$ , les moyennes

$$\frac{1}{N} \sum_{n=0}^{N-1} a_n b_n$$

convergent.

La proposition suivante montre qu'une certaine classe de suites vérifie la propriété  $\mathbf{P}$ , donc est universellement bonne d'après le théorème III.5.

**PROPOSITION III.7** (Proposition 7.1, [43]). *Soient  $(X, \mu, T)$  un système uniquement ergodique et  $k \geq 1$ . Soit  $Z_{k-1}$  le facteur défini par la définition II.8. Supposons que  $\pi_k : X \mapsto Z_{k-1}$  est une application facteur continue. Soient  $f$  un fonction Riemann-intégrable sur  $X$  et  $x \in X$ . Alors la suite  $(f(T^n x) : n \in \mathbb{Z})$  vérifie la propriété  $\mathbf{P}$ .*

En particulier, la suite  $(\{p(n)\} : n \in \mathbb{Z})$  et la suite  $(e(p(n)) : n \in \mathbb{Z})$ , où  $p$  est un polynôme généralisé, sont des suites universellement bonnes pour la convergence de moyennes ergodiques multiples le long de progressions arithmétiques (voir [43]). Un *polynôme généralisé* est une fonction à valeurs réelles obtenue à partir de la fonction identité et des constantes réelles en utilisant des opérations d'addition, de multiplication et de partie entière.

Dans cette section, nous allons étendre leurs résultats aux moyennes ergodiques multiples le long de polynômes :

**THÉORÈME III.8.** *Soient  $(Y, \nu, S)$  un système ergodique et  $\phi \in L^\infty(\nu)$ . Alors il existe  $Y_0 \subset Y$  avec  $\nu(Y_0) = 1$  tel que, pour tout  $y \in Y_0$ , tout système  $(X, \mu, T)$ , tout entier  $k \geq 1$ , tous polynômes*

à valeurs entières  $p_1, \dots, p_k$  et toutes fonctions  $f_1, \dots, f_k \in L^\infty(\mu)$ , les moyennes

$$(10) \quad \frac{1}{N} \sum_{n=0}^{N-1} \phi(S^n y) f_1(T^{p_1(n)} x) f_2(T^{p_2(n)} x) \dots f_k(T^{p_k(n)} x) .$$

convergent dans  $L^2(\mu)$ .

On en déduit que pour tout  $y \in Y_0$ , la suite  $(\phi(S^n y))$  est aussi universellement bonne pour la convergence de moyennes ergodiques multiples le long de polynômes.

Nous allons suivre une démarche parallèle à celle de [43]. Le but principal est d'établir un critère de convergence pour les moyennes (10) :

**THÉORÈME III.9.** *Pour tous  $k, b \in \mathbb{N}$ , il existe un entier  $K \geq 1$  avec la propriété suivante : pour toute suite bornée  $\mathbf{c} = (c_n : n \in \mathbb{Z})$ , si les moyennes*

$$\frac{1}{N} \sum_{n=0}^{N-1} c_n d_n$$

convergent lorsque  $N \rightarrow \infty$  pour toute nilsuite d'ordre  $K$ ,  $\mathbf{d} = (d_n : n \in \mathbb{Z})$ , alors pour tout système  $(X, \mu, T)$ , toutes fonctions  $f_1, \dots, f_k \in L^\infty(X)$ , et tous polynômes  $p_1, \dots, p_k$  à valeurs entières et de degré  $\leq b$ , les moyennes

$$(11) \quad \frac{1}{N} \sum_{n=0}^{N-1} c_n f_1(T^{p_1(n)} x) f_2(T^{p_2(n)} x) \dots f_k(T^{p_k(n)} x)$$

convergent dans  $L^2(X)$ .

En intégrant ce théorème et la généralisation du théorème de Wiener-Wintner (le théorème III.4), notre résultat principal (le théorème III.8) s'en déduit.

Il y a deux ingrédients pour la preuve du théorème III.9 : l'un est la proposition suivante qui montre que la suite polynomiale  $f(T^{p(n)} x_0)$  dans un nilsystème  $(X, T)$  est en fait une nilsuite, elle se déduit de deux propositions (la proposition 3.14 et la proposition 3.9) de l'article de Leibman [50] :

**PROPOSITION III.10.** *Soient  $(X = G/\Gamma, T)$  un nilsystème,  $x_0 \in X$ ,  $p(n)$  un polynôme à valeurs entières, et  $f \in \mathcal{C}(X)$ . Alors il existe un nilsystème  $(Y, S)$ ,  $y_0 \in Y$ , et une fonction  $h \in \mathcal{C}(Y)$ , tels que  $f(T^{p(n)} x_0) = h(S^n y_0)$  pour tout  $n$ .*

L'autre est une variante pondérée du théorème II.5 qui dit que les semi-normes contrôlent les moyennes multiples le long de polynômes :

THÉORÈME III.11. Soient  $(X, \mu, T)$  un système et  $p_1(n), p_2(n), \dots, p_k(n)$  des polynômes non constants à valeurs entières tels que pour tous  $1 \leq i \neq j \leq k$ , le polynôme  $p_i - p_j$  n'est pas une constante. Alors il existe un entier  $l \geq 1$ , tel que pour toutes fonctions bornées  $f_1, f_2, \dots, f_k$  sur  $X$  avec  $\|f_i\|_l = 0$  pour au moins une valeur des  $i = 1, \dots, k$ , et pour toute suite bornée  $\mathbf{c} = (c_n : n \in \mathbb{Z})$ ,

$$\lim_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=1}^N c_n f_1(T^{p_1(n)}x) f_2(T^{p_2(n)}x) \cdots f_k(T^{p_k(n)}x) \right\|_{L^2(\mu)} = 0 .$$

Autrement dit, il existe un entier  $l \geq 1$  tel que le facteur  $Z_{l-1}$  est un facteur caractéristique pour la moyenne (11).

Pour montrer le théorème III.9, d'après le théorème III.11, il suffit donc de le montrer en supposant que toutes les fonctions  $f_i$  sont mesurables par rapport à  $\mathcal{Z}_{l-1}$ . D'après le théorème de structure (le théorème II.9) avec des raisonnements standards (voir section 6.1 du chapitre II), on peut se restreindre au cas où  $(X, \mu, T)$  est un nil-système. D'après la proposition III.10, pour chaque  $(2k+1)$ -uplet  $(x, f_1, \dots, f_k,$

$p_1, \dots, p_k)$ , il existe  $k$  nilsystèmes  $(Y_i, S_i)$ ,  $k$  fonctions  $g_i$ , et  $k$  éléments  $y_i \in Y_i$ , tels que  $f_i(T^{p_i(n)}x) = g_i(S_i^n y_i)$ ,  $i = 1, \dots, k$ . En passant au système produit, puisque le produit des nilsystèmes est encore un nil-système, nous obtenons que la suite

$$\{f_1(T^{p_1(n)}x) \cdot f_2(T^{p_2(n)}x) \cdot \dots \cdot f_k(T^{p_k(n)}x)\}_{n \in \mathbb{Z}}$$

est une nilsuite. Nous concluons la preuve en appliquant l'hypothèse.

# CONVERGENCE OF WEIGHTED POLYNOMIAL MULTIPLE ERGODIC AVERAGES

QING CHU

ABSTRACT. In this article we study weighted polynomial multiple ergodic averages. A sequence of weights is called universally good if any polynomial multiple ergodic averages with this sequence of weights converges in  $L^2$ . We find a necessary condition and show that for any bounded measurable function  $\phi$  on an ergodic system, the sequence  $\phi(T^n x)$  is universally good for almost every  $x$ . The linear case was covered by Host and Kra.

## 1. INTRODUCTION

In his innovative proof of Szemerédi's Theorem via ergodic theory, Furstenberg introduced certain multiple ergodic averages. There have been many results on these and other nonconventional ergodic averages, including the multiple ergodic theorems of Host and Kra [2], [3], Leibman [5], Ziegler [10]. Recently Host and Kra studied weighted ergodic theorems for multiple averages along arithmetic progressions. We give a generalization of this result for polynomial averages, showing:

**Theorem 1.1.** *Let  $(Y, \nu, S)$  be an ergodic system and  $\phi \in L^\infty(\nu)$ . Then there exists  $Y_0 \subset Y$  with  $\nu(Y_0) = 1$  such that, for every  $y_0 \in Y_0$ , every system  $(X, \mu, T)$ , every  $r \geq 1$ , all integer polynomials  $p_1, \dots, p_r$  and all functions  $f_1, \dots, f_r \in L^\infty(\mu)$ , the averages*

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(S^n y_0) T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r$$

converge in  $L^2(\mu)$ .

Throughout this article, by *integer polynomial* we mean a polynomial all of whose coefficients are integers.

The case of  $p_i(n) = in$  was proved by Host and Kra [4].

---

2000 *Mathematics Subject Classification.* 37A05, 37A30.

*Key words and phrases.* Weighted ergodic averages, universally good sequences, Wiener-Wintner Ergodic Theorem, nilsequences.

*Remark 1.2.* One may wonder why in the theorem  $S^n$  is not replaced by  $S^{p(n)}$  for some integer polynomial  $p(n)$ . In fact, the latter would be much harder to prove, even in the simplest case that  $r = 1$  and  $p_1(n) = n$ . This problem is equivalent to showing the convergence of the averages of  $\phi(S^{p(n)}y_0)e^{2\pi int}$  for all  $t \in \mathbb{T}$  and all  $y_0 \in Y_0$ , where  $Y_0$  does not depend on  $t$ . But this problem reduces to a question of almost everywhere convergence of multiple ergodic averages along polynomials, and this question is out of reach for the moment.

Note that the set  $Y_0$  does not depend on  $X$  or on  $f_i, i = 1, \dots, r$ . We say that for every  $y_0 \in Y_0$ , the sequence  $\phi(S^n y_0)$  is *universally good for the convergence in the mean of polynomial multiple ergodic averages*.

For  $r = 1$  and  $p(n) = n$ , the result follows immediately from the classical Wiener-Winter ergodic Theorem [9] and a corollary of the Spectral Theorem. We follow a similar strategy, generalizing the proof in [4] along arithmetic progressions to polynomial progressions, but need to address some deeper technical issues.

We first recall some definitions, see [1] and [4] for details. Let  $G$  be a  $k$ -step nilpotent Lie group and  $\Gamma \subset G$  be a discrete, cocompact subgroup of  $G$ . The compact manifold  $X = G/\Gamma$  is called a  *$k$ -step nilmanifold*. The Haar measure  $\mu$  of  $X$  is the unique probability measure invariant under the left translations  $x \mapsto gx$  of  $G$  on  $X$ . Letting  $T$  denote left multiplication by a fixed element  $\alpha \in G$ , we call  $(X, \mu, T)$  a  *$k$ -step nilsystem*. Let  $f: X \rightarrow \mathbb{C}$  be a continuous function,  $x_0 \in X$ , then the sequence  $(f(\alpha^n x_0) : n \in \mathbb{Z})$  is called a *basic  $k$ -step nilsequence*. The family of basic  $k$ -step nilsequences forms a subalgebra of  $l^\infty$ . Under the uniform norm  $\|\cdot\|_\infty$  of  $l^\infty$ , we call a uniform limit of basic  $k$ -step nilsequences a  *$k$ -step nilsequence*.

The proof of Theorem 1.1 is broken down into two pieces. First we give a convergence criterion for weighted polynomial multiple ergodic averages.

**Theorem 1.3** (Convergence Criterion for Weighted Averages). *For any  $r, b \in \mathbb{N}$ , there exists an integer  $K \geq 1$  with the following property: for any bounded sequence  $\mathbf{c} = (c_n : n \in \mathbb{Z})$ , if the averages*

$$\frac{1}{N} \sum_{n=0}^{N-1} c_n d_n$$

*converge as  $N \rightarrow \infty$  for every  $K$ -step nilsequence  $\mathbf{d} = (d_n : n \in \mathbb{Z})$ . Then for every system  $(X, \mu, T)$ , all  $f_1, \dots, f_r \in L^\infty(X)$ , and all integer*

polynomials  $p_1, \dots, p_r$  of degree  $\leq b$ , the averages

$$(1) \quad \frac{1}{N} \sum_{n=0}^{N-1} c_n T^{p_1(n)} f_1 \cdot T^{p_2(n)} f_2 \cdots T^{p_r(n)} f_r$$

converge in  $L^2(X)$ .

The bulk of this paper is devoted to the proof of this theorem. Then our main result follows from the following Generalized Wiener-Wintner Theorem proved by Host and Kra in [4]. The case of a polynomial version of the Wiener-Wintner theorem was proved by Lesigne ([7], [8]).

**Theorem 1.4** (Generalized Wiener-Wintner Theorem [4]). *Let  $(X, \mu, T)$  be an ergodic system and  $\phi$  be a bounded measurable function on  $X$ . Then there exists  $X_0 \subset X$  with  $\mu(X_0) = 1$  such that for every  $x \in X_0$ , the averages*

$$\frac{1}{N} \sum_{n=0}^{N-1} \phi(T^n x) b_n$$

converge as  $N \rightarrow \infty$  for every  $x \in X_0$  and every nilsequence  $b = (b_n : n \in \mathbb{Z})$ .

While nilsequences do not appear in the statement of Theorem 1.1, they are used as tools in its proof. Both Theorems 1.3 and 1.4 are of interest on their own, as results on nilsequences.

**Acknowledgement.** We thank the referee for very useful remarks and the simplification of the proof.

## 2. PROOF OF THEOREM 1.3

Using a standard ergodic decomposition argument, it suffices to prove Theorem 1.3 for ergodic systems.

**2.1.** We first remind the reader of a definition from Leibman's paper [6]. We call a sequence  $\{g(n)\}_{n \in \mathbb{Z}}$  with values in a nilpotent group  $G$  a *polynomial sequence*, if  $g(n)$  is of the form  $g(n) = a_1^{p_1(n)} \cdots a_m^{p_m(n)}$ , where  $a_1, \dots, a_m \in G$  and  $p_1, \dots, p_m$  are polynomials taking integer values on the integers.

Before stating the next proposition, we explain briefly its meaning: we can view the sequence of values of a continuous function along a polynomial sequence on a nilmanifold as the sequence of values of some other continuous function along an ordinary orbit of some other nilsystem.

**Proposition 2.1.** Let  $(X = G/\Gamma, T)$  be a nilsystem,  $x_0 \in X$ ,  $p$  be an *integer polynomial*, and  $f \in \mathcal{C}(X)$ . Then there exists a nilsystem  $(Y, S)$ ,  $y_0 \in Y$ ,  $h \in \mathcal{C}(Y)$ , such that  $f(T^{p(n)}x_0) = h(S^n y_0)$  for every  $n$ .

*Proof.* Let  $(X = G/\Gamma, T)$  be a nilsystem. Suppose  $Tx := \alpha x$ , for some  $\alpha \in G$ . Then  $T^{p(n)}x = \alpha^{p(n)}x$ . Let  $g(n) := \alpha^{p(n)}$ , then  $g$  is a polynomial sequence in  $G$ . Let  $\pi : G \rightarrow X$  be the factorization mapping. We will assume that  $x_0 = \pi(\mathbf{1}_G)$ ; otherwise if  $x_0 = \pi(\gamma)$ ,  $\gamma \in G$ , we write  $g(n)x_0 = g(n)\gamma\pi(\mathbf{1}_G)$ , and replace  $g(n)$  by  $g(n)\gamma$ .

Now we have a nilpotent Lie group  $G$ , a discrete cocompact subgroup  $\Gamma$  and a polynomial sequence  $g$  in  $G$ . By Proposition 3.14 in Leibman's paper [6], there exist a nilpotent Lie group  $\tilde{G}$ , a discrete cocompact subgroup  $\tilde{\Gamma}$ , an epimorphism  $\eta : \tilde{G} \rightarrow G$  with  $\eta(\tilde{\Gamma}) \subseteq \Gamma$ , a unipotent automorphism  $\tilde{\tau}$  of  $\tilde{G}$  with  $\tilde{\tau}(\tilde{\Gamma}) = \tilde{\Gamma}$ , and an element  $\tilde{c} \in \tilde{G}$  such that

$$g(n) = \eta(\tilde{\tau}^n(\tilde{c})), n \in \mathbb{Z}.$$

Let  $\tilde{X} = \tilde{G}/\tilde{\Gamma}$  and let  $\tilde{\pi} : \tilde{G} \rightarrow \tilde{X}$  be the factorization mapping.

The epimorphism  $\eta : \tilde{G} \rightarrow G$  factors to a map  $\tilde{X} \rightarrow X$ , we also denote it by  $\eta$ , which is onto and satisfies,

$$\pi \circ \eta = \eta \circ \tilde{\pi}.$$

The map  $\tilde{\tau}$  induces a homomorphism  $\tilde{X} \rightarrow \tilde{X}$ , which we also denote it by  $\tilde{\tau}$ . It satisfies,

$$\tilde{\tau} \circ \tilde{\pi} = \tilde{\pi} \circ \tilde{\tau}.$$

Let  $\tilde{x}_0 = \tilde{\pi}(\mathbf{1}_{\tilde{G}})$ , then

$$\eta(\tilde{\tau}^n(\tilde{c} \tilde{x}_0)) = g(n)x_0, n \in \mathbb{Z}.$$

Let  $\hat{G}$  be the extension of  $\tilde{G}$  by  $\tilde{\tau}$ , then  $\hat{G}$  is a nilpotent Lie group (see Proposition 3.9 in [6]). Let  $\hat{\tau}$  be the element in  $\hat{G}$  representing  $\tilde{\tau}$ , so that  $\tilde{\tau}(\tilde{\alpha}) = \hat{\tau}\tilde{\alpha}\hat{\tau}^{-1}$  for any  $\tilde{\alpha} \in \tilde{G}$ , the multiplication of  $\hat{G}$  is given by this formula. We have  $\hat{G} = \{\tilde{g}\hat{\tau}^n : \tilde{g} \in \tilde{G}, n \in \mathbb{Z}\}$  and  $\tilde{G}$  is open in  $\hat{G}$ .

Let  $\hat{\Gamma}$  be the subgroup of  $\hat{G}$  spanned by  $\tilde{\Gamma}$  and  $\hat{\tau}$ . As  $\tilde{\tau}(\tilde{\Gamma}) = \tilde{\Gamma}$ , we have  $\hat{\Gamma} = \{\tilde{\gamma}\hat{\tau}^n : \tilde{\gamma} \in \tilde{\Gamma}, n \in \mathbb{Z}\}$  and  $\hat{\Gamma} \cap \tilde{G} = \tilde{\Gamma}$ . By the definition of the relative topology,  $\hat{\Gamma}$  is a discrete subgroup of  $\hat{G}$ . Moreover,  $\hat{G}/\hat{\Gamma} = (\tilde{G}\hat{\Gamma}/\hat{\Gamma})$  can be identified with  $\tilde{G}/(\hat{\Gamma} \cap \tilde{G}) = \tilde{G}/\tilde{\Gamma} = \tilde{X}$ . We write  $\hat{\pi} : \hat{G} \rightarrow \hat{X}$  for the quotient map.

Let  $\tilde{x} \in \tilde{X}$  and  $\tilde{g} \in \tilde{G}$  with  $\tilde{\pi}(\tilde{g}) = \tilde{x}$ . We have

$$\tilde{\tau}(\tilde{x}) = \tilde{\pi}(\tilde{\tau}(\tilde{g})) = \hat{\pi}(\hat{\tau}\tilde{g}\hat{\tau}^{-1}) = \hat{\pi}(\hat{\tau}\tilde{g}) = \hat{\tau}\tilde{x},$$

because  $\widehat{\tau}^{-1} \in \widehat{\Gamma}$ . So for every  $n$ ,

$$g(n)x_0 = \eta(\widetilde{\tau}^n(\tilde{c} \tilde{x}_0)) = \eta(\widehat{\tau}^n(\tilde{c} \tilde{x}_0)).$$

Let  $Y = (\widehat{G}/\widehat{\Gamma}, S) = (\widetilde{X}, S)$ ,  $S\tilde{x} = \widetilde{\tau}\tilde{x}$ , and let  $h = f \circ \eta$ , and  $y_0 = \tilde{c} \tilde{x}_0$ . This system and this function satisfy the announced properties.  $\square$

**2.2.** We recall a few properties of the seminorms and the factors introduced in [2]. Let  $(X, \mu, T)$  be an ergodic system. For an integer  $k \geq 0$ , we write  $X^{[k]} = X^{2^k}$  and  $T^{[k]}: X^{[k]} \rightarrow X^{[k]}$  for the map  $T \times T \times \dots \times T$ , taken  $2^k$  times. We define by induction a probability measure  $\mu^{[k]}$  on  $X^{[k]}$  that is invariant under  $T^{[k]}$ . Set  $\mu^{[0]} = \mu$ . For  $k \geq 0$ , let  $\mathcal{I}^{[k]}$  be the  $\sigma$ -algebra of  $T^{[k]}$ -invariant subsets of  $X^{[k]}$ . Then  $\mu^{[k+1]}$  is the relatively independent product of  $\mu^{[k]}$  over  $\mathcal{I}^k$ , which means for  $F, F' \in L^\infty(X^{[k]})$ ,

$$\int_{X^{[k+1]}} F \otimes F' d\mu^{[k+1]} = \int_{X^{[k]}} \mathbb{E}(F|\mathcal{I}^{[k]})\mathbb{E}(F'|\mathcal{I}^{[k]}) d\mu^{[k]}.$$

For a bounded measurable function  $f$ , we define

$$(2) \quad \|f\|_k = \left( \int_{X^{[k]}} \prod_{\varepsilon \in \{0,1\}^k} C^{|\varepsilon|} f(x_\varepsilon) d\mu^{[k]}(x) \right)^{1/2^k},$$

where  $C: \mathbb{C} \rightarrow \mathbb{C}$  is the conjugacy map  $z \mapsto \bar{z}$ ,  $\varepsilon = \varepsilon_1 \varepsilon_2 \dots \varepsilon_k$  with  $\varepsilon_i \in \{0, 1\}$  and  $|\varepsilon| = \varepsilon_1 + \varepsilon_2 + \dots + \varepsilon_k$ . It is shown in [2] that for every  $k \geq 1$ ,  $\|\cdot\|_k$  is a seminorm on  $L^\infty(\mu)$ .

Moreover, for every  $k \geq 2$ ,  $X$  admits a factor  $Z_{k-1}$  such that, for every  $f \in L^\infty(\mu)$ ,  $\|f\|_k = 0$  if and only if  $\mathbb{E}(f|Z_{k-1}) = 0$ . One of the main results of [2] is that, for every  $k$ ,  $Z_k$  is an inverse limit of  $k$ -step nilsystems. We call this result the Structure Theorem.

Let  $(Z = Z_1(X), m, T)$  be the Kronecker factor of  $(X, \mu, T)$ . For  $s \in Z$ , we define a measure  $\mu_s$  on  $X \times X$  by

$$\int_{X \times X} f(x)f'(x') d\mu_s(x, x') = \int_Z \mathbb{E}(f|Z)(z) \cdot \mathbb{E}(f'|Z)(sz) dm(z).$$

For every  $s \in Z$  the measure  $\mu_s$  is invariant under  $T \times T$  and is ergodic for  $m$ -almost every  $s$ . The ergodic decomposition of  $\mu \times \mu$  under  $T \times T$  is

$$\mu \times \mu = \int_Z \mu_s dm(s).$$

For each  $s \in Z$  such that  $(X \times X, \mu_s, T \times T)$  is ergodic, and for each integer  $k \geq 1$ , a measure  $(\mu_s)^{[k]}$  on  $(X \times X)^{[k]}$  can be defined in the same way as  $\mu^{[k]}$ . Furthermore, a seminorm  $\|\cdot\|_{k,s}$  on  $L^\infty(\mu_s)$  can

be associated to this measure in the same way as the seminorm  $\|\cdot\|_k$  is associated to  $\mu^{[k]}$ . It follows from the definition (2) that for every  $f \in L^\infty(\mu)$ ,

$$(3) \quad \|f\|_{k+1}^{2^{k+1}} = \int_Z \|f \otimes \bar{f}\|_{k,s}^{2^k} dm(s).$$

**2.3.** We return to the proof of theorem 1.3. We may assume that the polynomials  $p_1, \dots, p_r$  are nonconstant and essentially distinct, that is  $p_i - p_j \neq \text{constant}$  for  $i \neq j$ .

The following theorem will be proved in the next section.

**Theorem 2.2.** *For any  $r, b \in \mathbb{N}$ , there exists  $k \in \mathbb{N}$ , such that for any nonconstant essentially distinct polynomials  $p_1, \dots, p_r : \mathbb{Z} \rightarrow \mathbb{Z}$  of degree  $\leq b$ , for every ergodic system  $(X, \mu, T)$ , every  $f_1, \dots, f_r \in L^\infty(X)$  with  $\|f_1\|_k = 0$ , and any bounded sequence  $\mathbf{c} = (c_n : n \in \mathbb{Z})$ , one has*

$$(4) \quad \lim_{N \rightarrow \infty} \left\| \frac{1}{N} \sum_{n=0}^{N-1} c_n T^{p_1(n)} f_1 \dots T^{p_r(n)} f_r \right\|_{L^2(X)} = 0.$$

**2.4.** Now we give the proof of Theorem 1.3.

***Proof of Theorem 1.3 from Proposition 2.1 and Theorem 2.2.***

The proof is exactly the same as the proof of Theorem 2.24 in [4]. For any  $r, b \in \mathbb{N}$ , let  $k \in \mathbb{N}$  be the integer in Theorem 2.2, let  $Z_{k-1}$  be the  $(k-1)$ -th factor of  $(X, \mu, T)$ , as given by the Structure Theorem. By definition, if  $\mathbb{E}(f_1 | Z_{k-1}) = 0$ , then  $\|f_1\|_k = 0$ , and by Theorem 2.2, the averages (4) converge to zero in  $L^2(X)$ . We say that the factor  $Z_{k-1}$  is the characteristic for the convergence of these averages. Therefore, it suffices to prove the result when the functions are measurable with respect to the factor  $Z_{k-1}$ .

Since  $Z_{k-1}$  is an inverse limit of  $(k-1)$ -step nilsystems by density, we can assume that  $(X, \mu, T)$  is a  $(k-1)$ -step nilsystem and that the functions  $f_1, \dots, f_r$  are continuous.

But in this case, by Proposition 2.1, for every  $x \in X$ , and all polynomials  $p_1, \dots, p_r$ , there exist nilsystems  $(Y_1, S_1), \dots, (Y_r, S_r)$ ,  $y_i \in Y_i$ , and  $g_i \in \mathcal{C}(Y_i)$ , such that  $f_i(T^{p_i(n)} x) = g_i(S_i^n y_i)$ ,  $i = 1, \dots, r$ .

Let  $K$  be the maximal order of the nilsystems  $(Y_i, S_i)$ ,  $i = 1, \dots, r$ . Then the system  $(Y = Y_1 \times \dots \times Y_r, S = S_1 \times \dots \times S_r)$  is a  $K$ -step nilsystem. Let  $g : Y_1 \times \dots \times Y_r \rightarrow \mathbb{R}$  be given by  $g(y) = g(y_1, \dots, y_r) =$

$g_1(y_1) \cdot \dots \cdot g_r(y_r)$ . So the sequence

$$\begin{aligned} & \{f_1(T^{p_1(n)}x) \cdot f_2(T^{p_2(n)}x) \cdot \dots \cdot f_r(T^{p_r(n)}x)\}_{n \in \mathbb{Z}} \\ &= \{g_1(S_1^n y_1) \cdot g_2(S_2^n y_2) \cdot \dots \cdot g_r(S_r^n y_r)\}_{n \in \mathbb{Z}} \\ &= \{g(S^n y)\}_{n \in \mathbb{Z}} \end{aligned}$$

is a  $K$ -step nilsequence and by hypothesis, the averages (1) converge for every  $x \in X$ .  $\square$

### 3. PROOF OF THEOREM 2.2

Note that our goal is very similar to the main result in Leibman's paper [5], the only difference being that in our case we are dealing with the weighted averages. In fact, we can deduce Theorem 2.2 by following very closely the arguments of Leibman in his paper to cover our weighted case with some modifications. But here we adopt another way that allows us to deduce it directly from Leibman's result.

**Proof of Theorem 2.2.** We prove the result by several steps.

(i) The following assertion follows immediately from Theorem 3 in Leibman's paper [5]:

Let  $r, b \in \mathbb{N}$  be fixed. There exists an integer  $k = k(r, b)$  such that for every family of nonconstant essentially distinct polynomials  $p_1, \dots, p_r : \mathbb{Z}^2 \rightarrow \mathbb{Z}$  of degree  $\leq b$ , we have: for every ergodic system  $(X, \mathcal{B}, \mu, T)$ , and every  $f_1, \dots, f_r \in L^\infty(X)$  with  $\|f_1\|_k = 0$ , one has

$$\lim_{N \rightarrow \infty} \left\| \frac{1}{N^2} \sum_{0 \leq m, n < N} T^{p_1(m, n)} f_1 \dots T^{p_r(m, n)} f_r \right\|_{L^2(X)} = 0.$$

(ii) We use the notation of Section 2.2. It follows from (3) that  $\|f_1\|_{k+1} = 0$  implies  $\|f_1 \otimes \bar{f}_1\|_{k, s} = 0$  for  $m$ -almost every  $s$ . Using the previous result to the ergodic system  $(X \times X, \mu_s, T \times T)$ , one gets: if  $\|f_1\|_{k+1} = 0$ , then

$$\begin{aligned}
& \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{0 \leq m, n < N} \left| \int T^{p_1(m, n)} f_1 \cdots T^{p_r(m, n)} f_r \, d\mu \right|^2 \\
&= \lim_{N \rightarrow \infty} \frac{1}{N^2} \sum_{0 \leq m, n < N} \left| \int_Z \int_{X \times X} (T \times T)^{p_1(m, n)} f_1 \otimes \bar{f}_1 \cdots \right. \\
&\quad \left. (T \times T)^{p_r(m, n)} f_r \otimes \bar{f}_r \, d\mu_s \, dm(s) \right| \\
&\leq \int_Z \lim_{N \rightarrow \infty} \int_{X \times X} \left| \frac{1}{N^2} \sum_{0 \leq m, n < N} (T \times T)^{p_1(m, n)} f_1 \otimes \bar{f}_1 \cdots \right. \\
&\quad \left. (T \times T)^{p_r(m, n)} f_r \otimes \bar{f}_r \right| \, d\mu_s \, dm(s) \\
&\leq \int_Z \lim_{N \rightarrow \infty} \left\| \frac{1}{N^2} \sum_{0 \leq m, n < N} (T \times T)^{p_1(m, n)} f_1 \otimes \bar{f}_1 \cdots \right. \\
&\quad \left. (T \times T)^{p_r(m, n)} f_r \otimes \bar{f}_r \right\|_{L^2(\mu_s)} \, dm(s) = 0.
\end{aligned}$$

(iii) Expanding the square, one sees that

$$\left\| \frac{1}{N} \sum_{n=0}^{N-1} c_n T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r \right\|_{L^2}^2$$

is equal to

$$\frac{1}{N^2} \sum_{0 \leq m, n < N} c_n \bar{c}_m \int T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r \, T^{p_1(m)} \bar{f}_1 \cdots T^{p_r(m)} \bar{f}_r \, d\mu,$$

which is less than a constant times

$$(5) \quad \frac{1}{N^2} \sum_{0 \leq m, n < N} \left| \int T^{p_1(n)} f_1 \cdots T^{p_r(n)} f_r \, T^{p_1(m)} \bar{f}_1 \cdots T^{p_r(m)} \bar{f}_r \, d\mu \right|,$$

since by our assumption the sequence  $c_n$  is bounded. Notice that  $p_1(n), \dots, p_r(n), p_1(m), \dots, p_r(m)$  is a family of  $2r$  essentially distinct polynomials of degree at most  $b$ .

(iv) Combining the previous parts we see that if  $\|f_1\|_{k(2r, b)+1} = 0$ , where  $k(r, b)$  is defined in part (i), then for every family of nonconstant essentially distinct polynomials  $p_1, \dots, p_r : \mathbb{Z} \rightarrow \mathbb{Z}$  of degree  $\leq b$ , we have that the average (5) converges to 0. This proves Theorem 2.2.  $\square$

## REFERENCES

1. V. Bergelson, B. Host and B. Kra, with an Appendix by I. Ruzsa. Multiple recurrence and nilsequences. *Inventiones Math.* **160** (2005), 261–303.
2. B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.* **161** (2005), no. 1, 397–488.
3. B. Host and B. Kra. Convergence of polynomial ergodic averages. *Israel J. of Math.* **149** (2005), 1–20.
4. B. Host and B. Kra. Uniformity seminorms on  $\ell^\infty$  and applications. *J. Anal. Math.* **108** (2009), 219–276.
5. A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Israel J. Math.* **146** (2005), 303–315.
6. A. Leibman. Pointwise convergence of ergodic averages for polynomial sequence of translation on a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 201–213.
7. E. Lesigne. Un théorème de disjonction de systèmes dynamiques et une généralisation du théorème ergodique de Wiener-Wintner. *Ergodic Theory Dynam. Systems* **10** (1990), 513–521.
8. E. Lesigne. Spectre quasi-discret et théorème ergodique de Wiener-Wintner pour les polynômes. *Ergodic Theory Dynam. Systems* **13** (1993), 767–784.
9. N. Wiener and A. Wintner. *Harmonic analysis and ergodic theory*. Amer. J. Math. **63** (1941), 415–426.
10. T. Ziegler. A non-conventional ergodic theorem for a nilsystem. *Ergodic Theory Dynam. Systems* **25** (2005), no. 4, 1357–1370.

## CHAPITRE IV

### Convergence de moyennes ergodiques multiples sur les cubes pour plusieurs transformations qui commutent

L'étude présentée dans ce chapitre fait l'objet de l'article [18].

#### 1. Introduction

En vue de faciliter la lecture, nous réécrivons les résultats suivants qui ont été présentés dans le chapitre I :

Rappelons que nous écrivons  $\epsilon = \epsilon_1 \epsilon_2 \dots \epsilon_k$  avec  $\epsilon_i \in \{0, 1\}$  pour un élément dans  $\{0, 1\}^k$ .

**THÉORÈME (Host et Kra [39]).** *Soit  $(X, \mathcal{X}, \mu, T)$  un système. Soient  $k \geq 1$  un entier et  $f_\epsilon, \epsilon \in \{0, 1\}^k \setminus \{00 \dots 0\}$ ,  $2^k - 1$  fonctions bornées sur  $X$ . Alors les moyennes*

$$(12) \quad \frac{1}{N^k} \sum_{n_1, n_2, \dots, n_k=0}^{N-1} \prod_{\substack{\epsilon \in \{0,1\}^k \\ \epsilon \neq 00 \dots 0}} f_\epsilon(T^{n_1 \epsilon_1 + \dots + n_k \epsilon_k} x) .$$

convergent dans  $L^2(\mu)$ .

En conséquence, pour tout sous-ensemble  $A$  de  $X$ , la limite des moyennes

$$\frac{1}{N^k} \sum_{n_1, n_2, \dots, n_k=0}^{N-1} \mu\left( \bigcap_{\epsilon \in \{0,1\}^k} T^{n_1 \epsilon_1 + \dots + n_k \epsilon_k} A \right)$$

existe. D'après le fait que les moyennes (12) sont liées directement à la mesure  $\mu^{[k]}$  (le lemme II.3), on en déduit que cette limite est supérieure ou égale à  $\mu(A)^{2^k}$ . Le résultat combinatoire correspondant se déduit du principe de correspondance (voir la section 2.5 du chapitre I).

Une question naturelle est de savoir s'il existe des résultats analogues à ceux présentés ci-dessus pour un système muni de plusieurs transformations. Dans ce chapitre, nous établissons les résultats suivants :

D'abord un résultat de convergence pour des moyennes sur les cubes pour plusieurs transformations qui commutent :

THÉORÈME IV.1. *Soient  $k \geq 1$  un entier et  $(X, \mu, T_1, \dots, T_k)$  un système. Soient  $f_\epsilon, \epsilon \in \{0, 1\}^k \setminus \{00 \dots 0\}$ ,  $2^k - 1$  fonctions bornées sur  $X$ . Alors les moyennes*

$$(13) \quad \prod_{i=1}^k \frac{1}{N_i - M_i} \left( \sum_{\substack{n_i \in [M_i, N_i] \\ i=1, \dots, k}} \prod_{\substack{\epsilon \in \{0, 1\}^k \\ \epsilon \neq 00 \dots 0}} T_1^{n_1 \epsilon_1} \dots T_k^{n_k \epsilon_k} f_\epsilon \right)$$

convergent dans  $L^2(\mu)$  pour toutes suites d'intervalles  $[M_1, N_1), \dots, [M_k, N_k)$  dont la longueur  $N_i - M_i$  ( $1 \leq i \leq k$ ) tend vers  $\infty$ .

Par exemple, lorsque  $k = 2$ , les moyennes (13) avec toutes les fonctions égales sont

$$\frac{1}{N^2} \sum_{n=1}^N \sum_{m=1}^N f(T_1^n x) f(T_2^m x) f(T_1^n T_2^m x) .$$

En conséquence, pour tout sous-ensemble  $A$  de  $X$ , la limite des moyennes

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i] \\ i=1, \dots, k}} \int \prod_{\epsilon \in \{0, 1\}^k} T_1^{n_1 \epsilon_1} \dots T_k^{n_k \epsilon_k} \mathbf{1}_A d\mu$$

existe. D'après le lemme II.3, les moyennes (13) sont liées directement à la mesure  $\mu^*$ , on en déduit que cette limite est égale à  $\|\mathbf{1}_A\|_{T_1, \dots, T_k}^{2^k}$ , cette dernière expression est supérieure ou égale à  $\mu(A)^{2^k}$  :

THÉORÈME IV.2. *Soient  $(X, \mu, T)$  un système et  $A$  un sous-ensemble de  $X$ . Alors pour tout  $\delta > 0$ , l'ensemble*

$$\{n = (n_1, \dots, n_k) \in \mathbb{Z}^k : \mu \left( \bigcap_{\epsilon \in \{0, 1\}^k} T_1^{n_1 \epsilon_1} \dots T_k^{n_k \epsilon_k} A \right) \geq \mu(A)^{2^k} - \delta\}$$

est syndétique.

En vertu du principe de correspondance de Furstenberg (le théorème I.1), nous avons le résultat combinatoire correspondant :

THÉORÈME. *Soit  $E \subset \mathbb{N}^k$  un ensemble avec la densité  $d^*(E) > \delta > 0$ . Alors le sous-ensemble de  $\mathbb{Z}^k$*

$$\{n = (n_1, \dots, n_k) \in \mathbb{Z}^k : d^* \left( \bigcap_{\epsilon \in \{0, 1\}^k} \{E + (n_1 \epsilon_1, \dots, n_k \epsilon_k)\} \right) \geq \delta^{2^k}\}$$

est syndétique.

Nous allons utiliser les notations suivantes qui ont déjà été introduites dans la section 5 du chapitre II :

NOTATION. Nous écrivons  $[k] = \{1, \dots, k\}$  et identifions  $\{0, 1\}^k$  avec la famille de sous-ensembles de  $[k]$ . Par conséquent, pour  $\epsilon \in \{0, 1\}^k$  et  $1 \leq i \leq k$ , l'assertion " $\epsilon_i = 1$ " est équivalente à " $i \in \epsilon$ ". En particulier, nous écrivons  $\emptyset$  pour  $00 \dots 0 \in \{0, 1\}^k$ .

Par ailleurs, soit  $1 \leq j \leq k$  un entier, pour chaque  $\epsilon = \{i_1, \dots, i_j\} \subset [k]$ , nous écrivons

$$T_\epsilon^n := T_{i_1}^{n_{i_1}} \circ \dots \circ T_{i_j}^{n_{i_j}}.$$

Nous allons utiliser la machinerie des systèmes magiques pour la preuve du théorème IV.1. En effet, nous allons montrer que le théorème est valable sur son extension magique  $(X^*, \mu^*, T_1^*, \dots, T_k^*)$  construit dans la section 4.1 du chapitre I :

THÉORÈME IV.3. *Soient  $f_\epsilon$ ,  $\epsilon \subset [k]$ ,  $\epsilon \neq \emptyset$ ,  $2^k - 1$  fonctions bornées sur  $X^*$ . Alors les moyennes*

$$(14) \quad \prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i] \\ i=1, \dots, k}} \prod_{\substack{\epsilon \subset [k] \\ \epsilon \neq \emptyset}} T_\epsilon^{*n} f_\epsilon$$

convergent dans  $L^2(\mu^*)$  pour toutes suites d'intervalles  $[M_1, N_1), \dots, [M_k, N_k)$  dont la longueur  $N_i - M_i$  ( $1 \leq i \leq k$ ) tend vers  $\infty$ .

Pour ce faire, nous montrons d'abord que nous avons la borne supérieure suivante pour les moyennes (13), dont la preuve suit la même démarche que celle correspondant au cas d'une transformation (voir [41]) en utilisant le lemme de van der Corput II.22 :

THÉORÈME IV.4. *Soit  $(X, \mu, T_1, \dots, T_k)$  un système. Alors pour toutes fonctions  $f_\epsilon \in L^\infty(\mu)$ , avec  $\|f_\epsilon\|_{L^\infty(\mu)} \leq 1$ ,  $\epsilon \subset [k]$ ,  $\epsilon \neq \emptyset$ ,*

$$\limsup_{\substack{N_i - M_i \rightarrow \infty \\ i=1, \dots, k}} \left\| \prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \dots \times [M_k, N_k)} \prod_{\substack{\epsilon \subset [k] \\ \epsilon \neq \emptyset}} T_\epsilon^n f_\epsilon \right\|_{L^2(\mu)} \leq \min_{\substack{\epsilon \subset [k] \\ \epsilon \neq \emptyset}} \|f_\epsilon\|_{T_1 \dots T_k}.$$

D'après la définition d'un système magique (la définition II.16) et le fait que le système  $(X^*, \mu^*, T_1^*, \dots, T_k^*)$  est magique (le théorème II.17), on en déduit que la sous- $\sigma$ -algèbre

$$\mathcal{Z}^* := \bigvee_{i=1}^k \mathcal{I}(T_i^*)$$

est un facteur caractéristique pour les moyennes (14). Alors pour montrer le théorème IV.3, nous pouvons nous restreindre au cas où toutes les fonctions  $f_\epsilon$  sont mesurables par rapport à  $\mathcal{Z}^*$ . Par un argument de densité, nous pouvons supposer que

$$f_\epsilon = \prod_{i \in [k]} f_{\epsilon, i}$$

où la fonction  $f_{\epsilon, i}$  est  $T_i^*$ -invariante. Alors

$$T_\epsilon^{*n} f_\epsilon = \prod_{i \in [k]} T_{\epsilon \setminus \{i\}}^{*n} f_{\epsilon, i} ,$$

si  $i \notin \epsilon$ , on pose  $T_{\epsilon \setminus \{i\}}^{*n} = \text{Id}$ .

Nous constatons immédiatement que les moyennes (14) peuvent se réécrire autrement. En effet, nous nous ramenons à traiter les moyennes

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \dots \times [M_k, N_k)} \prod_{\substack{\epsilon \subset [k] \\ 0 < |\epsilon| \leq k-1}} T_\epsilon^n f_\epsilon .$$

Pour montrer que ces moyennes convergent, nous devons trouver un facteur caractéristique pour elles. Pour ce faire, nous montrons d'abord le résultat suivant qui est plus général que le théorème IV.4, mais dont la preuve dépend de ce dernier :

NOTATION. Pour  $\epsilon \subset [k]$ ,  $\epsilon \neq \emptyset$ , nous notons  $\|\cdot\|_\epsilon$  la semi-norme sur  $L^\infty(\mu)$  associée aux transformations  $T_i$ ,  $i \in \epsilon$ . Par exemple,  $\|\cdot\|_{110\dots 00}$  est la semi-norme associée aux transformations  $T_1, T_2$  puisque  $\epsilon = 110\dots 0 \in \{0, 1\}^d$  est identifié avec  $\{1, 2\} \subset [d]$ .

THÉORÈME IV.5. Soit  $1 \leq r \leq k$  un entier. Alors pour toutes fonctions  $f_\epsilon \in L^\infty(\mu)$ , avec  $\|f_\epsilon\|_{L^\infty(\mu)} \leq 1$ ,  $\epsilon \subset [k]$ ,  $\epsilon \neq \emptyset$ ,

$$\limsup_{\substack{N_i - M_i \rightarrow \infty \\ i=1, \dots, k}} \left\| \prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \dots \times [M_k, N_k)} \prod_{\substack{\epsilon \subset [k] \\ 0 < |\epsilon| \leq r}} T_\epsilon^n f_\epsilon \right\|_{L^2(\mu)} \leq \min_{\substack{\epsilon \subset [k] \\ |\epsilon|=r}} \|f_\epsilon\|_\epsilon$$

Ensuite, nous montrons qu'il existe un facteur qui correspond à chaque semi-norme  $\|\cdot\|_\epsilon$  dans un système magique.

THÉORÈME IV.6. Soit  $(X, \mu, T_1, \dots, T_k)$  un système magique. Pour  $\epsilon \subset [k]$ ,  $\epsilon \neq \emptyset$ , nous définissons une sous- $\sigma$ -algèbre

$$\mathcal{Z}_\epsilon := \bigvee_{i \in \epsilon} \mathcal{I}(T_i) .$$

Par exemple  $\mathcal{Z}_{\{1,2,d\}} = \mathcal{I}(T_1) \vee \mathcal{I}(T_2) \vee \mathcal{I}(T_d)$ . Alors pour toute fonction  $f \in L^\infty(\mu)$ ,

$$(15) \quad \text{si } \mathbb{E}_\mu(f \mid \mathcal{Z}_\epsilon) = 0, \text{ alors } \|f\|_\epsilon = 0.$$

Grâce à ces deux théorèmes, nous montrons, par récurrence sur  $r$ , le théorème suivant :

**THÉORÈME IV.7.** *Soit  $f_\epsilon, \epsilon \in [d]$ , des fonctions bornées sur  $X^*$ . Soit  $r$  un entier avec  $1 \leq r \leq d$ . Alors les moyennes*

$$\prod_{i=1}^k \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1] \times \dots \times [M_k, N_k]} \prod_{\substack{\epsilon \in [k] \\ 0 < |\epsilon| \leq r}} T_\epsilon^{*n} f_\epsilon$$

convergent dans  $L^2(\mu^*)$  pour toutes suites d'intervalles  $[M_1, N_1), \dots, [M_k, N_k)$  dont la longueur  $N_i - M_i$  ( $1 \leq i \leq k$ ) tend vers  $\infty$ .

En prenant  $r = d$ , le théorème IV.3 s'en déduit immédiatement de ce théorème. Puisque le système  $(X^*, \mu^*, T_1^*, \dots, T_k^*)$  admet le système  $(X, \mu, T_1, \dots, T_k)$  comme un facteur, notre résultat principal (le théorème IV.1) s'en déduit.

# CONVERGENCE OF MULTIPLE ERGODIC AVERAGES ALONG CUBES FOR SEVERAL COMMUTING TRANSFORMATIONS

QING CHU

ABSTRACT. We prove the norm convergence of multiple ergodic averages along cubes for several commuting transformations, and derive corresponding combinatorial results. The method we use relies primarily on the “magic extension” established recently by B. Host.

## 1. INTRODUCTION

**1.1. Results.** By a system, we mean a probability space endowed with a single or several commuting measure preserving transformations. We prove the following result regarding the convergence of multiple ergodic averages along cubes for several commuting transformations :

**Theorem 1.1.** *Let  $d \geq 1$  be an integer and  $(X, \mathcal{B}, \mu, T_1, \dots, T_d)$  be a system. Let  $f_\epsilon, \epsilon \in \{0, 1\}^d \setminus \{00 \dots 0\}$  be  $2^d - 1$  bounded measurable functions on  $X$ . Then the averages*

$$(1) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i) \\ i=1, \dots, d}} \prod_{\substack{\epsilon \in \{0, 1\}^d \\ \epsilon \neq 00 \dots 0}} T_1^{n_1 \epsilon_1} \dots T_d^{n_d \epsilon_d} f_\epsilon$$

converge in  $L^2(\mu)$  for all sequences of intervals  $[M_1, N_1), \dots, [M_d, N_d)$  whose lengths  $N_i - M_i$  ( $1 \leq i \leq d$ ) tend to  $\infty$ .

To illustrate, when  $d = 2$ , the average (1) is

$$(2) \quad \frac{1}{(N_1 - M_1) \times (N_2 - M_2)} \sum_{\substack{n_1 \in [M_1, N_1) \\ n_2 \in [M_2, N_2)}} T_1^{n_1} f_{10} \cdot T_2^{n_2} f_{01} \cdot T_1^{n_1} T_2^{n_2} f_{11}.$$

When Theorem 1.1 is restricted to the case that each function  $f_\epsilon$  is the indicator function of a measurable set, we have the following lower bound for these averages:

---

2000 *Mathematics Subject Classification.* 37A05, 37A30.

*Key words and phrases.* Averages along cubes, commuting transformations, magic system, ergodic seminorms.

**Theorem 1.2.** Let  $(X, \mathcal{B}, \mu, T_1, \dots, T_d)$  be a system and let  $A \in \mathcal{B}$ . Then the limit of the averages

$$(3) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i] \\ i=1, \dots, d}} \mu \left( \bigcap_{\epsilon \in \{0,1\}^d} T_1^{-n_1 \epsilon_1} \dots T_d^{-n_d \epsilon_d} A \right)$$

exists and is greater than or equal to  $\mu(A)^{2^d}$  for all sequences of intervals  $[M_1, N_1), \dots, [M_d, N_d)$  whose lengths  $N_i - M_i$  ( $1 \leq i \leq d$ ) tend to  $\infty$ .

Recall that the upper density  $d^*(E)$  of a set  $E \subset \mathbb{Z}^d$  is defined to be

$$d^*(E) = \limsup_{\substack{N_i \rightarrow \infty \\ 1 \leq i \leq d}} \prod_{i=1}^d \frac{1}{N_i} |E \cap [1, N_1] \times \dots \times [1, N_d]|.$$

A subset  $E$  of  $\mathbb{Z}^d$  is said to be syndetic if  $\mathbb{Z}^d$  can be covered by finitely many translates of  $E$ .

We have the following corresponding combinatorial result:

**Theorem 1.3.** Let  $E \subset \mathbb{Z}^d$  with  $d^*(E) > \delta > 0$ . Then the set of  $n = (n_1, \dots, n_d) \in \mathbb{Z}^d$  such that

$$d^* \left( \bigcap_{\epsilon \in \{0,1\}^d} \{E + (n_1 \epsilon_1, \dots, n_d \epsilon_d)\} \right) \geq \delta^{2^d}$$

is syndetic.

**1.2. History of the problem.** In the case where  $T_1 = T_2 = \dots = T_d = T$ , the average (1) is

$$(4) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i] \\ i=1, \dots, d}} \prod_{\substack{\epsilon \in \{0,1\}^d \\ \epsilon \neq 0 \dots 0}} T^{n_1 \epsilon_1 + \dots + n_d \epsilon_d} f_\epsilon.$$

The norm convergence of (4) was proved by Bergelson for  $d = 2$  in [4], and more generally, by Host and Kra for  $d > 2$  in [7]. The related pointwise convergence problem was studied by Assani and he showed that the averages (4) converge a.e. in [1].

The lower bound for the average (3) was firstly studied by Leibman, he provided some lower bounds for (3) in [8]. In the same paper, he gave an example showing that the average (2) can diverge if the transformations do not commute.

However, Assani showed in [1] that the averages

$$\frac{1}{N^2} \sum_{n,m=1}^N f(T_1^n x) g(T_2^m x) h(T_3^{n+m} x)$$

do converge a.e. even if the transformations do not necessarily commute. He extended this result to the case of six functions in [2].

The norm convergence of multiple ergodic averages with several commuting transformations of the form

$$(5) \quad \frac{1}{N} \sum_{n=1}^N T_1^n f_1 \cdots T_d^n f_d,$$

was proved by Conze and Lesigne [5] when  $d = 2$ . The general case was originally proved by Tao [6], and subsequent proofs were given by Austin [3], Host [6] and Towsner [10].

**1.3. Methods.** The main tools we use in this paper are the seminorms and the existence of “magic extensions” for commuting transformations established by Host [6]. The “magic extensions” can be viewed as a concrete form of the pleasant extensions built by Austin in [3].

**Acknowledgement.** This paper was written while the author was visiting the Mathematical Sciences Research Institute and the author is grateful for their kind hospitality. The author also thanks the referee of this paper for remarks.

## 2. SEMINORM AND UPPER BOUND

**2.1. Notation and definitions.** Given a probability space  $(X, \mathcal{B}, \mu)$ , in general, we omit the  $\sigma$ -algebra  $\mathcal{B}$  from our notation and write  $(X, \mu)$ .

For an integer  $d \geq 1$ , we write  $[d] = \{1, 2, \dots, d\}$  and identify  $\{0, 1\}^d$  with the family of subsets of  $[d]$ . Therefore, the assertion “ $i \in \epsilon$ ” is equivalent to  $\epsilon_i = 1$ . In particular,  $\emptyset$  is the same as  $00 \cdots 0 \in \{0, 1\}^d$ . We write  $|\epsilon| = \sum_i \epsilon_i$  for the number of elements in  $\epsilon$ .

Let  $(X, \mu, T_1, \dots, T_d)$  be a system. For each  $n = (n_1, \dots, n_d)$ ,  $\epsilon = \{i_1, \dots, i_k\} \subset [d]$ , and for each integer  $1 \leq k \leq d$ , we write

$$T_\epsilon^n = T_{i_1}^{n_{i_1}} \cdots T_{i_k}^{n_{i_k}}.$$

For any transformation  $S$  of some probability space, we denote by  $\mathcal{I}(S)$  the  $\sigma$ -algebra of  $S$ -invariant sets.

We define a measure  $\mu_1$  on  $X^2$  by

$$\mu_1 = \mu \times_{\mathcal{I}(T_1)} \mu.$$

This means that for  $f_0, f_1 \in L^\infty(\mu)$ , we have

$$\int (f_0 \otimes f_1)(x_0, x_1) d\mu_1(x_0, x_1) = \int \mathbb{E}(f_0 | \mathcal{I}(T_1)) \cdot \mathbb{E}(f_1 | \mathcal{I}(T_1)) d\mu.$$

For  $2 \leq k \leq d$ , we define a measure  $\mu_k$  (see [6]) on  $X^{2^k}$  by

$$\mu_k = \mu_{k-1} \times_{\mathcal{I}(T_k^\Delta)} \mu_{k-1},$$

where  $T_k^\Delta := \underbrace{T_k \times \cdots \times T_k}_{2^{k-1}}$ .

We write  $X^* = X^{2^d}$ , and points of  $X^*$  are written as  $x = (x_\epsilon : \epsilon \subset [d])$ . We write  $\mu^* := \mu_d$ .

For  $f \in L^\infty(\mu)$ , define

$$\|f\|_{T_1, \dots, T_d} := \left( \int \prod_{\epsilon \in \{0,1\}^d} f(x_\epsilon) d\mu^*(x) \right)^{1/2^d}.$$

It was shown in Proposition 2 in [6] that  $\|\cdot\|_{T_1, \dots, T_d}$  is a seminorm on  $L^\infty(\mu)$ . We call this the box seminorm associated to  $T_1, \dots, T_d$ .

For  $\epsilon \subset [d]$ ,  $\epsilon \neq \emptyset$ , we write  $\|\cdot\|_\epsilon$  for the seminorm on  $L^\infty(\mu)$  associated to the transformations  $T_i$ ,  $i \in \epsilon$ . For example,  $\|\cdot\|_{110\dots 00}$  is the seminorm associated to  $T_1, T_2$  because  $\epsilon = 110\dots 00 \in \{0,1\}^d$  is identified with  $\{1,2\} \subset [d]$ .

By Proposition 3 in [6], if we rearrange the order of the digits in  $\epsilon$ , the seminorm  $\|\cdot\|_\epsilon$  remains unchanged.

**2.2. Upper bound.** In the following, we assume that all functions  $f_\epsilon$ ,  $\epsilon \subset [d]$ , are real valued and satisfy  $|f_\epsilon| \leq 1$ .

**Proposition 2.1.** Maintaining the above notation and hypotheses,

(6)

$$\limsup_{\substack{N_i - M_i \rightarrow \infty \\ i=1, \dots, d}} \left\| \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1] \times \cdots \times [M_d, N_d]} \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} T_\epsilon^n f_\epsilon \right\|_{L^2(\mu)} \leq \min_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} \|f_\epsilon\|_{T_1 \dots T_d}.$$

*Proof.* We proceed by induction on  $d$ . For  $d = 1$ , we have

$$\left\| \frac{1}{N_1 - M_1} \sum_{n_1 \in [M_1, N_1]} T_1^{n_1} f_1 \right\|_{L^2(\mu)}^2 \rightarrow \int \mathbb{E}(f_1 | \mathcal{I}(T_1))^2 d\mu = \|f_1\|_{T_1}^2.$$

Let  $d \geq 2$  and assume that (6) is established for  $d-1$  transformations.

We show that for every  $\alpha \subset [d]$ ,  $\alpha \neq \emptyset$ , the limsup of the the left hand side of (6) is bounded by  $\|f_\alpha\|_{T_1, \dots, T_d}$ . By a permutation of digits

if needed, we can assume that  $\alpha \neq \underbrace{0 \dots 0}_{d-1} 1$ . The square of the norm in the left hand side of (6) is equal to

$$\left\| \frac{1}{N_d - M_d} \sum_{n_d \in [M_d, N_d]} T_d^{n_d} f_{0 \dots 01} \cdot \prod_{i=1}^{d-1} \frac{1}{N_i - M_i} \sum_{m \in [M_1, N_1] \times \dots \times [M_{d-1}, N_{d-1}]} \prod_{\substack{\eta \subset [d-1] \\ \eta \neq \emptyset}} T_\eta^m (f_{\eta 0} \cdot T_d^{n_d} f_{\eta 1}) \right\|_{L^2(\mu)}^2.$$

By the Cauchy-Schwartz Inequality, this is less than or equal to

$$(7) \quad \frac{1}{N_d - M_d} \sum_{n_d \in [M_d, N_d]} \left\| \prod_{i=1}^{d-1} \frac{1}{N_i - M_i} \sum_{m \in [M_1, N_1] \times \dots \times [M_{d-1}, N_{d-1}]} \prod_{\substack{\eta \subset [d-1] \\ \eta \neq \emptyset}} T_\eta^m (f_{\eta 0} \cdot T_d^{n_d} f_{\eta 1}) \right\|_{L^2(\mu)}^2.$$

By the induction hypothesis, when  $N_i - M_i \rightarrow \infty$ ,  $i = 1, \dots, d-1$ , the limsup of the square of the norm in (7) is less than or equal to

$$\min_{\substack{\eta \subset [d-1] \\ \eta \neq \emptyset}} \|f_{\eta 0} \cdot T_d^{n_d} f_{\eta 1}\|_{T_1, \dots, T_{d-1}}^2,$$

where  $\|\cdot\|_{T_1, \dots, T_{d-1}}$  is the seminorm associated to the  $d-1$  transformations  $T_1, \dots, T_{d-1}$ .

Note that  $\alpha$  is equal to  $\eta 0$  or  $\eta 1$  for some  $\eta \subset [d-1]$ , and by the Cauchy-Schwartz Inequality, we have

$$\begin{aligned}
& \lim_{N_d - M_d \rightarrow \infty} \frac{1}{N_d - M_d} \sum_{n_d \in [M_d, N_d]} \|f_{\eta 0} \cdot T_d^{n_d} f_{\eta 1}\|_{T_1, \dots, T_{d-1}}^{2^{d-1}} \\
&= \lim_{N_d - M_d \rightarrow \infty} \frac{1}{N_d - M_d} \sum_{n_d \in [M_d, N_d]} \int_{\eta \subset [d-1]} \bigotimes (f_{\eta 0} \cdot T_d^{n_d} f_{\eta 1}) d\mu_{d-1} \\
&= \int \mathbb{E} \left( \bigotimes_{\eta \subset [d-1]} f_{\eta 0} \middle| \mathcal{I}(T_d^\Delta) \right) \cdot \mathbb{E} \left( \bigotimes_{\eta \subset [d-1]} f_{\eta 1} \middle| \mathcal{I}(T_d^\Delta) \right) d\mu_{d-1} \\
&\leq \left( \int \left| \mathbb{E} \left( \bigotimes_{\alpha \subset [d]} f_\alpha \middle| \mathcal{I}(T_d^\Delta) \right) \right|^2 d\mu_{d-1} \right)^{1/2} \\
&= \left( \int \bigotimes_{\alpha \subset [d]} f_\alpha d\mu^* \right)^{1/2} = \|f_\alpha\|_{T_1, \dots, T_d}^{2^{d-1}}.
\end{aligned}$$

This completes the proof.  $\square$

The following proposition is a generalization of Proposition 2.1, although its proof depends upon Proposition 2.1.

**Proposition 2.2.** Let  $r$  be an integer with  $1 \leq r \leq d$ . Then

$$(8) \quad \limsup_{\substack{N_i - M_i \rightarrow \infty \\ i=1, \dots, d}} \left\| \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1] \times \dots \times [M_d, N_d]} \prod_{\substack{\epsilon \subset [d] \\ 0 < |\epsilon| \leq r}} T_\epsilon^n f_\epsilon \right\|_{L^2(\mu)} \leq \min_{\substack{\epsilon \subset [d] \\ |\epsilon|=r}} \|f_\epsilon\|_\epsilon$$

*Proof.* We show that for every  $\alpha \subset [d]$  with  $|\alpha| = r$ , the limsup in (8) is bounded by  $\|f_\alpha\|_\alpha$ .

By a permutation of digits we can restrict to the case that

$$\alpha = \underbrace{11 \dots 1}_r 00 \dots 0.$$

We show that

$$(9) \quad \limsup_{\substack{N_i - M_i \rightarrow \infty \\ i=1, \dots, d}} \left\| \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1] \times \dots \times [M_d, N_d]} \prod_{\substack{\epsilon \subset [d] \\ 0 < |\epsilon| \leq r}} T_\epsilon^n f_\epsilon \right\|_{L^2(\mu)} \leq \|f_\alpha\|_\alpha.$$

The norm in (9) is equal to

$$(10) \quad \left\| \prod_{i=r+1}^d \frac{1}{N_i - M_i} \sum_{m \in [M_{r+1}, N_{r+1}] \times \dots \times [M_d, N_d]} \prod_{\substack{\epsilon \subset \{r+1, \dots, d\} \\ \epsilon \neq \emptyset}} T_\epsilon^m f_\epsilon \cdot \prod_{j=1}^r \frac{1}{N_j - M_j} \sum_{n \in [M_1, N_1] \times \dots \times [M_r, N_r]} \left( \prod_{\substack{\eta \subset [r] \\ \eta \neq \emptyset}} T_\eta^n \prod_{\substack{\theta \subset [d-r] \\ |\eta\theta| \leq r}} T_{r+\theta}^n f_{\eta\theta} \right) \cdot (T_1^{n_1} \dots T_r^{n_r} f_\alpha) \right\|_{L^2(\mu)}.$$

where  $r + \theta = \{r + k : k \in \theta\}$ .

Let

$$g_\eta = \begin{cases} \prod_{\substack{\theta \subset [d-r] \\ |\eta\theta| \leq r}} T_{r+\theta}^n f_{\eta\theta} & 0 < |\eta| < r ; \\ f_\alpha & |\eta| = r . \end{cases}$$

Then (10) is equal to

$$(11) \quad \left\| \prod_{i=r+1}^d \frac{1}{N_i - M_i} \sum_{m \in [M_{r+1}, N_{r+1}] \times \dots \times [M_d, N_d]} \left( \prod_{\substack{\epsilon \subset \{r+1, \dots, d\} \\ \epsilon \neq \emptyset}} T_\epsilon^m f_\epsilon \right) \cdot \prod_{j=1}^r \frac{1}{N_j - M_j} \sum_{n \in [M_1, N_1] \times \dots \times [M_r, N_r]} \left( \prod_{\substack{\eta \subset [r] \\ \eta \neq \emptyset}} T_\eta^n g_\eta \right) \right\|_{L^2(\mu)}.$$

By the Cauchy-Schwartz Inequality, the square of (11) is less than or equal to

$$(12) \quad \prod_{i=r+1}^d \frac{1}{N_i - M_i} \sum_{m \in [M_{r+1}, N_{r+1}] \times \dots \times [M_d, N_d]} \left\| \prod_{j=1}^r \frac{1}{N_j - M_j} \sum_{n \in [M_1, N_1] \times \dots \times [M_r, N_r]} \prod_{\substack{\eta \subset [r] \\ \eta \neq \emptyset}} T_\eta^n g_\eta \right\|_{L^2(\mu)}^2.$$

By Proposition 2.1, the limsup of (12) as  $N_i - M_i \rightarrow \infty$ ,  $i = 1, \dots, r$  is bounded by

$$\prod_{i=r+1}^d \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i] \\ i=r+1, \dots, d}} \|f_\alpha\|_{T_1, \dots, T_r}^2 = \|f_\alpha\|_{T_1, \dots, T_r}^2.$$

This completes the proof.  $\square$

### 3. THE CASE OF THE MAGIC EXTENSION

We recall the definition of a “magic” system.

**Definition 3.1** (Host, [6]). A system  $(X, \mu, T_1, \dots, T_d)$  is called a “magic” system if  $f \in L^\infty(\mu)$  is such that  $\mathbb{E}(f | \bigvee_{i=1}^d \mathcal{I}(T_i)) = 0$ , then  $\|f\|_{T_1, \dots, T_d} = 0$ .

Given a system  $(X, \mu, T_1, \dots, T_d)$ , let  $X^*$  and  $\mu^*$  be defined as in Section 2.1. We denote by  $T_i^*$  the *side transformations* of  $X^*$ , given by

$$\text{for every } \epsilon \in \{0, 1\}^d, \quad (T_i^* x)_\epsilon = \begin{cases} T_i x_\epsilon & \text{if } \epsilon_i = 0; \\ x_\epsilon & \text{if } \epsilon_i = 1. \end{cases}$$

By Theorem 2 of [6],  $(X^*, \mu^*, T_1^*, \dots, T_d^*)$  is a “magic” system, and admits  $(X, \mu, T_1, \dots, T_d)$  as a factor.

For  $\epsilon \subset [d]$ ,  $\epsilon \neq \emptyset$ , we write  $\|\cdot\|_\epsilon^*$  for the seminorm on  $L^\infty(\mu^*)$  associated to the transformations  $T_i^*$ ,  $i \in \epsilon$ . Moreover, we define the  $\sigma$ -algebra

$$\mathcal{Z}_\epsilon^* := \bigvee_{i \in \epsilon} \mathcal{I}(T_i^*)$$

of  $(X^*, \mu^*)$ . For example,  $\mathcal{Z}_{\{1,2,d\}}^* = \mathcal{I}(T_1^*) \vee \mathcal{I}(T_2^*) \vee \mathcal{I}(T_d^*)$ .

We prove Theorem 1.1 for the magic system  $(X^*, \mu^*, T_1^*, \dots, T_d^*)$ .

**Theorem 3.2.** *Let  $f_\epsilon$ ,  $\epsilon \subset [d]$ , be functions on  $X^*$  with  $\|f_\epsilon\|_{L^\infty(\mu^*)} \leq 1$  for every  $\epsilon$ . Then the averages*

$$(14) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1] \times \dots \times [M_d, N_d]} \prod_{\substack{\epsilon \subset [d] \\ \epsilon \neq \emptyset}} T_\epsilon^{*n} f_\epsilon$$

converge in  $L^2(\mu^*)$  for all sequences of intervals  $[M_1, N_1], \dots, [M_d, N_d]$  whose lengths  $N_i - M_i$  ( $1 \leq i \leq d$ ) tend to  $\infty$ .

Since the system  $(X^*, \mu^*, T_1^*, \dots, T_d^*)$  admits  $(X, \mu, T_1, \dots, T_d)$  as a factor, Theorem 3.2 implies our main result Theorem 1.1.

**Theorem 3.3.** *For every  $\epsilon \subset [d]$ ,  $\epsilon \neq \emptyset$ , and every function  $f \in L^\infty(\mu^*)$ , we have:*

$$(15) \quad \text{If } \mathbb{E}_{\mu^*}(f | \mathcal{Z}_\epsilon^*) = 0, \text{ then } \|f\|_\epsilon^* = 0.$$

*Proof.* Assume  $|\epsilon| = r > 0$ . By a permutation of digits we can assume that

$$\epsilon = \{d - r + 1, d - r + 2, \dots, d\}.$$

We define a new system  $(Y, \nu, S_1, \dots, S_r)$ , where  $Y = X^{2^{d-r}}$ ,  $\nu = \mu_{d-r}$ , the  $d-r$  step measure associated to  $T_1^*, \dots, T_{d-r}^*$ . Define

$$S_i = \underbrace{T_{d-r+i} \times \dots \times T_{d-r+i}}_{2^{d-r}}$$

on  $Y$  for  $i = 1, \dots, r$ .

Note that by definition,  $Y^* = X^{2^d} = X^*$ , and

$$S_i^* = T_{d-r+i}^*, \quad S_i^\Delta = T_{d-r+i}^\Delta$$

for  $i = 1, \dots, r$ . Moreover,

$$\nu_1 = \nu \times_{\mathcal{I}(S_1)} \nu = \mu_{d-r} \times_{\mathcal{I}(T_{d-r+1}^\Delta)} \mu_{d-r} = \mu_{d-r+1}.$$

By induction,

$$\nu_{i+1} = \nu_i \times_{\mathcal{I}(S_{i+1}^\Delta)} \nu_i = \mu_{d-r+i} \times_{\mathcal{I}(T_{d-r+i+1}^\Delta)} \mu_{d-r+i} = \mu_{d-r+i+1},$$

for  $i = 1, \dots, r-1$ .

Therefore  $(X^*, \mu^*, T_{d-r+1}^*, \dots, T_d^*)$  is just the magic extension  $(Y^*, \nu^*, S_1^*, \dots, S_r^*)$  of  $(Y, \nu, S_1, \dots, S_r)$ . So

$$\mathcal{Z}_\epsilon^* = \bigvee_{i \in \epsilon} \mathcal{I}(T_i^*) = \bigvee_{i=1}^r \mathcal{I}(S_i^*) := \mathcal{W}_Y^*.$$

If  $f \in L^\infty(\mu^*)$  with  $\mathbb{E}_{\mu^*}(f | \mathcal{Z}_\epsilon^*) = 0$ , this is equivalent to  $\mathbb{E}_{\mu^*}(f | \mathcal{W}_Y^*) = 0$ , and by Theorem 2 in [6], we have  $\|f\|_{S_1^*, \dots, S_r^*}^* = 0$ . Thus  $\|f\|_\epsilon^* = \|f\|_{S_1^*, \dots, S_r^*}^* = 0$   $\square$

**Proposition 3.4.** Let  $f_\epsilon, \epsilon \subset [d]$ , be functions on  $X^*$  with  $\|f_\epsilon\|_{L^\infty(\mu^*)} \leq 1$  for every  $\epsilon$ . Let  $r$  be an integer with  $1 \leq r \leq d$ . Then the averages

$$(16) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{n \in [M_1, N_1) \times \dots \times [M_d, N_d)} \prod_{\substack{\epsilon \subset [d] \\ 0 < |\epsilon| \leq r}} T_\epsilon^{*n} f_\epsilon$$

converge in  $L^2(\mu^*)$  for all sequences of intervals  $[M_1, N_1), \dots, [M_d, N_d)$  whose lengths  $N_i - M_i$  ( $1 \leq i \leq d$ ) tend to  $\infty$ .

We remark that Theorem 3.2 follows immediately from this proposition when  $r = d$ .

*Proof.* We proceed by induction on  $r$ . When  $r = 1$ , the average (16) is

$$(17) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i) \\ i=1, \dots, d}} T_1^{*n_1} f_{10\dots 0} \cdots T_d^{*n_d} f_{0\dots 01}.$$

By the Ergodic Theorem, this converges to  $\mathbb{E}(f_{10\dots 0} | \mathcal{I}(T_1^*)) \cdots \mathbb{E}(f_{0\dots 01} | \mathcal{I}(T_d^*))$ .

Assume  $r > 1$ , and that the proposition is true for  $r - 1$  transformations.

For  $\alpha \subset [d]$ ,  $|\alpha| = r$ , if  $\mathbb{E}_{\mu^*}(f_\alpha \mid \mathcal{Z}_\alpha^*) = 0$ , then by Theorem 3.3, we have  $\|f_\alpha\|_\alpha^* = 0$ . By Proposition 2.1, the average (16) converges to 0. Otherwise, by a density argument, we can assume that

$$f_\alpha = \prod_{i \in \alpha} f_{\alpha,i}$$

where  $f_{\alpha,i}$  is  $T_i^*$ -invariant. Then

$$T_\alpha^{*n} f_\alpha = \prod_{i \in \alpha} T_{\alpha \setminus \{i\}}^{*n} f_{\alpha,i}.$$

Thus

$$\prod_{\substack{\epsilon \subset [d] \\ 0 < |\epsilon| \leq r}} T_\epsilon^{*n} f_\epsilon = \prod_{\substack{\eta \subset [d] \\ 0 < |\eta| \leq r-1}} T_\eta^{*n} g_\eta,$$

where

$$g_\eta = \begin{cases} f_\eta & |\eta| < r - 1; \\ f_\eta \prod_{i \notin \eta} f_{\eta \cup i,i} & |\eta| = r - 1. \end{cases}$$

Therefore (16) converges by the induction hypothesis.  $\square$

#### 4. COMBINATORIAL INTERPRETATION

*Proof of Theorem 1.2.* Apply Theorem 1.1 to the indicator function  $\mathbf{1}_A$ , we know that the limit of the averages

$$(18) \quad \prod_{i=1}^d \frac{1}{N_i - M_i} \sum_{\substack{n_i \in [M_i, N_i] \\ i=1, \dots, d}} \int \prod_{\epsilon \in \{0,1\}^d} T_1^{n_1 \epsilon_1} \dots T_d^{n_d \epsilon_d} \mathbf{1}_A d\mu$$

exist. By Lemma 1 in [6], if we take the limit as  $N_1 - M_1 \rightarrow \infty$ , then as  $N_2 - M_2 \rightarrow \infty$ , ... and then as  $N_d - M_d \rightarrow \infty$ , the average (18) converges to  $\|\mathbf{1}_A\|_{T_1, \dots, T_d}^{2^d}$ . Thus the limit of the average (18) is  $\|\mathbf{1}_A\|_{T_1, \dots, T_d}^{2^d}$ . Since

$$\begin{aligned} \|f\|_{T_1, \dots, T_d}^{2^d} &= \|\mathbb{E}(\bigotimes_{\epsilon \in [d-1]} f \mid \mathcal{I}(T_d^\Delta))\|_{L^2(\mu_{d-1})}^2 \\ &\geq \left( \int \bigotimes_{\epsilon \in [d-1]} f d\mu_{d-1} \right)^2 = \|f\|_{T_1, \dots, T_{d-1}}^{2^d}, \end{aligned}$$

we have  $\|\mathbf{1}_A\|_{T_1, \dots, T_d} \geq \|\mathbf{1}_A\|_{T_1} \geq \int \mathbf{1}_A d\mu = \mu(A)$ , and the result follows.  $\square$

Theorem 1.2 has the following corollary:

**Corollary 4.1.** Let  $(X, \mathcal{B}, \mu, T_1, \dots, T_d)$  be a system, where  $T_1, \dots, T_d$  are commuting measure preserving transformations, and let  $A \in \mathcal{B}$ . Then for any  $c > 0$ , the set of  $n \in \mathbb{Z}^k$  such that

$$\mu\left(\bigcap_{\epsilon \in \{0,1\}^d} T_1^{-n_1\epsilon_1} \dots T_d^{-n_d\epsilon_d} A\right) \geq \mu(A)^{2^d} - c$$

is syndetic.

The proof is exactly the same as Corollary 13.8 in [7].

Theorem 1.3 follows by combining Furstenberg's correspondence principle and Corollary 4.1.

#### REFERENCES

1. I. Assani. Pointwise convergence of ergodic averages along cubes. To appear, *J. d'Analyse Math.*
2. I. Assani. Averages along cubes for not necessarily commuting measure preserving transformations. *Contemp. Math.* vol. **430**, 1–19, 2007.
3. T. Austin. On the norm convergence of nonconventional ergodic averages. To appear, *Ergodic Theory Dynam. Systems.* arXiv:0805.0320v3.
4. V. Bergelson. The multifarious Poincaré recurrence theorem. *Descriptive set theory and dynamical systems (Marseille-Luminy, 1996)*, 31–57, London Math. Soc. Lecture Note Ser., 277, Cambridge Univ. Press, Cambridge, 2000.
5. J. P. Conze and E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France* **112** (1984), no. 2, 143–175.
6. B. Host. Ergodic seminorms for commuting transformations and applications. *Studia Math.* **195** (1) (2009), 31–49.
7. B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.* **161** (2005), no. 1, 397–488.
8. A. Leibman. Lower bounds for ergodic averages. *Ergodic Theory Dynam. Systems* **22** (2002), 863–872.
9. T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems.* **28** (2008), 657–688.
10. H. Towsner. Convergence of diagonal ergodic averages. *Ergodic Theory Dynam. Systems* **29** (2009), no. 4, 1309–1326.

## CHAPITRE V

### Récurrence multiple quantitative pour deux transformations qui commutent

L'étude présentée dans ce chapitre fait l'objet de l'article [17] qui a été accepté pour publication dans *Ergodic Theory Dynam. Systems*.

#### 1. Introduction

**1.1. Résultats de récurrence.** Nous avons vu dans le chapitre I qu'il y a deux types de résultats de récurrence : les résultats de récurrence qualitatifs représentés par le théorème de Poincaré, et les résultats de récurrence quantitatifs représentés par le théorème de Khintchine. Les théorèmes I.3, I.5, I.8 et I.10 sont des résultats de récurrence qualitatifs tandis que les théorèmes I.7, I.13, et I.14 sont des résultats de récurrence quantitatifs.

Puisque nous avons le théorème I.7 comme la contrepartie quantitative du théorème I.5, il est naturel de se demander s'il existe un résultat comme la contrepartie quantitative du théorème I.8. Dans ce chapitre, nous établissons le résultat suivant concernant deux transformations qui commutent :

**THÉORÈME V.1.** *Soit  $(X, \mathcal{X}, \mu, T_1, T_2)$  un système ergodique. Soit  $A \in \mathcal{X}$  un ensemble avec  $\mu(A) > 0$ . Alors pour tout  $\epsilon > 0$ , l'ensemble*

$$\{n \in \mathbb{Z} : \mu(A \cap T_1^n A \cap T_2^n A) > \mu(A)^4 - \epsilon\}$$

*est syndétique.*

Nous remarquons que d'après le même contre-exemple dans [14], l'hypothèse d'ergodicité est nécessaire pour le théorème.

A première vue, le fait qu'on a  $\mu(A)^4$  pour la borne inférieure n'est pas intuitif. On s'attend à  $\mu(A)^3$ , la même borne inférieure du théorème I.7 pour une seule transformation, mais on ne peut pas faire mieux à cause du théorème suivant qui indique que l'exposant 4 est optimal :

**THÉORÈME V.2.** *Pour tout  $0 < c \leq 1$ , il existe un système ergodique  $(X, \mathcal{X}, \mu, T_1, T_2)$  et un ensemble  $A \in \mathcal{X}$  avec  $\mu(A) > 0$  tels*

que

$$\mu(A \cap T_1^n A \cap T_2^n A) < c\mu(A)^3$$

pour tout entier  $n \neq 0$ .

Pourtant, nous avons l'exposant 3 pour la borne inférieure dans certains cas :

**THÉORÈME V.3.** *Soient  $(Y_1, \mathcal{Y}_1, \nu_1, S_1)$  et  $(Y_2, \mathcal{Y}_2, \nu_2, S_2)$  deux systèmes ergodiques. Soit  $(X, \mathcal{X}, \mu)$  l'espace mesuré produit de  $Y_1$  et  $Y_2$ . Soient  $T_1 = S_1 \times \text{Id}$ ,  $T_2 = \text{Id} \times S_2$ . Soit  $A \in \mathcal{X}$  un ensemble avec  $\mu(A) > 0$ . Alors pour tout  $\epsilon > 0$ , l'ensemble*

$$\{n \in \mathbb{Z} : \mu(A \cap T_1^n A \cap T_2^n A) > \mu(A)^3 - \epsilon\}$$

est syndétique.

**1.2. Résultats combinatoires.** La contrepartie combinatoire du théorème V.1 se déduit d'un principe de correspondance qui se base sur la densité de Banach. Rappelons que la *densité de Banach* d'un sous-ensemble  $E$  de  $\mathbb{Z}^2$  est définie par :

$$d^*(E) = \limsup_{\substack{N_1 - M_1 \rightarrow \infty \\ N_2 - M_2 \rightarrow \infty}} \frac{|E \cap [M_1, N_1] \times [M_2, N_2]|}{(N_1 - M_1) \times (N_2 - M_2)}.$$

La densité de Banach d'un sous-ensemble  $E$  de  $\mathbb{Z}^k$ ,  $k \geq 1$  se définit de la même manière.

Nous allons utiliser une variante du principe de correspondance (le théorème I.2) qui concerne seulement les systèmes ergodiques :

**THÉORÈME.** *Soit  $E \subset \mathbb{Z}^k$  un ensemble de densité positive. Alors il existe un système ergodique  $(X, \mathcal{X}, \mu, T_1, \dots, T_k)$  et un ensemble  $A \in \mathcal{X}$  avec  $\mu(A) = d^*(E)$  tels que*

$$\mu\left(\bigcap_{(m_1, \dots, m_k) \in F} T_1^{m_1} \dots T_k^{m_k} A\right) \leq d^*\left(\bigcap_{(m_1, \dots, m_k) \in F} (E + (m_1, \dots, m_k))\right)$$

pour tout sous-ensemble fini  $F$  de  $\mathbb{Z}^k$ .

La preuve de ce théorème se trouve dans [14] pour  $k = 1$ , et le cas général se déduit de la même démarche.

D'après ce principe de correspondance et le théorème V.1, on en déduit :

**COROLLAIRE V.4.** *Soit  $E \subset \mathbb{Z}^2$  un ensemble de densité de Banach positive. Alors pour tout  $\epsilon > 0$ , l'ensemble*

$$\{n \in \mathbb{Z} : d^*(E \cap (E + (n, 0)) \cap (E + (0, n))) > d^*(E)^4 - \epsilon\}$$

est syndétique.

**1.3. Contre-exemple.** Dans cette section, nous présentons la construction d'un contre-exemple pour montrer le théorème V.2.

Soit  $(Y, \nu, S)$  le schéma de Bernoulli  $B(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ . C'est-à-dire que  $\{0, 1, 2\}$  est muni de la mesure de probabilité uniforme  $\tau$ , et que  $Y = \{0, 1, 2\}^{\mathbb{Z}}$  est muni de la mesure produit  $\nu = \tau^{\mathbb{Z}}$ . Les points de  $Y$  s'écrivent comme  $y = (y_n : n \in \mathbb{Z})$  où  $y_n \in \{0, 1, 2\}$  pour chaque  $n$ . La transformation (appelée shift ou décalage)  $S$  est définie par  $(Sy)_n = y_{n+1}$  pour chaque  $n$ .

Définissons

$$(X, \mu, T_1, T_2) = (Y \times Y \times Y, \nu \times \nu \times \nu, S \times \text{Id} \times S, \text{Id} \times S \times S).$$

Puisque  $(Y, \nu, S)$  est faiblement mélangeant, le système  $(X, \mu, T_1, T_2)$  est ergodique.

On construit une fonction  $f$  sur  $\{0, 1, 2\}^3$  en donnant une valeur parmi 0 et 1 à  $f(i, j, k)$  pour chaque  $(i, j, k) \in \{0, 1, 2\}^3$ . Définissons la fonction  $F$  sur  $X$  par  $F(y, z, w) = f(y_0, z_0, w_0)$ . Alors  $F$  est une fonction indicatrice de certain sous-ensemble  $A \subset X$  par construction.

Si on a bien construit la fonction  $f$ , alors l'ensemble  $A$  obtenu de telle manière vérifie qu'il existe  $0 < c_0 < 1$  tel que pour tout  $n \neq 0$

$$\mu(A \cap T_1^n A \cap T_2^n A) < c_0(\mu(A))^3.$$

Pour  $0 < c \leq 1$ , il existe un entier  $l > 0$  tel que  $(c_0)^l < c$ . Soient  $(X', \mu', T'_1, T'_2)$  le système produit ( $l$ -fois) de  $(X, \mu, T_1, T_2)$ , et  $A'$  l'ensemble-produit  $A \times \cdots \times A$  ( $l$ -fois). Puisque le système  $Y^l$  est faiblement mélangeant, le système  $(X', \mu', T'_1, T'_2)$  est ergodique. Donc le système  $(X', \mu', T'_1, T'_2)$  et l'ensemble  $A'$  satisfont aux propriétés de l'énoncé.

**1.4. Méthodes.** Avant de présenter les méthodes, on introduit des notations :

NOTATION. On dit que *les moyennes de certaine suite  $(a_n)$  convergent uniformément* vers une limite  $L$ , et on écrit :

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n \in [M, N]} a_n = L$$

si les moyennes de  $a_n$  sur toutes suites d'intervalles  $[M_i, N_i)$  dont la longueur  $N_i - M_i$  tend vers l'infini convergent vers  $L$  :

$$\lim_{i \rightarrow \infty} \frac{1}{N_i - M_i} \sum_{n \in [M_i, N_i)} a_n = L.$$

Pour  $t \in \mathbb{T}$ , on utilise la notation standard  $e(t) = e^{2\pi it}$ .

Une méthode importante pour la preuve du théorème V.1 et du théorème V.3 est la stratégie suivante introduite par Frantzikinakis dans sa démonstration du même théorème I.7 (le théorème C de [25], il a étudié en fait le cas où la progression arithmétique  $\{n, 2n, 3n\}$  est remplacée par des polynômes  $\{p_1(n), p_2(n), p_3(n)\}$ ). Grâce à cette méthode, sa démonstration est plus simple et courte que celle de [14]. Étant donné une suite  $\{a_n\}$  et un nombre réel  $a > 0$ , pour montrer que l'ensemble

$$E := \{n \in \mathbb{Z} : a_n > a - \epsilon\}$$

est syndétique quel que soit  $\epsilon > 0$ , on introduit une rotation ergodique  $(Z, \alpha, \theta)$ , et on introduit une fonction  $\chi$  (dépendant de  $\epsilon$ ) non-négative, continue sur  $Z$  avec  $\int \chi d\theta = 1$ , et bien choisie de sorte que la limite inférieure uniforme des moyennes pondérées

$$\frac{1}{N - M} \sum_{n \in [M, N)} \chi(n\alpha) a_n$$

soit supérieure à  $a - \epsilon/2$ . Cette fonction étant construite, on en déduit que l'ensemble  $E$  est syndétique. Sinon, il existerait un  $\epsilon$  et une suite d'intervalles  $[M_i, N_i)$  dont la longueur tend vers l'infini, tels que  $a_n \leq a - \epsilon$  pour tout  $i$  et tout  $n \in [M_i, N_i)$ . Prenons les moyennes de  $\chi(n\alpha) a_n$  sur ces intervalles. Comme la rotation  $(Z, \alpha, \theta)$  est ergodique, la suite  $\{n\alpha\}$  est uniformément distribuée dans  $Z$ , ce qui implique que les moyennes de  $\chi(n\alpha)$  convergent vers 1. Il vient donc que

$$\liminf_{i \rightarrow \infty} \frac{1}{N_i - M_i} \sum_{n \in [M_i, N_i)} \chi(n\alpha) a_n \leq a - \epsilon,$$

ce qui est contradictoire.

Pour nous ici,  $a_n = \mu(A \cap T_1^n A \cap T_2^n A)$  et  $a = \mu(A)^4$ . D'après cette méthode, pour montrer le théorème V.1, l'objectif est de trouver une rotation ergodique  $Z$  et une fonction  $\chi$  satisfaisant aux conditions imposées. Avant cela, nous étudions d'abord les moyennes générales de la suite sous la forme  $e(nt)a_n$ .

Cette méthode est suffisante pour montrer le théorème V.3 dit "cas facile" tandis que pour montrer le théorème V.1, on a besoin de plus d'outils élaborés comme la machinerie des systèmes magiques introduite par Host [39]. En effet, d'après la première méthode, nous voulons étudier le comportement asymptotique des moyennes pondérées de  $e(nt)a_n$ , il est naturel de penser de transférer le problème à une extension magique du système original dans laquelle nous pourrions trouver des facteurs caractéristiques pour cette moyenne. Nous détaillerons cette méthode dans la section 1.6.

En outre, nous allons utiliser l'inégalité suivante, qui est liée à une classe d'inégalités étudiée par Sidorenko ([61], [62]). Un cas particulier a été montré par Atkinson, Watterson et Moran [7].

LEMME V.5. *Soient  $(X, \mathcal{X}, \mu)$  un espace probabilisé,  $k \geq 1$  un entier, et  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_k$ ,  $k$  sous- $\sigma$ -algèbres de  $\mathcal{X}$ . Alors pour toute fonction  $f$  bornée, non-négative sur  $X$ ,*

$$\int f \cdot \prod_{i=1}^k \mathbb{E}(f \mid \mathcal{X}_i) d\mu \geq \left( \int f d\mu \right)^{k+1} .$$

**1.5. Schéma de la preuve du théorème V.3.** Dans cette section, nous présentons l'organisation de la preuve du théorème V.3 afin d'illustrer comment on se sert de la première méthode. Une démarche similaire va être reprise pour la preuve du théorème V.1.

Dans la suite,  $(X, \mathcal{X}, T_1, T_2)$  est le système comme décrit dans le théorème. Nous montrons en fait le théorème suivant, le théorème V.3 s'en déduit avec  $f = \mathbf{1}_A$ .

THÉORÈME V.6. *Soit  $0 \leq f \leq 1$ . Alors pour tout  $\epsilon > 0$ , l'ensemble*

$$\{n \in \mathbb{Z} : \int f(x) \cdot f(T_1^n x) \cdot f(T_2^n x) d\mu(x) > \left( \int f d\mu \right)^3 - \epsilon\}$$

*est syndétique.*

Dans la suite, on fixe la fonction  $f$ , et pour chaque  $n$ , on écrit

$$I_n := \int f(x) \cdot f(T_1^n x) \cdot f(T_2^n x) d\mu(x) .$$

Désignons par  $(Z, \mathcal{Z}, \theta, R)$  le facteur commun des systèmes  $Y_1$  et  $Y_2$  engendré par les fonctions propres associées aux mêmes valeurs propres. Alors  $Z$  est un groupe compact abélien,  $\mathcal{Z}$  est sa  $\sigma$ -algèbre de Borel,  $\theta$  est sa mesure de Haar, et  $R$  est la translation par un élément fixé  $\alpha \in Z$ . On considère  $\mathcal{Z}$  comme une sous- $\sigma$ -algèbre de tous les  $(Y_1, \mathcal{Y}_1)$  et  $(Y_2, \mathcal{Y}_2)$ .

Posons

$$\hat{f} := \mathbb{E}(f \mid \mathcal{Z} \times \mathcal{Y}_2) \text{ et } \tilde{f} := \mathbb{E}(f \mid \mathcal{Y}_1 \times \mathcal{Z}) .$$

On considère les fonctions  $\hat{f}$  et  $\tilde{f}$  comme des fonctions définies respectivement sur  $Z \times Y_2$  et  $Y_1 \times Z$ .

Posons

$$J_n := \int f(x) \cdot \hat{f}(T_1^n x) \cdot \tilde{f}(T_2^n x) d\mu(x) .$$

**Étape 1 :** La première étape est de trouver des facteurs caractéristiques pour la moyenne de  $e(nt)I_n$ .

LEMME V.7. *Soit  $e(t)$  une valeur propre commune des systèmes  $Y_1$  et  $Y_2$ . Alors les moyennes de la différence  $e(nt)(I_n - J_n)$  convergent uniformément vers 0.*

D'après ce lemme, pour étudier le comportement asymptotique des moyennes pondérées de  $e(nt)I_n$ , il suffit d'étudier celui des moyennes de  $e(nt)J_n$ .

**Étape 2 :** Maintenant, nous avons une rotation ergodique  $Z$ , il reste à trouver une fonction  $\chi$  sur  $Z$  satisfaisant aux conditions imposées dans la section 1.4. Fixons  $\epsilon > 0$  dans le théorème. Par la continuité de la translation dans  $L^1(Z)$ , il existe un voisinage ouvert  $U$  de 0 dans  $Z$ , tel que

$$\text{pour tout } s \in U, \int |\widehat{f}(z + s, y_2) - \widehat{f}(z, y_2)| d\theta(z) d\nu_2(y_2) < \epsilon/4 ,$$

et

$$\text{pour tout } t \in U, \int |\widetilde{f}(y_1, z + t) - \widetilde{f}(y_1, z)| d\nu_1(y_1) d\theta(z) < \epsilon/4 .$$

Alors

$$(16) \quad |J_0 - J_n| < \epsilon/2 \quad \text{pour tout } n \text{ tel que } n\alpha \in U .$$

Choisissons une fonction  $\chi$  on  $Z$  continue et non-négative, vérifiant

$$\int \chi(z) d\theta(z) = 1 ,$$

et

$$\chi(z) = 0 \quad \text{if } z \notin U .$$

Par un argument de densité, il résulte du lemme V.7 le corollaire suivant :

COROLLAIRE V.8. *Les moyennes de la différence*

$$\chi(n\alpha)(I_n - J_n)$$

*convergent uniformément vers 0.*

**Étape 3 :** Par construction de la fonction  $\chi$ , plus précisément, d'après la relation (16) et le corollaire V.8, on en déduit le lemme suivant :

LEMME V.9.

$$\limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} \chi(n\alpha)(J_0 - I_n) \right| \leq \epsilon/2 .$$

Comme la suite  $\{n\alpha\}$  est uniformément distribuée dans  $Z$ , cela entraîne

$$\liminf_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n \in [M, N]} \chi(n\alpha) I_n \geq J_0 - \epsilon/2 .$$

**Étape 4 :** Nous avons construit la rotation ergodique  $Z$  et la fonction  $\chi$  vérifiant les conditions imposées pour la première méthode avec  $a = J_0$ . Pour conclure la preuve, il suffit de montrer

$$J_0 \geq \left( \int f d\mu \right)^3 .$$

Cette inégalité découle immédiatement du lemme V.5.

**1.6. Schéma de la preuve du théorème V.1.** Dans cette section, nous présentons l'organisation de la preuve du théorème V.1.

Dans la suite,  $(X, \mu, T_1, T_2)$  est un système ergodique.

Nous montrons en fait le théorème suivant, le théorème V.1 s'en déduit avec  $f = \mathbf{1}_A$ .

**THÉORÈME V.10.** *Soit  $0 \leq f \leq 1$ . Alors pour tout  $\epsilon > 0$ , l'ensemble*

$$\{n \in \mathbb{Z} : \int f(x) \cdot f(T_1^n x) \cdot f(T_2^n x) d\mu(x) > \left( \int f d\mu \right)^4 - \epsilon\}$$

*est syndétique.*

Dans la suite, on fixe la fonction  $f$ , et pour chaque  $n$ , on écrit

$$I_n := \int f(x) \cdot f(T_1^n x) \cdot f(T_2^n x) d\mu(x) .$$

**Étape 1 :** Comme nous avons expliqué plus haut, pour montrer le théorème, nous montrons en fait que le théorème est valable sur une extension magique de  $X$  afin que nous puissions utiliser la machinerie des systèmes magiques. Comme nous avons remarqué que l'hypothèse d'ergodicité est nécessaire pour le théorème, il faut que cette extension soit ergodique. Pourtant, cette condition n'est pas évidente, comme le système magique  $(X^*, \mu^*, T_1^*, T_2^*)$  construit dans la section 4.1 n'est jamais ergodique, sauf si  $X$  est trivial. Donc, dans un premier temps, nous montrons l'existence d'une extension magique et ergodique pour le système  $X$ , c'est une étape importante :

**THÉORÈME V.11.** *Chaque système ergodique admet une extension magique ergodique.*

Pour prouver ce théorème, on fait une désintégration ergodique pour le système  $(X^*, \mu^*, T_1^*, T_2^*)$ , on montre en fait que presque tout composant ergodique est magique.

**Étape 2 :** Dans cette partie, nous utilisons le théorème V.11 et la machinerie des systèmes magiques pour construire une extension magique ergodique de  $X$ , telle que notre problème se transforme en un nouveau problème sur cette extension. Appliquons le théorème V.11 trois fois pour différentes paires de transformations :

Posons  $T_3 = T_2 T_1^{-1}$ . D'abord on prend les transformations  $T_1, T_2$ , on fait une extension magique ergodique  $(X', \mu', T_1', T_2')$ . Notons  $\pi' : X' \rightarrow X$  l'application facteur et posons  $T_3' = T_2' T_1'^{-1}$ .

Ensuite, on prend les transformations  $T_1', T_3'$ , on fait une extension magique ergodique  $(X'', \mu'', T_1'', T_3'')$ . Notons  $\pi'' : X'' \rightarrow X'$  l'application facteur et posons  $T_2'' = T_1'' T_3''^{-1}$ .

Enfin, on prend les transformations  $T_2'', T_3''$ , on fait une extension magique ergodique  $(\bar{X}, \bar{\mu}, \bar{T}_2, \bar{T}_3)$ . Notons  $\bar{\pi} : \bar{X} \rightarrow X''$  l'application facteur et posons  $\bar{T}_1 = \bar{T}_2 \bar{T}_3^{-1}$ .

Alors les trois systèmes  $(\bar{X}, \bar{\mu}, \bar{T}_2, \bar{T}_3)$ ,  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_3)$  et  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_2)$  sont ergodiques, et  $\pi' \circ \pi'' \circ \bar{\pi}$  est une application facteur du système  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_2, \bar{T}_3)$  au système  $(X, \mu, T_1, T_2, T_3)$ .

Posons

$$\begin{aligned} h &: = f \circ \pi' \circ \pi'' \circ \bar{\pi} \\ g_0 &: = \mathbb{E}(h \mid \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_2)) , \\ g_1 &: = \mathbb{E}(h \mid \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_3)) , \\ g_2 &: = \mathbb{E}(h \mid \mathcal{I}(\bar{T}_2) \vee \mathcal{I}(\bar{T}_3)) . \end{aligned}$$

Pour chaque  $n$ , on écrit

$$J_n : = \int g_0(y) \cdot g_1(\bar{T}_1^n y) \cdot g_2(\bar{T}_2^n y) d\bar{\mu}(y) .$$

Nous avons le théorème suivant :

**THÉORÈME V.12.** *Pour tout  $t \in \mathbb{T}$ , les moyennes de la différence  $e(nt)(I_n - J_n)$  convergent uniformément vers 0.*

D'après ce théorème, nous pouvons désormais identifier les moyennes de  $e(n\alpha)I_n$  sur le système original  $X$  et les moyennes de  $e(n\alpha)J_n$  sur l'extension magique ergodique  $\bar{X}$  en passant à la limite. Maintenant, nous nous trouvons dans une extension de  $X$  : le système  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_2, \bar{T}_3)$ , et nous avons la suite  $J_n$  dans  $\bar{X}$  pour  $a_n$ , et  $(\int h d\bar{\mu})^4 = (\int f d\mu)^4$  pour  $a$ .

**Étape 3 :** En fait, nous pouvons nous restreindre au cas où les fonctions  $g_0, g_1, g_2$  sont mesurables par rapport à un facteur de  $\bar{X}$ . Posons

$$\bar{\mathcal{M}}: = \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_2) \vee \mathcal{I}(\bar{T}_3) ,$$

et

$$g: = \mathbb{E}(h \mid \bar{\mathcal{M}}) .$$

Nous remarquons que

$$\mathbb{E}(g \mid \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_2)) = g_0 ,$$

$$\mathbb{E}(g \mid \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_3)) = g_1 ,$$

$$\mathbb{E}(g \mid \mathcal{I}(\bar{T}_2) \vee \mathcal{I}(\bar{T}_3)) = g_2 .$$

De plus, nous avons

$$\int g d\bar{\mu} = \int h d\bar{\mu} = \int f d\mu .$$

Nous sommes maintenant dans un nouveau système  $(\bar{X}, \bar{\mathcal{M}}, \bar{\mu}, \bar{T}_1, \bar{T}_2, \bar{T}_3)$  et nous avons la suite  $J_n$  pour  $a_n$ , et  $(\int g d\bar{\mu})^4$  pour  $a$ .

**Étape 5 :** Il reste à donner des descriptions sur le système  $(\bar{X}, \bar{\mathcal{M}}, \bar{\mu}, \bar{T}_1, \bar{T}_2, \bar{T}_3)$ . En fait, nous nous retrouvons dans un contexte proche de celui du cas facile (le théorème V.3) : ce système est isomorphe à un système sur un espace produit  $Y = Y_1 \times Y_2 \times Y_2$ , muni de la mesure comme la mesure produit relativement indépendant au-dessus de  $Z \times Z \times Z$  et des transformations définies par :

$$S_1 = \text{Id} \times R_2 \times R_3 ,$$

$$S_2 = R_1 \times \text{Id} \times R_3 ,$$

$$S_3 = R_1 \times R_2^{-1} \times \text{Id} .$$

Où les systèmes  $(Y_i, \nu_i, R_i)$ ,  $i = 1, 2, 3$  sont ergodiques, et  $(Z, R)$  est le facteur commun des systèmes  $Y_1, Y_2$  et  $Y_3$  engendré par les fonctions propres associées aux mêmes valeurs propres.

À partir d'ici, nous utilisons une démarche similaire à celle pour la preuve du théorème V.3. En effet, nous avons le lemme suivant comme le lemme V.9 pour le cas facile :

LEMME V.13.

$$\limsup_{N \rightarrow \infty} \left| \frac{1}{N - M} \sum_{n \in [M, N]} \chi(n\alpha)(J_0 - J_n) \right| \leq \frac{\epsilon}{2} .$$

Et d'après le lemme V.5, nous avons l'inégalité

$$J_0 \geq \left( \int g d\nu \right)^4 = \left( \int f d\mu \right)^4 .$$

Ainsi nous concluons la preuve.

**1.7. Questions.** Nous présentons ici quelques questions concernant l'étude de ce chapitre :

**Question 1 :** Une question naturelle est : qu'est-ce qui se passe pour le cas de trois transformations qui commutent ? Autrement dit, on se demande s'il existe un entier  $s$  tel que pour tout système ergodique  $(X, \mathcal{X}, \mu, T_1, T_2, T_3)$  et pour tout sous-ensemble  $A \in \mathcal{X}$  avec  $\mu(A) > 0$ , l'ensemble  $\{n \in \mathbb{Z} : \mu(A \cap T_1^n A \cap T_2^n A \cap T_3^n A) > \mu(A)^s - \epsilon\}$  soit syndétique.

**Question 2 :** La récurrence polynomiale a été démontrée par Bergelson et Leibman (le théorème I.10). Un résultat quantitatif a été obtenu par Frantzikinakis and Kra (le théorème I.13) pour les systèmes munis d'une seule transformation sous l'hypothèse que les polynômes sont linéairement indépendants. Leur résultat peut-il être généralisé pour les systèmes munis de transformations qui commutent ? Autrement dit, soient  $(X, \mathcal{X}, \mu, T_1, \dots, T_k)$  un système,  $p_1, \dots, p_k$  des polynômes linéairement indépendants et  $A \in \mathcal{X}$ . Peut-on dire que l'ensemble  $\{n \in \mathbb{Z} : \mu(A \cap T_1^{p_1(n)} A \cap \dots \cap T_k^{p_k(n)} A) \geq \mu(A)^{k+1} - \epsilon\}$  est syndétique ? Récemment, Chu, Frantzikinakis, et Host [20] ont donné une réponse affirmative pour cette question lorsque  $p_i(n) = n^{d_i}$ ,  $i = 1, \dots, k$  pour des entiers positifs distincts  $d_1, \dots, d_k$ .

# MULTIPLE RECURRENCE FOR TWO COMMUTING TRANSFORMATIONS

QING CHU

ABSTRACT. This paper is devoted to a study of the multiple recurrence of two commuting transformations. We derive a result which is similar but not identical to that of one single transformation established by Bergelson, Host and Kra. We will use the machinery of “magic systems” established recently by B. Host for the proof.

## 1. INTRODUCTION

1.1. **History and results.** Let  $(X, \mathcal{X}, \mu, T)$  be an invertible measure preserving system, and  $A$  be a set of positive measure. The Khintchine’s Recurrence Theorem [14] states that for every  $\epsilon > 0$ , the set

$$\{n \in \mathbb{Z}: \mu(A \cap T^n A) > \mu(A)^2 - \epsilon\}$$

is syndetic. More recently, Furstenberg [9] proved a Multiple Recurrence Theorem, showing that under the same assumptions, the set

$$\{n \in \mathbb{Z}: \mu(A \cap T^n A \cap T^{2n} A \cap \dots \cap T^{kn} A) > 0\}$$

is syndetic for every integer  $k \geq 1$ . Aiming at a simultaneous extension of Khintchine’s and Furstenberg’s Recurrence theorems, Bergelson, Host and Kra [4] established the following result:

**Theorem** (Bergelson, Host, Kra). *Let  $(X, \mathcal{X}, \mu, T)$  be an ergodic invertible measure preserving system and  $A \in \mathcal{X}$  with  $\mu(A) > 0$ . Then for every  $\epsilon > 0$ , the sets*

$$\{n \in \mathbb{Z}: \mu(A \cap T^n A \cap T^{2n} A) > \mu(A)^3 - \epsilon\}$$

and

$$\{n \in \mathbb{Z}: \mu(A \cap T^n A \cap T^{2n} A \cap T^{3n} A) > \mu(A)^4 - \epsilon\}$$

are syndetic.

---

2000 *Mathematics Subject Classification.* 37A05, 37A30.

*Key words and phrases.* Multiple recurrence, commuting transformations, magic system, ergodic seminorms.

We recall that a subset  $E$  of  $\mathbb{Z}$  is said to be *syndetic* if there exists an integer  $N > 0$  such that  $E \cap [M, M + N) \neq \emptyset$  for every  $M \in \mathbb{Z}$ .

It was shown in [4] that an analogous result fails both if we remove the assumption of ergodicity and for longer arithmetic progressions.

Furstenberg and Katznelson [11] generalized Furstenberg's Recurrence theorem to commuting transformations. It is therefore natural to ask the question if a result analogous to the above theorem can be established for commuting transformations. We prove the following result regarding the case of two transformations:

**Theorem 1.1.** *Let  $(X, \mathcal{X}, \mu)$  be a probability space, and  $T_1, T_2$  be two commuting invertible measure preserving transformations. Assume that  $(X, \mathcal{X}, \mu, T_1, T_2)$  is ergodic. Let  $A \in \mathcal{X}$  with  $\mu(A) > 0$ . Then for every  $\epsilon > 0$ , the set*

$$\{n \in \mathbb{Z} : \mu(A \cap T_1^n A \cap T_2^n A) > \mu(A)^4 - \epsilon\}$$

*is syndetic.*

We remark that, by the same counterexample as in [4], the hypothesis of ergodicity is necessary for the theorem.

The fundamental difference with the case of a single transformation is that the exponent 4 can not be replaced by 3:

**Theorem 1.2.** *For every  $0 < c \leq 1$ , there exist a probability space  $(X, \mathcal{X}, \mu)$ , with two commuting invertible measure preserving transformations  $T_1, T_2$  such that  $(X, \mathcal{X}, \mu, T_1, T_2)$  is ergodic, and a measurable set  $A \in \mathcal{X}$ , with  $\mu(A) > 0$ , such that*

$$\mu(A \cap T_1^n A \cap T_2^n A) < c\mu(A)^3$$

*for every integer  $n \neq 0$ .*

However, we have the exponent 3 for some class of systems:

**Theorem 1.3.** *Let  $(Y_1, \mathcal{Y}_1, \nu_1, S_1)$  and  $(Y_2, \mathcal{Y}_2, \nu_2, S_2)$  be two ergodic invertible measure preserving systems. Let  $(X, \mathcal{X}, \mu)$  be the product measure space  $(Y_1 \times Y_2, \mathcal{Y}_1 \otimes \mathcal{Y}_2, \nu_1 \times \nu_2)$ , and let  $T_1 = S_1 \times \text{Id}$ ,  $T_2 = \text{Id} \times S_2$ . Let  $A \in \mathcal{X}$  with  $\mu(A) > 0$ . Then for every  $\epsilon > 0$ , the set*

$$\{n \in \mathbb{Z} : \mu(A \cap T_1^n A \cap T_2^n A) > \mu(A)^3 - \epsilon\}$$

*is syndetic.*

We recall a definition:

**Definition 1.4.** The *upper Banach density* of a subset  $E$  of  $\mathbb{Z}^2$  is:

$$d^*(E) = \limsup_{\substack{N_1 - M_1 \rightarrow \infty \\ N_2 - M_2 \rightarrow \infty}} \frac{|E \cap [M_1, N_1) \times [M_2, N_2)|}{(N_1 - M_1) \times (N_2 - M_2)}.$$

Using a variation (see [4]) of Multidimensional Furstenberg's Correspondence Principle [10] and Theorem 1.1, we deduce:

**Corollary 1.5.** *Let  $E \subset \mathbb{Z}^2$  be a subset with positive upper Banach density. Then for every  $\epsilon > 0$ , the set*

$$\{n \in \mathbb{Z}: d^*(E \cap (E + (n, 0)) \cap (E + (0, n))) > d^*(E)^4 - \epsilon\}$$

*is syndetic.*

**1.2. Questions.** We address here some questions which are related to this paper and remain open.

**Question 1:** A natural question is how about the case of three commuting transformations. We ask if there exists some integer  $s$ , such that, for every ergodic system  $(X, \mathcal{X}, \mu, T_1, T_2, T_3)$  and every set  $A \in \mathcal{X}$  with  $\mu(A) > 0$ , the set  $\{n \in \mathbb{Z}: \mu(A \cap T_1^n A \cap T_2^n A \cap T_3^n A) > \mu(A)^s - \epsilon\}$  is syndetic.

**Question 2:** The Polynomial Recurrence Theorem was proved by Bergelson and Leibman [3]. Frantzikinakis and Kra [8] provided a more precise result for any family of linearly independent integer polynomials:

**Theorem** (Frantzikinakis, Kra). *Let  $(X, \mathcal{X}, \mu, T)$  be an invertible measure preserving system, and  $p_1, \dots, p_k$  be linearly independent integer polynomials with  $p_i(0) = 0$  for  $i = 1, \dots, k$ , and  $A \in \mathcal{X}$ . Then for every  $\epsilon > 0$ , the set*

$$\{n \in \mathbb{Z}: \mu(A \cap T^{p_1(n)} A \cap \dots \cap T^{p_k(n)} A) \geq \mu(A)^{k+1} - \epsilon\}$$

*is syndetic.*

Can this result be generalized for commuting transformations? Let  $(X, \mathcal{X}, \mu)$  be a probability space, and  $T_1, \dots, T_k$  be commuting invertible measure preserving transformations. Let  $p_1, \dots, p_k$  be as in the above theorem, and  $A \in \mathcal{X}$ . It is true that the set  $\{n \in \mathbb{Z}: \mu(A \cap T_1^{p_1(n)} A \cap \dots \cap T_k^{p_k(n)} A) \geq \mu(A)^{k+1} - \epsilon\}$  is syndetic? Very recently, Chu, Frantzikinakis, and Host [6] gave an affirmative answer to this question when  $p_i(n) = n^{d_i}$ ,  $i = 1, \dots, k$  for distinct positive integers  $d_1, \dots, d_k$ .

### 1.3. Conventions and notation.

1.3.1. As usual, we can restrict to the case that all the probability spaces that we deal with are standard.

In general, we write  $(X, \mu)$  for a probability space, omitting the  $\sigma$ -algebra. When needed, the  $\sigma$ -algebra of a the probability space  $(X, \mu)$  is written  $\mathcal{X}$ .

We implicitly assume that the term “bounded function” means real-valued, bounded and measurable.

If  $S$  is a measure-preserving transformation of a probability space  $(X, \mathcal{X}, \mu)$  then we write  $\mathcal{I}(S)$  for the sub- $\sigma$ -algebra of  $\mathcal{X}$  consisting in  $S$ -invariant sets.

Let  $(X_1, \mathcal{X}_1, \mu_1)$ ,  $(X_2, \mathcal{X}_2, \mu_2)$  be two probability spaces. Let  $f_1 \in L^\infty(\mu_1)$ , and  $f_2 \in L^\infty(\mu_2)$ , we denote by  $f_1 \otimes f_2$  the function on  $X_1 \times X_2$  given by  $f_1 \otimes f_2(x_1, x_2) = f_1(x_1)f_2(x_2)$ .

1.3.2. Throughout this paper, by a *system*, we mean a probability space  $(X, \mathcal{X}, \mu)$  endowed with a single or several commuting measure preserving invertible transformations.

For a system  $(X, \mu, T_1, T_2)$ , a *factor* is a system  $(Y, \nu, S_1, S_2)$  and a measurable map  $\pi: X \rightarrow Y$  such that the image  $\pi(\mu)$  of  $\mu$  under  $\pi$  is equal to  $\nu$  and  $S_i \circ \pi = \pi \circ T_i$ ,  $i = 1, 2$ ,  $\mu$ -a.e.

If  $f$  is an integrable function on  $X$ , we write  $\mathbb{E}(f | Y)$  for the function on  $Y$  defined by

$$\text{for all } g \in L^\infty(\nu), \quad \int_X f \cdot g \circ \pi \, d\mu = \int_Y \mathbb{E}(f|Y) \cdot g \, d\nu .$$

1.3.3. We say that *the averages of some sequence*  $(a_n)$  *converge uniformly* to some limit  $L$ , and we write:

$$\lim_{N-M \rightarrow \infty} \frac{1}{N-M} \sum_{n \in [M, N]} a_n = L$$

if the averages of  $a_n$  on any sequences of intervals  $[M_i, N_i)$  whose lengths  $N_i - M_i$  tend to infinity converge to  $L$ :

$$\lim_{i \rightarrow \infty} \frac{1}{N_i - M_i} \sum_{n \in [M_i, N_i)} a_n = L.$$

In other words, for every  $\epsilon > 0$ , there exists an integer  $R > 0$  which depends only on  $\epsilon$ , such that

$$\left| \frac{1}{N-M} \sum_{n \in [M, N]} a_n - L \right| < \epsilon$$

for all  $N, M$  such that  $N - M \geq R$ .

For  $t \in \mathbb{T}$ , we use the standard notation  $e(t) = \exp(2\pi it)$ .

#### 1.4. Methods.

1.4.1. The first ingredient used in the proofs of Theorem 1.1 and Theorem 1.3 is a method introduced by Frantzikinakis [7]: in order to show some sequence  $I_n$  is large for  $n$  in a syndetic set, we show that the averages of  $I_n$  over an appropriately chosen sequence of intervals converge uniformly to a large limit. More precisely, we introduce an ergodic rotation  $(Z, \alpha)$ , and we show that for some well chosen non-negative continuous function  $\chi$  on  $Z$ , the weighted averages of  $\chi(n\alpha)I_n$  converge uniformly to a large limit. This strategy is sufficient for the proof of Theorem 1.3 (Section 2).

However, for the general case (Theorem 1.1), before using the first ingredient, we need first to make some reduction. We will use more elaborated tools such as the machinery of “magic systems” introduced recently by Host [12].

Ever since Tao [17] proved the norm convergence of multiple ergodic averages with several commuting transformations of the form

$$(1) \quad \frac{1}{N} \sum_{n=1}^N T_1^n f_1 \cdot \dots \cdot T_d^n f_d ,$$

several other proofs were given with different approaches by Austin [1], Host [12] and Towsner [18]. Among these proofs, those by Austin and Host were proceeded by building an *extension* of the original system with good properties, called “pleasant” system (using terminology from [1]) and “magic” system (using terminology from [12]). We will follow this idea of building a suitable extension system.

We first prove that every ergodic system has an ergodic magic extension. Then we use this result three times with different transformations, and we build an extension of the original system with good properties. We translate our problem in the original system to a question in this new system. By using the properties of magic systems, it suffices to prove the result on a factor of the new system. Finally, we give a description of this factor, and we find that we are in a situation very similar to that of the product case.

We will review “magic systems” and the corresponding properties needed in Section 3, and give the proof of Theorem 1.1 in Section 4.

1.4.2. Another important ingredient is the following inequality related to a class of inequalities studied by Sidorenko ([15], [16]). A particular case was proved by Atkinson, Watterson, and Moran [2].

**Lemma 1.6.** *Let  $(X, \mathcal{X}, \mu)$  be a probability space,  $k \geq 1$  be an integer, and  $\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_k$  be  $k$  sub- $\sigma$ -algebras of  $\mathcal{X}$ . For any bounded non-negative function  $f$  on  $X$ , we have*

$$(2) \quad \int f \cdot \prod_{i=1}^k \mathbb{E}(f \mid \mathcal{X}_i) d\mu \geq \left( \int f d\mu \right)^{k+1} .$$

*Proof.* We can restrict to the case that the function  $f$  is bounded below by some positive constant  $\epsilon$ . Indeed, for the general case, it suffices to apply the inequality to the function  $f + \epsilon$  and to take the limit of both sides when  $\epsilon$  tends to 0.

We write

$$(3) \quad f = \left( f \cdot \prod_{i=1}^k \mathbb{E}(f \mid \mathcal{X}_i) \right)^{1/k+1} \cdot \prod_{i=1}^k \left( \frac{f}{\mathbb{E}(f \mid \mathcal{X}_i)} \right)^{1/k+1} .$$

Let  $J$  be the integral on the left hand side of (2). By the Hölder Inequality and (3),

$$(4) \quad \left( \int f d\mu \right)^{k+1} \leq J \cdot \prod_{i=1}^k \int \frac{f}{\mathbb{E}(f \mid \mathcal{X}_i)} d\mu .$$

On the other hand, for  $1 \leq i \leq k$ ,

$$\int \frac{f}{\mathbb{E}(f \mid \mathcal{X}_i)} d\mu = 1 ,$$

and this proves the inequality.  $\square$

**Acknowledgement.** The author would like to thank her advisor, Bernard Host, for many helpful discussions and suggestions.

## 2. PROOF OF THEOREM 1.3

In this section, we assume that  $(Y_1, \nu_1, S_1)$  and  $(Y_2, \nu_2, S_2)$  are two ergodic systems, and that  $(X, \mu, T_1, T_2) = (Y_1 \times Y_2, \nu_1 \times \nu_2, S_1 \times \text{Id}, \text{Id} \times S_2)$ . Then  $(X, \mu, T_1, T_2)$  is ergodic.

For every  $n$ , and all bounded functions  $\eta, \varphi, \psi$  on  $X$ , we define

$$\begin{aligned} I_n(\eta, \varphi, \psi) &: = \int \eta \cdot T_1^n \varphi \cdot T_2^n \psi d\mu \\ &= \int \eta(y_1, y_2) \varphi(S_1^n y_1, y_2) \psi(y_1, S_2^n y_2) d\nu_1(y_1) d\nu_2(y_2) . \end{aligned}$$

We prove:

**Theorem 2.1.** *Let  $0 \leq f \leq 1$ . Then for every  $\epsilon > 0$ , the set*

$$\{n \in \mathbb{Z}: I_n(f, f, f) > (\int f d\mu)^3 - \epsilon\}$$

*is syndetic.*

Theorem 1.3 follows from Theorem 2.1 with  $f = \mathbf{1}_A$ .

**2.1. An ergodic rotation.** Let  $(Z, \mathcal{Z}, \theta, R)$  denote the common factor of systems  $Y_1$  and  $Y_2$  spanned by the eigenfunctions corresponding to the common eigenvalues of these two systems.

We recall that  $Z$  is a compact abelian group, endowed with a Borel  $\sigma$ -algebra  $\mathcal{Z}$  and Haar measure  $\theta$ .  $R$  is the translation by some fixed  $\alpha \in Z$ . We consider  $\mathcal{Z}$  as a sub- $\sigma$ -algebra of both  $(Y_1, \mathcal{Y}_1)$  and  $(Y_2, \mathcal{Y}_2)$ .

For a bounded function  $\eta$  on  $X$ , we write

$$\hat{\eta}: = \mathbb{E}(\eta \mid \mathcal{Z} \times \mathcal{Y}_2) \text{ and } \tilde{\eta}: = \mathbb{E}(\eta \mid \mathcal{Y}_1 \times \mathcal{Z}).$$

We begin with a lemma:

**Lemma 2.2.**

$$\mathcal{I}(S_1 \times S_2) \subset \mathcal{Z} \times \mathcal{Z}.$$

*Proof.* For  $i = 1, 2$ , let  $\mathcal{K}_i$  be the Kronecker factor of  $(Y_i, S_i)$ . It is known (see for example [10], Lemma 4.18) that any  $S_1 \times S_2$ -invariant function on  $Y_1 \times Y_2$  is measurable with respect to  $\mathcal{K}_1 \times \mathcal{K}_2$ .

Let  $\{f_i\}$  (resp.  $\{g_j\}$ ) be an orthonormal basis of  $L^2(\mathcal{K}_1, \nu_1)$  (resp.  $L^2(\mathcal{K}_2, \nu_2)$ ) consisting of eigenfunctions of  $S_1$  (resp.  $S_2$ ), then the set  $\{f_i \otimes \bar{g}_j\}$  is an orthonormal basis of  $L^2(\mathcal{K}_1 \times \mathcal{K}_2, \nu_1 \times \nu_2)$ .

Any  $F \in L^\infty(Y_1 \times Y_2)$  such that  $F$  is invariant under  $S_1 \times S_2$  can be written as

$$F = \sum_{i,j} c_{i,j} f_i \otimes \bar{g}_j.$$

for some constants  $c_{i,j}$  with  $\sum_{i,j} |c_{i,j}|^2 \leq \infty$ . Write  $S_1 f_i = e(t_i) f_i$ ,  $S_2 g_j = e(s_j) g_j$ . By invariance of  $F$  under  $S_1 \times S_2$ , we have

$$\sum_{i,j} c_{i,j} f_i \otimes \bar{g}_j = \sum_{i,j} c_{i,j} e(t_i - s_j) f_i \otimes \bar{g}_j.$$

Hence  $c_{i,j} \neq 0$  only when  $t_i = s_j$ . This completes the proof.  $\square$

**Lemma 2.3.** *Let  $\eta, \varphi, \psi$  be bounded functions on  $X$ . Then the averages of the difference  $I_n(\eta, \varphi, \psi) - I_n(\eta, \hat{\varphi}, \tilde{\psi})$  converge uniformly to 0.*

*Proof.* Without loss of generality, we can assume that  $\max\{|\eta|, |\varphi|, |\psi|\} \leq 1$ . By a density argument, we can suppose that  $\eta(y_1, y_2) = \alpha_1(y_1)\beta_1(y_2)$ ,  $\varphi(y_1, y_2) = \alpha_2(y_1)\beta_2(y_2)$ , and that  $\psi(y_1, y_2) = \alpha_3(y_1)\beta_3(y_2)$  for some

bounded functions  $\alpha_i$  on  $Y_1$  and  $\beta_i$  on  $Y_2$ ,  $i = 1, 2, 3$ , all of them being bounded by 1 in absolute value. Using the notation introduced above, we have

$$\widehat{\varphi}(y_1, y_2) = \mathbb{E}(\alpha_2 \mid \mathcal{Z})(y_1) \cdot \beta_2(y_2) \text{ and } \widetilde{\psi}(y_1, y_2) = \alpha_3(y_1) \cdot \mathbb{E}(\beta_3 \mid \mathcal{Z})(y_2) .$$

For every  $n$ ,

$$\begin{aligned} I_n(\eta, \varphi, \psi) - I_n(\eta, \widehat{\varphi}, \widetilde{\psi}) &= \int (\alpha_1 \alpha_3)(y_1) \cdot (\beta_1 \beta_2)(y_2) \cdot \left( \alpha_2(S_1^n y_1) \cdot \beta_3(S_2^n y_2) \right. \\ &\quad \left. - \mathbb{E}(\alpha_2 \mid \mathcal{Z})(S_1^n y_1) \cdot \mathbb{E}(\beta_3 \mid \mathcal{Z})(S_2^n y_2) \right) d\nu(y_1) d\nu(y_2) . \end{aligned}$$

By the Ergodic Theorem, the averages of this expression converge uniformly to

$$\begin{aligned} &\int (\alpha_1 \alpha_3) \otimes (\beta_1 \beta_2) \cdot \left[ \mathbb{E} \left( (\alpha_2 - \mathbb{E}(\alpha_2 \mid \mathcal{Z})) \otimes \beta_3 \mid \mathcal{I}(S_1 \times S_2) \right) + \right. \\ &\quad \left. \mathbb{E} \left( \mathbb{E}(\alpha_2 \mid \mathcal{Z}) \otimes (\beta_3 - \mathbb{E}(\beta_3 \mid \mathcal{Z})) \mid \mathcal{I}(S_1 \times S_2) \right) \right] d\nu \times \nu . \end{aligned}$$

By Lemma 2.2, this is equal to 0. This completes the proof.  $\square$

**Lemma 2.4.** *For every common eigenvalue  $e(t)$  of  $Y_1$  and  $Y_2$ , the averages of the difference*

$$e(nt)(I_n(f, f, f) - I_n(f, \widehat{f}, \widetilde{f}))$$

*converge uniformly to 0.*

*Proof.* Let  $e(t)$  be a common eigenvalue of  $Y_1$  and  $Y_2$ ,  $h$  be a corresponding eigenfunction of  $Y_1$  with  $|h| = 1$ . Taking  $\eta(y_1, y_2) = f(y_1, y_2)\overline{h}(y_1)$ ,  $\varphi(y_1, y_2) = f(y_1, y_2)h(y_1)$ , and  $\psi(y_1, y_2) = f(y_1, y_2)$ , we have

$$e(nt)I_n(f, f, f) = I_n(\eta, \varphi, \psi) .$$

Since  $\widehat{\varphi}(y_1, y_2) = h(y_1)\widehat{f}(y_1, y_2)$  and  $\widetilde{\psi}(y_1, y_2) = \widetilde{f}(y_1, y_2)$ , we have

$$e(nt)I_n(f, \widehat{f}, \widetilde{f}) = I_n(\eta, \widehat{\varphi}, \widetilde{\psi}) .$$

By Lemma 2.3, the averages of the difference  $I_n(\eta, \varphi, \psi) - I_n(\eta, \widehat{\varphi}, \widetilde{\psi})$  converge uniformly to 0. This completes the proof.  $\square$

**2.2. A well chosen function  $\chi$ .** In the sequel, we fix  $f$ ,  $\epsilon > 0$  as in Theorem 2.1. By continuity of the translation in  $L^1(Z)$ , we can choose an open neighborhood  $U$  of 0 in  $Z$ , such that

$$\text{for every } s \in U, \int |\widehat{f}(z + s, y_2) - \widehat{f}(z, y_2)| d\theta(z) d\nu_2(y_2) < \epsilon/4 ,$$

and

$$\text{for every } t \in U, \int |\widetilde{f}(y_1, z + t) - \widetilde{f}(y_1, z)| d\nu_1(y_1) \theta(z) < \epsilon/4 .$$

We have

$$(5) \quad |I_0(f, \widehat{f}, \widetilde{f}) - I_n(f, \widehat{f}, \widetilde{f})| < \epsilon/2 \text{ for every } n \text{ such that } n\alpha \in U .$$

We choose a continuous function  $\chi$  on  $Z$ , non-negative, satisfying

$$\int \chi(z) d\theta(z) = 1 ,$$

and

$$\chi(z) = 0 \text{ if } z \notin U .$$

Since the linear combinations of characters are dense in the space of continuous functions on  $Z$  in the uniform norm (Stone-Weierstrass Theorem), from Lemma 2.4, we immediately get

**Corollary 2.5.** *The averages of the difference*

$$\chi(n\alpha)(I_n(f, f, f) - I_n(f, \widehat{f}, \widetilde{f}))$$

*converge uniformly to 0.*

We have

**Lemma 2.6.**

$$(6) \quad \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} \chi(n\alpha)(I_0(f, \widehat{f}, \widetilde{f}) - I_n(f, f, f)) \right| \leq \epsilon/2 .$$

*Proof.* By Corollary 2.5, it suffices to prove

$$(7) \quad \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} \chi(n\alpha)(I_0(f, \widehat{f}, \widetilde{f}) - I_n(f, \widehat{f}, \widetilde{f})) \right| \leq \epsilon/2 .$$

If  $n\alpha \notin U$ , then  $\chi(n\alpha)(I_0(f, \widehat{f}, \widetilde{f}) - I_n(f, \widehat{f}, \widetilde{f})) = 0$ . If  $n\alpha \in U$ , then by (5),

$$\chi(n\alpha)|I_0(f, \widehat{f}, \widetilde{f}) - I_n(f, \widehat{f}, \widetilde{f})| < \chi(n\alpha) \cdot \epsilon/2 .$$

By ergodicity of  $(Z, \theta, R)$ , the sequence  $\{n\alpha\}$  is uniformly distributed in  $Z$ . Therefore the averages of  $\chi(n\alpha)$  converge uniformly to 1, and this gives (7).  $\square$

**2.3. End of the proof.** Applying Lemma 1.6 with  $k = 2$ ,  $\mathcal{X}_1 = \mathcal{Z} \times \mathcal{Y}_2$  and  $\mathcal{X}_2 = \mathcal{Y}_1 \times \mathcal{Z}$ , we immediately get the following estimate of  $I_0(f, \widehat{f}, \widetilde{f})$ :

**Corollary 2.7.**

$$I_0(f, \widehat{f}, \widetilde{f}) \geq \left( \int f d\mu \right)^3 .$$

We can now resume the proof of Theorem 2.1:

*Proof of Theorem 2.1.* By Corollary 2.7, it suffices to prove that the set  $E := \{n: I_n(f, f, f) > I_0(f, \widehat{f}, \widetilde{f}) - \epsilon\}$  is syndetic.

Suppose by contradiction that  $E$  is not syndetic, then there exists a sequence of intervals  $[M_i, N_i)$  whose lengths tend to infinity such that  $I_n(f) \leq I_0(f, \widehat{f}, \widetilde{f}) - \epsilon$  for every  $i$  and for every  $n \in [M_i, N_i)$ . Taking averages of  $\chi(n\alpha)(I_0(f, \widehat{f}, \widetilde{f}) - I_n(f))$  over these intervals, we have a contradiction with Lemma 2.6.  $\square$

### 3. PRELIMINARIES: MAGIC SYSTEMS

In this Section,  $(X, \mathcal{X}, \mu, T_1, T_2)$  is a system. We review some definitions and properties of [12] needed in sequel, and establish more properties that do not appear there.

The results of this section hold for any number of commuting transformations, but we state and prove them only for two transformations.

**3.1. The box measure and the box seminorm.** Let  $X^* = X^4$ . The points of  $X^*$  are written as  $x^* = (x_{00}, x_{01}, x_{10}, x_{11})$ .

3.1.1. Let  $\mu_1 = \mu \times_{\mathcal{I}(T_1)} \mu$  be the relatively independent square of  $\mu$  over  $\mathcal{I}(T_1)$ . This means that  $\mu_1$  is the measure on  $X^2$  characterized by:

For  $f_0, f_1 \in L^\infty(\mu)$ , we have

$$\int f_0 \otimes f_1 d\mu_1 = \int \mathbb{E}(f_0 | \mathcal{I}(T_1)) \cdot \mathbb{E}(f_1 | \mathcal{I}(T_1)) d\mu .$$

Then  $\mu_1$  is invariant under the transformations  $T_1 \times T_1$ ,  $T_1 \times \text{Id}$  and  $T_2 \times T_2$ . We define the measure  $\mu^*$  on  $X^* = X^4$  by

$$\mu^* = \mu_1 \times_{\mathcal{I}(T_2 \times T_2)} \mu_1 .$$

When needed, we write  $\mu_{T_1, T_2}^*$  instead of  $\mu^*$ . Then  $\mu^*$  is invariant under the transformations  $T_1^*$  and  $T_2^*$  given by

$$T_1^* = T_1 \times \text{Id} \times T_1 \times \text{Id} ,$$

$$T_2^* = T_2 \times T_2 \times \text{Id} \times \text{Id} .$$

The projection  $\pi_{00}: x \mapsto x_{00}$  is a factor map from  $(X^*, \mu^*, T_1^*, T_2^*)$  to  $(X, \mu, T_1, T_2)$ .

We remark that by construction,

$$(8) \quad \mu_{T_1, T_2}^* = \mu_{T_1^{-1}, T_2}^* = \mu_{T_1, T_2^{-1}}^* .$$

3.1.2. We recall the definition of the seminorm introduced in [12]:

For every  $f \in L^\infty(\mu)$  on  $X$ ,

$$(9) \quad \|f\|_{\mu, T_1, T_2} = \left( \int f(x_{00})f(x_{01})f(x_{10})f(x_{11}) d\mu^*(x^*) \right)^{1/4}.$$

We generally omit the subscript  $\mu$  when there is no ambiguity. By (8), we have

$$\|f\|_{T_1, T_2} = \|f\|_{T_1^{-1}, T_2},$$

and by Proposition 2 of [12], we have

$$\|f\|_{T_1, T_2} = \|f\|_{T_2, T_1}.$$

Here are some properties that we need:

**Lemma 3.1.** *Let  $f_0, f_1, f_2 \in L^\infty(\mu)$  with  $|f_i| \leq 1$   $i = 0, 1, 2$ , and for every  $n$ , we write*

$$I_n(f_0, f_1, f_2) := \int f_0 \cdot T_1^n f_1 \cdot T_2^n f_2 d\mu.$$

Then for every  $t \in \mathbb{T}$ ,

$$(10) \quad \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} e(nt) I_n(f_0, f_1, f_2) \right| \leq \|f_0\|_{T_1, T_2}.$$

*Proof.* For  $t = 0$ , this is Proposition 1 of [12]. We rewrite the absolute value in (10) as

$$\left| \int f_1 \cdot \left( \frac{1}{N-M} \sum_{n \in [M, N]} e(nt) \cdot T_1^{-n} f_0 \cdot (T_2 T_1^{-1})^n f_2 \right) d\mu \right|.$$

By Cauchy-Schwartz Inequality, this expression is less than or equal to

$$\|f_1\|_{L^\infty(\mu)} \left\| \frac{1}{N-M} \sum_{n \in [M, N]} e(nt) \cdot T_1^{-n} f_0 \cdot (T_2 T_1^{-1})^n f_2 \right\|_{L^2(\mu)}.$$

By Van der Corput Lemma, the lim sup in (10) is bounded by

$$\limsup_{H \rightarrow \infty} \sum_{h=0}^{H-1} \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} e(nt) \cdot \overline{e((n+h)t)} \int T_1^{-n} (f_0 T_1^{-h} \bar{f}_0) \cdot (T_2 T_1^{-1})^n (f_2 (T_2 T_1^{-1})^h \bar{f}_2) d\mu \right|.$$

Since  $|e(ht)| = 1$ , the expression above is less than or equal to

$$\limsup_{H \rightarrow \infty} \sum_{h=0}^{H-1} \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} \int T_2^{-n} (f_0 T_1^{-h} \bar{f}_0) \cdot (f_2 (T_2 T_1^{-1})^h \bar{f}_2) d\mu \right|.$$

This expression does not depend on  $t$ . Therefore, from this point, the proof given in [12] for the case  $t = 0$  can be used without any modification.  $\square$

Let  $f_0, f_1, f_2$  and  $I_n(f_0, f_1, f_2)$  be as in Lemma 3.1. For every  $n$ , we have

$$I_n(f_0, f_1, f_2) = \int T_2^{-n} f_0 \cdot (T_1 T_2^{-1})^n f_1 \cdot f_2 d\mu .$$

Applying Lemma 3.1 to the pair of transformations  $(T_2, T_1 T_2^{-1})$ , we have

$$(11) \quad \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} e(nt) I_n(f_0, f_1, f_2) \right| \leq \|f_1\|_{T_2, T_1 T_2^{-1}} .$$

and by the same argument,

$$(12) \quad \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} e(nt) I_n(f_0, f_1, f_2) \right| \leq \|f_2\|_{T_1, T_1 T_2^{-1}} .$$

**Lemma 3.2** ([12], Lemma 1). *For every  $f \in L^\infty(\mu)$ ,*

$$\|f\|_{T_1, T_2} = \left( \lim_{N_2 \rightarrow +\infty} \frac{1}{N_2} \sum_{n_2=0}^{N_2-1} \left( \lim_{N_1 \rightarrow +\infty} \frac{1}{N_1} \sum_{n_1=0}^{N_1-1} \int f \cdot T_1^{n_1} f \cdot T_2^{n_2} f \cdot T_1^{n_1} T_2^{n_2} f d\mu \right) \right)^{1/4} .$$

In fact, we can take the simultaneous instead of iterated limit by Theorem 1.1 of [5].

By the Hölder Inequality, Lemma 3.2 implies that, for every  $f \in L^\infty(\mu)$ ,

$$(13) \quad \|f\|_{T_1, T_2} \leq \|f\|_{L^4(\mu)} .$$

Let  $\pi: (X, \mu, T_1, T_2) \rightarrow (Y, \nu, S_1, S_2)$  be a factor map, and  $f \in L^\infty(\nu)$ , by Lemma 3.2, we have

$$(14) \quad \|f \circ \pi\|_{\mu, T_1, T_2} = \|f\|_{\nu, S_1, S_2} .$$

Moreover, by applying Lemma 3.2 to the measure  $\mu$  and to each element of its ergodic decomposition, we get:

**Lemma 3.3.** *If  $\mu = \int \mu_x d\mu(x)$  is the ergodic decomposition of  $\mu$  under  $T_1$  and  $T_2$  then, for every  $f \in L^\infty(\mu)$ ,*

$$\|f\|_{\mu, T_1, T_2}^4 = \int \|f\|_{\mu_x, T_1, T_2}^4 d\mu(x) .$$

3.1.3. We recall some properties about some  $\sigma$ -algebra on  $X$  introduced in [12]:

**Lemma 3.4** ([12], Lemma 4). *Let  $\mathcal{G}$  be the  $\sigma$ -algebra on  $X$  consisting of sets  $B$  such that there exists a subset  $A$  of  $X^3$  satisfying the relation*

$$(15) \quad \mathbf{1}_B(x_{00}) = \mathbf{1}_A(x_{01}, x_{10}, x_{11})$$

for  $\mu^*$ -almost every  $x = (x_{00}, x_{01}, x_{10}, x_{11}) \in X^4$ . Then for every  $f \in L^\infty(\mu)$ , we have

$$(16) \quad \|f\|_{T_1, T_2} = 0 \text{ if and only if } \mathbb{E}_\mu(f | \mathcal{G}) = 0 .$$

**Lemma 3.5.**

$$\mathcal{G} \supseteq \mathcal{I}(T_1) \vee \mathcal{I}(T_2) .$$

*Proof.* Let  $f \in L^\infty(\mu)$ . If  $f$  is invariant under  $T_1$  then, by the construction of the measure  $\mu^*$ , we have  $f(x_{00}) = f(x_{10})$   $\mu^*$ -almost everywhere, and  $f$  is measurable with respect to  $\mathcal{G}$ . If  $f$  is invariant under  $T_2$  then, by a similar reason,  $f(x_{00}) = f(x_{01})$   $\mu^*$ -almost everywhere, and  $f$  is measurable with respect to  $\mathcal{G}$ . This proves the lemma.  $\square$

Combining Lemma 3.4 and Lemma 3.5, we immediately get:

**Corollary 3.6.** *Let  $(X, \mu, T_1, T_2)$  be a system, then for every  $f \in L^\infty(\mu)$ , we have*

$$\text{if } \|f\|_{T_1, T_2} = 0 \text{ then } \mathbb{E}(f | \mathcal{I}(T_1) \vee \mathcal{I}(T_2)) = 0 .$$

3.2. **Magic systems.** We recall the definition of the magic system:

**Definition 3.7** ([12]). A system  $(X, \mu, T_1, T_2)$  is called a *magic system* if for any  $f \in L^\infty(\mu)$  satisfying  $\mathbb{E}(f | \mathcal{I}(T_1) \vee \mathcal{I}(T_2)) = 0$ , we have  $\|f\|_{T_1, T_2} = 0$ .

**Corollary 3.8** (of Corollary 3.6). *If  $(X, \mu, T_1, T_2)$  is a magic system, then*

$$\|f\|_{T_1, T_2} = 0 \text{ if and only if } \mathbb{E}(f | \mathcal{I}(T_1) \vee \mathcal{I}(T_2)) = 0 .$$

The next Lemma plays an important role in Section 4.1.

**Lemma 3.9.** *Let  $p: (Y, \nu, S_1, S_2) \rightarrow (X, \mu, T_1, T_2)$  be a factor map and assume that  $(X, \mu, T_1, T_2)$  is magic. Then for every  $f \in L^\infty(\mu)$ ,*

$$\|f \circ p - \mathbb{E}(f \circ p | \mathcal{I}(S_1) \vee \mathcal{I}(S_2))\|_{S_1, S_2} = 0 .$$

*Proof.* We write  $f = f_1 + f_2$ , where  $f_1$  is  $\mathcal{I}(T_1) \vee \mathcal{I}(T_2)$ -measurable, and  $\mathbb{E}(f_2 | \mathcal{I}(T_1) \vee \mathcal{I}(T_2)) = 0$ . Since  $(X, \mu, T_1, T_2)$  is magic, we have  $\|f_2\|_{T_1, T_2} = 0$ . By (14),  $\|f_2 \circ p\|_{S_1, S_2} = 0$ . By Corollary 3.6,  $\mathbb{E}(f_2 \circ p | \mathcal{I}(S_1) \vee \mathcal{I}(S_2)) = 0$ . On the other hand,  $f_1 \circ p$  is measurable with respect to  $\mathcal{I}(S_1) \vee \mathcal{I}(S_2)$ . Therefore  $\mathbb{E}(f \circ p | \mathcal{I}(S_1) \vee \mathcal{I}(S_2)) = f_1 \circ p$  and  $\|f \circ p - \mathbb{E}(f \circ p | \mathcal{I}(S_1) \vee \mathcal{I}(S_2))\|_{S_1, S_2} = \|f_2 \circ p\|_{S_1, S_2} = 0$ .  $\square$

One of the main results of [12] is the following theorem:

**Theorem 3.10** (Theorem 1, [12]). *The system  $(X^*, \mu^*, T_1^*, T_2^*)$  built in Section 3.1 is magic.*

Since  $\pi_{00}: (X^*, \mu^*, T_1^*, T_2^*) \rightarrow (X, \mu, T_1, T_2)$  is a factor map, we have:

**Corollary 3.11.** *Every system  $(X, \mu, T_1, T_2)$  admits a magic extension.*

We can not use this result directly because we need to deal with ergodic systems, and the system  $(X^*, \mu^*, T_1^*, T_2^*)$  is never ergodic, except if  $X$  is trivial. Indeed, the projection  $x^* \mapsto x_{11}$  from  $X^*$  to  $X$  is invariant under  $T_1^*$  and  $T_2^*$ .

**3.3. The existence of ergodic magic extension.** In this Section we show:

**Theorem 3.12.** *Every ergodic system  $(X, \mathcal{X}, \mu, T_1, T_2)$  admits an ergodic magic extension.*

Recall that we assume that  $(X, \mathcal{X}, \mu)$  is a standard probability space. Then  $X$  admits a dense countably generated  $\sigma$ -algebra  $\mathcal{X}'$ . Substituting this algebra for  $\mathcal{X}$ , we reduce to the case that the  $\sigma$ -algebra  $\mathcal{X}$  is countably generated. In the sequel,

$$\mu = \int \mu_x d\mu(x)$$

denotes the ergodic decomposition of  $\mu$  under  $T_1$  and  $T_2$ .

Theorem 3.12 follows from the following proposition:

**Proposition 3.13.** *Let  $(X, \mathcal{X}, \mu, T_1, T_2)$  be a magic system. Then for  $\mu$ -almost every  $x$ , the system  $(X, \mathcal{X}, \mu_x, T_1, T_2)$  is magic.*

*Proof of Theorem 3.12 assuming Proposition 3.13.* By Corollary 3.11, there exist a magic system  $(X^*, \mu^*, T_1^*, T_2^*)$  and a factor map  $\pi: X^* \rightarrow X$ . Let  $\mu^* = \int \mu_x^* d\mu^*(x)$  be the ergodic decomposition of  $\mu^*$  under  $T_1^*$  and  $T_2^*$ . Then  $\mu = \int \pi(\mu_x^*) d\mu^*(x)$ . Since the measures  $\pi(\mu_x^*)$  are invariant under  $T_1$  and  $T_2$  and since  $\mu$  is ergodic, then for  $\mu^*$ -almost every  $x$  we have  $\pi(\mu_x^*) = \mu$ , and thus  $\pi: (X^*, \mu_x^*, T_1^*, T_2^*) \rightarrow (X, \mu, T_1, T_2)$  is a factor map.  $\square$

Before giving the proof of Proposition 3.13, we need the following two lemmas. We begin with the notation. We denote

$$\mathcal{A} = \mathcal{I}(T_1) \wedge \mathcal{I}(T_2) \text{ and } \mathcal{W} = \mathcal{I}(T_1) \vee \mathcal{I}(T_2) .$$

We recall that for  $\mu$ -almost every  $x$ , the measure  $\mu_x$  is invariant and ergodic under  $T_1$  and  $T_2$ . The map  $x \mapsto \mu_x$  is  $\mathcal{A}$ -measurable and for

every bounded function  $f$  on  $X$ ,

$$\mathbb{E}_\mu(f \mid \mathcal{A})(x) = \int f \, d\mu_x \text{ for } \mu\text{-almost every } x.$$

**Lemma 3.14.** *There exists a countable family  $\Phi$  of bounded  $\mathcal{W}$ -measurable functions on  $X$ , which is dense in  $L^p(\mathcal{W}, \nu)$  for every  $p \in [1, +\infty)$  and every probability measure  $\nu$  on  $(X, \mathcal{X})$  which is invariant under  $T_1$  and  $T_2$ .*

*Proof.* Let  $\mathcal{F} = \{f_n : n \in \mathbb{N}\}$  be a countable family of bounded  $\mathcal{X}$ -measurable functions on  $X$  which is dense in  $L^p(\nu)$  for every  $p \in [1, +\infty)$  and every probability measure  $\nu$  on  $(X, \mathcal{X})$ . For each  $n$ , define

$$g_n(x) = \begin{cases} \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=0}^{K-1} f_n(T_1^k x) & \text{if this limit exists;} \\ 0 & \text{otherwise.} \end{cases}$$

By ergodic theorem, the limit exists  $\nu$ -almost everywhere for every invariant measure  $\nu$ . Then each function  $g_n$  is bounded and invariant under  $T_1$ . For every  $n$ , let  $h_n$  be the function associated to  $T_2$  by the same construction.

Let  $\mathcal{F}_1 = \{g_n : n \in \mathbb{N}\}$  and  $\mathcal{F}_2 = \{h_n : n \in \mathbb{N}\}$ . Then, for  $i = 1, 2$ ,  $\mathcal{F}_i$  is dense in  $L^p(\mathcal{I}(T_i), \nu)$  for every  $p \in [1, +\infty)$  and every  $T_i$ -invariant probability measure  $\nu$  on  $(X, \mathcal{X})$ . Let  $\Phi$  be family of functions consisting in finite sums of functions of the form  $gh$  with  $g \in \mathcal{F}_1$  and  $h \in \mathcal{F}_2$ , then  $\Phi$  satisfies the lemma.  $\square$

**Lemma 3.15.** *Let  $f$  be a bounded function on  $X$  and let  $\tilde{f}$  be a  $\mathcal{W}$ -measurable representative of  $\mathbb{E}_\mu(f \mid \mathcal{W})$ . Then, for  $\mu$ -almost every  $x$ , we have*

$$\mathbb{E}_{\mu_x}(f \mid \mathcal{W}) = \tilde{f} \quad \mu_x\text{-almost everywhere.}$$

*Proof.* Let  $\Phi = \{\phi_n : n \in \mathbb{N}\}$  be the family of functions given by Lemma 3.14. For every  $n$ ,  $\mathbb{E}_\mu((f - \tilde{f})\phi_n \mid \mathcal{W}) = 0$  and, since  $\mathcal{A} \subset \mathcal{W}$ ,  $\mathbb{E}_\mu((f - \tilde{f})\phi_n \mid \mathcal{A}) = 0$   $\mu$ -almost everywhere. That is,  $\int (f - \tilde{f})\phi_n \, d\mu_x = 0$  for  $\mu$ -almost every  $x$ . Thus there exists a set  $A_n$  with  $\mu(A_n) = 1$  such that  $\int (f - \tilde{f})\phi_n \, d\mu_x = 0$  for every  $x \in A_n$ .

Let  $A$  be the intersection of the sets  $A_n$  for  $n \in \mathbb{N}$ . Then  $\mu(A) = 1$ . Let  $x \in A$ . We have  $\int (f - \tilde{f})\phi_n \, d\mu_x = 0$  for every  $n$  and, by density of the family  $\{\phi_n : n \in \mathbb{N}\}$  in  $L^1(\mathcal{W}, \mu_x)$ , we have  $\mathbb{E}_{\mu_x}(f - \tilde{f} \mid \mathcal{W}) = 0$ . Since  $\tilde{f}$  is measurable with respect to  $\mathcal{W}$ , we have  $\mathbb{E}_{\mu_x}(f \mid \mathcal{W}) = \tilde{f}$ .  $\square$

*Proof of Proposition 3.13.* Let  $\mathcal{F} = \{f_n : n \in \mathbb{N}\}$  be a countable family of bounded  $\mathcal{X}$ -measurable functions on  $X$  which is dense in  $L^p(\nu)$  for every  $p \in [1, +\infty)$  and every probability measure  $\nu$  on  $(X, \mathcal{X})$ .

For each  $n$ , let  $\tilde{f}_n$  be a  $\mathcal{W}$ -measurable representative of  $\mathbb{E}_\mu(f_n | \mathcal{W})$ . By Lemma 3.15, for every  $n$  we have that  $\tilde{f}_n = \mathbb{E}_{\mu_x}(f_n | \mathcal{W})$   $\mu_x$ -almost everywhere for  $\mu$ -almost every  $x$ , and there exists a set  $A_n$  such that  $\mu(A_n) = 1$  and,

$$\text{for every } x \in A_n, \quad \mathbb{E}_{\mu_x}(f_n | \mathcal{W}) = \tilde{f}_n \quad \mu_x\text{-a.e.}$$

Fix an integer  $n \in \mathbb{N}$ . Since  $\mathbb{E}_\mu(f_n - \tilde{f}_n | \mathcal{W}) = 0$  and  $(X, \mu, T_1, T_2)$  is magic, we have  $\|f_n - \tilde{f}_n\|_{\mu, T_1, T_2} = 0$ .

By Lemma 3.3, we have

$$0 = \|f_n - \tilde{f}_n\|_{\mu, T_1, T_2}^4 = \int \|f_n - \tilde{f}_n\|_{\mu_x, T_1, T_2}^4 d\mu(x).$$

So  $\|f_n - \tilde{f}_n\|_{\mu_x, T_1, T_2} = 0$  for  $\mu$ -almost every  $x$ . There exists a set  $B_n \subset X$  such that  $\mu(B_n) = 1$  and

$$\text{for every } x \in B_n, \quad \|f_n - \tilde{f}_n\|_{\mu_x, T_1, T_2} = 0.$$

Define

$$C = \bigcap_{n \in \mathbb{N}} (A_n \cap B_n).$$

We have  $\mu(C) = 1$ . We check that for every  $x \in C$  the system  $(X, \mu_x, T_1, T_2)$  is magic.

Let  $x \in C$  and  $f$  be a bounded measurable function on  $X$ , we have to show that  $\|f - \mathbb{E}_{\mu_x}(f | \mathcal{W})\|_{\mu_x, T_1, T_2} = 0$ .

Since the family  $\{f_n\}$  is dense in  $L^4(\mu_x)$ , there exists a sequence  $(n_i)$  of integers such that  $f_{n_i}$  converges to  $f$  in  $L^4(\mu_x)$ . By (13) applied to the measure  $\mu_x$ , we have

$$(17) \quad \|f_{n_i} - f\|_{\mu_x, T_1, T_2} \rightarrow 0.$$

On the other hand, since  $f_{n_i}$  converges to  $f$  in  $L^4(\mu_x)$ , we have that  $\mathbb{E}_{\mu_x}(f_{n_i} | \mathcal{W})$  converges to  $\mathbb{E}_{\mu_x}(f | \mathcal{W})$  in  $L^4(\mu_x)$  and, as above,

$$(18) \quad \|\mathbb{E}_{\mu_x}(f_{n_i} | \mathcal{W}) - \mathbb{E}_{\mu_x}(f | \mathcal{W})\|_{\mu_x, T_1, T_2} \rightarrow 0.$$

For every  $i$ , since  $x \in A_{n_i}$  we have that  $\mathbb{E}_{\mu_x}(f_{n_i} | \mathcal{W}) = \tilde{f}_{n_i}$  ( $\mu_x$ -a.e.). Since  $x \in B_{n_i}$ , we have  $\|f_{n_i} - \tilde{f}_{n_i}\|_{\mu_x, T_1, T_2} = 0$  and thus

$$(19) \quad \|f_{n_i} - \mathbb{E}_{\mu_x}(f_{n_i} | \mathcal{W})\|_{\mu_x, T_1, T_2} = 0.$$

For every  $i$ ,

$$\begin{aligned} & \|f - \mathbb{E}_{\mu_x}(f \mid \mathcal{W})\|_{\mu_x, T_1, T_2} \\ & \leq \|f_{n_i} - f\|_{\mu_x, T_1, T_2} + \|\mathbb{E}_{\mu_x}(f_{n_i} \mid \mathcal{W}) - \mathbb{E}_{\mu_x}(f \mid \mathcal{W})\|_{\mu_x, T_1, T_2} \\ & \quad + \|f_{n_i} - \mathbb{E}_{\mu_x}(f_{n_i} \mid \mathcal{W})\|_{\mu_x, T_1, T_2} . \end{aligned}$$

Combining (17), (18), (19), we get  $\|f - \mathbb{E}_{\mu_x}(f \mid \mathcal{W})\|_{\mu_x, T_1, T_2} = 0$  for every  $x \in C$ . This finishes the proof.  $\square$

#### 4. PROOF OF THEOREM 1.1

In this section, we prove:

**Theorem 4.1.** *Let  $(X, \mu, T_1, T_2)$  be an ergodic system and  $0 \leq f \leq 1$ . Then for every  $\epsilon > 0$ , the set*

$$\{n \in \mathbb{Z} : \int f \cdot T_1^n f \cdot T_2^n f d\mu > (\int f d\mu)^4 - \epsilon\}$$

*is syndetic.*

Theorem 1.1 follows from Theorem 4.1 with  $f = \mathbf{1}_A$ .

In this section, we fix the ergodic system  $(X, \mu, T_1, T_2)$ , and the function  $f$  as in Theorem 4.1, and for every  $n$ , we write

$$I_n := \int f \cdot T_1^n f \cdot T_2^n f d\mu .$$

**4.1. Construction of ergodic magic extensions.** We write  $T_3 = T_1 T_2^{-1}$ . We use Theorem 3.12 three times:

Let  $(X', \mu', T'_1, T'_2)$  be an ergodic magic extension of  $(X, \mu, T_1, T_2)$  and let  $\pi' : X' \rightarrow X$  be the factor map. Write  $T'_3 = T'_1 T'^{-1}_2$ . Then the system  $(X', \mu', T'_1, T'_3)$  is ergodic.

Let  $(X'', \mu'', T''_1, T''_3)$  be an ergodic magic extension of  $(X', \mu', T'_1, T'_3)$  and let  $\pi'' : X'' \rightarrow X'$  be the factor map. Write  $T''_2 = T''_1 T''^{-1}_3$ . Then the system  $(X'', \mu'', T''_2, T''_3)$  is ergodic.

Let  $(\bar{X}, \bar{\mu}, \bar{T}_2, \bar{T}_3)$  be an ergodic magic extension of  $(X'', \mu'', T''_2, T''_3)$  and let  $\bar{\pi} : \bar{X} \rightarrow X''$  be the factor map. Write  $\bar{T}_1 = \bar{T}_2 \bar{T}_3$ . Then the three systems  $(\bar{X}, \bar{\mu}, \bar{T}_2, \bar{T}_3)$ ,  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_3)$  and  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_2)$  are all ergodic, and  $\pi' \circ \pi'' \circ \bar{\pi}$  is a factor map from  $(\bar{X}, \bar{\mu}, \bar{T}_1, \bar{T}_2, \bar{T}_3)$  to  $(X, \mu, T_1, T_2, T_3)$ .

In the sequel, we write

$$\begin{aligned}
 h &=: f \circ \pi' \circ \pi'' \circ \bar{\pi} \\
 g_0 &=: \mathbb{E}(h \mid \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_2)) , \\
 g_1 &=: \mathbb{E}(h \mid \mathcal{I}(\bar{T}_1) \vee \mathcal{I}(\bar{T}_3)) , \\
 g_2 &=: \mathbb{E}(h \mid \mathcal{I}(\bar{T}_2) \vee \mathcal{I}(\bar{T}_3)) .
 \end{aligned}
 \tag{20}$$

For every  $n$ , and all bounded functions  $\eta, \varphi, \psi$  on  $\bar{X}$ , we define

$$\bar{I}_n(\eta, \varphi, \psi) := \int \eta \cdot \bar{T}_1^n \varphi \cdot \bar{T}_2^n \psi \, d\bar{\mu} .$$

In particular,

$$\bar{I}_n(h, h, h) = I_n . \tag{21}$$

We write also

$$J_n := \bar{I}_n(g_0, g_1, g_2) .$$

We have

**Proposition 4.2.** *Let  $t \in \mathbb{T}$ . Then the averages of the difference  $e(nt)(I_n - J_n)$  converge uniformly to 0.*

*Proof.* Since  $(\bar{X}, \bar{\mu}, \bar{T}_2, \bar{T}_3)$  is magic, by Definition 3.7 of magic systems and the definition of  $g_2$ ,

$$\|h - g_2\|_{\bar{T}_2, \bar{T}_3} = 0 .$$

Then by Lemma 3.1, the averages of the difference of

$$e(nt)(\bar{I}_n(h, h, h) - \bar{I}_n(h, h, g_2))$$

converge uniformly to 0. Since  $(X'', \mu'', T_1'', T_3'')$  is magic, by Lemma 3.9,

$$\|h - g_1\|_{\bar{T}_1, \bar{T}_3} = 0 .$$

Then by Lemma 3.1, the averages of the difference of

$$e(nt)(\bar{I}_n(h, h, g_2) - \bar{I}_n(h, g_1, g_2))$$

converge uniformly to 0. Since  $(X', \mu', T_1', T_2')$  is magic, by the same argument as above, the averages of the difference of

$$e(nt)(\bar{I}_n(h, g_1, g_2) - J_n)$$

converge uniformly to 0. Summing up all the results above, we get the announced result.  $\square$

In the sequel, we write

$$\overline{\mathcal{M}} := \mathcal{I}(\overline{T}_1) \vee \mathcal{I}(\overline{T}_2) \vee \mathcal{I}(\overline{T}_3) ,$$

and

$$g := \mathbb{E}(h \mid \overline{\mathcal{M}}) .$$

We remark that

$$\mathbb{E}(g \mid \mathcal{I}(\overline{T}_1) \vee \mathcal{I}(\overline{T}_2)) = g_0 ,$$

$$\mathbb{E}(g \mid \mathcal{I}(\overline{T}_1) \vee \mathcal{I}(\overline{T}_3)) = g_1 ,$$

$$\mathbb{E}(g \mid \mathcal{I}(\overline{T}_2) \vee \mathcal{I}(\overline{T}_3)) = g_2 .$$

Furthermore

$$(22) \quad \int g d\bar{\mu} = \int h d\bar{\mu} = \int f d\mu .$$

**4.2. Reduction to a special system.** For  $i = 1, 2, 3$ , let  $(Y_i, \mathcal{Y}_i, \nu_i)$  be a factor of  $\overline{X}$  which is isomorphic to  $(\overline{X}, \mathcal{I}(\overline{T}_i), \bar{\mu})$ . Denote by  $\pi_i : \overline{X} \rightarrow Y_i$  the factor map. Then  $\nu_i := \pi_i(\bar{\mu})$ .

Since for  $j = 1, 2, 3$ , every  $\sigma$ -algebra  $\mathcal{I}(\overline{T}_i)$  is invariant under the transformation  $\overline{T}_j$ , then each of these transformations induces a measure preserving transformation on each factor  $Y_i$ .

The transformation  $\overline{T}_1$  induces the identity on  $Y_1$  and the transformations  $\overline{T}_2$  and  $\overline{T}_3$  induce the same transformation on  $Y_1$ , which we write  $R_1$ . Therefore we have

$$R_1 \circ \pi_1 = \pi_1 \circ \overline{T}_2 = \pi_1 \circ \overline{T}_3 .$$

By ergodicity of  $(\overline{X}, \overline{\mathcal{X}}, \bar{\mu}, \overline{T}_1, \overline{T}_2)$ , the system  $(\overline{X}, \mathcal{I}(\overline{T}_1), \bar{\mu}, \overline{T}_2)$  is ergodic and thus  $(Y_1, \nu_1, R_1)$  is ergodic.

In the same way, we get a measure preserving transformation  $R_2$  on  $Y_2$  and a measure preserving transformation  $R_3$  on  $Y_3$  with

$$R_2 \circ \pi_2 = \pi_2 \circ \overline{T}_1 = \pi_2 \circ \overline{T}_3^{-1} ,$$

$$R_3 \circ \pi_3 = \pi_3 \circ \overline{T}_1 = \pi_3 \circ \overline{T}_2 .$$

As above, we have  $(Y_2, \nu_2, R_2)$  and  $(Y_3, \nu_3, R_3)$  are ergodic.

Let  $Y = Y_1 \times Y_2 \times Y_3$  and  $\mathcal{Y}$  be the product  $\sigma$ -algebra on  $Y$ . Define  $\pi : (\overline{X}, \overline{\mathcal{M}}) \rightarrow (Y, \mathcal{Y})$  by  $\pi = \pi_1 \times \pi_2 \times \pi_3$ . Let  $\nu$  be the image of  $\bar{\mu}$  under  $\pi$ . Define

$$(23) \quad \begin{aligned} S_1 &= \text{Id} \times R_2 \times R_3 , \\ S_2 &= R_1 \times \text{Id} \times R_3 , \\ S_3 &= R_1 \times R_2^{-1} \times \text{Id} . \end{aligned}$$

By construction,  $\pi$  is a factor map, so  $\pi^{-1}(\mathcal{Y}) \subseteq \overline{\mathcal{M}}$ . On the other hand,  $\pi^{-1}(\mathcal{Y})$  contains the  $\sigma$ -algebras  $\mathcal{I}(\overline{T}_1)$ ,  $\mathcal{I}(\overline{T}_2)$  and  $\mathcal{I}(\overline{T}_3)$ , so is equal to  $\overline{\mathcal{M}} = \mathcal{I}(\overline{T}_1) \vee \mathcal{I}(\overline{T}_2) \vee \mathcal{I}(\overline{T}_3)$ . We get

**Lemma 4.3.**  $\pi$  is an isomorphism from  $(\overline{X}, \overline{\mathcal{M}}, \overline{\mu}, \overline{T}_1, \overline{T}_2, \overline{T}_3)$  to  $(Y, \mathcal{Y}, \nu, S_1, S_2, S_3)$ .

We can thus identify these two systems and consider  $g$  as a function on  $Y$ . Then we have

$$\begin{aligned} g_0 &= \mathbb{E}(g \mid Y_1 \times Y_2) , \\ g_1 &= \mathbb{E}(g \mid Y_1 \times Y_3) , \\ g_2 &= \mathbb{E}(g \mid Y_2 \times Y_3) . \end{aligned}$$

By (22), we have

$$(24) \quad \int g d\nu = \int f d\mu .$$

**4.3. Description of the special system.** We recall the properties that will be used in the sequel:

- (a) For  $i = 1, 2, 3$ ,  $(Y_i, \mathcal{Y}_i, \nu_i, R_i)$  is an ergodic system.
- (b)  $Y = Y_1 \times Y_2 \times Y_3$  is endowed with the product  $\sigma$ -algebra  $\mathcal{Y}$ . The projection  $Y \rightarrow Y_i$  is written  $\pi_i$ .
- (c)  $\nu$  is a joining of  $\nu_1, \nu_2$  and  $\nu_3$ , meaning that, for  $i = 1, 2, 3$ , the projection of  $\nu$  on  $Y_i$  is equal to  $\nu_i$ . Moreover  $\nu$  is invariant under the transformations  $S_1, S_2, S_3$  defined in (23).
- (d) For  $i = 1, 2, 3$ ,  $\mathcal{Y}_i = \mathcal{I}(S_i)$ .
- (e)  $(Y, \mathcal{Y}, \nu, S_1, S_2)$  is ergodic.

**Remark.** For  $i \neq j$ ,  $Y$  is ergodic under  $S_i$  and  $S_j$ . By (d), the  $\sigma$ -algebra  $\mathcal{Y}_i$  and  $\mathcal{Y}_j$  are independent. In other words, the projection of  $\nu$  on  $Y_i \times Y_j$  is equal to  $\nu_i \times \nu_j$ .

The following proposition is a more precise description of  $\nu$ :

**Proposition 4.4.** Let  $(Z, \theta, R)$  be the common factor of the systems  $Y_1, Y_2, Y_3$  spanned by the eigenfunctions corresponding to the common eigenvalues of these three systems. Then for any  $\alpha_1 \in L^\infty(Y_1)$ ,  $\alpha_2 \in L^\infty(Y_2)$  and  $\alpha_3 \in L^\infty(Y_3)$ , we have

$$(25) \quad \int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = \int \mathbb{E}(\alpha_1 \mid Z)(z_1)\mathbb{E}(\alpha_2 \mid Z)(z_2)\mathbb{E}(\alpha_3 \mid Z)(z_3) dm$$

where  $m$  is the image of  $\nu$  on  $Z \times Z \times Z$ . In other words,  $\nu$  is the conditionally independent product measure over  $Z \times Z \times Z$ .

*Proof.* For  $i = 1, 2, 3$  let  $\mathcal{K}_i$  be the Kronecker factor of  $Y_i$ . Since  $\nu$  is invariant under  $S_1$ , we have

$$\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = \frac{1}{N} \sum_{n=1}^N \int \alpha_1(y_1) \cdot R_2^n \alpha_2(y_2) \cdot R_3^n \alpha_3(y_3) d\nu ,$$

for every  $N$ . Since  $Y_2, Y_3$  are independent, taking the limit as  $N \rightarrow \infty$ , we have

$$\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = \int \alpha_1(y_1) \cdot \mathbb{E}_{\nu_2 \times \nu_3}(\alpha_2 \otimes \alpha_3 \mid \mathcal{I}(R_2 \times R_3))(y_2, y_3) d\nu .$$

Note that the functions  $\alpha_2 \otimes \alpha_3$  and  $\mathbb{E}(\alpha_2 \mid \mathcal{K}_2) \otimes \mathbb{E}(\alpha_3 \mid \mathcal{K}_3)$  have the same conditional expectation on  $\mathcal{I}_{\nu_2 \times \nu_3}(R_2 \times R_3)$ . Therefore we have

$$\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = \int \alpha_1(y_1)\mathbb{E}(\alpha_2 \mid \mathcal{K}_2)(y_2)\mathbb{E}(\alpha_3 \mid \mathcal{K}_3)(y_3) d\nu .$$

By the same computation, substituting  $S_2$  for  $S_1$ , we obtain

(26)

$$\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = \int \mathbb{E}(\alpha_1 \mid \mathcal{K}_1)(y_1)\mathbb{E}(\alpha_2 \mid \mathcal{K}_2)(y_2)\mathbb{E}(\alpha_3 \mid \mathcal{K}_3)(y_3) d\nu .$$

On the other hand, for  $i = 1, 2, 3$ , the right hand side of (25) remains unchanged if  $\mathbb{E}(\alpha_i \mid \mathcal{K}_i)$  is substituted for  $\alpha_i$ . Therefore we can reduce to the case that  $\alpha_i$  is measurable with respect to  $\mathcal{K}_i$ . By density, we can suppose that  $\alpha_i$  is an eigenfunction of  $Y_i$  corresponding to the eigenvalue  $\lambda_i$ . By invariance of  $\nu$  under  $S_1$  again, we get

$$\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = e(\lambda_2 + \lambda_3) \int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu ,$$

thus  $\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = 0$  except if  $\lambda_2 = -\lambda_3$ .

By the same argument, we have that  $\int \alpha_1(y_1)\alpha_2(y_2)\alpha_3(y_3) d\nu = 0$  except if  $\lambda_1 = \lambda_2 = -\lambda_3$ . The announced conclusion follows.  $\square$

We have the following description of the measure  $m$ :

**Proposition 4.5.** *The measure  $m$  is the Haar measure on the compact subgroup  $H$  of  $Z \times Z \times Z$ , where*

$$H = \{(z_1, z_2, z_3), z_1 + z_2 - z_3 = 0\} .$$

*Proof.* Let  $\alpha \in Z$  be the element defining the translation  $R$ . Since  $\nu$  is invariant and ergodic under the transformations  $S_1$  and  $S_2$ , then the measure  $m$  is invariant and ergodic under the transformations  $\text{Id} \times R \times R$  and  $R \times \text{Id} \times R$ , that is under the translations by  $(0, \alpha, \alpha)$  and  $(\alpha, 0, \alpha)$ . The function  $z_1 + z_2 - z_3$  is invariant under these translations and by ergodicity it is equal to a constant  $c$ .

Since the map  $z \mapsto z - c$  is an isomorphism from  $(Z, \theta, R)$  to itself, we can thus reduce to the case that  $c = 0$  and  $m$  is concentrated on  $H$ . Since the subset  $\mathbb{Z}\alpha$  is dense in  $Z$ ,  $m$  is invariant under the translations by  $(0, z, z)$  and  $(z', 0, z')$  for any  $z, z' \in Z$ . Since these elements span  $H$ ,  $m$  is invariant under translation by any element of  $H$ , and thus  $m$  is the Haar measure on  $H$ .  $\square$

**4.4. Study of  $J_n$ .** In this section, let  $(Y, \nu, S_1, S_2)$  be as described in Section 4.3. We keep using the notation  $g, g_0, g_1, g_2$  and  $J_n = \bar{I}_n(g_0, g_1, g_2)$  introduced above. From this point, the proof is parallel to that in Section 2.2.

**Proposition 4.6.** *For every  $\epsilon > 0$ , there exists an open neighborhood  $U$  of 0 in  $Z$ , such that*

$$|J_0 - J_n| < \epsilon .$$

for every  $n$  such that  $n\alpha \in U$ .

*Proof.* More generally, we show for all bounded functions  $h_0$  on  $Y_1 \times Y_2$ ,  $h_1$  on  $Y_1 \times Y_3$ , and  $h_2$  on  $Y_2 \times Y_3$ , that there exists an open neighborhood  $U$  of 0 in  $Z$ , such that

$$\left| \int h_0 \cdot h_1 \cdot h_2 \, d\nu - \int h_0 \cdot S_1^n h_1 \cdot S_2^n h_2 \, d\nu \right| < \epsilon$$

for every  $n$  such that  $n\alpha \in U$ .

Without loss of generality, we can suppose that  $\max\{|h_0|, |h_1|, |h_2|\} \leq 1$ . We first suppose that  $h_0, h_1, h_2$  are product functions:  $h_0(y_1, y_2) = \alpha(y_1)\beta(y_2)$ ,  $h_1(y_1, y_3) = \alpha'(y_1)\gamma(y_3)$ , and  $h_2(y_2, y_3) = \beta'(y_2)\gamma'(y_3)$ , for some functions  $\alpha, \alpha'$  on  $Y_1$ ,  $\beta, \beta'$  on  $Y_2$ , and  $\gamma, \gamma'$  on  $Y_3$ , all of them being bounded by 1 in absolute value. Then, for every  $n$

$$\int h_0 \cdot S_1^n h_1 \cdot S_2^n h_2 \, d\nu = \int (\alpha\alpha')(y_1) \cdot (\beta\beta')(y_2) \cdot (\gamma\gamma')(R_3^n y_3) \, d\nu(y_1, y_2, y_3) .$$

By Proposition 4.4, this is equal to

$$\int \mathbb{E}(\alpha\alpha' \mid Z)(z_1) \mathbb{E}(\beta\beta' \mid Z)(z_2) \mathbb{E}(\gamma\gamma' \mid Z)(z_3 + n\alpha) \, dm(z_1, z_2, z_3) .$$

Set

$$U = \{t : \int |\mathbb{E}(\gamma\gamma' \mid Z)(z_3 + t) - \mathbb{E}(\gamma\gamma' \mid Z)(z_3)| \, d\theta(z_3) < \epsilon\},$$

then  $U$  is an open neighborhood of 0 in  $Z$  and  $|J_0 - J_n| < \epsilon$  when  $n\alpha \in U$ .

By linearity, the neighborhood  $U$  with the announced property exists when each  $h_i$ ,  $i = 1, 2, 3$ , is a finite sum of product functions of the type described above. The general case follows by a density argument.  $\square$

Let  $\epsilon > 0$  be as in Theorem 4.1, and let  $U$  be the neighborhood of 0 in  $Z$  defined by Proposition 4.6 with  $\epsilon/2$  substituted for  $\epsilon$ . Let  $\chi$  be a non-negative continuous function on  $Z$ , such that

$$\int \chi d\theta = 1 ,$$

and

$$\chi(z) = 0 \text{ if } z \notin U .$$

We keep the notation  $\epsilon, U, \chi$  in the sequel.

By Proposition 4.2 and a density argument, we have

**Corollary 4.7.** *The averages of the difference of*

$$\chi(n\alpha)(I_n - J_n)$$

*converge uniformly to zero.*

**Lemma 4.8.**

$$(27) \quad \limsup_{N-M \rightarrow \infty} \left| \frac{1}{N-M} \sum_{n \in [M, N]} \chi(n\alpha)(J_0 - J_n) \right| \leq \frac{\epsilon}{2} .$$

*Proof.* If  $n\alpha \notin U$ , then  $\chi(n\alpha)(J_0 - J_n) = 0$ . If  $n\alpha \in U$ , then by Proposition 4.6,

$$\chi(n\alpha)|J_0 - J_n| < \frac{\epsilon}{2} \cdot \chi(n\alpha) .$$

Since  $\{n\alpha\}$  is uniformly distributed in  $Z$ , the averages of  $\chi(n\alpha)$  converge uniformly to 1, this gives (27).  $\square$

**4.5. End of the proof.** We have the following estimate of  $J_0$ :

**Corollary 4.9** (of Lemma 1.6).

$$J_0 \geq \left( \int g d\nu \right)^4 = \left( \int f d\mu \right)^4 .$$

*Proof.* Since  $0 \leq g \leq 1$ , we have

$$J_0 = \int g_0 \cdot g_1 \cdot g_2 d\nu \geq \int g \cdot g_0 \cdot g_1 \cdot g_2 d\nu .$$

Applying Lemma 1.6 to the probability space  $(Y, \nu)$  with  $k = 3$ ,  $\mathcal{X}_1 = Y_1 \times Y_2$ ,  $\mathcal{X}_2 = Y_1 \times Y_3$  and  $\mathcal{X}_3 = Y_2 \times Y_3$ , the announced conclusion follows.  $\square$

We can now resume the proof of Theorem 4.1:

*Proof of Theorem 4.1.* By Corollary 4.9, it suffices to prove that the set  $E := \{n \in \mathbb{Z} : I_n > J_0 - \epsilon\}$  is syndetic. Suppose by contradiction that  $E$  is not syndetic. Then there exists a sequence of intervals  $[M_i, N_i]$  whose lengths  $N_i - M_i$  tend to infinity and such that  $I_n \leq J_0 - \epsilon$  for every  $i$  and every  $n \in [M_i, N_i]$ . Therefore

$$(28) \quad \frac{1}{N_i - M_i} \sum_{n \in [M_i, N_i]} \chi(n\alpha)(J_0 - I_n) \geq \frac{1}{N_i - M_i} \sum_{n \in [M_i, N_i]} \chi(n\alpha) \cdot \epsilon$$

for every  $i$ . Taking the lim sup of both sides, since the averages of  $\chi(n\alpha)$  converge uniformly to 1, we get that the lim sup of the left hand side of (28) is greater than or equal to  $\epsilon$ .

On the other hand, it follows from Corollary 4.7 and Lemma 4.8 that the lim sup of the left hand side of (28) is less than or equal to  $\epsilon/2$ . We have a contradiction.  $\square$

## 5. PROOF OF THEOREM 1.2

Let  $(Y, \nu, S)$  be the two-side  $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ -Bernoulli-shift. This means that  $\{0, 1, 2\}$  is endowed with the uniform probability measure  $\tau$ , and that  $Y = \{0, 1, 2\}^{\mathbb{Z}}$  is endowed with the product measure  $\nu = \tau^{\mathbb{Z}}$ . The points of  $Y$  are written  $y = (y_n : n \in \mathbb{Z})$  where  $y_n \in \{0, 1, 2\}$  for every  $n$ . The shift map  $S$  is given by  $(Sy)_n = y_{n+1}$  for every  $n$ .

Define

$$(X, \mu, T_1, T_2) = (Y \times Y \times Y, \nu \times \nu \times \nu, S \times \text{Id} \times S, \text{Id} \times S \times S).$$

Since  $(Y, \nu, S)$  is weakly mixing, the system  $(X, \mu, T_1, T_2)$  is ergodic.

For  $(i, j, k) \in \{0, 1, 2\}^3$ , let  $f(i, j, k)$  be defined by the following table:

	i=0	i=1	i=2
j=0	0 0 1	0 0 1	1 1 1
j=1	0 1 1	0 0 1	1 1 0
j=2	1 1 0	1 1 1	1 0 0
k=	0 1 2	0 1 2	0 1 2

Let  $F$  be the function on  $Y \times Y \times Y$  defined by  $F(y, z, w) = f(y_0, z_0, w_0)$ . Then  $F$  is an indicator function of some measurable subset  $A \subset X$ . For every  $n \neq 0$ , we have

$$\begin{aligned} \mu(A \cap T_1^n A \cap T_2^n A) &= \int F \cdot T_1^{-n} F \cdot T_2^{-n} F \, d\mu \\ &= \int F(y, z, w) \cdot F(S^{-n}y, z, S^{-n}w) \cdot F(y, S^{-n}z, S^{-n}w) \, d\mu(y, z, w) \\ &= \int f(y_0, z_0, w_0) \cdot f(y_{-n}, z_0, w_{-n}) \cdot f(y_0, z_{-n}, w_{-n}) \, d\mu(y, z, w). \end{aligned}$$

Since  $n \neq 0$ , the variables at the position 0 are independent with those at the position  $-n$ , we have

$$\begin{aligned} & \mu(A \cap T_1^n A \cap T_2^n A) \\ &= \int f(y, z, w') f(y', z, w) f(y, z', w) d\tau(y) d\tau(y') d\tau(z) d\tau(z') d\tau(w) d\tau(w') \\ &= \frac{1}{3^6} \sum_{i,j,k,i',j',k'} f(i, j, k') f(i, j', k) f(i, j', k) . \end{aligned}$$

By computation, for every integer  $n \neq 0$ , we have

$$\begin{aligned} \mu(A \cap T_1^n A \cap T_2^n A) &= \frac{145}{729} < 0.96 \left(\frac{16}{27}\right)^3 \\ &= 0.96 \left(\frac{1}{3^3} \sum_{i,j,k} f(i, j, k)\right)^3 = 0.96(\mu(A))^3 . \end{aligned}$$

Let  $0 < c \leq 1$ . There exists an integer  $l > 0$  such that  $(0.96)^l < c$ . Let  $(X', \mu', T'_1, T'_2)$  be the  $l$ -times product system of  $(X, \mu, T_1, T_2)$ , and  $A'$  be the  $l$ -times product of set  $A$ . Then the system  $(X', \mu', T'_1, T'_2)$  is ergodic and the set  $A'$  satisfies the announced property.

#### REFERENCES

1. T. Austin. On the norm convergence of nonconventional ergodic averages. To appear, *Ergodic Theory Dynam. Systems*.
2. F. V. Atkinson, G. A. Watterson and P. A. P. Moran. A matrix inequality. *Quart. J. Math. Oxford Ser. (2)* **11** (1960) 137–140.
3. V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.* **9** (1996), no. 3, 725–753.
4. V. Bergelson, B. Host and B. Kra, with an Appendix by I. Ruzsa. Multiple recurrence and nilsequences. *Inventiones Math.* **160** (2005), 261–303.
5. Q. Chu. Convergence of multiple ergodic averages along cubes for several commuting transformations. *Studia Math.* **196** (1) (2009), 13–22.
6. Q. Chu, N. Frantzikinakis and B. Host. Commuting averages with polynomial iterates of distinct degrees. *Preprint*.
7. N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.* **360** (2008), no. 10, 5435–5475.
8. N. Frantzikinakis and B. Kra. Ergodic averages for independent polynomials and applications. *J. London Math. Soc. (2)* **74** (2006), no. 1, 131–142.
9. H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31**, 204–256 (1977).
10. H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory*. M. B. Porter Lectures. Princeton University Press, Princeton, N.J., (1981).
11. H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.* **34** (1978), 275–291 (1979).
12. B. Host. Ergodic seminorms for commuting transformations and applications. *Studia Math.* **195** (1) (2009), 31–49.

13. B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.* **161** (2005), no. 1, 397-488.
14. A.Y. Khintchine. Eine Verschärfung des Poincaréschen “Wiederkehr-Satzes”. *Compos. Math.* **1**, (1934), 177-179.
15. A. Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.* **9** (1993), no. 2, 201-204.
16. A. Sidorenko. An analytic approach to extremal problems for graphs and hypergraphs. *Extremal problems for finite sets (Visegrád, 1991)*, 423-455, Bolyai Soc. Math. Stud., **3**, János Bolyai Math. Soc., Budapest, 1994.
17. T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems.* **28** (2008), 657-688.
18. H. Towsner. Convergence of diagonal ergodic averages. *Ergodic Theory Dynam. Systems* **29** (2009), no. 4, 1309-1326.

## Bibliographie

- [1] I. Assani. Pointwise convergence of ergodic averages along cubes. À paraître, *J. d'Analyse Math.*
- [2] I. Assani. Averages along cubes for not necessarily commuting measure preserving transformations. *Contemp. Math.* vol. **430**, 1–19, 2007.
- [3] T. Austin. On the norm convergence of nonconventional ergodic averages. À paraître, *Ergodic Theory Dynam. Systems.* arXiv :0805.0320v3.
- [4] T. Austin. Pleasant extensions retaining algebraic structure, I. *Preprint.* arXiv :0905.0518v3.
- [5] T. Austin. Pleasant extensions retaining algebraic structure, II. *Preprint.* arXiv :0910.0907v2.
- [6] T. Austin. Pleasant extensions retaining algebraic structure, III. *Preprint.* arXiv :0910.0909v2.
- [7] F. V. Atkinson, G. A. Watterson et P. A. P. Moran. A matrix inequality. *Quart. J. Math. Oxford Ser. (2)* **11** (1960) 137–140.
- [8] V. Bergelson. Weakly mixing PET. *Ergodic Theory Dynam. Systems* **7** (1987), no. 3, 337–349.
- [9] V. Bergelson. Ergodic Ramsey theory — an update. Ergodic theory of  $Z^d$  actions (Warwick, 1993–1994), 1–61, *London Math. Soc. Lecture Note Ser.*, **228**, Cambridge Univ. Press, Cambridge, 1996.
- [10] V. Bergelson. The multifarious Poincaré recurrence theorem. *Descriptive set theory and dynamical systems (Marseille-Luminy, 1996)*, 31–57, London Math. Soc. Lecture Note Ser., **277**, Cambridge Univ. Press, Cambridge, 2000.
- [11] V. Bergelson et A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.* **9** (1996), no. 3, 725–753.
- [12] V. Bergelson et R. McCutcheon. An ergodic IP polynomial Szemerédi theorem. *Mem. Amer. Math. Soc.* **146** (2000), no. 695.
- [13] V. Bergelson et A. Leibman. A nilpotent Roth theorem. *Invent. Math.* **147** (2002), no. 2, 429–470.
- [14] V. Bergelson, B. Host et B. Kra, with an Appendix by I. Ruzsa. Multiple recurrence and nilsequences. *Inventiones Math.* **160** (2005), 261–303.

- [15] J. Bourgain. Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. *Inst. Hautes Études Sci. Publ. Math.* No. **69** (1989), 5–45.
- [16] J. Bourgain. Double recurrence and almost sure convergence. *J. Reine Angew. Math.* **404** (1990), 140–161.
- [17] Q. Chu. Convergence of weighted polynomial multiple ergodic averages. *Proc. Amer. Math. Soc.* **137** (2009), no. 4, 1363–1369.
- [18] Q. Chu. Convergence of multiple ergodic averages along cubes for several commuting transformations. *Studia Math.* **196** (1) (2009), 13–22.
- [19] Q. Chu. Multiple recurrence for two commuting transformations. À paraître, *Ergodic Theory Dynam. Systems*.
- [20] Q. Chu, N. Frantzikinakis et B. Host. Commuting averages with polynomial iterates of distinct degrees. *Preprint*. [arXiv :0912.2641v1](https://arxiv.org/abs/0912.2641v1).
- [21] J. P. Conze et E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France* **112** (1984), no. 2, 143–175.
- [22] J. P. Conze et E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *Probabilités, 1-31, Publ. Inst. Rech. Math. Rennes, 1987-1, Univ. Rennes I, Rennes*, 1988.
- [23] J. P. Conze et E. Lesigne. Sur un théorème ergodique pour des mesures diagonales. *C. R. Acad. Sci. Paris Sér. I Math.* **306** (1988), no. 12, 491–493.
- [24] P. Erdős et T. Turán. On some sequences of integers. *J. London Math. Soc.* **11** (1936), 261–264.
- [25] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.* **360** (2008), no. 10, 5435–5475.
- [26] N. Frantzikinakis et R. McCutcheon. Ergodic Theory : Recurrence. [arXiv :0705.0033v2](https://arxiv.org/abs/0705.0033v2).
- [27] N. Frantzikinakis et B. Kra. Polynomial averages converge to the product of integrals. Probability in mathematics. *Israel J. Math.* **148** (2005), 267–276.
- [28] N. Frantzikinakis et B. Kra. Ergodic averages for independent polynomials and applications. *J. London Math. Soc.* (2) **74** (2006), no. 1, 131–142.
- [29] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.* **31**, 204–256 (1977).
- [30] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory*. M. B. Porter Lectures. Princeton University Press, Princeton, N.J., 1981.
- [31] H. Furstenberg et Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.* **34** (1978), 275–291 (1979).
- [32] H. Furstenberg et B. Weiss. A mean ergodic theorem for  $(1/N) \sum_{n=1}^N f(T^n x)g(T^{n^2} x)$ . *Convergence in ergodic theory and probability*

(Columbus, OH, 1993), 193–227, Ohio State Univ. Math. Res. Inst. Publ., 5, de Gruyter, Berlin, 1996.

- [33] H. Furstenberg, Y. Katznelson et B. Weiss. Ergodic theory and configurations in sets of positive density. *Mathematics of Ramsey theory*, 184–198, Algorithms Combin., **5**, Springer, Berlin, 1990.
- [34] W. T. Gowers. A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.* **11** (2001), no. 3, 465–588.
- [35] W. T. Gowers. Hypergraph regularity and the multidimensional Szemerédi theorem. *Ann. of Math. (2)* **166** (2007), no. 3, 897–946.
- [36] B. Green et T. Tao. Yet another proof of Szemerédi’s theorem. *Preprint*. arXiv :1002.2254v1.
- [37] B. Host. Convergence of multiple ergodic averages. *Information and randomness : cours donnés par le CIMPA ; école du C.M.M. de l’Université de Santiago de Chili*. Éditeurs A. Maass, S. Martinez, J. San Martin. Hermann, Paris, Travaux en cours **66** (2006). arXiv :math/0606362v1.
- [38] B. Host. Progressions arithmétiques dans les nombres premiers (d’après B. Green et T. Tao). Séminaire Bourbaki. Vol. 2004/2005. *Astérisque* No. **307** (2006), Exp. No. 944, viii, 229–246
- [39] B. Host. Ergodic seminorms for commuting transformations and applications. *Studia Math.* **195** (1) (2009), 31–49.
- [40] B. Host et B. Kra. Convergence of Conze-Lesigne averages. *Ergodic Theory Dynam. Systems* **21** (2001), no. 2, 493–509.
- [41] B. Host et B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.* **161** (2005), no. 1, 397–488.
- [42] B. Host et B. Kra. Convergence of polynomial ergodic averages. *Israel J. of Math.* **149** (2005), 1–20.
- [43] B. Host et B. Kra. Uniformity seminorms on  $\ell^\infty$  and applications. *J. Anal. Math.* **108** (2009), 219–276.
- [44] M. Johnson. Convergence of polynomial ergodic averages for some commuting transformations. À paraître, *Illinois J. Math.*
- [45] A.Y. Khintchine. Eine Verschärfung des Poincaréschen “Wiederkehr-Satzes”. *Compos. Math.* **1**, (1934), 177–179.
- [46] B. Kra. Ergodic methods in additive combinatorics. *Additive combinatorics*, 103–143, CRM Proc. Lecture Notes, **43**, Amer. Math. Soc., Providence, RI, 2007.
- [47] B. Kra. From combinatorics to ergodic theory and back again. *International Congress of Mathematicians*. Vol. III, 57–76, Eur. Math. Soc., Zürich, 2006.
- [48] L. Kuipers et H. Niederreiter. *Uniform distribution of sequences*. Pure and Applied Mathematics. Wiley-Interscience [John Wiley & Sons], New York-London-Sydney, 1974.

- [49] A. Leibman. Lower bounds for ergodic averages. *Ergodic Theory Dynam. Systems* **22** (2002), 863–872.
- [50] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequence of translation on a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 201–213.
- [51] A. Leibman. Pointwise convergence of ergodic averages for polynomial actions of  $\mathbb{Z}^d$  by translations on a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 215–225.
- [52] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables. *Israel J. Math.* **146** (2005), 303–315.
- [53] E. Lesigne. Un théorème de disjonction de systèmes dynamiques et une généralisation du théorème ergodique de Wiener-Wintner. *Ergodic Theory Dynam. Systems* **10** (1990), 513–521.
- [54] E. Lesigne. Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. *Ergodic Theory Dynam. Systems* **11** (1991), no. 2, 379–391.
- [55] E. Lesigne. Spectre quasi-discret et théorème ergodique de Wiener-Wintner pour les polynômes. *Ergodic Theory Dynam. Systems* **13** (1993), 767–784.
- [56] E. Lesigne. Équations fonctionnelles, couplages de produits gauches et théorèmes ergodiques pour mesures diagonales. *Bull. Soc. Math. France* **121** (1993), no. 3, 315–351.
- [57] B. Nagle, V. Rödl et M. Schacht. The counting lemma for regular  $k$ -uniform hypergraphs. *Random Structures Algorithms* **28** (2006), no. 2, 113–179.
- [58] V. Rödl et M. Schacht. Regularity lemma for  $k$ -uniform hypergraphs. *Random Structures Algorithms* **25** (2004), no. 1, 142.
- [59] A. Sárközy. On difference sets of sequences of integers. I. *Acta Math. Acad. Sci. Hungar.* **31** (1978), no. 1–2, 125–149.
- [60] E. Szemerédi. On sets of integers containing no  $k$  elements in arithmetic progression. *Acta Arith.* **27** (1975), 199–245.
- [61] A. Sidorenko. A correlation inequality for bipartite graphs. *Graphs Combin.* **9** (1993), no. 2, 201–204.
- [62] A. Sidorenko. An analytic approach to extremal problems for graphs and hypergraphs. *Extremal problems for finite sets (Visegrád, 1991)*, 423–455, Bolyai Soc. Math. Stud., **3**, János Bolyai Math. Soc., Budapest, 1994.
- [63] T. Tao. A variant of the hypergraph removal lemma. *J. Combin. Theory Ser. A* **113** (2006), no. 7, 1257–1280.
- [64] T. Tao. A quantitative ergodic theory proof of Szemerédi’s theorem. *Electron. J. Combin.* **13** (2006), no. 1, Research Paper 99, 49 pp. (electronic).

- [65] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems.* **28** (2008), 657–688.
- [66] J. P. Thouvenot. La démonstration de Furstenberg du théorème de Szemerédi sur les progressions arithmétiques. *Séminaire Bourbaki, 30e année (1977/78)*, Exp. No. 518, pp. 221–232, Lecture Notes in Math., 710, Springer, Berlin, 1979.
- [67] H. Towsner. Convergence of diagonal ergodic averages. *Ergodic Theory Dynam. Systems* **29** (2009), no. 4, 1309–1326.
- [68] N. Wiener and A. Wintner. *Harmonic analysis and ergodic theory*. Amer. J. Math. **63** (1941), 415–426.
- [69] Q. Zhang. On convergence of the averages  $\frac{1}{N} \sum_{n=1}^N f_1(R^n x) f_2(S^n x) f_3(T^n x)$ . *Monatsh. Math.* **122** (1996), 275–300.
- [70] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.* **20** (2007), no. 1, 53–97 (electronic).
- [71] T. Ziegler. Nilfactors of  $\mathbb{R}^m$ -actions and configurations in sets of positive upper density in  $\mathbb{R}^m$ . *J. Anal. Math.* **99** (2006), 249–266.

**NOM** : Qing CHU

---

**TITRE** : Quelques problèmes de convergence et de récurrence multiple en théorie ergodique

---

**RÉSUMÉ**

Cette thèse est consacrée à l'étude de certaines questions de convergence et de récurrence multiple en théorie ergodique.

Nous distinguons les systèmes munis d'une transformation et ceux munis de plusieurs transformations qui commutent. Dans les premiers, le mécanisme de facteurs caractéristiques et les nilsystèmes jouent un rôle important. À l'aide de ces outils, nous étendons les résultats sur la convergence de moyennes ergodiques multiples pondérées de Host et Kra pour le cas linéaire au cas polynomial. En conséquence, nous montrons que pour toute fonction  $f$  mesurable bornée sur un système ergodique, la suite  $(f(T^n x))$  est universellement bonne pour presque tout  $x$ .

Quand il y a plusieurs transformations qui commutent, à l'aide de la machinerie des systèmes magiques introduite récemment par Host et développée dans cette thèse, nous étendons les résultats sur la convergence de moyennes ergodiques multiples sur les cubes de Host et Kra avec une transformation à plusieurs transformations qui commutent. Nous obtenons aussi un résultat de récurrence multiple quantitatif pour deux transformations qui commutent, similaire en faveur du cas d'une transformation établi par Bergelson, Host et Kra, mais avec une conclusion différente.

---

**MOTS CLÉS** : Théorie ergodique, moyennes ergodiques multiples, récurrence multiple, transformations qui commutent, facteur caractéristique, nilsystèmes, systèmes magiques.

---

**TITLE** : Some problems of multiple convergence and recurrence in ergodic theory

---

**ABSTRACT**

This thesis is devoted to the study of some questions of multiple convergence and recurrence in ergodic theory.

We distinguish between systems endowed with a single transformation and systems endowed with several commuting transformations. In the former, characteristic factors and nilsystems play an important role in the study of multiple convergence and recurrence. Using these tools, we extend results on convergence of weighted multiple ergodic averages of Host and Kra for the linear case to the polynomial case. As a consequence, we show that for any bounded measurable function  $f$  on an ergodic system, the sequence  $f(T^n x)$  is universally good for almost every  $x$ .

In systems endowed with several commuting transformations, we use the machinery of magic systems introduced recently by Host and further properties of magic systems developed in this thesis, to extend results of Host and Kra on convergence of multiple ergodic averages along cubes with a single transformation to commuting transformations. We obtain a quantitative multiple recurrence result for two commuting transformations, similar in flavour to that of a single transformation established by Bergelson, Host and Kra, but with a different conclusion.

---

**KEY WORDS** : Ergodic theory, multiple ergodic averages, multiple recurrence, commuting transformations, characteristic factor, nilsystems, magic systems.

---

**CLASSIFICATION AMS** : [2000] 37A05, 37A30

---

Laboratoire d'Analyse et de Mathématiques Appliquées de l'Université Paris-Est-Marne-la-Vallée