Convergence of multiple ergodic averages along polynomials of several variables

A. Leibman
Department of Mathematics
The Ohio State University
Columbus, OH 43210, USA
e-mail: leibman@math.ohio-state.edu

February 15, 2004

Abstract

Let T be an invertible measure preserving transformation of a probability measure space X. Generalizing a recent result of Host and Kra, we prove that the averages $\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} T^{p_1(u)}f_1\cdot\ldots\cdot T^{p_r(u)}f_r$ converge in $L^1(X)$ for any $f_1,\ldots,f_r\in L^\infty(X)$, any polynomials $p_1,\ldots,p_r\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ and any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$.

Throughout the paper, (X,μ) is a probability measure space and T is an invertible measure preserving transformation of X. Our goal is to prove the following:

Theorem 1. *For any $f_1,\ldots,f_r\in L^\infty(X)$, any polynomials $p_1,\ldots,p_r\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ and any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$ the averages*

$$\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} T^{p_1(u)}f_1\cdot T^{p_2(u)}f_2\cdot\ldots\cdot T^{p_r(u)}f_r \tag{1}$$

converge in $L^1(X)$ as $N\to\infty$.

The "multiple ergodic averages"

$$\frac{1}{N}\sum_{n=M_N+1}^{M_N+N} T^n f_1\cdot T^{2n}f_2\cdot\ldots\cdot T^{rn}f_r,\quad f_1,\ldots,f_r\in L^\infty(X), \tag{2}$$

were introduced by H. Furstenberg in his ergodic theoretical proof of Szemerédi's theorem ([F]). In the situation where T is weakly mixing the L^1-convergence of the averages (2) as $N\to\infty$ was proved in [F]. For general T, the L^1-convergence of these averages was proved for $r=2$ in [F]; for $r=3$ in the case of a totally ergodic T by Conze and Lesigne ([CL1],

Supported by NSF grant DMS-0345350.

[CL2]), in the case of general T by Furstenberg and Weiss ([FW]), and by Host and Kra ([HK1]); for $r=4$ by Ziegler ([Z1]). Finally, the L^1-convergence of the averages (2) for arbitrary r was proved by Host and Kra ([HK3]), and independently by Ziegler ([Z2]).

The L^1-convergence of the "polynomial" multiple ergodic averages

$$\frac{1}{N}\sum_{n=M_N+1}^{M_N+N} T^{p_1(n)}f_1\cdot T^{p_2(n)}f_2\cdot\ldots\cdot T^{p_r(n)}f_r,\quad f_1,\ldots,f_r\in L^\infty(X),\ p_1,p_2,\ldots,p_r\in\mathbb{Z}[n], \tag{3}$$

in the case of a weakly mixing T was established by Bergelson in [B1]. For general T, the convergence of the simplest nonlinear multiple ergodic averages $\frac{1}{N}\sum_{n=M_N+1}^{M_N+N} T^{n^2}f_1\cdot T^n f_2$ was proved by Furstenberg and Weiss ([FW]). The 2-parameter multiple ergodic averages $\frac{1}{N^2}\sum_{n_1,n_2=M_N+1}^{M_N+N} f_{0,0}\cdot T^{n_1}f_{1,0}\cdot T^{n_2}f_{0,1}\cdot T^{n_1+n_2}f_{1,1}$ were introduced and proven to converge by Bergelson ([B2]). Host and Kra proved the convergence of the d-parameter averages of this sort, $\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{\epsilon\in\{0,1\}^d} T^{\epsilon\cdot u}f_\epsilon$, where $\Phi_N\subset\mathbb{Z}^d$ is a sequence of rectangles whose sizes tend to infinity in all directions, for $d=3$ in [HK2] and for arbitrary d in [HK3]. The theory of "nilpotent factors" established by Host and Kra in [HK3] allowed the authors to prove in [HK4] the L^1-convergence of the polynomial (one-parameter) multiple ergodic averages (3) (in the case T is totally ergodic, or under some negligible restrictions on the system of polynomials $\{p_1,\ldots,p_r\}$.)

Our proof of Theorem 1 is very similar to the proof in [HK4]; dealing with the multi-parameter situation simplifies it a little bit and allows one to extend it to the cases missed in [HK4].

We first remind the reader some elements of the Host-Kra theory from [HK3]. The measure preserving systems $(X^{[k]},\mu^{[k]},T^{[k]})$, $k=0,1,2,\ldots$, are constructed inductively; one puts $(X^{[0]},\mu^{[0]},T^{[0]})=(X,\mu,T)$. When $(X^{[k]},\mu^{[k]},T^{[k]})$ has already been defined for certain k, let $\mathcal{I}_k$ be the σ-algebra of measurable subsets of $X^{[k]}$ invariant under the action of $T^{[k]}$, and let I_k be the factor of $X^{[k]}$ associated with $\mathcal{I}_k$. Then $(X^{[k+1]},\mu^{[k+1]})$ is the relative product $(X^{[k]},\mu^{[k]})\times_{I_k}(X^{[k]},\mu^{[k]})$, with $T^{[k+1]}=T^{[k]}\times T^{[k]}$ naturally acting on $X^{[k+1]}$. For $F,G\in L^\infty(X^{[k]})$ this means that

$$\int_{X^{[k+1]}} F\otimes G\,d\mu^{[k+1]}=\int_{I_k} E(F|I_k)\cdot E(G|I_k)\,d\mu^{[k]}.$$

For $k\geq 0$, let $\mathcal{Z}_k$ be the minimal σ-algebra on X such that $\mathcal{I}_k\subseteq\mathcal{Z}_k^{\otimes 2^k}$. *The k-th Host-Kra factor* Z_k is the factor of X associated with $\mathcal{Z}_k$. In particular, Z_0 is the trivial (one-point) factor and Z_1 is the Kronecker factor of X. The factors Z_k form an increasing sequence: for any $k\geq 1$, Z_k is an extension of Z_{k-1}. *A k-step nilmanifold* is a homogeneous space of a nilpotent Lie group of nilpotency class k equipped with the Haar measure, and *a k-step pro-nilmanifold* is the inverse limit of a sequence of k-step nilmanifolds. The central result of the Host-Kra theory is that, for any k, Z_k possesses a natural structure of a compact k-step pro-nilmanifold such that T acts on Z_k as a translation.

For a bounded measurable real-valued function f on X and $k=0,1,2,\ldots$ one defines

$$|||f|||_k=\Bigl(\int_{X^{[k]}} f^{\otimes 2^k}\,d\mu^{[k]}\Bigr)^{1/2^k}.$$

In particular, $\|\|f\|\|_0 = \int_X f\,d\mu$. The seminorms $\|\|f\|\|_k$ form a nondecreasing sequence: $\|\|f\|\|_0 \leq \|\|f\|\|_1$ and $0 \leq \|\|f\|\|_1 \leq \|\|f\|\|_2 \leq \ldots \leq \|f\|_{L^\infty(X)}$, and are T-invariant: $\|\|Tf\|\|_k = \|\|f\|\|_k$ for any k. By the definition of $\mu^{[k+1]}$, $\|\|f\|\|_{k+1}^{2^{k+1}} = \int_{I_k} E\bigl(f^{\otimes 2^k}|I_k\bigr)^2 d\mu^{[k]}$. Thus, $\|\|f\|\|_{k+1} = 0$ if $E(f|Z_k) = 0$, that is, if $f \perp L^2(Z_k)$ in $L^2(X)$.

For $k \geq 0$ and $n \in \mathbb{Z}$ one has

$$\|\|f \cdot T^n f\|\|_k^{2^k} = \int_{X^{[k]}} f^{\otimes 2^k} \cdot (T^{[k]})^n f^{\otimes 2^k} d\mu^{[k]}.$$

By the ergodic theorem, $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N (T^{[k]})^n f^{\otimes 2^k} = E\bigl(f^{\otimes 2^k}|I_k\bigr)$ in $L^1(X^{[k]})$, and thus

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \|\|f \cdot T^n f\|\|_k^{2^k} &= \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N \int_{X^{[k]}} f^{\otimes 2^k} \cdot (T^{[k]})^n f^{\otimes 2^k} d\mu^{[k]} \\ &= \int_{I_k} E\bigl(f^{\otimes 2^k}|I_k\bigr)^2 d\mu^{[k]} = \|\|f\|\|_{k+1}^{2^{k+1}}.\end{aligned}$$

This provides one with an inductive definition of the seminorms $\|\|\cdot\|\|_k$ that is extremely convenient in applications.

Let us return to Theorem 1. Fix $K \in \mathbb{N}$. Because of the multilinearity of the expression in (1), it suffices to prove the theorem only in the case where each f_i either belongs to $L^2(Z_K)$ or is orthogonal to this space in $L^2(X)$. If all $f_1,\ldots,f_r \in L^2(Z_K)$ one may replace X by Z_K and assume that X is a pro-nilmanifold, or even a nilmanifold. In this situation Theorem 1 is a corollary of the following fact:

Theorem 2. ([Le1]) *Let N be a compact homogeneous space of a nilpotent Lie group G, let $T_1,\ldots,T_r \in G$ and let $p_1,\ldots,p_r$ be polynomials $\mathbb{Z}^d \longrightarrow \mathbb{Z}$. Then as $N \to \infty$ the averages $\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} T_1^{p_1(u)}\ldots T_r^{p_r(u)} f$ converge pointwise for any $f \in C(X)$ and any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$.*

Applying Theorem 2 to $N = X^r$, $T_1 = T\times \mathrm{Id}_X \times\ldots\times \mathrm{Id}_X,\ldots,T_r = \mathrm{Id}_X\times\ldots\times\mathrm{Id}_X \times T$ and $f = f_1\otimes\ldots\otimes f_r$ we obtain the pointwise convergence of the averages (1) for continuous $f_1,\ldots,f_r$; the L^1-convergence of the averages (1) for arbitrary $f_1,\ldots,f_r \in L^\infty(X)$ follows.

The problem is therefore reduced to the case where one of f_i, say f_1, is orthogonal to $L^2(Z_K)$; we then have $\|\|f_1\|\|_{K+1} = 0$. Clearly, we may assume that the polynomials $p_1,\ldots,p_r$ in the formulation of Theorem 1 are nonconstant and *essentially distinct*, that is, $p_i - p_j \neq \mathrm{const}$ for $i \neq j$. We will prove the following:

Theorem 3. *For any $r, b \in \mathbb{N}$ there exists $k \in \mathbb{N}$ such that for any system of nonconstant essentially distinct polynomials $p_1,\ldots,p_r\colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ of degree $\leq b$ and any $f_1,\ldots,f_r \in L^\infty(X)$ with $\|\|f_1\|\|_k = 0$ one has $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} T^{p_1(u)}f_1\cdot\ldots\cdot T^{p_r(u)}f_r = 0$ in $L^1(X)$ for any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$.*

Remark. The integer k in Theorem 3 depends on neither the measure preserving system (X,T) nor d.

In the proof of Theorem 3 we will use the following version of the van der Corput lemma:

Lemma 4. (Cf. [BMQ], Lemma 4.2) *Let $\{g_u\}_{u\in G}$ be a bounded family of elements of a Hilbert space indexed by elements of a finitely generated abelian group G and let $\{\Phi_N\}_{N=1}^\infty$ be a Følner sequence in G.*
(i) *For any finite set $F \subseteq G$,*

$$\limsup_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_u\Big\|^2 \le \limsup_{N\to\infty}\frac{1}{|F|^2}\sum_{v,w\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\langle g_{u+v}, g_{u+w}\rangle \in \mathbb{R}.$$

(ii) *There exists a Følner sequence $\{\Theta_M\}_{M=1}^\infty$ in G^3 such that*

$$\limsup_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_u\Big\|^2 \le \limsup_{M\to\infty}\frac{1}{|\Theta_M|}\sum_{(u,v,w)\in\Theta_M}\langle g_{u+v}, g_{u+w}\rangle \in \mathbb{R}.$$

Proof. (i) Let $F \subseteq G$, $|F| < \infty$. For every $u \in \mathbb{Z}^d$ and $v \in F$ put $g_{u,v} = g_u$. For any $N \in \mathbb{N}$ we have

$$\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_u = \frac{1}{|F|}\sum_{v\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_{u,v} = \Big(\frac{1}{|F|}\sum_{v\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_{u+v,v}\Big) - A_N + B_N,$$

where $A_N = \frac{1}{|F|}\sum_{v\in F}\frac{1}{|\Phi_N|}\sum_{\substack{u\in\Phi_N\\ u+v\notin\Phi_N}} g_{u+v}$ and $B_N = \frac{1}{|F|}\sum_{v\in F}\frac{1}{|\Phi_N|}\sum_{\substack{u\notin\Phi_N\\ u+v\in\Phi_N}} g_{u+v}$. Since $\{\Phi_N\}_{N=1}^\infty$ is a Følner sequence and $\{g_u\}_{u\in G}$is a bounded set, $\|A_N\|, \|B_N\| \to 0$ as $N \to \infty$. Thus,

$$\limsup_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_u\Big\| = \limsup_{N\to\infty}\Big\|\frac{1}{|F|}\sum_{v\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_{u+v}\Big\|.$$

And by the Cauchy-Schwarz inequality,

$$\begin{aligned}\Big\|\frac{1}{|F|}\sum_{v\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_{u+v}\Big\|^2 &= \frac{1}{|F|^2}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\sum_{v\in F} g_{u+v}\Big\|^2\\ &\le \frac{1}{|F|^2}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\Big\|\sum_{v\in F} g_{u+v}\Big\|^2 = \frac{1}{|F|^2}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\sum_{v,w\in F}\langle g_{u+v}, g_{u+w}\rangle.\end{aligned}$$

(ii) Put $S = \limsup_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N} g_u\Big\|^2$. Choose any Følner sequence $\{\Psi_M\}_{M=1}^\infty$ in G and, using (i), find an increasing sequence $N_1, N_2, \ldots \in \mathbb{N}$ such that for each $M \in \mathbb{N}$

$$\frac{1}{|\Psi_M|^2}\sum_{v,w\in\Psi_M}\frac{1}{|\Phi_{N_M}|}\sum_{u\in\Phi_{N_M}}\langle g_{u+v}, g_{u+w}\rangle > S - \tfrac{1}{M}.$$

Define $\Theta_M = \Phi_{N_M} \times \Psi_M^2$, $M = 1, 2, \ldots$. Then $\{\Theta_M\}_{M=1}^\infty$ is a Følner sequence in G^3 and

$$\limsup_{M\to\infty}\frac{1}{|\Theta_M|}\sum_{(u,v,w)\in\Theta_M}\langle g_{u+v}, g_{u+w}\rangle = \limsup_{M\to\infty}\frac{1}{|\Psi_M|^2\cdot|\Phi_{N_M}|}\sum_{\substack{v,w\in\Psi_M\\ u\in\Phi_{N_M}}}\langle g_{u+v}, g_{u+w}\rangle \ge S.$$

■

Agreement. For simplicity, starting from this point we will assume all functions on X we deal with to be real-valued.

We first prove Theorem 3 for polynomials of degree 1, which we will call *linear functions.*

Proposition 5. *Let $p_1,\ldots,p_r\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ be nonconstant essentially distinct linear functions. There exists a constant C such that*

$$\limsup_{N\to\infty}\Bigl\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p_1(u)}f_1\cdot\ldots\cdot T^{p_r(u)}f_r\Bigr\|_{L^2(X)}\le C|||f_1|||_{r+1}\cdot\prod_{i=2}^r\|f_i\|_{L^\infty(X)}$$

for any $f_1,\ldots,f_r\in L^\infty(X)$ and any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$.

Corollary 6. *Let $p_1,\ldots,p_r\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ be nonconstant essentially distinct linear functions. For any $f_1,\ldots,f_r\in L^\infty(X)$ with $|||f_1|||_{r+1}=0$ one has $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p_1(u)}f_1\cdot\ldots\cdot T^{p_r(u)}f_r=0$ in $L^1(X)$ for any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$.*

Remark. Actually, if $r\ge 2$, for $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p_1(u)}f_1\cdot\ldots\cdot T^{p_r(u)}f_r=0$ it is enough that $|||f_1|||_r=0$, but proving this fact requires a more careful investigation. (See [Le2].)

Lemma 7. *Let $p\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ be a nonconstant linear function. There exists a constant c such that for any $f\in L^\infty(X)$ and any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$ one has $\lim_{N\to\infty}\bigl\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p(u)}f\bigr\|_{L^2(X)}\le c|||f|||_2$.*

Proof. In coordinates, let $p(u)=a_1u_1+\ldots+a_du_d+a_0$, $u=(u_1,\ldots,u_d)\in\mathbb{Z}^d$, with $a_1,\ldots,a_d\in\mathbb{Z}$. After replacing f by $T^{a_0}f$ we may assume that $a_0=0$. Put $a=\gcd(a_1,\ldots,a_d)$. Then, in $L^1(X)$, $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p(u)}f=E(f|J_a)$ where J_a is the factor of X associated with the σ-algebra of T^a-invariant measurable subsets of X. Recalling that $|||\cdot|||_0\le|||\cdot|||_1$ and $|||\cdot|||_1\ge 0$, we get

$$\begin{aligned}\lim_{N\to\infty}\Bigl\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p(u)}f\Bigr\|_{L^2(X)}^2&=\bigl\|E(f|J_a)\bigr\|_{L^2(X)}^2=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N\int_X f\cdot T^{an}f\,d\mu\\&=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N|||f\cdot T^{an}f|||_0\le\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N|||f\cdot T^{an}f|||_1\le a\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N|||f\cdot T^{n}f|||_1\\&\le a\Bigl(\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N|||f\cdot T^nf|||_1^2\Bigr)^{1/2}=a|||f|||_2^2.\end{aligned}$$

∎

Lemma 8. *Let $p\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ be a nonconstant linear function. There exists a constant c such that for any $f\in L^\infty(X)$, any $k\ge 1$ and any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$ one has $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}|||f\cdot T^{p(u)}f|||_k^{2^k}\le c|||f|||_{k+1}^{2^{k+1}}$.*

Proof. Let, again, $p(u) = a_1u_1 + \ldots + a_du_d$ and $a = \gcd(a_1, \ldots, a_d)$. Denote by $J_{k,a}$ the factor of $X^{[k]}$ associated with the σ-algebra of $(T^{[k]})^a$-invariant measurable subsets of $X^{[k]}$. We have

$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} |||f \cdot T^{p(u)} f|||_k^{2^k} = \lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \int_{X^{[k]}} f^{\otimes 2^k} \cdot (T^{[k]})^{p(u)} f^{\otimes 2^k}\, d\mu^{[k]}$$
$$= \int_{X^{[k]}} E\big(f^{\otimes 2^k} | J_{k,a}\big)^2 d\mu^{[k]} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \int_{X^{[k]}} f^{\otimes 2^k} \cdot (T^{[k]})^{an} f^{\otimes 2^k}\, d\mu^{[k]}$$
$$= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |||f \cdot T^{an} f|||_k^{2^k} \leq a \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} |||f \cdot T^{n} f|||_k^{2^k} = a|||f|||_{k+1}^{2^{k+1}}.$$

■

Proof of Proposition 5. We proceed by induction on r. For $r = 1$ the statement is given by Lemma 7. Let $r \geq 2$, let $f_1, \ldots, f_r \in L^\infty(X)$ and let $\{\Phi_N\}_{N=1}^\infty$ be a Følner sequence in $\mathbb{Z}^d$. We will assume that $|f_2|, \ldots, |f_r| \leq 1$. We will also assume that $p_1(0) = \ldots = p_r(0) = 0$. By Lemma 4(i), applied to the elements $g_u = T^{p_1(u)} f_1 \cdot \ldots \cdot T^{p_r(u)} f_r$, $u \in \mathbb{Z}^d$, of $L^2(X)$, for any finite $F \subset \mathbb{Z}^d$ we get

$$\limsup_{N\to\infty} \Big\| \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \prod_{i=1}^{r} T^{p_i(u)} f_i \Big\|_{L^2(X)}^2$$
$$\leq \limsup_{N\to\infty} \frac{1}{|F|^2} \sum_{v,w\in F} \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \int_X \prod_{i=1}^{r} T^{p_i(u+v)} f_i \cdot \prod_{i=1}^{r} T^{p_i(u+w)} f_i \, d\mu$$
$$= \limsup_{N\to\infty} \frac{1}{|F|^2} \sum_{v,w\in F} \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \int_X \prod_{i=1}^{r} T^{p_i(u)} \big(T^{p_i(v)} f_i \cdot T^{p_i(w)} f_i\big)\, d\mu$$
$$= \limsup_{N\to\infty} \frac{1}{|F|^2} \sum_{v,w\in F} \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \int_X \Big(\prod_{i=1}^{r-1} T^{(p_i - p_r)(u)} \big(T^{p_i(v)} f_i \cdot T^{p_i(w)} f_i\big)\Big)$$
$$\cdot \big(T^{p_r(v)} f_r \cdot T^{p_r(w)} f_r\big)\, d\mu$$
$$\leq \frac{1}{|F|^2} \sum_{v,w\in F} \limsup_{N\to\infty} \Big\| \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \prod_{i=1}^{r-1} T^{(p_i - p_r)(u)} \big(T^{p_i(v)} f_i \cdot T^{p_i(w)} f_i\big) \Big\|_{L^2(X)}.$$

By the induction hypothesis there exists a constant C', independent on $f_1, \ldots, f_r$ and $\{\Phi_N\}_{N=1}^\infty$, such that

$$\limsup_{N\to\infty} \Big\| \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \prod_{i=1}^{r-1} T^{(p_i - p_r)(u)} \big(T^{p_i(v)} f_i \cdot T^{p_i(w)} f_i\big) \Big\|_{L^2(X)} \leq C' |||T^{p_1(v)} f_1 \cdot T^{p_1(w)} f_1|||_r$$

for all $v,w\in\mathbb{Z}^d$. Thus, for any finite set $F\subset\mathbb{Z}^d$,

$$\begin{aligned}\limsup_{N\to\infty}\Bigl\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{i=1}^{r}T^{p_i(u)}f_i\Bigr\|_{L^2(X)}&\leq\Bigl(\frac{C'}{|F|^2}\sum_{v,w\in F}\bigl|\!\bigl|\!\bigl|T^{p_1(v)}f_1\cdot T^{p_1(w)}f_1\bigr|\!\bigr|\!\bigr|_r\Bigr)^{1/2}\\ &=C'^{1/2}\Bigl(\frac{1}{|F|^2}\sum_{v,w\in F}\bigl|\!\bigl|\!\bigl|f_1\cdot T^{p_1(w-v)}f_1\bigr|\!\bigr|\!\bigr|_r\Bigr)^{1/2}\\ &\leq C'^{1/2}\Bigl(\frac{1}{|F|^2}\sum_{v,w\in F}\bigl|\!\bigl|\!\bigl|f_1\cdot T^{p_1(w-v)}f_1\bigr|\!\bigr|\!\bigr|_r^{2^r}\Bigr)^{(1/2)^{r+1}}.\end{aligned}\tag{4}$$

Let $\{\Psi_M\}_{M=1}^{\infty}$ be any Følner sequence in $\mathbb{Z}^d$. Then $\{\Psi_M^2\}_{M=1}^{\infty}$ is a Følner sequence in $\mathbb{Z}^{2d}$, and since $(v,w)\mapsto p_1(w-v)$ is a nonconstant linear function on $\mathbb{Z}^{2d}$, by Lemma 8 we have

$$\limsup_{M\to\infty}\frac{1}{|\Psi_M|^2}\sum_{v,w\in\Psi_M}\bigl|\!\bigl|\!\bigl|f_1\cdot T^{p_1(w-v)}f_1\bigr|\!\bigr|\!\bigr|_r^{2^r}\leq c|\!|\!|f_1|\!|\!|_{r+1}^{2^{r+1}},$$

with c independent on f_1. Substituting the sets Ψ_M, $M\in\mathbb{N}$, for F in (4) we obtain

$$\limsup_{N\to\infty}\Bigl\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{i=1}^{r}T^{p_i(u)}f_i\Bigr\|_{L^2(X)}\leq C'^{1/2}c^{(1/2)^{r+1}}|\!|\!|f_1|\!|\!|_{r+1}.$$

■

We now turn to the case of nonlinear p_i. We will call *a system* any finite set of polynomials on a space $\mathbb{Z}^d$. *The degree*, $\deg P$, of a system P is the maximum of the degrees of its elements. *The weight*, $\omega(P)$, of a system P is defined in the following way. We will say that polynomials p, q *are equivalent* if $\deg p=\deg q$ and $\deg(p-q)<\deg p$; the degree of a class of equivalent polynomials is the degree of its elements. P is partitioned into equivalence classes; for each positive integer $l\leq\deg P$ let ω_l be the number of classes of degree l in P. Then $\omega(P)$ is the vector $\bigl(\omega_1,\ldots,\omega_{\deg P}\bigr)$. For two integer vectors $\omega=(\omega_1,\ldots,\omega_m)$ and $\omega'=(\omega'_1,\ldots,\omega'_{m'})$ we will write $\omega<\omega'$ if either $m<m'$, or $m=m'$ and there is $n\leq m$ such that $\omega_n<\omega'_n$ and $\omega_l=\omega'_l$ for $l=n+1,\ldots,m$. Under this relation the set of weights of systems of polynomials becomes well ordered. The *PET-induction*, introduced in [B1], is an induction on this well ordered set.

An ordered system $P=\{p_1,\ldots,p_r\}$ will be said to be *standard* if all p_i are nonconstant and essentially distinct (that is, $p_i-p_j\neq\text{const}$ for $i\neq j$), and $\deg p_1=\deg P$. We will be proving the following:

Proposition 9. *For any $r\in\mathbb{N}$ and any integer vector $\omega=(\omega_1,\ldots,\omega_l)$ there is $k\in\mathbb{N}$ such that for any standard system $\bigl\{p_1,\ldots,p_r:\mathbb{Z}^d\longrightarrow\mathbb{Z}\bigr\}$ of weight ω and any $f_1,\ldots,f_r\in L^\infty(X)$ with $|\!|\!|f_1|\!|\!|_k=0$ one has $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{p_1(u)}f_1\cdot\ldots\cdot T^{p_r(u)}f_r=0$ in $L^1(X)$ for any Følner sequence $\{\Phi_N\}_{N=1}^{\infty}$ in $\mathbb{Z}^d$.*

We will say that a certain property holds *for almost all $v\in\mathbb{Z}^d$* if the set of elements of $\mathbb{Z}^d$ for which it does not hold is contained in the set of zeroes of a nontrivial polynomial on $\mathbb{Z}^d$ (or in the union of such sets, which is the same). Note that the set of zeroes of a nontrivial polynomial has zero density with respect to any Følner sequence in $\mathbb{Z}^d$.

Proof of Proposition 9. We will proceed by PET-induction. For systems of degree 1 the proposition is given by Corollary 6. Let $P=\{p_1,\ldots,p_r\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}\}$ be a standard system of degree ≥ 2 and of weight ω. There are only finitely many integer vectors $\omega'<\omega$ which are the weights of systems with $s<2r$ elements. By our PET-induction hypothesis there exists $k\in\mathbb{N}$ such that for any standard system $\bigl\{q_1,\ldots,q_s\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}\bigr\}$ with $s\leq 2r$ of weight $\omega'<\omega$ and any $h_1,\ldots,h_s\in L^\infty(X)$ with $|||h_1|||_k=0$ one has $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}T^{q_1(u)}h_1\cdot\ldots\cdot T^{q_s(u)}h_s=0$ in $L^1(X)$ for any Følner sequence $\{\Phi_N\}_{N=1}^\infty$ in $\mathbb{Z}^d$.

Let $I_1=\bigl\{i\in\{1,\ldots,r\}:\deg p_i=1\bigr\}$ and $I_2=\bigl\{i\in\{1,\ldots,r\}:\deg p_i\geq 2\bigr\}$. Choose $i_0\in\{2,\ldots,r\}$ such that p_{i_0} has the minimal degree in P; if all polynomials in P have the same degree, choose i_0 so that p_{i_0} is not equivalent to p_1; if all polynomials in P are equivalent, choose i_0 arbitrarily. For each $v,w\in\mathbb{Z}^d$ define

$$P_{v,w}=\bigl\{p_i(u+v),p_i(u+w):i\in I_2\bigr\}\bigcup\bigl\{p_i(u+w):i\in I_1\bigr\}$$

(where $p_i(u+v),p_i(u+w)$ are considered as polynomials in u), and order the system $P_{v,w}=\{q_{v,w,1},\ldots,q_{v,w,s}\}$ so that $q_{v,w,1}(u)=p_1(u+v)$ and $q_{v,w,s}(u)=p_{i_0}(u+w)$. Then $P_{v,w}$ is a standard system for almost all $(v,w)\in\mathbb{Z}^{2d}$. Since for any $v,w\in\mathbb{Z}^d$ and $i\in\{1,\ldots,r\}$ the polynomials $p_i(u+v)$ and $p_i(u+w)$ are equivalent to $p_i(u)$, we have $\omega(P_{v,w})=\omega(P)=\omega$ for all $v,w\in\mathbb{Z}^d$.

For $v,w\in\mathbb{Z}^d$ define

$$P'_{v,w}=\bigl\{q_{v,w,1}-q_{v,w,s},\ldots,q_{v,w,s-1}-q_{v,w,s}\bigr\}.$$

Then for almost all $(v,w)\in\mathbb{Z}^d$, $P'_{v,w}$ is a standard system. (Indeed, the polynomials $q_{v,w,j}-q_{v,w,s}$, $j=1,\ldots,s-1$, are nonconstant and essentially distinct whenever $q_{v,w,j}$ are. If p_{i_0} is not equivalent to p_1, then $\deg(q_{v,w,1}-q_{v,w,s})=\deg\bigl(p_1(u+v)-p_{i_0}(u+w)\bigr)=\deg p_1=\deg P_{v,w}$ for all v,w; otherwise $\deg(q_{v,w,1}-q_{v,w,s})=\deg p_1-1=\deg P_{v,w}$ for almost all (v,w).) Also, for all $(v,w)\in\mathbb{Z}^{2d}$, $\omega(P'_{v,w})<\omega$. (Indeed, the equivalence classes in $P'_{v,w}$ and their degrees remain the same as in $P_{v,w}$, except that the class in $P_{v,w}$ containing q_s splits into several new classes of less degree.)

Now let $f_1,\ldots,f_r\in L^\infty(X)$ with $|||f_1|||_k=0$, and let $\{\Phi_N\}_{N=1}^\infty$ be a Følner sequence in $\mathbb{Z}^d$. We will assume that $|f_2|,\ldots,|f_r|\leq 1$. By Lemma 4(i), for any finite set $F\subset\mathbb{Z}^d$ we get

$$
\begin{aligned}
\limsup_{N\to\infty}&\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{i=1}^{r}T^{p_i(u)}f_i\Big\|^2_{L^2(X)}\\
&\le \limsup_{N\to\infty}\frac{1}{|F|^2}\sum_{v,w\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\int_X\prod_{i=1}^{r}T^{p_i(u+v)}f_i\cdot\prod_{i=1}^{r}T^{p_i(u+w)}f_i\,d\mu\\
&=\limsup_{N\to\infty}\frac{1}{|F|^2}\sum_{v,w\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\int_X\prod_{i\in I_2}T^{p_i(u+v)}f_i\cdot\prod_{i\in I_2}T^{p_i(u+w)}f_i\cdot\\
&\qquad\qquad\prod_{i\in I_1}T^{p_i(u+w)}(f_i\cdot T^{p_i(v)-p_i(w)}f_i)\,d\mu\\
&=\limsup_{N\to\infty}\frac{1}{|F|^2}\sum_{v,w\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\int_X\prod_{j=1}^{s}T^{q_{v,w,j}(u)}h_{v,w,j}\,d\mu\\
&=\limsup_{N\to\infty}\frac{1}{|F|^2}\sum_{v,w\in F}\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\int_X\Big(\prod_{j=1}^{s-1}T^{(q_{v,w,j}-q_{v,w,s})(u)}h_{v,w,j}\Big)\cdot h_{v,w,s}\,d\mu\\
&\le\frac{1}{|F|^2}\sum_{v,w\in F}\limsup_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{j=1}^{s-1}T^{(q_{v,w,j}-q_{v,w,s})(u)}h_{v,w,j}\Big\|_{L^1(X)},
\end{aligned}
$$

where, for $v,w\in\mathbb{Z}^d$, $q_{v,w,1},\dots,q_{v,w,s}$ are the elements of the system $P_{v,w}$, and $h_{v,w,j}$ is either f_i for certain $i\in I_2$ or $f_i\cdot T^{p_i(v)-p_i(w)}f_i$ for certain $i\in I_1$; note that, since $\deg p_1=\deg P\ge 2$, $1\in I_2$ and $h_{v,w,1}=f_1$. By the induction hypothesis applied to the systems $P'_{v,w}$,

$$
\lim_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{j=1}^{s-1}T^{(q_{v,w,j}-q_{v,w,s})(u)}h_{v,w,j}\Big\|_{L^1(X)}=0
$$

for all $(v,w)\in\mathbb{Z}^{2d}$ for which $P'_{v,w}$ is standard, that is, for almost all (v,w). Since for all other (v,w) this norm is bounded by 1,

$$
\inf_F\frac{1}{|F|^2}\sum_{(v,w)\in F}\limsup_{N\to\infty}\Big\|\frac{1}{|\Phi_N|}\sum_{u\in\Phi_N}\prod_{j=1}^{s-1}T^{(q_{v,w,j}-q_{v,w,s})(u)}h_{v,w,j}\Big\|_{L^1(X)}=0.
$$

∎

Proof of Theorem 3. Proposition 9 implies Theorem 3 for standard systems, and our goal is to reduce the general case to this one. Let $P=\{p_1,\dots,p_r\}$ be a (nonstandard) system of nonconstant essentially distinct polynomials $\mathbb{Z}^d\longrightarrow\mathbb{Z}$ of degree $\le b$, let $f_1,\dots,f_r\in L^\infty(X)$ and let $\{\Phi_N\}_{N=1}^\infty$ be a Følner sequence in $\mathbb{Z}^d$. By Lemma 4(ii) there exists a Følner sequence $\{\Theta_M\}_{M=1}^{\infty}$ in $\mathbb{Z}^{3d}$ such that

$$\begin{aligned}
\limsup_{N\to\infty} \Big\| \frac{1}{|\Phi_N|} \sum_{u\in\Phi_N} \prod_{i=1}^{r} T^{p_i(u)} f_i \Big\|^2_{L^2(X)} & \\
\leq \limsup_{M\to\infty} \frac{1}{|\Theta_M|} \sum_{(u,v,w)\in\Theta_M} \int \prod_{i=1}^{r} T^{p_i(u+v)} f_i \cdot \prod_{i=1}^{r} T^{p_i(u+w)} f_i \, d\mu & \\
= \limsup_{M\to\infty} \frac{1}{|\Theta_M|} \sum_{(u,v,w)\in\Theta_M} \int \prod_{i=1}^{r} T^{p_i(u+v)+q(u)} f_i \cdot \prod_{i=1}^{r} T^{p_i(u+w)+q(u)} f_i \, d\mu & \\
\leq \limsup_{M\to\infty} \Big\| \frac{1}{|\Theta_M|} \sum_{(u,v,w)\in\Theta_M} \prod_{i=1}^{r} T^{p_i(u+v)+q(u)} f_i \cdot \prod_{i=1}^{r} T^{p_i(u+w)+q(u)} f_i \Big\|_{L^1(X)} &
\end{aligned}$$

where q is any polynomial $\mathbb{Z}^d \longrightarrow \mathbb{Z}$ of degree b. The set

$$\big\{p_1(u+v)+q(u), \ldots, p_r(u+v)+q(u), p_1(u+w)+q(u), \ldots, p_r(u+w)+q(u)\big\}$$

of polynomials $\mathbb{Z}^{3d} \longrightarrow \mathbb{Z}$ is a standard system of degree b with $2r$ elements, thus there exists $k \in \mathbb{N}$ (depending on r and b only) such that

$$\lim_{M\to\infty} \frac{1}{|\Theta_M|} \sum_{(u,v,w)\in\Theta_M} \prod_{i=1}^{r} T^{p_i(u+v)+q(u)} f_i \cdot \prod_{i=1}^{r} T^{p_i(u+w)+q(u)} f_i = 0$$

in $L^1(X)$. ∎

Acknowledgment. I thank B. Kra for her comments on the preprint.

Bibliography

[B1] V. Bergelson, Weakly mixing PET, *Erg. Th. and Dyn. Sys.* **7** (1987), 337–349.

[B2] 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.

[BMQ] V. Bergelson, R. McCutcheon and Q. Zhang, A Roth theorem for amenable groups, *Amer. J. Math.* **119** (1997), 1173–1211.

[CL1] J.-P. Conze and E. Lesigne, Théorèmes ergodiques pour des mesures diagonales, *Bull. Soc. Math. France* **112** (1984), 143-175.

[CL2] J.-P. Conze and E. Lesigne, Sur un théorème ergodique pour des mesures diagonales, *Publications de l'Institut de Recherche de Mathématiques de Rennes, Probabilités*, 1987.

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

[FW] H. Furstenberg and B. Weiss, A mean ergodic theorem for $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)g(T^{n^2}x)$, *Convergence in Ergodic Theory and Probability*, Walter de Gruyter, 1996, 193–227.

[HK1] B. Host and B. Kra, Convergence of Conze-Lesigne averages, *Erg. Th. and Dyn. Sys.* **21** (2001), 493-509.

[HK2] B. Host and B. Kra, *Averaging along cubes*, Dynamical Systems and Related Topics, Cambridge University Press.

[HK3] B. Host and B. Kra, Non-conventional ergodic averages and nilmanifolds, to appear in *Annals of Math.*

[HK4] B. Host and B. Kra, Convergence of polynomial ergodic averages, to appear in *Israel J. of Math.*

[Le1] A. Leibman, Pointwise convergence of ergodic averages for polynomial actions of $\mathbb{Z}^d$ by translations on a nilmanifold, to appear in *Ergodic Theory and Dynamical Systems.* Available at *http://www.math.ohio-state.edu/∼leibman/preprints*

[Le2] A. Leibman, Host-Kra factors for the powers of a transformation, preprint.

[Z1] T. Ziegler, *Nonconventional Ergodic Averages*, Ph.D. Thesis, Technion, 2002.

[Z2] T. Ziegler, Universal characteristic factors and non-conventional ergodic averages, preprint.