GENERALIZED POLYNOMIALS AND MILD MIXING SYSTEMS

RANDALL MCCUTCHEON AND ANTHONY QUAS

Abstract. An unsettled conjecture of V. Bergelson and I. Håland proposes that if $(X, \mathcal{A}, \mu, T)$ is an invertible weak mixing measure preserving system, where $\mu(X) < \infty$, then if $p_1, p_2, \ldots, p_k$ are generalized polynomials (functions built out of regular polynomials via iterated use of the greatest integer or floor function) having the property that no p_i, nor any $p_i - p_j$, $i \neq j$, is constant on a set of positive density, then for any measurable sets $A_0, A_1, \ldots, A_k$, there exists a zero-density set $E \subset \mathbf{Z}$ such that $\lim_{n\to\infty, n\notin E} u(A_0 \cap T^{p_1(n)}A_1 \cap \cdots \cap T^{p_k(n)}A_k) = \prod_{i=0}^{k} \mu(A_i)$. We formulate and prove a faithful version of this conjecture for mild mixing systems and partially characterize, in the degree two case, the set of families $\{p_1, p_2, \ldots, p_k\}$ satisfying the hypotheses of this theorem.

1. Introduction

A single operator, strongly mixing measure preserving system may be defined as a quadruple $(X, \mathcal{A}, \mu, T)$, where $(X, \mathcal{A}, \mu)$ is a probability space and T is an invertible measure preserving transformation of X having the property that for every $A, B \in \mathcal{A}$, $\lim_{n\to\infty} \mu(A \cap T^n B) = \mu(A)\mu(B)$. An outstanding open problem in ergodic theory is that of whether such systems must exhibit "strong mixing of higher orders". So, for example, it is unknown whether the above entails $\lim_{m,n,m-n\to\infty} \mu(A \cap T^n B \cap T^m C) = \mu(A)\mu(B)\mu(C)$. More (or less, if one is betting against) modestly, it is unknown whether $\lim_{n\to\infty} \mu(A \cap T^n B \cap T^{2n} C) = \mu(A)\mu(B)\mu(C)$ must hold for mixing systems.

On the other hand, there are other notions of mixing for which such questions do have satisfying answers, for example *weak mixing* and *mild mixing*. A system $(X, \mathcal{A}, \mu, T)$ is weakly mixing if $L^2(X)$ has no non-constant eigenfunctions under the unitary action $Tf(x) = f(Tx)$. Another characterization of weak mixing, running more along the lines of the prior characterization of strong mixing, goes as follows: for a set $E \subset \mathbf{Z}$, define the *upper density* of E by $\overline{d}(E) =$

$\limsup_{n\to\infty} \frac{|E\cap[-n,n]|}{2n+1}$. Next, for x and a sequence (x_n) in a topological space, we write $\text{D-lim}_n x_n = x$ if for every neighborhood U of x, the complement of the set $\{n : x_n \in U\}$ has zero upper density. We may now characterize weak mixing as follows: a system $(X, \mathcal{A}, \mu, T)$ is weakly mixing if and only if for every $A, B \in \mathcal{A}$, $\text{D-lim}_n \mu(A \cap T^n B) = \mu(A)\mu(B)$.

In light of this characterization, it becomes natural to ask if one may in fact obtain, e.g., $\text{D-lim}_n \mu(A \cap T^n B \cap T^{2n} C) = \mu(A)\mu(B)\mu(C)$. The answer is yes. H. Furstenberg has shown in [6] that in weak mixing systems, for any $k \in \mathbf{N}$ and measurable $A_0, A_1, \ldots, A_k$,

$$\text{D-lim}_n\ \mu(A_0 \cap T^n A_1 \cap \cdots \cap T^{kn} A_k) = \prod_{i=0}^{k} \mu(A_i). \tag{1.1}$$

A "relativized" version of this consequence of weak mixing formed an important part of his proof, via ergodic theory, of Szemerédi's theorem on arithmetic progressions, and is often referred to as "weak mixing of all orders", a designation that is perhaps somewhat misleading, as we shall now see. In [1], V. Bergelson proved a result involving polynomial powers of T for which (1.1) forms the linear case. His theorem states that in weak mixing systems and for any polynomials $p_i \in \mathbf{Z}[x]$, $1 \le i \le k$, having the property that no p_i and no $p_i - p_j$ is constant, $1 \le i \ne j \le k$, one obtains

$$\text{D-lim}_n\ \mu(A_0 \cap T^{p_1(n)} A_1 \cdots \cap T^{p_k(n)} A_k) = \prod_{i=0}^{k} \mu(A_i). \tag{1.2}$$

Thus we see that weak mixing implies, not merely weak mixing of higher linear orders, but of higher polynomial orders as well. The next question that arises is this: do such polynomial functions constitute a suitably "most general class" of integer sequences along which weak mixing systems are well behaved? In other words, are polynomial orders all orders?

Again, the answer seems to be no. Bergelson and I. Håland have proved (unpublished) that for some interesting and non-trivial classes of *generalized polynomials* p_i, (1.2) remains valid under suitable hypotheses. Generalized polynomials $\mathbf{Z} \to \mathbf{Z}$ may be defined as follows. For $r \in \mathbf{R}$, let $[r]$ denote the integer part of r, i.e. the greatest integer less than or equal to r. Put also $\{r\} = r - [r]$, the fractional part of r. The set of *generalized polynomials* $\mathbf{Z} \to \mathbf{Z}$ is the smallest set $\mathcal{G}$ that is a function algebra (i.e. is closed under sums and products) containing $\mathbf{Z}[x]$ and having the additional property that for all $m \in \mathbf{N}$, $c_1, \ldots, c_m \in \mathbf{R}$ and $p_1, \ldots, p_m \in \mathcal{G}$, the mapping $n \to [\sum_{i=1}^m c_i p_i(n)]$ is in $\mathcal{G}$.

Though we cannot explain here (the definitions are quite technical) precisely which classes prove amenable to the Bergelson-Håland analysis, it is simple to understand why (1.2) cannot possibly hold for arbitrary generalized polynomials p_i under the hypothesis that each p_i and each $p_i - p_j$ fails to be constant. Consider for example $p(n) = [2\{\pi n\}]$, a generalized polynomial that, while not constant, is finite-valued. Or, for an only slightly more exotic example, $q(n) = np(n)$, which, while taking on infinitely many values, is zero for many (roughly half, in the sense of asymptotic density) n. Note that these examples violate (1.2) for the same rather pedestrian reason: they are constant on a set E having positive upper density. Indeed as Bergelson and Håland note, (1.2) cannot possibly hold across all weak mixing systems if there exists a set E of positive upper density on which some p_i or some $p_i - p_j$ is constant. Accordingly, any k-tuple $(p_1, \ldots, p_k)$ for which they are able to prove (1.2) has *a fortiori* the following property $\mathcal{P}$: on no set E having positive upper density is any p_i or $p_i - p_j$ constant, $1 \leq i \neq j \leq k$.

Interestingly, Bergelson and Håland could find no counterexample to (1.2) having property $\mathcal{P}$. On the other hand, they were unable to prove that $\mathcal{P}$ entails (1.2). Accordingly, they formulate the following conjecture.

Conjecture A. *(Bergelson-Håland). If p_i are generalized polynomials, $1 \leq i \leq k$, such that no p_i and no $p_i - p_j$, $1 \leq i \neq j \leq k$, is constant on a set of positive upper density, then for any weak mixing system $(X, \mathcal{A}, \mu, T)$ and any $A_i \in \mathcal{A}$, $0 \leq i \leq k$, (1.2) holds.*

We do not directly address the foregoing conjecture in this paper. We do, however, offer a faithful rendition of the conjecture in the context of mild mixing, then provide an affirmative answer to this recasting of the problem. First, some background.

Let $(X, \mathcal{A}, \mu, T)$ be an invertible measure preserving system, where $\mu(X) = 1$, and suppose that $f \in L^2(X)$. If there is some sequence of natural numbers (n_k) such that $T^{n_k} f \to f$ (weakly or strongly, in measure or pointwise), then f is said to be a *rigid function.* $(X, \mathcal{A}, \mu, T)$ is mildly mixing if the only rigid functions in $L^2(X)$ are the constants. Any mildly mixing system is weakly mixing, for eigenfunctions are rigid. To see this, suppose that $Tf = \alpha f$. Since T acts unitarily on $L^2(X)$, $|\alpha| = 1$, and we may choose a sequence (n_k) of natural numbers with $\alpha^{n_k} \to 1$, etc. On the other hand, there are weakly mixing systems that fail to be mildly mixing, and mildly mixing systems that fail to be strongly mixing (see [9] for details).

We have seen that weak and strong mixing each have characterizations running roughly as follows: for any system $(X, \mathcal{A}, \mu, T)$, it is both necessary and sufficient for (insert version of, i.e. weak or strong) mixing that for any $A, B \in \mathcal{A}$ and any $\epsilon > 0$ the set of n for which $|\mu(A \cap T^n B) - \mu(A)\mu(B)| < \epsilon$ is (in an appropriate sense) "large". Here "large" means "co-finite" in the case of strong mixing, and "complement of a set of density zero" in the case of weak mixing. Mild mixing has a characterization of this form as well; the sense of "large" appropriate to it is "IP*". (A subset E of an additive semigroup S is IP* if for any sequence (x_i) in S, there is some finite, non-empty $\alpha \subset \mathbf{N}$ such that $\sum_{i\in\alpha} x_i \in E$. See below.)

One may now guess that a mild mixing analogue of (1.1) would say that for every $\epsilon > 0$, the set $M = \{n : |\mu(A_0 \cap T^n A_1 \cdots \cap T^{kn} A_k) - \prod_{i=0}^k \mu(A_i)| < \epsilon\}$ is IP*. (By IP* here and elsewhere, we mean IP* as a subset of $\mathbf{N}$, not as a subset of $\mathbf{Z}$. The reason this distinction is important in that any set that is IP* as a subset of $\mathbf{Z}$ must contain zero; M does not.) This is in fact true for mild mixing systems, as is shown in [7](Section 9.5), entitled "Mild mixing of all orders." Here again, however, linear orders proved not to be "all": indeed, Bergelson already in [1] (Theorem 4.8) states as an unproved corollary to (the proof of) his main result the following mild mixing version.

Theorem B. *Let $(X, \mathcal{A}, \mu, T)$ be a mildly mixing system, and suppose $p_i(x) \in \mathbf{Z}[x]$ are polynomials with no p_i and no $p_i - p_j$ constant, $1 \le i \ne j \le k$. Then for any $\epsilon > 0$ the set $\{n : |\mu(A_0 \cap T^{p_1(n)} A_1 \cdots \cap T^{p_k(n)} A_k) - \prod_{i=0}^k \mu(i)| < \epsilon\}$ is IP*.*

Let us consider now how to best formulate a version of Theorem B for generalized polynomials. By analogy with the discussion leading up to Conjecture A, the hypotheses must be strengthened to preclude the possibility of some p_i or $p_i - p_j$ being constant on a non-trivial set, for some appropriate interpretation of "non-trivial". (In the weak mixing case, "non-trivial set" meant "positive upper density set.") A moment's reflection indicates that the correct interpretation of "non-trivial set" in the mild mixing case is "IP set." In other words, given generalized polynomials p_i, $1 \le i \le k$, and $\epsilon > 0$, an obvious necessary condition for $\{n : |\mu(A_0 \cap T^{p_1(n)} A_1 \cdots \cap T^{p_k(n)} A_k) - \prod_{i=0}^k \mu(A_i)| < \epsilon\}$ to be IP* across mildly mixing systems is for p_i and $p_i - p_j$ to be non-constant on every IP set in $\mathbf{N}$. (For example, if $p_i - p_j = C$ on an IP set R, then taking A_j to be the complement of $T^C A_i$ yields a zero-measure intersection on R.) A faithful mild mixing version of Conjecture A, then, would assert that this necessary condition is also sufficient.

Theorem C. *If p_i are generalized polynomials, $1 \leq i \leq k$, such that no p_i and no $p_i - p_j$, $1 \leq i < j \leq k$, is constant on an IP set of natural numbers, then for any mild mixing system $(X, \mathcal{A}, \mu, T)$ and any $A_0, A_1, \ldots, A_k \in \mathcal{A}$, the set $\{n : |\mu(A_0 \cap T^{p_1(n)}A_1 \cdots \cap T^{p_k(n)}A_k) - \prod_{i=0}^{k} \mu(A_i)| < \epsilon\}$ is IP*.*

The structure of the paper is as follows. In Section 2, we give our proof of Theorem C. Then, in Section 3, we offer a partial characterization, in the degree 2 case, of the set of families of generalized polynomials meeting the hypotheses of Theorem C. Finally, in Section 4, we discuss some $\mathbf{Z}^r$ extensions of our results.

Remark. The class of *totally ergodic* systems is also suitable for multiple mixing, at least for families of polynomials that are independent over the rationals, as N. Franzikinakis and B. Kra established in [5]. The question as to which families of generalized polynomials this may work for is an interesting one that we shall not attempt to address here.

2. Proof of main theorem.

We denote by $\mathcal{F}$ the family of non-empty finite subsets of $\mathbf{N}$. An *$\mathcal{F}$-sequence* in a set G is a function $v : \mathcal{F} \to G$. There is a special sense of convergence applicable to $\mathcal{F}$-sequences when G is a topological space, of which more in a moment. For $\alpha, \beta \in \mathcal{F}$, we write $\beta < \alpha$ if $\max \beta < \min \alpha$. Suppose $\alpha_i \in \mathcal{F}$, $i \in \mathbf{N}$, with $\alpha_1 < \alpha_2 < \cdots$. The set $\mathcal{F}^{(1)}$ of non-empty finite unions of the α_i's, is called an *IP ring.* Note that $\mathcal{F}^{(1)}$ is the isomorphic image of $\mathcal{F}$ under the map $\beta \to \bigcup_{i \in \beta} \alpha_i$. The restriction of a given $\mathcal{F}$ sequence to an IP ring $\mathcal{F}^{(1)}$ plays a role analogous to that of a subsequence of a given sequence.

Having defined IP rings, we are in a position to state the following well-known result of N. Hindman [10](Corollary 3.3): given any finite coloring of an IP ring $\mathcal{F}^{(1)}$, there exists an monochromatic IP *subring* $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$. (In saying that $\mathcal{F}^{(2)}$ is a subring of $\mathcal{F}^{(1)}$, we mean nothing over and above set-theoretic inclusion.) Now suppose G is an abelian (semi-)group. If v is an $\mathcal{F}$ sequence into G satisfying $v(\alpha \cup \beta) = v(\alpha)v(\beta)$ when $\alpha \cap \beta = \emptyset$ then we say that v is an *IP system,* and we refer to its range $v(\mathcal{F}) \subset G$ as an *IP set.* An *IP** subset of G is a set $E \subset G$ that intersects every IP set in G nontrivially.

As for the notion of convergence, here it is: if G is a topological space, $x \in G$, v is an $\mathcal{F}$-sequence in G and $\mathcal{F}^{(1)}$ is an IP ring, we write IP-$\lim_{\alpha \in \mathcal{F}^{(1)}} v(\alpha) = x$, or say that $v(\alpha) \to x$, $\alpha \in \mathcal{F}^{(1)}$, if for every neighborhood U of x there exists $\alpha_0 \in \mathcal{F}$ such that for every $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$, $v(\alpha) \in U$. If $p : \mathbf{N} \to G$, we write IP*- $\lim_{n\in\mathbf{N}} p(n) = x$ if for every neighborhood U of x, the set $\{n \in \mathbf{N} : p(n) \in U\}$ is IP*.

The following can be proved by substituting Hindman's theorem for the pigeonhole principle in the proof of the Bolzano-Weierstrass theorem.

Theorem 2.1. *Let X be a compact metric space and suppose v is an $\mathcal{F}$-sequence in X. Then for some IP ring $\mathcal{F}^{(1)}$ and some $x \in X$, IP-lim$_{\alpha\in\mathcal{F}^{(1)}} v(\alpha) = x$.*

The following result is well known. (See [8], Theorem 1.7.)

Lemma 2.2. *Let U be an IP system into a commutative group of unitary operators of a Hilbert space $\mathcal{H}$. If $\mathcal{F}^{(1)}$ is an IP ring such that IP-lim$_{\alpha\in\mathcal{F}^{(1)}} U(\alpha)f = Pf$ exists weakly for all $f \in \mathcal{H}$, then P is the orthogonal projection onto a closed subspace of $\mathcal{H}$.*

At first glance it might appear that the hypotheses of Lemma 2.2 could be rarely satisfied, however this is not the case. Indeed, with the help of Theorem 2.1 and a standard diagonalization argument, one can always find an IP ring $\mathcal{F}^{(1)}$ such that IP-lim$_{\alpha\in\mathcal{F}^{(1)}} U(\alpha)f$ exists weakly for all $f \in \mathcal{H}$, provided $\mathcal{H}$ is separable.

We now pass to *VIP systems*, which are polynomial-like variants of IP systems (for more information, see [2] and [12]). Suppose again that G is an abelian group (from now on, however, we shall write the operation of G additively). For a set α, let $|\alpha|$ denote its cardinality. Suppose there exist $d \in \mathbf{N}$ and a function $f : \{\emptyset\} \cup \{\alpha \in \mathcal{F} : |\alpha| \le d\} \to G$ satisfying $f(\emptyset) = 0$ and $f(\gamma) \neq 0$ for some $|\gamma| = d$. Then, letting $v(\alpha) = \sum_{\gamma\subset\alpha,|\gamma|\le d} f(\gamma)$, v is called a VIP system of degree d. f is called the *generating function* of v. It is a simple exercise to show that the generating function of a VIP system is unique, so that in particular the degree is well defined.

Given a VIP system v into $\mathbf{Z}$ and some fixed $\beta \in \mathcal{F}$, we define, for $\alpha > \beta$, $v^\beta(\alpha) = v(\alpha \cup \beta) - v(\alpha) - v(\beta)$. The family of VIP systems into $\mathbf{Z}$ itself forms an abelian group under addition, and one may check that the map $\beta \to v^\beta$ is, modulo a certain technicality we shall address presently, a VIP system into this group. The technicality involves the fact that v^β is only defined for $\alpha > \beta$. One may overcome this difficulty as follows: given an IP ring $\mathcal{F}^{(1)}$, let $\Omega' = \Omega'(\mathbf{Z}, \mathcal{F}^{(1)})$ be the group of all VIP systems $\mathcal{F}^{(1)} \to \mathbf{Z}$ and define an equivalence relation on Ω' whereby v_1 is equivalent to v_2 if there exists some $\alpha \in \mathcal{F}$ such that $v_1(\alpha) = v_2(\alpha)$ for every $\alpha \in \mathcal{F}^{(1)}$ with $\alpha > \alpha_0$. One easily checks that the group structure lifts in a well-defined way to the set of equivalence classes, which we may denote by $\Omega = \Omega(\mathbf{Z}, \mathcal{F}^{(1)})$. Now it is easy to see that $\beta \to v^\beta$ has a natural interpretation as a VIP system $\mathcal{F}^{(1)} \to \Omega$, and moreover that $\deg v^\beta = \deg v - 1$. We shall utilize this construction in the proof of Theorem 2.7 below.

What we require of the connection between VIP systems and generalized polynomials is summarized in Proposition 2.3, which is a weak form of [3] (Theorem 2.9).

Proposition 2.3. *Let $p(x)$ be a generalized polynomial and suppose n is an IP system in $\mathbf{N}$. Then for every IP ring $\mathcal{F}^{(1)}$ there exists an IP ring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ and $c \in \mathbf{Z}$ such that the restriction of $v(\alpha) = p(n(\alpha)) + c$ to $\mathcal{F}^{(2)}$ is a VIP system.*

The following lemma is standard (see [8] Lemma 5.3).

Lemma 2.4. *Suppose that $(x_\alpha)_{\alpha\in\mathcal{F}}$ is a bounded $\mathcal{F}$-sequence in a Hilbert space and $\mathcal{F}^{(1)}$ is an IP-ring. If*

$$\underset{\beta\in\mathcal{F}^{(1)}}{\textit{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\textit{IP-lim}} \langle x_\alpha, x_{\alpha\cup\beta}\rangle = 0$$

then along some subring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$, $x_\alpha \to 0$ in the weak topology.

Proposition 2.5. *Suppose $(X, \mathcal{A}, \mu, T)$ is a mildly mixing system. If $f, g \in L^\infty(X)$ and n is an IP system into $\mathbf{Z}$ that is not identically zero when restricted to any IP subring of a given IP ring $\mathcal{F}^{(1)}$, then there exists a refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that*

$$\underset{\alpha\in\mathcal{F}^{(2)}}{\textit{IP-lim}} \int fT^{n(\alpha)}g\, d\mu = \left(\int f\, d\mu\right)\left(\int g\, d\mu\right).$$

Proof. As remarked earlier, by passing to a refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$, we may assume that in the weak operator topology, $T^{n(\alpha)}h$ converges to a limit Qh for all $h \in L^2(X)$. By Lemma 2.2, Q is an orthogonal projection. In particular, Q is idempotent, so that $T^{v(\alpha)}Qg \to Q^2 g = Qg$. Since Qg is rigid and T is mildly mixing, Qg is constant. In fact, as Q is an orthogonal projection, we have $Qg = \int g\, d\mu$. In other words, $T^{n(\alpha)}g$ converges weakly to $\int g\, d\mu$ as $\alpha \to \infty$, $\alpha \in \mathcal{F}^{(2)}$. The result follows. □

The following theorem is a special case of [11] (Lemma 1.2).

Proposition 2.6. *Let G be a commutative group with identity I and suppose that v is a VIP system into G. If $v(\alpha) = g \in G$ for every α in an IP ring $\mathcal{F}^{(1)}$ then $g = I$.*

The next theorem forms the bulk of the work required for our main result. It uses an inductive scheme originally used in [1], under the moniker "PET-induction", and which runs in our case as follows: given two VIP systems v and w, we write $v \sim w$ if $\deg v = \deg w > \deg(v - w)$. One easily checks that $\sim$ is an equivalence relation. Given a finite set $A = \{v_1, \ldots, v_k\}$ of VIP systems, define the *weight* of A by $\mathbf{w}(A) = (w_1, w_2, \ldots)$, where w_i is the number of equivalence classes of degree i VIP systems represented in A. Finally for distinct weights $\mathbf{w} = (w_1, w_2, \ldots)$ and $\mathbf{u} = (u_1, u_2, \ldots)$, one writes $\mathbf{w} > \mathbf{u}$ if $w_d > u_d$, where d is the largest j satisfying $w_j \neq u_j$. This is a well-ordering of the set of weights, and PET-induction is simply induction on it.

We shall use one other combinatorial fact as well. Recall that Hindman's theorem states that for any finite coloring of an IP ring $\mathcal{F}^{(1)}$, there exists a monochromatic refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$. In fact, for any finite coloring of the pairs $\{(\beta, \alpha) : \beta, \alpha \in \mathcal{F}^{(1)}, \beta < \alpha\}$, there exists a refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that $\{(\beta, \alpha) : \beta, \alpha \in \mathcal{F}^{(2)}, \beta < \alpha\}$ is monochromatic. (This is a special case of the Milliken-Taylor theorem. See [13, 14].)

Theorem 2.7. *Suppose $(X, \mathcal{A}, \mu, T)$ is a mildly mixing system, $k \in \mathbf{N}$, and let $v_1, \ldots, v_k$ be VIP systems into $\mathbf{Z}$ such that neither v_i nor $v_i - v_j$ is identically zero on any IP subring of a given IP ring $\mathcal{F}^{(1)}$, $1 \leq i \neq j \leq k$. If $f_0, \ldots, f_k \in L^\infty(X)$ then there exists a refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that*

$$\underset{\alpha \in \mathcal{F}^{(2)}}{\text{IP-lim}} \int f_0 T^{v_1(\alpha)} f_1 \cdots T^{v_k(\alpha)} f_k \, d\mu = \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big).$$

Proof. The proof is by induction on the weight vector $\mathbf{w}(A)$ of $A = \{v_1, \ldots, v_k\}$. By Proposition 2.5, the result holds when the weight vector is $(1, 0, 0, \ldots)$. Suppose for induction that the result holds for families having weight vector $\mathbf{w} < \mathbf{w}(A)$.

For standard reasons, we may assume without loss of generality that $\int f_a \, d\mu = 0$ for some a, $0 \leq a \leq k$. We reduce the general case to this special case by employing the identity

$$\prod_{i=0}^{k} a_i - \prod_{i=0}^{k} b_i = (a_0 - b_0) \prod_{i=1}^{k} b_i + a_0(a_1 - b_1) \prod_{i=2}^{k} b_i + \cdots + \Big(\prod_{i=0}^{k-1} a_i \Big)(a_t - b_t)$$

under the integral, with $a_i = T^{v_i(\alpha)} f_i$ and $b_i = \int f_i \, d\mu$, $0 \leq i \leq k$. Indeed, by composing through by $T^{-v_i(\alpha)}$, if necessary, where v_i is of minimal degree (this does not change the weight vector), we may in fact assume that $1 \leq a \leq k$. Also without loss of generality we may assume that $\|f_i\|_\infty \leq 1$, $0 \leq i \leq k$.

We shall complete the proof by showing that for some refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$, $\prod_{i=1}^{k} T^{v_i(\alpha)} f_i \to 0$ weakly. Using Lemma 2.4, with $x_\alpha =$

$\prod_{i=1}^k T^{v_i(\alpha)} f_i$, it will suffice for this purpose to show that (by passing to a refinement if necessary),

$$\begin{aligned}
&\underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \langle x_\alpha, x_{\alpha\cup\beta}\rangle \\
&= \underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \prod_{i=1}^k T^{v_i(\alpha)} f_i \prod_{i=1}^k T^{v_i(\alpha\cup\beta)} f_i \, d\mu \qquad (2.1)\\
&= \underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \prod_{i=1}^k T^{v_i(\alpha)} f_i \prod_{i=1}^k T^{v_i(\alpha)+v_i^\beta(\alpha)} (T^{v_i(\beta)} f_i)\, d\mu = 0,
\end{aligned}$$

where $v_i^\beta(\alpha) = v_i(\alpha\cup\beta) - v_i(\alpha) - v_i(\beta)$. Recall that $\deg(v_i^\beta) < \deg v_i$, so that in particular $v_i + v_i^\beta \sim v_i$.

We claim we can pass now to a refinement (we continue to call it $\mathcal{F}^{(1)}$) having the property that for all $\beta, \alpha \in \mathcal{F}^{(1)}$ with $\beta < \alpha$, $v_i^\beta(\alpha) \neq v_j(\alpha) - v_i(\alpha)$ for all i, j. For, otherwise, by Milliken-Taylor, we could pass to a refinement for which there exist i, j such that for all $\beta, \alpha \in \mathcal{F}^{(1)}$ with $\beta < \alpha$, $v_i^\beta(\alpha) = v_j(\alpha) - v_i(\alpha)$, which would require that the VIP system $\beta \to v_i^\beta$ from $\mathcal{F}^{(1)}$ into $\Omega(\mathbf{Z}, \mathcal{F}^{(1)})$ take on the constant value $v_j - v_i$, which by hypothesis is not equal to I in violation of Proposition 2.6. It follows that for any $\beta \in \mathcal{F}^{(1)}$ there exists no IP subring of $\mathcal{F}^{(1)}$ restricted to which $v_i^\beta = v_j - v_i$. Similar considerations achieve the same conclusion regarding the equations $v_i^\beta - v_j^\beta = v_j - v_i$, $1 \le i \neq j \le k$.

Again by a similar consideration, we may assume that v_i^β is either the identity for all $\beta \in \mathcal{F}^{(1)}$ (as will happen when v_i is of degree one), or for no $\beta \in \mathcal{F}^{(1)}$. Let w be the number of indices for which the former occurs. Permuting indices so that $\deg v_i$ is non-decreasing with i, we may assume $v_i^\beta = I$ if $1 \le i \le w$ and $v_i^\beta \neq I$ if $w < i \le k$. For $\beta \in \mathcal{F}^{(1)}$, write $A^\beta = \{v_1, \cdots, v_k, v_{w+1} + v_{w+1}^\beta, \cdots, v_k + v_k^\beta\}$. By the facts obtained in the previous paragraph, A^β is a mixing set, meaning there is no IP subring of $\mathcal{F}^{(1)}$ on which some member of A^β, or some difference of two members, is constant. Moreover, $\mathbf{w}(A^\beta) = \mathbf{w}(A)$.

The double limit in the last line of (2.1) may be rewritten as

$$\begin{aligned}
&\underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \prod_{i=1}^w T^{v_i(\alpha)} (f_i T^{v_i(\beta)} f_i) \qquad (2.2)\\
&\qquad \prod_{i=w+1}^k T^{v_i(\alpha)} f_i T^{v_i(\alpha)+v_i^\beta(\alpha)} (T^{v_i(\beta)} f_i)\, d\mu.
\end{aligned}$$

For fixed $\beta \in \mathcal{F}^{(1)}$ the set $B^\beta = \{v_2 - v_1, \ldots, v_w - v_1, v_{w+1} - v_1, \ldots, v_k - v_1, v_{w+1} + v^\beta_{w+1} - v_1, \cdots, v_k + v^\beta_k - v_1\}$ precedes A^β. The reason for this is that v_1 is of minimal weight, so that subtracting throughout by v_1 will decrease the degree of every VIP system that is equivalent to v_1, while failing to change the degrees of the other ones. Moreover, if neither v_i nor v_j is equivalent to v_1, then $v_i \sim v_j$ if and only if $v_i - v_1 \sim v_j - v_1$. These considerations imply that B^β has one less equivalence class under $\sim$ of degree $\deg v_1$ and the same number of equivalence classes at any degree greater than $\deg v_1$.

At any rate, by the fact that T is measure preserving, and making use of the induction hypothesis to pass to the limit in α, (2.2) may now be rewritten

$$\underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}}\ \underset{\alpha\in\mathcal{F}^{(1)}}{\text{IP-lim}} \int \prod_{i=1}^{w} T^{v_i(\alpha)-v_1(\alpha)}(f_i T^{v_i(\beta)} f_i)$$
$$\prod_{i=w+1}^{k} T^{v_i(\alpha)-v_1(\alpha)} f_i T^{v_i(\alpha)+v_i^\beta(\alpha)-v_1(\alpha)}(T^{v_i(\beta)} f_i)\, d\mu$$
$$= \underset{\beta\in\mathcal{F}^{(1)}}{\text{IP-lim}} \prod_{i=1}^{w} \Big(\int f_i T^{v_i(\beta)} f_i\, d\mu\Big) \prod_{i=w+1}^{k} \Big(\int f_i\, d\mu\Big)^2 = \prod_{i=1}^{k} \Big(\int f_i\, d\mu\Big)^2 = 0,$$

as required.

□

We now come to our main result. Denote by $\mathcal{G}_{\mathrm{NC}}$ the set of those $p \in \mathcal{G}$ that are constant on no IP set in $\mathbf{N}$.

Theorem 2.8. *Suppose $(X, \mathcal{A}, \mu, T)$ is a mildly mixing system, $k \in \mathbf{N}$, and let $p_1, \ldots, p_k$ be generalized polynomials such that $p_i \in \mathcal{G}_{NC}$ and $p_i - p_j \in \mathcal{G}_{NC}$, $1 \le i \ne j \le k$. If $f_0, \ldots, f_k \in L^\infty(X)$ then*

$$\underset{n\in\mathbf{N}}{IP^*\text{-lim}} \int f_0 T^{p_1(n)} f_1 \cdots T^{p_k(n)} f_k\, d\mu = \prod_{i=0}^{k} \Big(\int f_i\, d\mu\Big).$$

Proof. Let n be an arbitrary IP system into $\mathbf{N}$ and let $\epsilon > 0$ be arbitrary. By Proposition 2.3 there exist an IP ring $\mathcal{F}^{(1)}$ and constants c_i, $1 \le i \le k$, such that $v_i(\alpha) = p_i\big(n(\alpha)\big) - c_i$ define VIP systems, and since moreover $p_i \in \mathcal{G}_{\mathrm{NC}}$ and $p_i - p_j \in \mathcal{G}_{\mathrm{NC}}$, we may assume as well (invoking Theorem 2.1 in the extended integers) that $|v_i(\alpha)| \to \infty$ and $|v_i(\alpha) - v_j(\alpha)| \to \infty$, $\alpha \in \mathcal{F}^{(1)}$, $1 \le i \ne j \le k$. Thus Theorem 2.7

applies, so there exists a refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ with

$$\operatorname*{IP-lim}_{\alpha\in\mathcal{F}^{(2)}} \int f_0 T^{v_1(\alpha)} f_1 \cdots T^{v_k(\alpha)} f_k \, d\mu = \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big).$$

This means that for some $\alpha \in \mathcal{F}^{(2)}$,

$$\begin{aligned}
&\Big| \int f_0 T^{p_1(n(\alpha))} f_1 \cdots T^{p_k(n(\alpha))} f_k \, d\mu - \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big) \Big| \\
=&\Big| \int f_0 T^{p_1(n(\alpha))-c_1} (T^{c_1} f_1) \cdots T^{p_k(n(\alpha))-c_k} (T^{c_k} f_k) \, d\mu - \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big) \Big| \\
=&\Big| \int f_0 T^{v_1(\alpha)} (T^{c_1} f_1) \cdots T^{v_k(\alpha)} (T^{c_k} f_k) \, d\mu - \prod_{i=0}^{k} \Big(\int T^{c_i} f_i \, d\mu \Big) \Big| < \epsilon,
\end{aligned}$$

as required. □

Theorem C from the introduction follows from Theorem 2.8 by letting $f = 1_A$.

3. A partial characterization of $\mathcal{G}_{\mathrm{NC}}$ in the degree 2 case.

In this section we offer conditions for identifying whether or not a generalized polynomial of degree two belongs to $\mathcal{G}_{\mathrm{NC}}$. The most general form we consider is

$$\begin{aligned}
H(n) = &\sum_{i=1}^{n_1} [a_i n [b_i n + u_i] + v_i] + \sum_{i=1}^{n_2} [c_i n + w_i][d_i n + x_i] \\
&+ \sum_{i=1}^{n_3} [e_i n + y_i] + \sum_{i=1}^{n_4} [p_i n^2 + z_i].
\end{aligned} \tag{3.1}$$

Any generalized polynomial of order at most 2 can be reduced, up to a bounded error term, to this form, excepting those which contain iterated expressions of the type

$$G(n) = [a_1[a_2[\cdots[a_k n]]\cdots]][b_1[b_2[\cdots[b_l n]]\cdots]].$$

Cases of this type are usually considered pathological; Bergelson and Håland, for example, exclude them from their treatment.

We will write $[x]$ for the integer part of x and $\{x\}$ for the fractional part of x. We also put $[\![x]\!] = [x+\frac{1}{2}]$ for x rounded to the nearest integer and we let $\{\!\{x\}\!\} = x - [\![x]\!]$. We will use $\otimes$ to denote the $\mathbf{Q}$-linear tensor product. We will consider $\mathbf{R}\otimes\mathbf{R}$ so that $1\otimes 2 = 2\otimes 1 = 2(1\otimes 1)$, but $\sqrt{2}\otimes\sqrt{3} \neq \sqrt{3}\otimes\sqrt{2}$.

For the moment we shall restrict attention to the subclass obtained by requiring $u_i = v_i = w_i = x_i = y_i = z_i = \frac{1}{2}$ in (3.1); in other words, where $[\![\cdot]\!]$ replaces $[\cdot]$ and these constants are missing. Using generalized polynomials with rounding will cause us a few problems we shall have to deal with when we return to the general form (3.1) later.

There are however some advantages. For example, given real numbers a_i, $1 \le i \le k$, and any IP system n into $\mathbf{N}$, there exists an IP ring $\mathcal{F}^{(1)}$ such that

$$\text{IP-}\lim_{\alpha \in \mathcal{F}^{(1)}} \{\!\{a_i n(\alpha)\}\!\} = 0, \quad 1 \le i \le k. \tag{3.2}$$

The corresponding statement for $\{\cdot\}$ fails (consider a case where, for example, $a_2 = -a_1$). If $|x_i| < 1/(2|n|)$ for $i = 1, \ldots, n$, we see that $\{\!\{x_1 + \ldots + x_n\}\!\} = \{\!\{x_1\}\!\} + \ldots + \{\!\{x_n\}\!\}$. It follows that if $|x| < 1/(2|n|)$ then $\{\!\{nx\}\!\} = n\{\!\{x\}\!\}$ (where n is an integer). A corollary of these observations is the following fact that we exploit in several places:

Lemma 3.1. *Suppose that $(a_i)_{1\le i\le m}$ and $(b_j)_{1\le j\le n}$ are related by the equalities $a_i = \sum_j C_{ij} b_j$. Then given any IP-system n into $\mathbf{N}$, there is an IP-ring $\mathcal{F}^{(1)}$ such that*

$$\{\!\{a_i n(\alpha)\}\!\} = \sum_j C_{ij} \{\!\{b_j n(\alpha)\}\!\} \quad \textit{for all } 1 \le i \le m \textit{ and } \alpha \in \mathcal{F}^{(1)}.$$

Suppose we are given sequences of reals $(a_i)_{i=1}^{n_1}$, $(b_i)_{i=1}^{n_1}$, $(c_i)_{i=1}^{n_2}$, $(d_i)_{i=1}^{n_2}$, $(e_i)_{i=1}^{n_3}$ and $(p_i)_{i=1}^{n_4}$. Define

$$F(n) = \sum_{i=1}^{n_1} [\![a_i n [\![b_i n]\!]]\!] + \sum_{i=1}^{n_2} [\![c_i n]\!][\![d_i n]\!] + \sum_{i=1}^{n_3} [\![e_i n]\!] + \sum_{i=1}^{n_4} [\![p_i n^2]\!]. \tag{3.3}$$

Then one has the following:

Theorem 3.2. *Let F be as in (3.3). Suppose there is an IP system n into $\mathbf{N}$ such that $F \circ n$ is constant. Then the following three conditions must be satisfied:*

$$\sum_{i=1}^{n_1} a_i b_i + \sum_{i=1}^{n_2} c_i d_i + \sum_{i=1}^{n_4} p_i = 0; \tag{1}$$

$$\sum_{i=1}^{n_3} e_i = 0; \tag{2}$$

$$\sum_{i=1}^{n_1} a_i \otimes b_i + \sum_{i=1}^{n_2} (c_i \otimes d_i + d_i \otimes c_i) \in \mathbf{R} \otimes \mathbf{Q}. \tag{3}$$

Conversely, suppose (1), (2) and (3) are satisfied. Then for any IP system n into $\mathbf{N}$, there exists an IP ring $\mathcal{F}^{(1)}$ such that $F \circ n$ is constant on $\mathcal{F}^{(1)}$.

We shall prove Theorem 3.2 with the help of the following lemma, whose proof is, for the moment, deferred.

Lemma 3.3. *Let $(a_i)_{i=1}^k$ and $(b_i)_{i=1}^k$ be real numbers. Then the following are equivalent*

- **(i):** *There is an IP system n into $\mathbf{N}$ such that $\sum_i a_i \{\!\{b_i n(\alpha)\}\!\} = 0$ for all $\alpha \in \mathcal{F}$.*
- **(ii):** *Given any IP system n into $\mathbf{N}$, there exists an IP ring $\mathcal{F}^{(1)}$ such that $\sum_i a_i \{\!\{b_i n(\alpha)\}\!\} = 0$ for all $\alpha \in \mathcal{F}^{(1)}$.*
- **(iii):** $\sum_i a_i \otimes b_i \in \mathbf{R} \otimes \mathbf{Q}$.

Proof of Theorem 3.2. We start by observing the following:

(3.4)

$$\begin{aligned} F(n) &= \sum_{i=1}^{n_1} [\![a_i b_i n^2 - a_i n \{\!\{b_i n\}\!\}]\!] + \sum_{i=1}^{n_3} (e_i n - \{\!\{e_i n\}\!\}) \\ &\quad + \sum_{i=1}^{n_4} (p_i n^2 - \{\!\{p_i n^2\}\!\}) \qquad + \sum_{i=1}^{n_2} (c_i n - \{\!\{c_i n\}\!\})(d_i n - \{\!\{d_i n\}\!\}) \\ &= n^2 \left(\sum_{i=1}^{n_1} a_i b_i + \sum_{i=1}^{n_2} c_i d_i + \sum_{i=1}^{n_4} p_i \right) + n \sum_{i=1}^{n_3} e_i + f(n) \\ &\qquad + n \left(\sum_{i=1}^{n_1} a_i \{\!\{b_i n\}\!\} + \sum_{i=1}^{n_2} (c_i \{\!\{d_i n\}\!\} + d_i \{\!\{c_i n\}\!\}) \right), \end{aligned}$$

where $f(n)$ is bounded.

Suppose that F is constant on the range of some IP system n. We must establish (1), (2) and (3). Since the coefficient of n in the above expression is bounded, it is immediately clear that in order for the expression to be constant, the coefficient of n^2 must be zero, yielding (1). Restricting now to an IP ring $\mathcal{F}^{(1)}$, we may assume that the $\{\!\{b_i n\}\!\}$, $\{\!\{c_i n\}\!\}$, $\{\!\{d_i n\}\!\}$ terms converge to 0. Hence in order for $F \circ n$ to be constant, it is required that $\sum e_i = 0$, yielding (2).

Define $G(n)$ by

$$G(n) = \sum_{i=1}^{n_1} a_i \{\!\{b_i n\}\!\} + \sum_{i=1}^{n_2} (c_i \{\!\{d_i n\}\!\} + d_i \{\!\{c_i n\}\!\}).$$

We require $nG(n)$ to be bounded on $S = n(\mathcal{F}^{(1)})$ so that $G(n) \leq \frac{C}{n}$ for some C. Choose pairwise disjoint $\alpha_k \in \mathcal{F}^{(1)}$ and let $n_k = n(\alpha_k)$, so that $|\{\!\{b_i n_k\}\!\}|$, $|\{\!\{c_i n_k\}\!\}|$, $|\{\!\{d_i n_k\}\!\}|$ are less than 2^{-k} for each i. Fix $n \in S$. Now for large k, $m = n + n_k \in S$ so that $G(m) \leq \frac{C}{m}$. On the other hand, for large k, $G(n+n_k) = G(n)+G(n_k)$; it follows that $G \circ n$ is identically zero on $\mathcal{F}^{(1)}$ and so by Lemma 3.3 we get (3).

Conversely, suppose (1), (2) and (3) are satisfied and let n be an IP system into $\mathbf{N}$. Looking at (3.4), we see that the only non-trivial, potentially unbounded term of $F \circ n$ is $n\big(G \circ n\big)$. But according to Lemma 3.3, $G \circ n$ is 0 restricted to some IP ring $\mathcal{F}^{(1)}$. A routine application of Hindman's theorem finishes the proof.

□

We need the following for the proof of Lemma 3.3.

Lemma 3.4. *Let $f_1, \ldots, f_\ell$ be independent over $\mathbf{Q}$. For $n \in \mathbf{N}$, define the vector $\mathbf{g}_n$ by $(\mathbf{g}_n)_i = [\![f_i n]\!]$. Then for any infinite set $S \subset \mathbf{N}$, $\{\mathbf{g}_n : n \in S\}$ spans $\mathbf{Q}^\ell$.*

Proof. Suppose not. Then the vectors $(\mathbf{g}_n)_{n\in S}$ span a proper subspace of $\mathbf{Q}^\ell$. Pick $\xi \in \mathbf{Q}^\ell$ such that $\langle \xi, \mathbf{g}_n \rangle = 0$ for each $n \in S$. Dividing by n and taking the limit, we get $\xi_1 f_1 + \cdots + \xi_\ell f_\ell = 0$, a contradiction. □

Proof of Lemma 3.3. Adding an index $k+1$, if necessary, with $a_{k+1} = 0$ and $b_{k+1} = 1$, doesn't change the truth value of any of the conditions, so we may assume without loss of generality that 1 is in the rational linear span of the (b_i). Let $e_1, e_2, \ldots, e_m$ be a basis (over $\mathbf{Q}$) for the subspace of $\mathbf{R}$ spanned by (a_i). Similarly, let $f_1, \ldots, f_\ell$ be a basis for the subspace of $\mathbf{R}$ spanned by b_i. These bases may be chosen in such a way that the a_i and b_i are integer combinations of the e_j and f_j. Specifically, we will have $a_i = \sum_j A_{ij} e_j$ and $b_i = \sum_k B_{ik} f_k$, where the A_{ij} and B_{ik} are integers. As noted above, 1 is assumed to be in the rational linear span of the (b_i) and hence of the (f_i). We will assume that $f_1 \in \mathbf{Q}$. Write $C_{jk} = \sum_i A_{ij} B_{ik}$, so that C is an $m \times \ell$ matrix with integer coefficients.

Assume that condition (i) holds. We shall establish (iii). Using Lemma 3.1, choose a subsystem n of the system guaranteed by (i) such that for every $\alpha \in \mathcal{F}$, $\{\!\{b_i n(\alpha)\}\!\} = \{\!\{\sum_k B_{ik} f_k n(\alpha)\}\!\} = \sum_k B_{ik} \{\!\{f_k n(\alpha)\}\!\}$. Put S for the range of n.

For $n \in S$ one has

$$\begin{aligned} 0 = \sum_i a_i \{\!\{ b_i n \}\!\} &= \sum_i \sum_j A_{ij} e_j \sum_k B_{ik} \{\!\{ f_k n \}\!\} \\ &= \sum_{j,k} \left(\sum_i A_{ij} B_{ik} \right) e_j \{\!\{ f_k n \}\!\} \\ &= \sum_{j,k} C_{jk} e_j \{\!\{ f_k n \}\!\}. \end{aligned}$$

Let $\mathbf{e}$ denote the column vector having coordinates $e_1, \dots, e_m$, let $\mathbf{f}$ denote the column vector having coordinates $f_1, \dots, f_\ell$, and let $\mathbf{g}_n$ denote the column vector whose ith coordinate is $[\![n f_i]\!]$.

By assumption, we have $\mathbf{e}^T C(n\mathbf{f} - \mathbf{g_n}) = 0$ for all $n \in S$. Let $\mathbf{h} = \mathbf{e}^T C$. We consider two cases: either $\langle \mathbf{h}, \mathbf{f} \rangle = 0$ or not. In the latter case, we write $\langle \mathbf{h}, \mathbf{f} \rangle = \gamma$ and we set $\tilde{\mathbf{h}} = \frac{\mathbf{h}}{\gamma}$ so that $\langle \tilde{\mathbf{h}}, \mathbf{f} \rangle = 1$.

If $\langle \mathbf{h}, \mathbf{f} \rangle = 0$, then we must have $\langle \mathbf{h}, \mathbf{g}_n \rangle = 0$ for all $n \in S$. However since by Lemma 3.4 the $\mathbf{g}_n$ span $\mathbf{Q}^\ell$, as real vectors they span $\mathbf{R}^\ell$. It follows that $\mathbf{h} = 0$. Since the e_i were assumed to be independent, it then follows that the matrix C was 0 so that $\sum_i a_i \otimes b_i = \sum_{j,k} C_{j,k} e_j \otimes f_k = 0 \in \mathbf{R} \otimes \mathbf{Q}$.

If $\langle \tilde{\mathbf{h}}, \mathbf{f} \rangle = 1$, then we must have $\langle \tilde{\mathbf{h}}, \mathbf{g}_n \rangle = n$ for all $n \in S$. Since the $\mathbf{g}_n$ span $\mathbf{Q}^\ell$, it follows that the map $\mathbf{x} \mapsto \langle \tilde{\mathbf{h}}, \mathbf{x} \rangle$ maps points in $\mathbf{Q}^\ell$ to rational values, so that $\tilde{\mathbf{h}}$ is a rational vector. Since $f_1, f_2, \dots, f_\ell$ are independent over $\mathbf{Q}$, and $\langle \tilde{\mathbf{h}}, \mathbf{f} \rangle$ is a rational multiple of f_1, the (rational) vector $\tilde{\mathbf{h}}$ must have a non-zero first entry and zero in all other entries and so the same is true of $\mathbf{h} = \mathbf{e}^T C$. Since the coordinates of $\mathbf{e}$ were assumed to be rationally independent and C has integer entries, it follows that all columns of C except for the first must be 0.

This yields

$$\begin{aligned}\sum_i a_i \otimes b_i &= \sum_i \left(\sum_j A_{ij} e_j\right) \otimes \left(\sum_k B_{ik} f_k\right)\\ &= \sum_{j,k} \left(\sum_i A_{ij} B_{ik}\right) e_j \otimes f_k\\ &= \sum_{j,k} C_{jk} e_j \otimes f_k\\ &= \sum_j C_{j1} e_j \otimes f_1\\ &= \left(\sum_j C_{j1} e_j\right) \otimes f_1 \in \mathbf{R} \otimes \mathbf{Q},\end{aligned}$$

where in the second equality, we used the fact that $\otimes$ is $\mathbf{Q}$-linear in both factors. Thus we have shown that (i) implies (iii).

We now show that (iii) implies (ii). Assume that condition (iii) holds. Let the denominator of f_1 (the rational element of the basis) be q. We have

$$\sum_i a_i \otimes b_i = \sum_{j,k} C_{jk} e_j \otimes f_k = \sum_k \left(\sum_j C_{jk} e_j\right) \otimes f_k.$$

Since this sum was assumed to be in $\mathbf{R} \otimes \mathbf{Q}$, it follows that the terms involving $f_2, \ldots, f_\ell$ in this sum must vanish, so that for $k \geq 2$, $\sum_j C_{jk} e_j = 0$. Since the e_j are assumed to be independent over $\mathbf{Q}$ and the C_{jk} are integers, it follows that $C_{jk} = 0$ for $k \geq 2$, so that C has non-zero entries only in the first column.

Let n be an IP system into $\mathbf{N}$. Choose an IP ring $\mathcal{F}^{(1)}$ such that $n(\alpha) \in q\mathbf{N}$ for all $\alpha \in \mathcal{F}^{(1)}$ and note that for all $\alpha \in \mathcal{F}^{(1)}$ one has $\{\!\{ f_1 n(\alpha) \}\!\} = 0$. Now using Lemma 3.1, pass to a further sub-ring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that for all $\alpha \in \mathcal{F}^{(2)}$,

$$\{\!\{ b_i n(\alpha) \}\!\} = \{\!\{ \sum_k B_{ik} f_k n(\alpha) \}\!\} = \sum_k B_{ik} \{\!\{ f_k n(\alpha) \}\!\}.$$

Thus for $\alpha \in \mathcal{F}^{(2)}$,

$$\sum_i a_i \{\!\{ b_i n \}\!\} = \sum_{j,k} C_{jk} e_j \{\!\{ f_k n \}\!\} = \sum_j C_{j1} e_j \{\!\{ f_1 n \}\!\} = 0,$$

as required. Thus we have shown that (iii) implies (ii). Since (ii) obviously implies (i), we are done.

□

We now prepare to apply Theorem 3.2 to generalized polynomials expressed with the greatest integer function (not rounding). Let $a_1, \dots, a_k \in \mathbf{R} \setminus \mathbf{Q}$. Choose a basis $e_1, \dots, e_s$ for the $\mathbf{Q}$-linear span of $\{a_1, \dots, a_k\}$ over $\mathbf{Q}$ in such a way that $a_i = \sum_j A_{ij} e_j$, $1 \le i \le k$, where the A_{ij} are integers. Write A for the $k \times s$ matrix (A_{ij}).

Proposition 3.5. *Let $\epsilon_i \in \{-1, 1\}$, $1 \le i \le k$. The following are equivalent.*

(1) *There exists an IP system n into $\mathbf{N}$ such that $sgn \{\!\{ a_i n(\alpha) \}\!\} = \epsilon_i$, $1 \le i \le k$ and $\alpha \in \mathcal{F}$.*
(2) *There exists a column vector $\mathbf{u} \in \mathbf{R}^s$ such that, letting $\mathbf{v} = (v_1, \dots, v_k)^T = A\mathbf{u}$, one has $sgn\, v_i = \epsilon_i$, $1 \le i \le k$.*

Proof. Suppose (1) holds. One has $na_i = \sum_j A_{ij}(ne_j)$. Again, using Lemma 3.1, choose an IP-ring $\mathcal{F}^{(1)}$ such that for all $\alpha \in \mathcal{F}^{(1)}$, $\{\!\{ n(\alpha) a_i \}\!\} = \sum_j A_{ij} \{\!\{ n(\alpha) e_j \}\!\}$. Fixing such an α and letting $\mathbf{u}$ be the column vector whose coordinates are given by $\mathbf{u}_j = \{\!\{ n(\alpha) e_j \}\!\}$, $1 \le j \le s$, one has $A\mathbf{u} = \mathbf{v} = (v_1, \dots, v_k)$, where $v_i = \{\!\{ n(\alpha) a_i \}\!\}$, $1 \le i \le k$. Thus (2) follows from (1).

Next suppose (2) holds. We may assume without loss of generality that both $\mathbf{u}$ and $\mathbf{v}$ have sufficiently small modulus to guarantee the linearity we presume in the expressions to follow. For $m \in \mathbf{N}$, choose by density of $(ne_1, \dots, ne_s)$ on the torus some $n_m \in \mathbf{N}$ such that the vector $\mathbf{h} = (\{\!\{ n_m e_1 \}\!\}, \dots, \{\!\{ n_m e_s \}\!\})^T$ is sufficiently close to $2^{-m}\mathbf{u}$ to ensure that $A\mathbf{h} = (\{\!\{ n_m a_1 \}\!\}, \dots, \{\!\{ n_m a_k \}\!\})^T$ is very close to $2^{-m}\mathbf{v}$, in particular, close enough that all coordinates have the same sign. Now for $\alpha \in \mathcal{F}$, let $n(\alpha) = \sum_{m \in \alpha} n_m$. □

Remark. Proposition 3.5 restricts the (a_i) to irrational values; note that $\{\!\{ a_i n \}\!\} = 0$ can always be arranged along IP systems n for rational values of a_i.

Now consider the case of a generalized polynomial having the form

$$F(n) = \sum_{i=1}^{n_1} [a_i n [b_i n]] + \sum_{i=1}^{n_2} [c_i n][d_i n] + \sum_{i=1}^{n_3} [e_i n] + \sum_{i=1}^{n_4} [p_i n^2]. \tag{3.5}$$

We would like to determine whether or not there is an IP system n into $\mathbf{N}$ such that $F \circ n$ is constant. Our first (obvious) observation is that we can change three sets of outer brackets to rounding brackets, changing the value by a uniform bounded amount.

$$F(n)=\sum_{i=1}^{n_1}[\![a_in[b_in]]\!]+\sum_{i=1}^{n_2}[c_in][d_in]+\sum_{i=1}^{n_3}[\![e_in]\!]+\sum_{i=1}^{n_4}[\![p_in^2]\!]+B(n),$$

where $B(n)$ is bounded. (By Hindman's theorem we can make B constant along relevant IP systems so we'll just omit the $B(n)$ term in what follows.) In changing the remaining brackets to rounding brackets, we do introduce potentially significant new terms:

$$F(n)=\sum_{i=1}^{n_1}[\![a_in([\![b_in]\!]+\epsilon_i(n))]\!]+\sum_{i=1}^{n_2}([\![c_in]\!]+\delta_i(n))([\![d_in]\!]+\gamma_i(n))$$
$$+\sum_{i=1}^{n_3}[\![e_in]\!]+\sum_{i=1}^{n_4}[\![p_in^2]\!],$$

where now $\epsilon_i(n)=[b_in]-[\![b_in]\!]$, $\delta_i(n)=[c_in]-[\![c_in]\!]$ and $\gamma_i(n)=[d_in]-[\![d_in]\!]$ take values in $\{-1,0\}$. Now, again using Hindman's theorem, we can choose constants x_i, y_i and z_i in $\{-1,0\}$ and an IP system n into $\mathbf{N}$ such that $\epsilon_i\circ n=x_i$, $\delta_i\circ n=y_i$ and $\gamma_i\circ n=z_i$ on $\mathcal{F}$. Now for $n=n(\alpha)$ one has:

$$\begin{aligned}F(n)=&\sum_{i=1}^{n_1}[\![a_in([\![b_in]\!]+x_i)]\!]+\sum_{i=1}^{n_2}([\![c_in]\!]+y_i)([\![d_in]\!]+z_i)\\&+\sum_{i=1}^{n_3}[\![e_in]\!]+\sum_{i=1}^{n_4}[\![p_in^2]\!]\\=&\sum_{i=1}^{n_1}[\![a_in[\![b_in]\!]]\!]+\sum_{i=1}^{n_1}[\![x_ia_in]\!]+\sum_{i=1}^{n_2}[\![c_in]\!][\![d_in]\!]+\sum_{i=1}^{n_2}[\![d_iy_in]\!]\\&+\sum_{i=1}^{n_2}[\![c_iz_in]\!]+\sum_{i=1}^{n_3}[\![e_in]\!]+\sum_{i=1}^{n_4}[\![p_in^2]\!]+B'(n),\end{aligned}$$

where B' is bounded and hence negligible. Applying now the condition of Theorem 3.2 to the above form yields the following.

Theorem 3.6. *Let F be given by* (3.5). *There exists an IP system n into $\mathbf{N}$ such that $F\circ n$ is constant if and only if there are choices x_i, y_i and z_i in $\{-1,0\}$ having the following properties:*

(i) For some IP set $S\subset\mathbf{N}$, one has, for all $n\in S$, $[b_in]-[\![b_in]\!]=x_i$, $[c_in]-[\![c_in]\!]=y_i$ and $[d_in]-[\![d_in]\!]=z_i$.
(ii) These hold:

$$\sum_{i=1}^{n_1} a_i b_i + \sum_{i=1}^{n_2} c_i d_i + \sum_{i=1}^{n_4} p_i = 0; \tag{1}$$

$$\sum_{i=1}^{n_3} e_i + \sum_{i=1}^{n_1} x_i a_i + \sum_{i=1}^{n_2} (y_i d_i + z_i c_i) = 0; \tag{2}$$

$$\sum_{i=1}^{n_1} a_i \otimes b_i + \sum_{i=1}^{n_2} (c_i \otimes d_i + d_i \otimes c_i) \in \mathbf{R} \otimes \mathbf{Q}. \tag{3}$$

Notice that, since $[x] - [\![x]\!]$ is -1 when $sgn\ \{\!\{x\}\!\} = -1$ and 0 otherwise, condition (i) can be determinately checked for any candidate values of x_i, y_i and z_i using Proposition 3.5 and the remark following it; hence the necessary and sufficient conditions given in Theorem 3.6 are wholly explicit.

We now return to the general form $H(n)$ given in (3.1). Our first observation is that by moving integer parts outside of brackets we may assume without loss of generality that u_i, v_i, w_i, x_i, y_i and z_i are all in $[0, 1)$. Next, we may assume $v_i = y_i = z_i = \frac{1}{2}$ while changing $H(n)$ by a bounded amount. Finally, by (3.2), given any IP system n, every IP ring contains a subring along which the value of $H \circ n$ will be unaffected by setting all the non-zero u_i, w_i and x_i equal to $\frac{1}{2}$ (in other words, we may assume that $u_i, w_i, x_i \in \{0, \frac{1}{2}\}$). The most general form we must therefore consider is much as before, namely

(3.6)

$$F(n) = \sum_{i=1}^{n_1} [\![a_i n([\![b_i n]\!] + \epsilon_i(n))]\!] + \sum_{i=1}^{n_2} ([\![c_i n]\!] + \delta_i(n))([\![d_i n]\!] + \gamma_i(n))$$
$$+ \sum_{i=1}^{n_3} [\![e_i n]\!] + \sum_{i=1}^{n_4} [\![p_i n^2]\!],$$

where ϵ_i is either given by $\epsilon_i(n) = [b_i n] - [\![b_i n]\!]$ or is identically zero, δ_i is either given by $\delta_i(n) = [c_i n] - [\![c_i n]\!]$ or is identically zero, and γ_i is either given by $\gamma_i(n) = [d_i n] - [\![d_i n]\!]$ or is identically zero.

Theorem 3.7. *Let F be given by* (3.6). *There exists an IP system n into $\mathbf{N}$ such that $F \circ n$ is constant if and only if there are choices x_i, y_i and z_i in $\{-1, 0\}$ having the following properties:*

(i) For some IP set $S \subset \mathbf{N}$, one has, for all $n \in S$, $\epsilon_i(n) = x_i$, $\delta_i(n) = y_i$ and $\gamma_i(n) = z_i$.
(ii) (1), (2) and (3) above hold.

Theorem 3.7 gives an explicit answer to the question when a generalized polynomial F having form (3.6) has the property that $F \circ n$ is constant for some IP system n. On the other hand, as we have argued, if H has form (3.1) then there is some readily computable F having form (3.6) such that for any IP system n and any IP ring there is a subring on which $|H \circ n - F \circ n|$ is bounded. Invoking Hindman, the verdict that 3.7 gives in regard to F will apply equally to H.

4. A note on mild mixing $\mathbf{Z}^r$ actions

Up to now, we have restricted ourselves to $\mathbf{Z}$-actions, however our main theorem does have a $\mathbf{Z}^r$ version. Given a measure preserving $\mathbf{Z}^r$-action $\{T_{\mathbf{n}}\}$ of a probability space $(X, \mathcal{A}, \mu)$, and $f \in L^2(X)$, f is said to be rigid if there is a sequence $(\mathbf{n}_k) \subset \mathbf{Z}^r$ with $|\mathbf{n}_k| \to \infty$ such that $T_{\mathbf{n}_k} f \to f$. $\{T_{\mathbf{n}}\}$ is mild mixing if there are no non-constant rigid functions.

For fixed $l \in \mathbf{N}$, the set of *generalized polynomials* $\mathbf{Z}^l \to \mathbf{Z}$ is the smallest set $\mathcal{G}^{(l,1)}$ that is a function algebra (i.e. is closed under sums and products) containing $\mathbf{Z}[x_1, \ldots, x_l]$ and having the additional property that for all $m \in \mathbf{N}$, $c_1, \ldots, c_m \in \mathbf{R}$ and $p_1, \ldots, p_m \in \mathcal{G}^{(l,1)}$, the mapping $\mathbf{n} \to [\sum_{i=1}^m c_i p_i(\mathbf{n})]$ is in $\mathcal{G}^{(l,1)}$. A map $p : \mathbf{Z}^l \to \mathbf{Z}^r$ is a generalized polynomial if its coordinate functions are generalized polynomials, and we write $p \in \mathcal{G}^{(l,r)}$.

We now indicate how one would get a version of Theorem C for mild mixing $\mathbf{Z}^r$-actions and generalized polynomials $\mathbf{Z}^l \to \mathbf{Z}^r$. One ingredient is the following strengthening of Proposition 2.3.

Proposition 4.1. *(cf.* [3]*, Theorem 2.9). Let $p(x)$ be a generalized polynomial $\mathbf{Z}^l \to \mathbf{Z}^r$ and suppose n is a VIP system in $\mathbf{Z}^l$. Then for every IP ring $\mathcal{F}^{(1)}$ there exists an IP ring $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ and some $\mathbf{c} \in \mathbf{Z}^r$ such that the restriction to $\mathcal{F}^{(2)}$ of $v(\alpha) = p(n(\alpha)) + \mathbf{c}$ is a VIP system.*

Next, we state (without proof) a $\mathbf{Z}^r$ version of Theorem 2.7.

Theorem 4.2. *Suppose $(X, \mathcal{A}, \mu, \{T_{\mathbf{n}} : \mathbf{n} \in \mathbf{Z}^r\})$ is a mildly mixing system, $k \in \mathbf{N}$, and let $v_1, \ldots, v_k$ be VIP systems into $\mathbf{Z}^r$ such that neither v_i nor $v_i - v_j$ is identically zero on any IP subring of a given IP ring $\mathcal{F}^{(1)}$, $1 \le i \neq j \le k$. If $f_0, \ldots, f_k \in L^\infty(X)$ then there exists a refinement $\mathcal{F}^{(2)} \subset \mathcal{F}^{(1)}$ such that*

$$\operatorname*{IP\text{-}lim}_{\alpha \in \mathcal{F}^{(2)}} \int f_0 T_{v_1(\alpha)} f_1 \cdots T_{v_k(\alpha)} f_k \, d\mu = \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big).$$

Deriving a satisfactory (for our purposes) analog of Theorem 2.8 is a bit more of an adventure. Notice that Theorem B from the introduction contains no proviso vis-à-vis degeneration along IP sets, only one precluding constant polynomials from consideration. This is because non-constant polynomials $\mathbf{Z} \to \mathbf{Z}$ do not degenerate along IP sets, however such is not the case for polynomials $\mathbf{Z}^l \to \mathbf{Z}^r$. For example, let $p(x_1, x_2) = x_2 - x_1$ and let n be any IP system into $\mathbf{N}$. Then $\mathbf{n}(\alpha) \to \big(n(\alpha), n(\alpha)\big)$ defines an IP system into $\mathbf{Z}^2$ having the property that $p\big(\mathbf{n}(\alpha)\big) = 0$ for every $\alpha \in \mathcal{F}$.

This is all somewhat annoying, as we wish to formulate questions concerning potential degeneration of *generalized* polynomials $\mathbf{Z}^l \to \mathbf{Z}^r$, and we would like to have regular polynomials acting as a kind of ideal base case free of degeneration, as before. In other words, we are interested in which ways generalized polynomials $\mathbf{Z}^l \to \mathbf{Z}^r$ can degenerate along IP rings, however, we wouldn't like to count silly examples such as the foregoing one as "legitimate." This perspective can be adopted at minimal cost, as we now outline.

Note that the cause of degeneration in the example just considered is a linear dependence existing between the coordinate functions of $\mathbf{N}$. Accordingly, we agree to call an IP system $\mathbf{n}$ into $\mathbf{Z}^l$ "degenerate" if for some non-trivial vector $\mathbf{c} \in \mathbf{Z}^l$ and some IP ring $\mathcal{F}^{(1)}$, $\mathbf{c} \cdot \mathbf{n}(\alpha) = 0$ for every $\alpha \in \mathcal{F}^{(1)}$. (Here $\cdot$ denotes ordinary dot product.) An IP system that is not degenerate in this sense will be called an *NIP system.* NIP* subsets of $\mathbf{Z}^l$ and NIP*-limits are defined in the obvious ways.

Now by [4](Lemma 6.9), if $p : \mathbf{Z}^l \to \mathbf{Z}^r$ is a non-constant polynomial and $\mathbf{n}$ is an NIP system into $\mathbf{Z}^l$, the restriction of $\alpha \to p(\mathbf{n})$ to any IP ring $\mathcal{F}^{(1)}$ cannot be constant. This leads to the following theorem, which is a special case of [4](Theorem 6.10).

Theorem 4.3. *Suppose* $(X, \mathcal{A}, \mu, \{T_{\mathbf{n}} : \mathbf{n} \in \mathbf{Z}^r\})$ *is a mildly mixing system,* $k \in \mathbf{N}$*, and let* $p_1, \ldots, p_k$ *be polynomials* $\mathbf{Z}^l \to \mathbf{Z}^r$ *such that neither* p_i *nor* $p_i - p_j$ *is constant,* $1 \le i \ne j \le k$*. If* $f_0, \ldots, f_k \in L^2(X)$ *then*

$$\textit{NIP}^*\text{-}\lim_{\mathbf{n} \in \mathbf{Z}^l} \int f_0 T_{p_1(\mathbf{n})} f_1 \cdots T_{p_k(\mathbf{n})} f_k \, d\mu = \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big).$$

Now, denote by $\mathcal{G}^{(l,r)}_{\mathrm{NC}}$ the set of all $p \in \mathcal{G}^{(l,r)}$ having the property that $p(n(\alpha))$ isn't constant for any NIP system n into $\mathbf{Z}^l$. (Hence, as desired, polynomials $\mathbf{Z}^l \to \mathbf{Z}^r$ comprise a subclass of $\mathcal{G}^{(l,r)}_{\mathrm{NC}}$.) The following extension of Theorem 4.3 can now be established from Theorem 4.2, using Proposition 4.1.

Theorem 4.4. *Suppose* $(X, \mathcal{A}, \mu, \{T_{\mathbf{n}} : \mathbf{n} \in \mathbf{Z}^r\})$ *is a mildly mixing system,* $k \in \mathbf{N}$*, and let* $p_1, \ldots, p_k \in \mathcal{G}^{(l,r)}$ *such that* $p_i \in \mathcal{G}_{NC}^{(l,r)}$ *and* $p_i - p_j \in \mathcal{G}_{NC}^{(l,r)}$*,* $1 \le i \ne j \le k$*. If* $f_0, \ldots, f_k \in L^2(X)$ *then*

$$NIP^*\text{-}\lim_{\mathbf{n}\in\mathbf{Z}^l} \int f_0 T_{p_1(\mathbf{n})} f_1 \cdots T_{p_k(\mathbf{n})} f_k \, d\mu = \prod_{i=0}^{k} \Big(\int f_i \, d\mu \Big).$$

In order to appreciate Theorem 4.4 more fully, it would be nice to know something about $\mathcal{G}_{\mathrm{NC}}^{(l,r)}$ in some simple cases other than $l = r = 1$.

References

[1] V. Bergelson. Weakly mixing PET. *Ergodic Theory and Dynamical Systems*, 7:337–349, 1987.

[2] V. Bergelson, H. Furstenberg, and R. McCutcheon. IP-sets and polynomial recurrence. *Ergodic Theory and Dynamical Systems*, 16:963–974, 1996.

[3] V. Bergelson, I. Håland-Knutson, and R. McCutcheon. IP-systems, generalized polynomials and recurrence. *Ergodic Theory and Dynamical Systems*, 26:999–1020, 2006.

[4] V. Bergelson and R. McCutcheon. *An ergodic IP polynomial Szemerédi theorem.* Number 695 in Mem. Amer. Math. Soc. 2000.

[5] N. Franzikinakis and B. Kra. Polynomial averages converge to the product of integrals. *Israel J. Math.*, 148:267–276, 2005.

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

[7] H. Furstenberg. *Recurrence in ergodic theory and combinatorial number theory.* Princeton Univ. Press, 1981.

[8] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for IP-systems and combinatorial theory. *J. d'Anal. Math.*, 45:117–168, 1985.

[9] H. Furstenberg and B. Weiss. *The finite multipliers of infinite ergodic transformations*, pages 61–85. Springer, 1978.

[10] N. Hindman. Finite sums from sequences within cells of a partition of **N**. *J. Comb. Th. (Series A)*, 17:1–11, 1974.

[11] R. McCutcheon. FVIP systems and multiple recurrence. *Israel J. Math.*, 146:157–188, 2005.

[12] R. McCutcheon. An infinitary van der Waerden theorem. *J. Comb. Th. (Series A)*, 86:214–231, 1999.

[13] K. Milliken. Ramsey's theorem with sums or unions. *J. Comb. Th. (Series A)*, 18:276–290, 1975.

[14] A. Taylor. A canonical partition relation for finite subsets of ω. *J. Comb. Th. (Series A)*, 21:137–146, 1976.

E-mail address: `rmccutch(a)memphis.edu`

Department of Mathematical Sciences, University of Memphis

E-mail address: `aquas(a)math.uvic.ca`

Department of Mathematics and Statistics, University of Victoria