Cubic averages and large intersections

V. Bergelson and A. Leibman

January 22, 2018

Abstract

We discuss the nature of the phenomenon of "large limits along cubic averages" and obtain some rather general new results. We also prove a general version of Furstenberg's correspondence principle that applies to uncountable amenable discrete groups and utilize this to derive combinatorial corollaries of ergodic results involving large limits.

0. Introduction

Let $(X,\mathcal{B},\mu,T)$ be an invertible probability measure preserving system. The classical Poincaré recurrence theorem (see [Po], Theorem I) states that for any $A\in\mathcal{B}$ with $\mu(A)>0$ there exists $n\in\mathbb{N}$ such that $\mu(A\cap T^nA)>0$. For mixing systems one has $\lim_{n\to\infty}\mu(A\cap T^nA)=\mu(A)^2$, and it is natural to ask whether for any $(X,\mathcal{B},\mu,T)$, any $A\in\mathcal{B}$, and any $\delta>0$ one can find $n\neq 0$ such that

$$\mu(A\cap T^nA)\geq\mu(A)^2-\delta,$$

and, if so, how "large" the set

$$R_\delta(A)=\bigl\{n\in\mathbb{Z}:\mu(A\cap T^nA)\geq\mu(A)^2-\delta\bigr\}$$

can be. These questions were addressed in [Kh], where the following result, often quoted as "Khintchine's recurrence theorem", was obtained.

Theorem 0.1. *For any invertible probability measure preserving system $(X,\mathcal{B},\mu,T)$, any $A\in\mathcal{B}$, and any $\delta>0$, the set $R_\delta(A)$ is syndetic.*[(1)]

(We are taking the liberty to formulate Khintchine's result for powers of a single measure preserving transformation. Khintchine's paper dealt with continuous measure preserving flows.)

In [Kh], Theorem 0.1 was derived from the following stronger result. (Again, we formulate it for $\mathbb{Z}$- rather than for $\mathbb{R}$-actions.)

Partially supported by NSF grants DMS-0901106 and DMS-1162073.

(1) A set S in a topological group G is said to be *(left) syndetic* if there exists a compact set $F\subseteq G$ such that $G=FS$. If G is a discrete group, F has to be finite.

Theorem 0.2. *For any invertible probability measure preserving system* $(X, \mathcal{B}, \mu, T)$ *and any* $A \in \mathcal{B}$*, one has*

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu(A \cap T^n A) \geq \mu(A)^2.$$

By iterating Khintchine's theorem, one immediately observes that for any $n \in R_\delta(A)$ there exists a syndetic set of $m \in \mathbb{Z}$ such that

$$\mu\big((A \cap T^n A) \cap T^m(A \cap T^n A)\big) = \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A) > \mu(A)^4 - 2\delta. \tag{0.1}$$

This leads to the natural question of whether, for any $\delta > 0$, the set of pairs (n, m) satisfying (0.1) is syndetic in $\mathbb{Z}^2$. The affirmative answer follows from the following theorem.

Theorem 0.3. ([B1]) *For any invertible probability measure preserving system* $(X, \mathcal{B}, \mu, T)$ *and any* $A \in \mathcal{B}$*, one has*

$$\lim_{N_1-M_1, N_2-M_2\to\infty} \frac{1}{(N_1-M_1)(N_2-M_2)} \sum_{\substack{M_1\leq n<N_1\\ M_2\leq m<N_2}} \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A) \geq \mu(A)^4.$$

Corollary 0.4. *For any invertible probability measure preserving system* $(X, \mathcal{B}, \mu, T)$*, any* $A \in \mathcal{B}$*, and any* $\delta > 0$*, the set*

$$\big\{(n, m) \in \mathbb{Z}^2 : \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A) > \mu(A)^4 - \delta\big\}$$

is syndetic.

Theorem 0.3 was extended in [HK1] and [HK2] to multiparameter expressions of the form $\mu\big(\bigcap_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} T^{\varepsilon_1 n_1+\ldots+\varepsilon_k n_k} A\big)$, $n = (n_1, \ldots, n_k) \in \mathbb{Z}^k$, which in turn implies, for any $\delta > 0$, the syndeticity of the set

$$\Big\{n \in \mathbb{Z}^k : \mu\Big(\bigcap_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} T^{\varepsilon_1 n_1+\ldots+\varepsilon_k n_k} A\Big) > \mu(A)^{2^k} - \delta\Big\}.$$

The proofs of these results in [HK1] and [HK2] utilize in a rather crucial way knowledge about *characteristic factors* responsible for the limiting behavior of the expressions $\frac{1}{(N_1-M_1)\cdots(N_k-M_k)} \sum_{\substack{M_1\leq n_1<N_1\\ \vdots \\ M_k\leq n_k<N_k}} \prod_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} T^{\varepsilon_1 n_1+\ldots+\varepsilon_k n_k} f$ [(2)], and it is of interest to attempt to establish the results on large intersections in the situations where there is (so far) no sufficient knowledge on the structure of the corresponding characteristic factors. In particular, one would like to know if, given commuting invertible measure preserving

---

[(2)] A factor Y of X is *characteristic* if the limit, as $N_i - M_i \to \infty$, of the expressions under consideration depends not on f, but only on $E(f|Y)$ – the conditional expectation of f with respect to Y (the orthogonal projection of f on $L^2(Y) \subseteq L^2(X)$).

transformations $T_1, \ldots, T_k$ of a probability space $(X, \mathcal{B}, \mu)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$, the set

$$R_\delta^{(k)}(A) = \Big\{(n_1, \ldots, n_k) \in \mathbb{Z}^k : \mu\Big(\bigcap_{\varepsilon_1,\ldots,\varepsilon_k \in \{0,1\}} T_1^{\varepsilon_1 n_1} \ldots T_k^{\varepsilon_k n_k} A\Big) > \mu(A)^{2^k} - \delta\Big\} \tag{0.2}$$

is syndetic.

The goal of this short paper is to show that the phenomenon of large limits along cubic configurations has a very general scope and hinges on a simple "Fubini principle" (Lemma 1.1 below), which, roughly, states that the limits of uniform multiparameter Cesàro averages can be replaced by iterated limits of uniform Cesàro averages. In what follows we will establish a general scheme that can be viewed as a machine for obtaining various results on large limits of cubic averages. This machine has an input and an output. The input consists of two convergence statements:
(i) a general multiparameter limiting theorem that stipulates the existence of certain uniform Cesàro limits (but does not require any description of these limits);
(ii) an ergodic limiting result pertaining to single or multiple recurrence that guarantees a "large" uniform Cesàro limit. An example of such a result is given by Theorem 0.2. (We will provide below a few more examples of various degrees of generality.)
Then the machine's output is a theorem on large limits of cubic averages. An example of such a theorem is Theorem 0.3, but, as we will see, significantly more general theorems hold true as well.

To help the reader to get a feeling about the general results which we obtain in this paper, let us give a sketch of the proof of the following fact, which is a rather special case of Theorem 0.8 below.

Theorem 0.5. *Let $T_1, \ldots, T_k$ be invertible transformations of a probability measure space $(X, \mathcal{B}, \mu)$, which commute (or, more generally, generate a nilpotent group). Then for any $A \in \mathcal{B}$,*

$$\lim_{N_1-M_1,\ldots,N_k-M_k\to\infty} \frac{1}{(N_1-M_1)\cdots(N_k-M_k)} \sum_{\substack{M_1\le n_1<N_1\\ \vdots \\ M_k\le n_k<N_k}} \mu\Big(\bigcap_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} T_1^{\varepsilon_1 n_1}\cdots T_k^{\varepsilon_k n_k} A\Big) \ge \mu(A)^{2^k},$$

and hence the set $R_\delta^{(k)}(A)$ defined in (0.2) is syndetic.

Proof sketch. The input statement of the form (i) that is needed for the proof is the following fact (see [W], Theorem 5.1):

- *For any $A \in \mathcal{B}$, the limit*

$$\lim_{N_1-M_1,\ldots,N_k-M_k\to\infty} \frac{1}{(N_1-M_1)\cdots(N_k-M_k)} \sum_{\substack{M_1\le n_1<N_1\\ \vdots \\ M_k\le n_k<N_k}} \mu\Big(\bigcap_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} T_1^{\varepsilon_1 n_1}\cdots T_k^{\varepsilon_k n_k} A\Big)$$

exists.

The input statement of type (ii) is Theorem 0.2 above:

- *For any* $A\in\mathcal{B}$, $\lim_{N-M\to\infty}\frac{1}{N-M}\sum_{n=M}^{N-1}\mu(A\cap T^nA)\geq\mu(A)^2$.

Take $k=2$. Replacing the uniform double Cesàro limit by the iterated uniform Cesàro limits (this is an instance of the Fubini principle alluded to above), we have:

$$\lim_{N_1-M_1,N_2-M_2\to\infty}\frac{1}{(N_1-M_1)(N_2-M_2)}\sum_{\substack{M_1\leq n<N_1\\ M_2\leq m<N_2}}\mu\big(A\cap T_1^{n_1}A\cap T_2^{n_2}A\cap T_1^{n_1}T_2^{n_2}A\big)$$
$$=\lim_{N_1-M_1\to\infty}\frac{1}{N_1-M_1}\sum_{n_1=M_1}^{N_1-1}\lim_{N_2-M_2\to\infty}\frac{1}{N_2-M_2}\sum_{n_2=M_2}^{N_2-1}\mu\big((A\cap T_1^{n_1}A)\cap T_2^{n_2}(A\cap T_1^{n_1}A)\big)$$
$$\geq\lim_{N_1-M_1\to\infty}\frac{1}{N_1-M_1}\sum_{n_1=M_1}^{N_1-1}\mu(A\cap T_1^{n_1}A)^2\geq\Big(\lim_{N_1-M_1\to\infty}\frac{1}{N_1-M_1}\sum_{n_1=M_1}^{N_1-1}\mu(A\cap T_1^{n_1}A)\Big)^2$$
$$\geq(\mu(A)^2)^2.$$

Inductively repeating this argument for $k=3$, 4, …, gives us the desired result.

In order to formulate our results in proper generality we have to introduce some notation. For a (left) amenable group G with left Haar measure τ and a mapping $g\mapsto v_g$, $g\in G$, from G to a Banach space V, let us write $\text{UC-lim}_{g\in G}v_g=v$ if for every left Følner net (Φ_α) [(3)] in G one has
$$\lim_{\alpha\to\infty}\frac{1}{\tau(\Phi_\alpha)}\int_{\Phi_\alpha}v_g\,d\tau(g)=v.$$
(If G is σ-compact (that is, is a countable union of compact sets), Følner nets can be replaced by Følner sequences; see [P].)

In what follows, let $(X,\mathcal{B},\mu)$ be a probability measure space. We will prefer to deal with functions from $L^2(X)$ instead of subsets of X; to obtain results about sets $A\in\mathcal{B}$ one takes $f=1_A$. Notice also that to get a left action of a non-commutative group G on $L^2(X)$ we have to assume that G acts on X from the right; of course, this does not affect the results. (In Section 3 we pass back to a left action of G on X.) In Section 2 we will prove the following:

Theorem 0.6. *Let $G_1,\dots,G_k$ be locally compact amenable groups and, for each $i=1,\dots,k$, let $T_{i,1},\dots,T_{i,r_i}$ be (not necessarily homomorphic or continuous) mappings from G_i to the set of measure preserving transformations of X. Assume that the following two conditions are satisfied:*
(i) *for any collection $f_{j_1,\dots,j_k}$, $1\leq j_1\leq r_1$, …, $1\leq j_k\leq r_k$, of functions from $L^\infty(X)$, the limit*
$$\underset{(g_1,\dots,g_k)\in G_1\times\dots\times G_k}{\text{UC-lim}}\int_X\prod_{\substack{1\leq j_1\leq r_1\\ \vdots\\ 1\leq j_k\leq r_k}}\Big(T_{1,j_1}(g_1)\cdots T_{k,j_k}(g_k)f_{j_1,\dots,j_k}\Big)d\mu$$

(3) A *left Følner net* in a group G is a family $(\Phi_\alpha)_{\alpha\in A}$ of compact non-null subsets of G, indexed by a directed set A, such that $\lim_\alpha\tau(\Phi_\alpha\triangle(g\Phi_\alpha))/\tau(\Phi_\alpha)=0$.

exists;
(ii) *for any non-negative function* $f \in L^2(X)$ *and any* $i \in \{1, \ldots, k\}$,

$$\underset{g_i \in G_i}{\text{UC-lim}} \int_X \prod_{j_i=1}^{r_i} \Big(T_{i,j_i}(g_i) f \Big) d\mu \geq \Big(\int_X f \, d\mu \Big)^{r_i}.$$

Then for any non-negative function $f \in L^\infty(X)$,
(a)

$$\underset{(g_1,\ldots,g_k) \in G_1 \times \cdots \times G_k}{\text{UC-lim}} \int_X \prod_{\substack{1 \leq j_1 \leq r_1 \\ \vdots \\ 1 \leq j_k \leq r_k}} \Big(T_{1,j_1}(g_1) \cdots T_{k,j_k}(g_k) f \Big) d\mu \geq \Big(\int_X f \, d\mu \Big)^{r_1 \cdots r_k};$$

(b) *for any* $\delta > 0$, *the set*

$$\Big\{ (g_1, \ldots, g_k) \in G_1 \times \cdots \times G_k : \int_X \prod_{\substack{1 \leq j_1 \leq r_1 \\ \vdots \\ 1 \leq j_k \leq r_k}} \Big(T_{1,j_1}(g_1) \cdots T_{k,j_k}(g_k) f \Big) d\mu > \Big(\int_X f \, d\mu \Big)^{r_1 \cdots r_k} - \delta \Big\}$$

is syndetic in $G_1 \times \cdots \times G_k$.

In order to apply this result we only need to verify, in concrete situations, whether conditions (i) and (ii) are satisfied. A very general theorem guaranteeing the fulfillment of condition (i) is the following:

Theorem 0.7. ([Z]) *Let* G *be a locally compact amenable group and let* $S_1, \ldots, S_k$ *be continuous polynomial mappings*[(4)] *from* G *to a nilpotent group of measure preserving transformations of* X. *Then, for any* $f_1, \ldots, f_l \in L^\infty(X)$, $\text{UC-lim}_{g \in G} \prod_{i=1}^{l} S_i(g) f_i$ *exists in* $L^2(X)$.

This theorem will allow us to obtain two kinds of applications. First, we will get a general result (Theorem 0.8 below) on convergence and recurrence. One can then use a variant of Furstenberg's corresponcence principle for combinatorial actions (see [BF]) to obtain some combinatorial applications. There is however a much more interesting vista that Theorem 0.7 opens up. Namely, it applies, in particular, to discrete uncountable amenable groups, which – via an appropriate generalization of Furstenberg's correspondence principle – will allow us to obtain somewhat unexpected combinatorial applications (see Section 3).

As for condition (ii), it holds, for example, in the following situations:

(4) A mapping P from a group G to a (nilpotent) group H is said to be *polynomial* if there is $d \in \mathbb{N}$ such that for any $h_1, \ldots, h_d \in G$ one has $D_{h_1} \cdots D_{h_d} P = 1$, where the "derivative" D_h is defined by $(D_h P)(g) = P(g)^{-1} P(gh)$.

(1) Let T be a measure preserving action on X of a locally compact amenable group G. For any $f \in L^2(X)$, by the classical von Neumann ergodic theorem, adapted to limits along Følner nets, one has $\text{UC-lim}_{g\in G} T_g f = Pf$, where P is the orthogonal projection onto the space $\{f \in L^2(X) : T_g f = f \text{ for all } g \in G\}$. For $f \geq 0$ we have $Pf \geq 0$ and $\int_X Pf\, d\mu = \int_X f\, d\mu$. Hence,

$$\underset{g\in G}{\text{UC-lim}} \int_X f{\cdot}T_g f\, d\mu = \int_X f{\cdot}Pf\, d\mu = \int_X Pf{\cdot}Pf\, d\mu = \|Pf\|^2 \geq \Big(\int_X Pf\, d\mu\Big)^2 = \Big(\int_X f\, d\mu\Big)^2.$$

Thus, condition (ii) is satisfied for $T_{i,1} = \text{Id}$ and $T_{i,2} = T$. This gives us the following result:

Theorem 0.8. *Let $G_1, \ldots, G_k$ be locally compact amenable groups and let T_i, $i = 1, \ldots, k$, be (continuous) homomorphisms from G_i to a nilpotent group of measure preserving transformations of X. Then for any non-negative function $f \in L^\infty(X)$,*
(a)

$$\underset{(g_1,\ldots,g_k)\in G_1\times\cdots\times G_k}{\text{UC-lim}} \int_X \prod_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} \Big(T_1(g_1)^{\varepsilon_1}\cdots T_k(g_k)^{\varepsilon_k} f\Big) d\mu \geq \Big(\int_X f\, d\mu\Big)^{2^k};$$

(b) *for any $\delta > 0$, the set*

$$\Big\{(g_1,\ldots,g_k) \in G_1\times\cdots\times G_k : \int_X \prod_{\varepsilon_1,\ldots,\varepsilon_k\in\{0,1\}} \Big(T_1(g_1)^{\varepsilon_1}\cdots T_k(g_k)^{\varepsilon_k} f\Big) d\mu > \Big(\int_X f\, d\mu\Big)^{2^k} - \delta\Big\}$$

is syndetic in $G_1 \times \cdots \times G_k$.

(2) Let F be a (discrete) field and let V be a finite dimensional vector space over F. We call a mapping T from V to a group $\mathcal{T}$ *polynomial* if $T = S \circ P$, where P is a polynomial (in the usual sense) mapping from V to a finite dimensional vector space W over F and S is a homomorphism from W to $\mathcal{T}$. The following theorem, extending a result of Larick ([La]), can be found in [BLM]:

Theorem 0.9. *Let V be a finite dimensional vector space and let U be a polynomial unitary action of V on a Hilbert space $\mathcal{H}$. Then for any $f \in \mathcal{H}$,* $\text{UC-lim}_{v\in V} U(v)f$ *exists and is the orthogonal projection of f on the space of $U(V)$-invariant vectors.*

(In [BLM], this theorem is stated and proved for the case where the base field F is countable, but it remains true, with the same proof, when F is uncountable.)

It follows that if T is a polynomial measure preserving action on X of a finite dimensional vector space V, then for any non-negative $f \in L^2(X)$, $\text{UC-lim}_{v\in V} \int_X f \cdot T(v) f\, d\mu \geq \big(\int_X f\, d\mu\big)^2$. Based on this fact, Theorem 0.6 acquires the following form:

Theorem 0.10. *Let $V_1,\dots,V_k$ be finite dimensional vector spaces over a field F and let T_i, $i=1,\dots,k$, be polynomial mappings from V_i to a nilpotent group of measure preserving transformations of X. Then for any non-negative function $f\in L^\infty(X)$,*
(a)

$$\operatorname*{UC-lim}_{(v_1,\dots,v_k)\in V_1\oplus\dots\oplus V_k}\int_X \prod_{\varepsilon_1,\dots,\varepsilon_k\in\{0,1\}}\Big(T_1(v_1)^{\varepsilon_1}\cdots T_k(v_k)^{\varepsilon_k}f\Big)d\mu \ge \Big(\int_X f\,d\mu\Big)^{2^k};$$

(b) *for any $\delta>0$, the set*

$$\Big\{(u_1,\dots,u_k)\in V_1\oplus\dots\oplus V_k :\\ \int_X \prod_{\varepsilon_1,\dots,\varepsilon_k\in\{0,1\}}\Big(T_1(v_1)^{\varepsilon_1}\cdots T_k(v_k)^{\varepsilon_k}f\Big)d\mu > \Big(\int_X f\,d\mu\Big)^{2^k}-\delta\Big\}$$

is syndetic in the (discrete) group $V_1\oplus\dots\oplus V_k$.

(3) Let us say that polynomials $p_1,\dots,p_r$ are *essentially linearly independent* if the polynomials $p_1-p_1(0),\dots,p_r-p_r(0)$ are linearly independent. The following theorem can be found in [FK]:

Theorem 0.11. *Let T be a totally ergodic measure preserving transformation of X and let $p_1,\dots,p_r$ be integer valued, essentially linearly independent polynomials of one or several integer variables. Then for any $f_1,\dots,f_r\in L^\infty(X)$, $\operatorname{UC-lim}_{n\in\mathbb{Z}}\prod_{j=1}^r T^{p_j(n)}f_j = \prod_{j=1}^r\int_X f_j\,d\mu$.*

(In [FK], this theorem was stated and proved for the case of polynomials of one variable only, but the same argument works in the case of polynomials of several variables as well.)

Via Theorem 0.6, this result implies the following:

Theorem 0.12. *Let $T_1,\dots,T_k$ be totally ergodic measure preserving transformation of X generating a nilpotent group, and let for each $i=1,\dots,k$, $p_{i,j}$, $j=1,\dots,r_k$, be essentially linearly independent polynomials $\mathbb{Z}^{d_i}\longrightarrow\mathbb{Z}$. Then for any non-negative function $f\in L^\infty(X)$,*
(a)

$$\operatorname*{UC-lim}_{(n_1,\dots,n_k)\in\mathbb{Z}^{d_1}\times\dots\times\mathbb{Z}^{d_k}}\int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots\\ 1\le j_k\le r_k}}\Big(T_1^{p_{1,j_1}(n_1)}\cdots T_k^{p_{k,j_k}(n_k)}f\Big)d\mu \ge \Big(\int_X f\,d\mu\Big)^{r_1\cdots r_k};$$

(b) *for any $\delta>0$, the set*

$$\Big\{(n_1,\dots,n_k)\in\mathbb{Z}^{d_1}\times\dots\times\mathbb{Z}^{d_k} :\\ \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots\\ 1\le j_k\le r_k}}\Big(T_1^{p_{1,j_1}(n_1)}\cdots T_k^{p_{k,j_k}(n_k)}f\Big)d\mu > \Big(\int_X f\,d\mu\Big)^{r_1\cdots r_k}-\delta\Big\}$$

is syndetic in $\mathbb{Z}^{d_1+\dots+d_k}$.

(4) Theorem 0.11, stated for $\mathbb{Z}$-actions, implies a similar result for $\mathbb{R}$-actions (see [BLMo]). Moreover, for $\mathbb{R}$-flows, the total ergodicity is no longer an issue ($\mathbb{R}$-flows are totally ergodic if ergodic, and the ergodic decomposition allows us to give up the condition of ergodicity as well). We therefore also get the following theorem:

Theorem 0.13. *Let $(T_1)_{t\in\mathbb{R}}, \dots, (T_k)_{t\in\mathbb{R}}$ be flows of measure preserving transformations of X generating a nilpotent group, and let, for each $i = 1, \dots, k$, $p_{i,j}$, $j = 1, \dots, r_k$, be essentially linearly independent polynomials $\mathbb{R}^{d_i} \longrightarrow \mathbb{R}$. Then for any non-negative function $f \in L^\infty(X)$,*
(a)

$$\operatorname*{UC-lim}_{(t_1,\dots,t_k)\in\mathbb{R}^{d_1}\times\cdots\times\mathbb{R}^{d_k}} \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots \\ 1\le j_k\le r_k}} \Big(T_1^{p_{1,j_1}(t_1)} \cdots T_k^{p_{k,j_k}(t_k)} f\Big)\, d\mu \ge \Big(\int_X f\, d\mu\Big)^{r_1\cdots r_k};$$

(b) *for any $\delta > 0$, the set*

$$\Big\{(t_1,\dots,t_k) \in \mathbb{R}^{d_1}\times\cdots\times\mathbb{R}^{d_k} : \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots \\ 1\le j_k\le r_k}} \Big(T_1^{p_{1,j_1}(t_1)} \cdots T_k^{p_{k,j_k}(t_k)} f\Big)\, d\mu > \Big(\int_X f\, d\mu\Big)^{r_1\cdots r_k} - \delta\Big\}$$

is syndetic in (the topological) group $\mathbb{R}^{d_1+\cdots+d_k}$.

The fundamental nature of cubic averages is also manifested by the fact that the multiple recurrence results, such as Theorem 0.8, lead to new sharp combinatorial applications involving large sets in uncountable amenable groups. For example, one has the following result (see Theorem 3.2 in Section 3 for a more general statement):

Theorem 0.14. *Let $d, k \in \mathbb{N}$ and let m be an invariant mean*[(5)] *on the group $\mathbb{R}^d$, considered as a discrete group. Then for any $E \subseteq \mathbb{R}^k$ with $m(1_E) > 0$ and any $\delta > 0$, the set*

$$R_\delta = \Big\{(u_1,\dots,u_k) \in (\mathbb{R}^d)^k : m\Big(\bigcap_{\varepsilon_1,\dots,\varepsilon_k\in\{0,1\}} \Big(E - (\varepsilon_1 u_1 + \cdots + \varepsilon_k u_k)\Big)\Big) > m(E)^{2^k} - \delta\Big\}$$

is syndetic in (the discrete) group $(\mathbb{R}^d)^k$ (meaning that finitely many shifts of R_δ cover $(\mathbb{R}^d)^k$).

---

(5) *A left-invariant mean* is a linear functional m on the space $BC(G)$ of bounded continuous real-valued functions on G such that $m(1) = 1$, $m(f) \ge 0$ if $f \ge 0$, and $m(f(tx)) = m(f(x))$ for all f and all $t \in G$.

In order to establish this result one needs a variant of Furstenberg's correspondence principle that holds for general, possibly uncountable, discrete amenable groups. (See Theorem 3.1 in Section 3.)

The structure of the rest of the paper is as follows. In Section 1 we state and prove the Fubini principle alluded to above. In Section 2 we prove Theorem 0.8 and obtain its corollaries. Section 3 is devoted to establishing a general form of Furstenberg's correspondence principle and utilizing it to derive some combinatorial applications.

1. The Fubini principle – double and repeated Cesàro limits

Lemma 1.1. *Let G, H be amenable groups and let $(h,g) \mapsto v_{h,g}$, $(h,g) \in H \times G$, be a bounded continuous mapping from $H \times G$ to a Banach space V. Assume that* $\text{UC-lim}_{(h,g)\in H\times G} v_{h,g}$ *exists and for every $g \in G$,* $\text{UC-lim}_{h\in H} v_{h,g}$ *exists; then*

$$\underset{g\in G}{\text{UC-lim}}\ \underset{h\in H}{\text{UC-lim}}\, v_{h,g} = \underset{(h,g)\in H\times G}{\text{UC-lim}} v_{h,g}.$$

Proof. Let $v = \text{UC-lim}_{(h,g)\in H\times G} v_{h,g}$, and for every $g \in G$, let $v_g = \text{UC-lim}_{h\in H} v_{h,g}$. Let (Φ_α) be a Følner net in G and (Ψ_β) be a Følner net in H. Then $(\Psi_\beta \times \Phi_\alpha)$ is a Følner net in $H \times G$ (where by $(\beta_1, \alpha_1) < (\beta_2, \alpha_2)$ we mean $\beta_1 < \beta_2$ and $\alpha_1 < \alpha_2$). Let $\delta > 0$. Let τ, τ' be left Haar measures on G and H respectively, and let $\tau'' = \tau' \times \tau$. Find β_0, α_0 such that

$$\frac{1}{\tau''(\Psi_\beta \times \Phi_\alpha)} \int_{\Psi_\beta\times\Phi_\alpha} v_{h,g}\, d\tau''(h,g) \overset{\delta}{\approx} v$$

whenever $\beta > \beta_0$ and $\alpha > \alpha_0$. It then follows that for any $\alpha > \alpha_0$,

$$\begin{aligned}\frac{1}{\tau(\Phi_\alpha)} \int_{\Phi_\alpha} v_g\, d\tau(g) &= \frac{1}{\tau(\Phi_\alpha)} \int_{\Phi_\alpha} \lim_\beta \frac{1}{\tau(\Psi_\beta)} \Big(\int_{\Psi_\beta} v_{h,g}\, d\tau'(h)\Big) d\tau(g)\\ &= \lim_\beta \frac{1}{\tau''(\Psi_\beta\times\Phi_\alpha)} \int_{\Psi_\beta\times\Phi_\alpha} v_{h,g}\, d\tau''(h,g) \overset{\delta}{\approx} v.\end{aligned}$$

■

For σ-compact amenable groups one has a version of Lemma 1.1 that involves Følner sequences instead of Følner nets. For such a group G and a mapping $g \mapsto v_g$, $g \in G$, from G to a Banach space V, we have $\text{UC-lim}_{g\in G} v_g = v$ iff $\lim_{N\to\infty} \frac{1}{\tau(\Phi_n)} \int_{\Phi_n} v_g\, d\tau(g) = v$ for every Følner sequence (Φ_n) in G.

Lemma 1.2. *Let G, H be σ-compact amenable groups and let $v_{h,g}$, $(h,g) \in H \times G$, be a continuous mapping from $H \times G$ to a Banach space V. Assume that* $\text{UC-lim}_{(h,g)\in H\times G} v_{h,g}$ *exists and for every $h \in H$,* $\text{UC-lim}_{g\in G} v_{h,g}$ *exists; then*

$$\underset{h\in H}{\text{UC-lim}}\ \underset{g\in G}{\text{UC-lim}}\, v_{h,g} = \underset{(h,g)\in H\times G}{\text{UC-lim}} v_{h,g}.$$

Proof. Let $v = \text{UC-lim}_{(h,g)\in H\times G} v_{h,g}$, and for every $g \in G$, let $v_g = \text{UC-lim}_{h\in H} x_{h,g}$. Let (Φ_n) be any Følner sequence in G and (Ψ_m) be a Følner sequence in H. Let τ, τ' be left Haar measures on G and H respectively, and let $\tau'' = \tau' \times \tau$. Choose an increasing sequence of integers (m_n) such that for every n,

$$\Big\|\frac{1}{\tau'(\Psi_{m_n})}\int_{\Phi_{m_n}} v_{h,g}\,d\tau'(h) - v_g\Big\| < \frac{1}{n}$$

for all $g \in \Phi_n$. Then for any n,

$$\Big\|\frac{1}{\tau''(\Psi_{m_n}\times\Phi_n)}\int_{\Psi_{m_n}\times\Phi_n} v_{h,g}\,d\tau''(h,g) - \frac{1}{\tau(\Psi_n)}\int_{\Phi_n} v_g\,d\tau(g)\Big\| < \frac{1}{n}.$$

$(\Psi_{m_n}\times\Phi_n)$ is a Følner sequence in $H\times G$, so $\frac{1}{\tau''(\Psi_{m_n}\times\Phi_n)}\int_{\Psi_{m_n}\times\Phi_n} v_{h,g}\,d\tau''(h,g) \longrightarrow v$ as $n\longrightarrow\infty$. Thus, $\frac{1}{\tau(\Phi_n)}\int_{\Phi_n} v_g\,d\tau(g)\longrightarrow v$. $\blacksquare$

To deduce, from the largeness of uniform Cesàro limits, the syndeticity of the "sets of large intersection" (part (b) of Theorem 0.6), we will need the following fact:

Lemma 1.3. *Let G be a locally compact amenable group and let a_g, $g\in G$, be a mapping $G\longrightarrow\mathbb{R}$ with* $\text{UC-lim}_{g\in G} a_g = a$. *Then for any $\delta>0$ the set $R=\big\{g\in G : a_g > a-\delta\big\}$ is syndetic in G.*

Proof. Choose a Følner net (Φ_α) in G. If R is not syndetic, then there is a net (g_α) in G such that for any α, $g_\alpha\Phi_\alpha\cap R=\emptyset$. But $(g_\alpha\Phi_\alpha)$ is also a Følner net in G, and for this net we have $\limsup_{\alpha\to\infty}\frac{1}{|g_\alpha\Phi_\alpha|}\sum_{g\in\Phi_\alpha} a_g \le a-\delta$, which contradicts $\text{UC-lim}_{g\in G} a_g = a$. $\blacksquare$

2. Large cubic averages: a proof of Theorem 0.6

Proof of Theorem 0.6. Part (b) of the theorem follows from part (a) and Lemma 1.3, so we only have to prove (a). Notice that, putting $f_{j_1,\ldots,j_{k-1},1}=1$ and $f_{j_1,\ldots,j_{k-1},j_k} = f_{j_1,\ldots,j_{k-1}}$ for $j_k\ge 2$ for all $j_1,\ldots,j_{k-1}$, we have from (i) that the limit

$$\underset{(g_1,\ldots,g_{k-1})\in G_1\times\cdots\times G_{k-1}}{\text{UC-lim}}\int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots\\ 1\le j_{k-1}\le r_{k-1}}}\Big(T_{1,j_1}(g_1)\cdots T_{k-1,j_{k-1}}(g_{k-1})f_{j_1,\ldots,j_{k-1}}\Big)d\mu$$

exists for any collection $f_{j_1,\ldots,j_{k-1}}$ of functions from $L^\infty(X)$.

Let $f\in L^\infty(X)$, $f\ge 0$. For any $g_1\in G_1$, $\ldots$, $g_k\in G_k$, we can rewrite

$$\int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots\\ 1\le j_k\le r_k}}\Big(T_{1,j_1}(g_1)\cdots T_{k-1,j_{k-1}}(g_{k-1})T_{k,j_k}(g_k)f\Big)d\mu$$
$$= \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots\\ 1\le j_{k-1}\le r_{k-1}}}\Big(T_{1,j_1}(g_1)\cdots T_{k-1,j_{k-1}}(g_{k-1})\Big(\prod_{j_k=1}^{r_k}\big(T_{k,j_k}(g_k)f\big)\Big)\Big)d\mu.$$

By assumption (i), the limit

$$\underset{(g_1,\dots,g_k)\in G_1\times\cdots\times G_k}{\text{UC-lim}} \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots \\ 1\le j_k\le r_k}} \Big(T_{1,j_1}(g_1)\cdots T_{k,j_k}(g_k)f\Big)d\mu$$

exists, also for any $g_k \in G_k$ the limit

$$\underset{(g_1,\dots,g_{k-1})\in G_1\times\cdots\times G_{k-1}}{\text{UC-lim}} \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots \\ 1\le j_{k-1}\le r_{k-1}}} \Big(T_{1,j_1}(g_1)\cdots T_{k-1,j_{k-1}}(g_{k-1})\Big(\prod_{j_k=1}^{r_k}\big(T_{k,j_k}(g_k)f\big)\Big)\Big)d\mu$$

exists, and, by induction on k, is $\ge \big(\int_X \prod_{j_k=1}^{r_k}\big(T_{k,j_k}(g_k)f\big)\,d\mu\big)^{r_1\cdots r_{k-1}}$. Thus Lemma 1.1 applies and, combined with Hölder's inequality and assumption (ii), implies

$$\begin{aligned}
&\underset{(g_1,\dots,g_k)\in G_1\times\cdots\times G_k}{\text{UC-lim}} \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots \\ 1\le j_k\le r_k}} \Big(T_{1,j_1}(g_1)\cdots T_{k,j_k}(g_k)f\Big)d\mu\\
&\underset{(g_1,\dots,g_{k-1})\in G_1\times\cdots\times G_{k-1}}{\text{UC-lim}} \int_X \prod_{\substack{1\le j_1\le r_1\\ \vdots \\ 1\le j_{k-1}\le r_{k-1}}} \Big(T_{1,j_1}(g_1)\cdots T_{k-1,j_{k-1}}(g_{k-1})\Big(\prod_{j_k=1}^{r_k}\big(T_{k,j_k}(g_k)f\big)\Big)\Big)d\mu\\
&\ge \underset{g_k\in G_k}{\text{UC-lim}}\Big(\int_X \prod_{j_k=1}^{r_k}\big(T_{k,j_k}(g_k)f\big)\,d\mu\Big)^{r_1\cdots r_{k-1}} \ge \Big(\underset{g_k\in G_k}{\text{UC-lim}}\int_X \prod_{j_k=1}^{r_k}\big(T_{k,j_k}(g_k)f\big)\,d\mu\Big)^{r_1\cdots r_{k-1}}\\
&\ge \Big(\int_X f\,d\mu\Big)^{r_1\cdots r_k}.
\end{aligned}$$

■

3. Furstenberg's correspondence principle for general amenable groups and applications

When G is a countable amenable group, one can get combinatorial corollaries of Theorem 0.6 by invoking a version of Furstenberg's correspondence principle for amenable groups (see [B3], Theorem 6.4.17). For general amenable locally compact groups, a variant of correspondence principle was obtained in [BF]. This variant allows us to obtain combinatorial corollaries of multiple recurrence results for *continuous* actions and guarantees existence of combinatorial patterns only in properly dilated large sets in topological amenable groups (see the details in [BF], Section 1). Since some of the results obtained in Theorem 0.6 hold true for actions of discrete uncountable groups, it is natural to inquire whether there exists a properly general version of Furstenberg's correspondence principle which would guarantee the existence (and abundance) of "cubic" patterns in any large set of, say, $\mathbb{R}^d$, considered with the discrete topology. The goal of this section is to establish such a general principle and to derive some of its corollaries. To make the discussion precise, one has, of course, to define first what is meant by a "large" set.

If G is a countable amenable group, the standard way of defining the basic notion of largness for a set $E \subseteq G$ is to declare E to be large if it has positive upper density $\bar{d}_{(\Phi_n)}(E) := \limsup_{n\to\infty} |E \cap \Phi_n|/|\Phi_n|$ with respect to some Følner sequence (Φ_n) in G. It is not hard to show that $\bar{d}_{(\Phi_n)}(E) > 0$ if and only if there exists an invariant mean on the space $B(G)$ of bounded real-valued functions on G such that $m(1_E) > 0$. If G is an uncountable discrete amenable group, one still has at ones disposal both approaches to defining the notion of largeness. However, a word of caution is in order here. Namely, in uncountable discrete groups one no longer has the luxury of working with Følner sequences and has to switch to Følner nets in order to define the notion of upper dentisty and establish its translation invariance (see [HS]). After that, in full analogy with the case of countable amenable groups, one can show that a set $E \subseteq G$ has positive upper density if and only if for some invariant mean m on $B(G)$ one has $m(E) > 0$. (When convenient, we will write $m(E)$ for $m(1_E)$.) We prefer to work with the invariant means from the outset. The following version of Furstenberg's correspondence principle has the more familiar correspondence principle for countable amenable groups ([B3]) as a special case.

Theorem 3.1. *Let G be a discrete amenable group. Let m be a left-invariant mean on $B(G)$. Let $E \subseteq G$, $m(E) > 0$. Then there exists a probability measure preserving system $(X, \mathcal{B}, \mu, (T_g)_{g\in G})$, where X is a compact space, $\mathcal{B}$ is the Borel σ-algebra on X, and $(T_g)_{g\in G}$ is an action of G on X by homeomorphisms, and a set $A \in \mathcal{B}$ with $\mu(A) = m(E)$, such that for any $k \in \mathbb{N}$ and any $g_1, \ldots, g_k \in G$ one has*

$$m\big(E \cap g_1^{-1}E \cap \ldots \cap g_k^{-1}E\big) = \mu\big(A \cap T_{g_1}^{-1}A \cap \ldots \cap T_{g_k}^{-1}A\big). \tag{3.1}$$

Proof. The perspicacious reader has already guessed that we will set X to be βG, the Stone-Čech compactification of G. Let μ be the unique probability measure on the Borel σ-algebra of βG corresponding to m. The correspondence is implemented by the formula $m(f) = \int_{\beta G} \hat{f}\, d\mu$, where $f \in B(G)$ and $\hat{f}$ denotes the continuous extension of f to βG. For any set $E \subseteq G$, let $\overline{E}$ be the closure if $E \in \beta G$. Sets of the form $\overline{E}$, $E \subseteq G$, are closed and open and form the basis of open sets in βG. For any $g \in G$, the map $h \mapsto gh$, $h \in G$, has a unique continuous extension to βG, which we will denote by T_g. The maps T_g, $g \in G$, are μ-preserving homeomorphisms of βG. (The μ-invariance follows from the invariance of the mean m.) Now, for any $f_0, f_1, \ldots, f_k \in B(G)$ on has

$$m\Big(\prod_{i=0}^{k} f_i\Big) = \int_{\beta G} \prod_{i=0}^{k} \hat{f}_i\, d\mu.$$

Applying this to $f_0 = 1_E$ and $f_i = 1_{g_i^{-1}E}$, $i = 1, \ldots, k$, we get (3.1). ∎

The following combinatorial result immediately follows from Theorem 3.1 and Theorem 0.8:

Theorem 3.2. *Let G be a (discrete) nilpotent group and let m be an invariant mean on $B(G)$. If $E \subseteq G$ satisfies $m(E) > 0$, then, for any $k \in \mathbb{N}$ and $\delta > 0$, the set*

$$\Big\{(g_1, \ldots, g_k) \in G^k : m\Big(\bigcap_{\varepsilon_1,\ldots,\varepsilon_k \in \{0,1\}} (g_1^{\varepsilon_1} \cdots g_k^{\varepsilon_k})^{-1} E\Big) > m(E)^{2^k} - \delta\Big\}$$

is syndetic in G^k (that is, finitely many translates of this set cover G^k).

Theorem 0.10, in its turn, implies the following combinatorial fact:

Theorem 3.3. *Let $V_1, \dots, V_k$ be finite dimensional vectors spaces over a (discrete) field F, let $P_1, \dots, P_k$ be polynomial mappings from V_i to a finite dimensional vector space W over F, let m be an invariant mean on $B(W)$, and let $E \subseteq W$ satisfies $m(E) > 0$. Then for any $\delta > 0$, the set*

$$\Big\{(u_1, \dots, u_k) \in V_1 \oplus \dots \oplus V_k : m\Big(\bigcap_{\varepsilon_1, \dots, \varepsilon_k \in \{0,1\}} \big(\varepsilon_1 P_1(u_1) + \dots + \varepsilon_k P_k(u_k)\big)E\Big) > m(E)^{2^k} - \delta\Big\}$$

is syndetic in the group $V_1 \oplus \dots \oplus V_k$.

Remark. A more familiar form of Furstenberg's correspondence principle deals with large sets in countable groups (see, for example, [B2], [B3]), and in this case one can guarantee that the resulting measure preserving system $(X, \mathcal{B}, \mu, (T_g)_{g \in G})$ is regular, meaning that X is a compact metric space and $\mathcal{B}$ is the Borel σ-algebra on X. The regularity of $(X, \mathcal{B}, \mu, (T_g)_{g \in G})$ plays an instrumental role in ergodic proofs of Szemerédi's theorem and its extensions (see [Fu1], [FuKa], [BL], [BM]). To quote from [Fu2], p.103: "For certain of the constructions to be carried out it will be necessary to choose between equivalent measure spaces, confining ones attention to *regular* spaces.". The goal of the next proposition is to show that when G is countable, one can replace the measure preserving system $(X, \mathcal{B}, \mu, (T_g)_{g \in G})$, appearing in the proof of Theorem 3.1, by a regular one.

Proposition 3.4. *Let $(T_g)_{g \in G}$ be a measure preserving action of a countable group G on a probability space $(X, \mathcal{B}, \mu)$ and let $A \in \mathcal{B}$. Then there exists a regular probabilty measure preserving system $(\widetilde{X}, \widetilde{\mathcal{B}}, \tilde{\mu}, (\widetilde{T}_g)_{g \in G})$ and a set $\widetilde{A} \in \mathcal{B}$ such that, for any $k \in \mathbb{N}$ and any $g_1, \dots, g_k \in G$, one has*

$$\tilde{\mu}\big(\widetilde{A} \cap \widetilde{T}_{g_1}^{-1} \widetilde{A} \cap \dots \cap \widetilde{T}_{g_k}^{-1} \widetilde{A}\big) = \mu\big(A \cap T_{g_1}^{-1} A \cap \dots \cap T_{g_k}^{-1} A\big).$$

Proof. Let $f = 1_A$ and let $\mathcal{A}$ be the closure in $L^\infty(X, \mathcal{B}, \mu)$ of the algebra generated by the functions $T_g f$, $g \in G$, and their complex conjugates. $\mathcal{A}$ is a separable commutative C^*-algebra, and so, by Gelfand's theorem, there exists a compact metric space $\widetilde{X}$ such that $\mathcal{A}$ is isomorphic to $C(\widetilde{X})$. Let $\tilde{f} \in C(\widetilde{X})$ be the image of f under this isomorphism. Since $f = 1_A$ is an idempotent, $\tilde{f}$ is also an idempotent, and thus is of the form $\tilde{f} = 1_{\widetilde{A}}$ for some (clopen) set $\widetilde{A} \in \widetilde{\mathcal{B}}$, where $\widetilde{\mathcal{B}}$ is the Borel σ-algebra on $\widetilde{X}$. The measure μ, interpreted as a linear functional on $\mathcal{A}$, gives rise to a positive linear functional on $C(\widetilde{X})$, which, by Riesz's theorem, can be represented by a measure $\tilde{\mu}$ on $\widetilde{\mathcal{B}}$. Clearly, $\tilde{\mu}(\widetilde{A}) = \mu(A)$.

The isometric operators induced on $\mathcal{A}$ by T_g, $g \in G$, give rise to isometries of $C(\widetilde{X})$, which, by a classical theorem of Banach, determine (or, rather, are determined) by homomorphisms $\widetilde{T}_g \colon \widetilde{X} \longrightarrow \widetilde{X}$, which in our case are also $\tilde{\mu}$-preserving. Let $g_0 = e$, $g_1, \dots, g_k \in G$. It is clear that the functions $T_{g_i} f = 1_{T_{g_i}^{-1} A}$ and their products correspond to the functions $1_{\widetilde{T}_{g_i}^{-1} \widetilde{A}}$ and the products thereof; hence,

$$\tilde{\mu}\big(\widetilde{A} \cap \widetilde{T}_{g_1}^{-1} \widetilde{A} \cap \dots \cap \widetilde{T}_{g_k}^{-1} \widetilde{A}\big) = \int_{\widetilde{X}} \prod_{i=0}^{k} 1_{\widetilde{T}_{g_i}^{-1} \widetilde{A}} d\tilde{\mu} = \int_X \prod_{i=0}^{k} 1_{T_{g_i}^{-1} A} d\mu = \mu\big(A \cap T_{g_1}^{-1} A \cap \dots \cap T_{g_k}^{-1} A\big).$$

■

Acknowledgment. We thank Donald Robertson for helpful comments.

Bibliography

[B1] V. Bergelson, The multifarious Poincaré recurrence theorem, *Descriptive set theory and dynamical systems*, 31–57, London Math. Soc. Lecture Note Ser. **277**, Cambridge Univ. Press, Cambridge, 2000.

[B2] V. Bergelson, Ergodic Ramsey theory, *Logic and combinatorics*, 63–87, Contemp. Math. **65**, Amer. Math. Soc., Providence, RI, 1987.

[B3] V. Bergelson, Ergodic theory and Diophantine problems., Topics in symbolic dynamics and applications (Temuco, 1997), 167–205, London Math. Soc. Lecture Note Ser., **279**, Cambridge Univ. Press, Cambridge, 2000.

[BF] V. Bergelson and H. Furstenberg, WM groups and Ramsey theory, *Topology Appl.* **156** (2009), no. 16, 2572–2580.

[BL] V. Bergelson and A. Leibman, Polynomial extensions of van der Waerden's and Szemerédi's theorems, *J. of AMS* **9** (1996), 725–753.

[BLM] V. Bergelson, A. Leibman, and R. McCutcheon, Polynomial Szemerédi theorems for countable modules over integral domains and finite fields, *J. d'Analyse Math.* **95** (2005), 243–296.

[BLMo] V. Bergelson, A. Leibman, and C.G. Moreira, From discrete- to continuous-time ergodic theorems, *Ergodic Theory and Dynam. Systems* **32** (2012), no. 2, 383–426.

[BM] V. Bergelson and R. McCutcheon, An ergodic IP polynomial Szemerédi theorem, *Mem. Amer. Math. Soc* **146** (2000), no. 695, viii+106pp.

[FK] N. Frantzikinakis and B. Kra, Polynomial averages converge to the product of integrals, *Israel J. Math* **148** (2005), 267–276.

[Fu1] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. D'Analyse Mathematique* **31** (1977), 204–256.

[Fu2] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, 1981.

[FuKa] H. Furstenberg and Y. Katznelson, An ergodic Szemerédi theorem for commuting transformations, *J. D'Analyse Mathematique* **34** (1978), 275–291.

[HS] N. Hindman and D. Strauss, Density in arbitrary semigroups, *Semigroup Forum* **73** (2006), no. 2, 273–300.

[HK1] B. Host and B. Kra, Non-conventional ergodic averages and nilmanifolds, *Annals of Math.* **161** (2005), no. 1, 397–488.

[HK2] B. Host and B. Kra, Averaging along cubes, *Modern dynamical systems and applications*, Cambridge University Press, 2004.

[Kh] A.Y. Khintchine, Eine Verschärfung des Poincaréschen "Wiederkehrsatzes", *Comp. Math* **1** (1934), 177–179.

[La] P. Larick, Results in polynomial recurrence for actions of fields, PhD Thesis, the Ohio State University, 1998.

[P] A.L.T. Paterson, *Amenability*, AMS, Providence, RI, 1988.

[Po] H. Poincaré, Sur le problème des trois corps et les équations de la Dynamique, *Acta Mathematica* **13** (1890), 1–270.

[W] M.N. Walsh, Norm convergence of nilpotent ergodic averages, *Annals of Math* **175** (2012), 1667–1688.

[Z] P. Zorin-Kranich, Norm convergence of multiple ergodic averages on amenable groups, *J. Anal. Math.* **130** (2016), 219–241.