MODEL THEORY AND METRIC CONVERGENCE II:
AVERAGES OF UNITARY POLYNOMIAL ACTIONS

EDUARDO DUEÑEZ AND JOSÉ N. IOVINO

Abstract. Using Henson’s formalism of metric structures, we prove the pointwise convergence of averages of a polynomial sequence \( \{T_n\} \) (in Leibman’s sense) of unitary operators on a Hilbert space. (As a special, more concrete case, this applies to unitary sequences \( \{U^{p(n)}\} \) where \( p \) is a polynomial \( \mathbb{Z} \to \mathbb{Z} \) and \( U \) a fixed unitary operator.) Via a general principle of uniform metastability (implied by the compactness of Henson’s logic), such averages converge with universal rates of metastability. We also indicate how these results generalize to arbitrary Følner averages of unitary polynomial actions of abelian groups.

Introduction

The first result on “mean” convergence of averages was von Neumann’s 1932 Mean Ergodic Theorem [vN32]:

Mean Ergodic Theorem (MET). For any unitary operator \( U \) on a Hilbert space \( \mathcal{H} \) and any \( x \in \mathcal{H} \), the sequence \( AV_\bullet(x) = (AV_n(x) : n \in \mathbb{N}) \) of pointwise averages

\[
AV_n(x) = \frac{1}{n} \sum_{i=1}^{n} U^i(x)
\]

converges as \( n \to \infty \). The limit is equal to the orthogonal projection of \( x \) on the space of vectors fixed by \( U \).

Historically, generalizations of von Neumann’s theorem have largely followed a path influenced by a measure-theoretic viewpoint that is completely absent from the formulation above as a statement about convergence in Hilbert spaces. We provide further historical background below. Leaving history and measure theory aside for the moment, one may suggest the following different possible directions of generalization for MET:

1. Replace the sequence \( (U^i : i \in \mathbb{N}) \) with a “higher-degree” sequence \( (U^{p(i)} : i \in \mathbb{N}) \) where \( p \) is a fixed polynomial.

2. The sequence \( (T_i) = (U^{p(i)}) \) above necessarily satisfies the commutativity condition \( T_i \circ T_j = T_j \circ T_i \) for all \( i, j \). To what extent (and in what sense) can such commutativity requirement be removed?

3. What conditions on a family \( (T_i) \) of unitary operators indexed by a semigroup other than \( \mathbb{N} \) ensure the pointwise convergence of suitable averages?
Theorem 4 in this manuscript is arguably the most natural generalization of von Neumann’s result simultaneously in all three directions above. (For technical reasons, Theorem 4 is proved in the context of (polynomial) actions of groups rather than semigroups.)

Theorem 1 below is a very particular case of more general results (Theorems 2, 3 and 4). However, it is easiest to formulate and already generalizes MET all the way in direction (1) and beyond.

**Theorem 1** (MET for abelian unitary polynomial actions of $\mathbb{Z}$). Fix $d \in \mathbb{N}$. Let $\mathcal{H}$ be a Hilbert space, and let $U_0, U_1, \ldots, U_d$ be pairwise-commuting unitary operators on $\mathcal{H}$. For every $x \in \mathcal{H}$, the sequence
\[
\left( \text{AV}_n(x) = (AV_n(x) : n \in \mathbb{N}) \right)
\]
converges as $n \to \infty$.

In particular, if $p : \mathbb{Z} \to \mathbb{Z}$ is a polynomial of degree at most $d$ and $U$ is a unitary operator on $\mathcal{H}$, then \(\left( \sum_{0 \leq k \leq n} U_p^k(x)/(n+1) : n \in \mathbb{N} \right)\) converges.

Furthermore, there exists a universal metastability rate (depending only on $d$) that applies uniformly to all sequences of averages of arbitrary $x$ in the unit ball of an arbitrary Hilbert space $\mathcal{H}$ under arbitrary unitary operators $U_0, U_1, \ldots, U_d$ on $\mathcal{H}$.

The notion of uniformly metastable convergence above was first introduced in ergodic theory by Tao. It is a main theme of our prior manuscript, but shall presently play a minor role [DnI17, Tao08, Tao12].

Taking a step in direction (2), pairwise commutativity is not a necessary assumption; the sequence of averages under a family $(T_i)$ converges provided $i \mapsto T_i$ is a Leibman polynomial sequence in the group $U\mathcal{H}$ of unitary operators on $\mathcal{H}$ (Theorems 2 and 3), but the range of this sequence need not generate an abelian group. The definition of Leibman polynomial sequence (Definition 2.3) is motivated by the familiar fact that degree-$d$ polynomials $\mathbb{R} \to \mathbb{R}$ are characterized as those functions having $(d+1)$-iterated finite differences equal to zero. The same essential definition gives the notion of Leibman polynomial mapping from an arbitrary group $G$ into $U\mathcal{H}$ [Lei02]. Theorem 4 generalizes von Neumann’s result in direction (3) for Leibman polynomials $(T_i : i \in G) \subset U\mathcal{H}$ on abelian groups $G$ endowed with a notion of averaging provided by a countable Følner net.

Continuing our historical remarks, the formulation of von Neumann’s result above hides its conceptual genesis via the study of convergence of averages of square-integrable functions $f \in L^2(\Omega)$ on a probability space $(\Omega, \mu)$ under the action of a measure-preserving transformation $T$ of $\Omega$. In this setting, MET asserts that the sequence $\text{AV}_n(f)$ of averages
\[
\text{AV}_n(f) = \frac{1}{n} \sum_{i=1}^{n} f \circ T^i
\]
converges in $L^2(\Omega)$ (after all, $f \mapsto f \circ T$ is a unitary transformation of $L^2(\Omega)$). This particular case of von Neumann’s result explains why it is called a convergence result “in mean”, i.e., in the mean-square (“$L^2$”) sense. (By contrast, Birkhoff’s Ergodic Theorem asserts the almost-everywhere pointwise convergence of the averages $\text{AV}_n(f)$ for any $f \in L^1(\Omega)$ [Bir31].) The $L^2$ setting entails no loss of generality since every Hilbert space $\mathcal{H}$ is realized as a space of square-integrable functions. However, this viewpoint is artificial for

\[\binom{k}{j} = k(k-1) \cdots (k-j+1)/j!\] is the $j$-th binomial coefficient.
purposes of studying convergence under unitary actions (at least insofar as simple actions are concerned, in contrast to multiple actions mentioned below).

Although generalizations of MET in direction (1) seem very natural, we are not aware of direct proofs of Theorem 1 but only of indirect proofs as byproduct of results on mean convergence of “multiple” ergodic averages. Starting in the 1970’s, Furstenberg pioneered the ergodic study of actions of multiple simultaneous transformations; equivalently, the study of convergence of “multiple averages” of the product of two or more measurable bounded functions on a probability space $\Omega$ as acted upon by powers of measure-preserving transformations. As an application of multiple averages, Furstenberg obtained a purely ergodic proof of Szemerédi’s Theorem on the existence of arbitrary long arithmetic progressions in positive-density subsets of the integers [Fur77, Sze75]. However, Furstenberg’s seminal results from the seventies did not extend von Neumann’s theorem in either of the directions (1)–(3). It was Bergelson who, in 1987, first extended some of Furstenberg’s results to multiple ergodic averages of (plus quam linear) polynomial powers of a fixed measure-preserving transformation acting on products of functions [Ber87]. When specialized to simple measure-preserving actions, Bergelson’s results are a step toward generalizing von Neumann’s MET in direction (1). However, there is no purely Hilbert-theoretical formulation of Bergelson’s weak mixing hypothesis, so even the convergence of pointwise averages of $(U^{n(n)})$ stated in Theorem 1 only follows unconditionally from 2005 results for multiple ergodic averages of Host and Kra, and of Leibman (which depend on no mixing assumptions) [HK05, Lei05].

To our knowledge, Walsh’s theorem [Wal12] on mean convergence of nilpotent ergodic averages is the first result in the literature from which Theorem 1 follows as a corollary. (Pointwise convergence of averages of $(U^n \circ V^{n^2})$ under the assumption $U \circ V = V \circ U$ is a special case of 2009 results of Austin [Aus15a, Aus15b].) Thus, Walsh’s theorem actually implies the convergence of averages asserted in the more general Theorem 2 but only under the additional explicit hypothesis that $(T_i)$ generates a nilpotent subgroup of $U_{3\mathbb{C}}$. However, our methods do not require a nilpotence hypothesis, but only the more intrinsic property that $(T_i)$ be a Leibman sequence in the sense of Definition 2.3 (or in Leibman’s more general sense of polynomial mapping used in Theorem 4). In Example 2.5 we construct a quadratic Leibman sequence whose range does not generate a nilpotent subgroup of $U_{3\mathbb{C}}$.

Generalizations of Walsh’s theorem by Austin and Zorin-Kranich imply steps in direction (3) [Aus16, ZK16]. However, Theorems 2, 3 and 4 appear to be new in the general form stated. Nevertheless, given the close relation of our results to others in the existing literature, the main novelty is our “soft” direct approach to proving pointwise convergence of polynomial averages in Hilbert spaces using the framework of Henson metric structures. Our viewpoint is heavily influenced by Tao’s outline [Tao12] of a nonstandard proof à la Robinson of Walsh’s theorem (although we use only standard real numbers, and none of Robinson’s apparatus as such). A significant part of the manuscript consists of natural definitions and basic results on model-theoretic notions of integration and convergence that parallel classical ones; nevertheless, we capture, refine, and in some cases extend those results in Henson’s framework. Section 1 contains the rather long definition of the Henson class of PET structures over $\mathbb{Z}$. Section 2 introduces the ergodic averages that will be studied in the rest of the manuscript and defines the key concept of Leibman polynomial sequence of unitary operators on a Hilbert space. In Section 3, we state and prove Theorems 2 and 3 on metastable convergence of polynomial unitary averages for Leibman sequences (over $\mathbb{Z}$), and also explain how Theorem 1 follows as an immediate corollary. In Section 4 we state and
prove the most general of our results in the form of Theorem 4, which generalizes MET in all three directions (1)–(3). A number of foundational results are contained in the Appendix, which bears a close relation to our prior manuscript [DnI17]. These results pertain to measure theory and integration of real functions, as well as abstract notions of integration of functions taking values in Banach spaces. In this way we obtain a Dominated Convergence Theorem for notions of integration in an ad hoc Henson class of Banach integration frameworks (Theorem 5). We also show that the compactness of Henson’s logic implies a Uniform Metastability Principle for convergence in models of any Henson theory (Proposition A.10). Via this principle, all our results on convergence of averages admit refinements to convergence with metastability rates that are universal. These are gratis refinements thanks to the model-theoretic approach.

1. PET Structures

1.1. Classical PET Structures.

Notation 1.1. Below we list a number of formal symbols \( \mathbb{R}, \mathbb{Z}, \mathbb{N}, \mathcal{H}, \ldots \) that will eventually become sort descriptors for a Henson language of metric structures. However, throughout this subsection, these symbols have the following classical interpretations:

- \( \mathbb{R}, \mathbb{Z}, \mathbb{N} \) shall denote the sets of real numbers, integers and naturals.
- \( \mathcal{A}_Z \) shall denote the Boolean algebra of all subsets of \( \mathbb{Z} \).
- \( \mathcal{H} \) shall denote a real Hilbert space.
- \( \mathcal{B} \) shall denote the real Banach algebra \( \mathcal{B}(\mathcal{H}, \mathcal{H}) \) of bounded operators on \( \mathcal{H} \).
- \( \mathcal{M} \) shall denote the real Banach space of signed finite measures on \( \mathbb{Z} \) (i.e., on the measure space \( (\mathbb{Z}, \mathcal{A}_Z) \)).
- \( L^\infty_{\mathbb{Z}, \mathbb{R}} \) shall denote the Banach space \( L^\infty(\mathbb{Z}, \mathbb{R}) \) of bounded real functions on \( \mathbb{Z} \).
- \( L^\infty_{\mathbb{Z}, \mathcal{H}} \) shall denote the Banach space \( L^\infty(\mathbb{Z}, \mathcal{H}) \) of bounded functions \( \mathbb{Z} \to \mathcal{H} \).
- \( L^\infty_{\mathbb{Z}, \mathcal{B}} \) shall denote the Banach space \( L^\infty(\mathbb{Z}, \mathcal{B}) \) of bounded functions \( \mathbb{Z} \to \mathcal{B} \).
- \( L^\infty_{\mathbb{Z} \times \mathbb{Z}, \mathbb{R}} \) shall denote the Banach space \( L^\infty(\mathbb{Z} \times \mathbb{Z}, \mathbb{R}) \) of bounded real functions on \( \mathbb{Z} \times \mathbb{Z} \).
- \( L^\infty_{\mathbb{Z} \times \mathbb{Z}, \mathcal{H}} \) shall denote the Banach space \( L^\infty(\mathbb{Z} \times \mathbb{Z}, \mathcal{H}) \) of bounded functions \( \mathbb{Z} \times \mathbb{Z} \to \mathcal{H} \).
- \( L^\infty_{\mathbb{Z} \times \mathbb{Z}, \mathcal{B}} \) shall denote the Banach space \( L^\infty(\mathbb{Z} \times \mathbb{Z}, \mathcal{B}) \) of bounded functions \( \mathbb{Z} \times \mathbb{Z} \to \mathcal{B} \).

From a model-theoretic viewpoint, the sets \( \mathbb{R}, \mathbb{N}, \mathbb{Z}, \ldots \) denoted by the formal symbols above are the sorts of a metric Henson structure \( \mathcal{M} \). (Discrete sorts \( \mathbb{N}, \mathbb{Z}, \mathcal{A}_Z \) are still viewed as metric spaces endowed with the discrete metric.) In addition, \( \mathcal{M} \) is endowed with a number of distinguished elements (“constants”) and continuous functions between sorts. The distinguished elements include:

- All elements of \( \mathbb{N} \) and \( \mathbb{Z} \).
- All rational numbers in \( \mathbb{R} \).
- The zero element of each real Banach space above (\( \mathcal{H}, \mathcal{B}, \mathcal{M}, L^\infty_{\mathbb{Z}, \mathbb{R}}, \ldots \)).
- The identity operator \( I \in \mathcal{B} \).
- The zero (empty set \( \emptyset \)) and unity (improper subset \( \mathbb{Z} \subseteq \mathbb{Z} \)) of the Boolean algebra \( \mathcal{A}_Z \).

The distinguished functions between sorts include:

- The discrete metric in each the discrete sorts \( \mathbb{Z}, \mathbb{N}, \mathcal{A}_Z \).
- The operations of addition, subtraction, multiplication, absolute value, and lattice operations (binary minimum and maximum) on \( \mathbb{R} \).
The order \( \leq \) of \( \mathbb{N} \), identified with its characteristic function \([- \leq : \mathbb{N} \times \mathbb{N} \rightarrow \{0, 1\}\).

- The membership relation from \( \mathbb{Z} \) to \( \mathcal{A}_\mathbb{Z} \), identified with its characteristic function \([\cdot \in \cdot] : \mathbb{Z} \times \mathcal{A}_\mathbb{Z} \rightarrow \{0, 1\}\).
- The group operations (unary negation, binary addition and subtraction) of \( \mathbb{Z} \).
- The operations of union, intersection and complementation on \( \mathcal{A}_\mathbb{Z} \).
- The Hilbert space operations (addition, scalar multiplication, and inner product \( (x, y) \mapsto x \cdot y \)) on \( \mathcal{H} \). For convenience, also the norm \( \|x\| = \sqrt{x \cdot x} \).
- The operations of addition and scalar product, and the Banach norm \( \|\cdot\| \) on each Banach sort \( \mathfrak{B}, \mathfrak{M}, \mathcal{L}_{\infty, \infty} \).

For \( f \in \mathcal{L}_{\infty, \infty} \), the Banach norm is \( \|f\| = \sup_{x \in X} \|f(x)\| \), where \( \|f(x)\| \) is the norm of \( f(x) \) as an element of Banach sort \( Y \). The Banach norm on \( \mathfrak{B} \) is \( \|T\| = \sup \{\|T(x)\| : x \in \mathcal{H}, \|x\| \leq 1\} \). The Banach norm on \( \mu \in \mathfrak{M} \) is “total variation”: Recall that \( \mu \) has an atomic decomposition \( \mu = \sum_{i \in \mathbb{Z}} c_i \delta_i \) where \( \delta_i \) is the unit mass at \( i \) and \( c_i = \mu\{\{i\}\} \). With this notation, \( \|\mu\| = \sum_{i} |c_i| \).

(To abbreviate the long list of distinguished functions, above and in what follows we use \( X \) to denote either of the “domain” discrete sets \( \mathbb{Z}, \mathbb{Z}^2 \) of the various sorts \( \mathcal{L}^\infty \), and \( Y \) to denote the “codomain” Banach sorts \( \mathbb{R}, \mathcal{H}, \mathfrak{B}, \mathfrak{M} \).)

The list of distinguished functions continues as follows:

- The operations \( \mathcal{L}_{\infty, \infty} \times \mathcal{L}_{\infty, \infty} \rightarrow \mathcal{L}_{\infty, \infty} \) induced by (pointwise) application of the inner product of \( \mathcal{H} \).
- The unary operation of pointwise absolute value \( \cdot \) and the binary lattice operations (pointwise max and min) on sorts \( \mathcal{L}_{\infty, \mathbb{R}} \).
- The unary operation \( \cdot \) of measure of total variation and the binary lattice operations (“pointwise” max and min) on \( \mathfrak{M} \) (i.e., \( |\mu| = \sum_i |a_i| \delta_i \), max(\( \mu, \nu \)) = \( \sum_i \max\{a_i, b_i\} \delta_i \), and min(\( \mu, \nu \)) = \( \sum_i \min\{a_i, b_i\} \delta_i \) if \( \mu = \sum_i a_i \delta_i \) and \( \nu = \sum_i b_i \delta_i \)).
- The operation of pointwise magnitude \( |\cdot| : \mathcal{L}_{\infty, \mathbb{Y}} \rightarrow \mathcal{L}_{\infty, \mathbb{R}} \), namely \( |f| : x \mapsto \|f(x)\| \) for any \( f \in \mathcal{L}_{\infty, \mathbb{Y}} \).

- The unary adjoint operation \( T \mapsto T^* \) on \( \mathfrak{B} \), and the corresponding induced operations (pointwise adjoint) on sorts \( \mathcal{L}_{\infty, \mathfrak{B}} \).
- The binary operation \( (S, T) \mapsto S \circ T \) of composition on \( \mathfrak{B} \), and the corresponding induced operations of pointwise composition on sorts \( \mathcal{L}_{\infty, \mathfrak{B}} \).
- The inclusions:
  - \( \mathbb{Z} \hookrightarrow \mathcal{A}_\mathbb{Z} : i \mapsto \{i\} \).
  - \( \mathcal{A}_\mathbb{Z} \hookrightarrow \mathcal{L}_{\infty, \mathbb{R}} : A \mapsto \chi_A \) where \( \chi_A \) is the characteristic function of the subset \( A \subseteq \mathbb{Z} \).
  - \( \mathbb{Z} \hookrightarrow \mathfrak{M} \) given by \( i \mapsto \delta_i \) (the unit point mass at \( i \)).
  - \( \mathbb{Y} \hookrightarrow \mathcal{L}_{\infty, \mathbb{Y}} \), with \( y \in \mathbb{Y} \) identified with the constant function \( y(\square) : x \mapsto y \) in \( \mathcal{L}_{\infty, \mathbb{Y}} \).
  - The right inclusion map \( \mathcal{L}_{\mathbb{Y}, \infty} \hookrightarrow \mathcal{L}_{\mathbb{Y}, \infty} \) whereby \( f \in \mathcal{L}_{\mathbb{Y}, \infty} \) is identified with \( f(\square, \cdot) : (w, x) \mapsto f(x) \); also, the analogous left inclusion map identifying \( f \) with \( f(\cdot, \square) : (w, x) \mapsto f(w) \).
- The function-evaluation maps
  - \( (T, x) \mapsto T(x) \) from \( \mathfrak{B} \times \mathcal{H} \) to \( \mathcal{H} \).
  - \( (f, x) \mapsto f(x) \) from \( \mathcal{L}_{\infty, \infty} \times \mathbb{X} \) to \( \mathbb{Y} \).

Also, the maps \( \mathcal{L}_{\infty, \mathfrak{B}} \times \mathcal{L}_{\infty, \mathfrak{B}} \rightarrow \mathcal{L}_{\infty, \mathfrak{B}} \) induced by pointwise evaluation.
- The partial evaluation maps:
We may use integral notation and write \( L \int \langle \delta \rangle \) for all bounded functions \( Z \rightarrow L^{\infty}_{Z,Y} \). We also have a different identification of \( \mathcal{L}^{\infty}(Z, L^{\infty}_{Z,Y}) \) with \( \mathcal{L}^{\infty}(Z^2, Y) \) via right evaluation.

- The \textit{Følner-measure map} \( \sigma : \mathbb{N} \rightarrow \mathcal{M} \), where

\[
\sigma_n = \frac{1}{n + 1} \sum_{0 \leq i \leq n} \delta_i \quad \text{for all } n \in \mathbb{N}.
\]

- The translation action of \( Z \) on \( \mathcal{L}^{\infty}_{Z,Y} \). We regard this action as a function \( \mathcal{L}^{\infty}_{Z,Y} \rightarrow \mathcal{L}^{\infty}_{Z,Y} = \mathcal{L}^{\infty}(Z, L^{\infty}_{Z,Y}) \) (with the latter identification by partial evaluation on the left). The action is denoted \( f \mapsto \mu \) where \( \mu \in \mathcal{L}^{\infty}_{Z,Y} \) is the function \( i \mapsto i \cdot \mu \) with \( i, \mu \in \mathcal{L}^{\infty}_{Z,Y} \) the function \( j \mapsto f(i + j) \).

- The \textit{shear transformation} \( \mathcal{L}^{\infty}_{Z,Y} \rightarrow \mathcal{L}^{\infty}_{Z,Y} \), specifically \( F \mapsto \tilde{F} \) where \( \tilde{F} : (i, j) \mapsto F(i, i + j) \).

- The \textit{integration operations}

  - \( \mathcal{L}^{\infty}_{Z,Y} \times \mathcal{M} \rightarrow Y : (f, \mu) \mapsto \langle f, \mu \rangle = \sum_{i \in \mathbb{Z}} c_i f(i) \) for \( \mu = \sum_{i \in \mathbb{Z}} c_i \delta_i \).

  - (Left integral) \( \mathcal{M} \times \mathcal{L}^{\infty}_{Z,Y} \rightarrow \mathcal{L}^{\infty}_{Z,Y} : \langle \mu, F \rangle \mapsto \langle \mu, F \rangle \), where \( \langle \mu, F \rangle \in \mathcal{L}^{\infty}_{Z,Y} \) is the function \( i \mapsto \langle F(i, \cdot), \mu \rangle = \sum_{j \in \mathbb{Z}} c_j F(i, j) \).

  - (Right integral) \( \mathcal{L}^{\infty}_{Z,Y} \times \mathcal{M} \rightarrow \mathcal{L}^{\infty}_{Z,Y} : (F, \mu) \mapsto \langle F, \mu \rangle \), where \( \langle F, \mu \rangle \in \mathcal{L}^{\infty}_{Z,Y} \) is the function \( i \mapsto \langle F(\cdot, i), \mu \rangle = \sum_{j \in \mathbb{Z}} c_j F(i, j) \).

For visual convenience, we may use integral notation and write \( \int f \, d\mu \) or \( \int f(i) \, d\mu(i) \) for \( \langle f, \mu \rangle \), and \( \int F(i, \cdot) \, d\mu(i) \) for \( \langle \mu, F \rangle \) (resp., \( \int F(\cdot, j) \, d\mu(j) \)) for \( \langle \mu, F \rangle \).

\textbf{Remarks 1.2.}

- There are redundancies on the list of functions above. For instance, the \( L^2 \)-norm on \( \mathcal{H} \) is implicitly defined by its inner product: \( \|x\|^2 = x \cdot x \). As a less trivial example, the action of \( Z \) on \( \mathcal{L}^{\infty}_{Z,Y} \) is obtained from the right inclusion \( \mathcal{L}^{\infty}_{Z,Y} \hookrightarrow \mathcal{L}^{\infty}_{Z,Y} \) followed by the shear transformation. However, for reasons of exposition we make no effort to present a minimal list of distinguished functions. The model-theoretic approach fundamentally requires that all sorts, functions and constants that are relevant to the problem at hand be part of the structures under study.
• The nonstrict order relations (≤ and ≥) of \( \mathbb{R} \) are the only predicate symbols of a Henson language. However, any discrete predicate \( P \) may be identified with a \( \{0, 1\} \)-valued function \( \chi_P \) (the characteristic function of the truth set of \( P \)), so the usual interpretation of \( P(x) \) (resp., of \( \neg P(x) \)) agrees with the interpretation of the Henson formula \( \chi_P(x) \geq 1/2 \) (resp., of \( \chi_P(x) \leq 1/2 \)).

**Definition 1.3** (Classical PET structure over \( \mathbb{Z} \)). A classical PET structure (over \( \mathbb{Z} \)) is a triple \( \mathcal{M} = (S, C, F) \) where

\[
S = (\mathbb{R}, \mathbb{N}, \mathbb{Z}, A_{\mathbb{Z}}, H, B, M, L_{\mathbb{Z}, R}, L_{\mathbb{Z}, 3}, L_{\mathbb{Z}, R}, L_{\mathbb{Z}, B}, L_{\mathbb{Z}, M}, L_{\mathbb{Z}^2, R}, L_{\mathbb{Z}^2, 3}, L_{\mathbb{Z}^2, B})
\]

is a collection of sorts, \( C \) is a collection of distinguished elements \( (\text{constants}) \), and \( F \) is a collection of distinguished functions between sorts, provided these sorts, constants and functions are obtained in the manner prescribed by Notation 1.1.

1.2. Abstract PET structures.

**Definition 1.4** (Henson signature and language for PET structures over \( \mathbb{Z} \)). The Henson signature for PET structures over \( \mathbb{Z} \) consists of three ingredients:

• A collection of formal symbols, called sort descriptors \( (\text{or sort names}) \) in one-to-one correspondence with the collection \( S \) of sorts of a classical PET structure. For definiteness, the collection of descriptors is taken to be

\[
(\mathbb{R}, \mathbb{N}, \mathbb{Z}, A_{\mathbb{Z}}, H, B, M, L_{\mathbb{Z}, R}, L_{\mathbb{Z}, 3}, L_{\mathbb{Z}, R}, L_{\mathbb{Z}, B}, L_{\mathbb{Z}, M}, L_{\mathbb{Z}^2, R}, L_{\mathbb{Z}^2, 3}, L_{\mathbb{Z}^2, B})
\]

its members regarded as purely formal symbols.

• A collection of lexical constant symbols containing a unique symbol \( c \) for each of the distinguished elements in Definition 1.3 with each such symbol endowed with a sort descriptor \( s \) naming that sort to which the element \( c \) named by \( c \) belongs per Definition 1.3.

• A collection of lexical function symbols containing a unique symbol \( f \) for each of the functions named in Definition 1.3 with each such symbol endowed with a sort specification of the form \( s_1 \times \cdots \times s_n \to s_0 \) where \( s_0, s_1, \ldots, s_n \) are sort descriptors chosen in accordance with the specification of the domain (Cartesian product of sorts named by \( s_1, \ldots, s_n \)) and codomain (sort named by \( s_0 \)) of the function \( f \) named by the symbol \( f \).

The Henson language \( L \) for PET structures over \( \mathbb{Z} \) is the Henson language (of positive bounded formulas) whose signature is the one just described [HI02, Lov14, DnI17].

**Definition 1.5** (PET structure over \( \mathbb{Z} \)). Let \( L \) be the Henson language for PET structures. Let \( \text{PET} \) be the class of all classical PET structures over \( \mathbb{Z} \) per Definition 1.3 and let \( \text{Th}_{\text{PET}} \) be the \( L \)-theory of \( \text{PET} \) in Henson’s logic of approximate satisfaction of positive bounded formulas. An (abstract) PET structure over \( \mathbb{Z} \) is a model of \( \text{Th}_{\text{PET}} \).

The class \( \text{PET} \) of abstract PET structures obviously extends PET.

**Remarks 1.6.** • In principle, one may provide an explicit axiomatization in positive bounded Henson formulas of the class \( \text{PET} \). However, given the large number of sorts and functions in a PET structure this task is impractical. We refer the reader to our prior manuscript in which we provide explicit Henson axiomatizations of certain classes of structures somewhat more general than \( \text{PET} \) [DnI17]. Nevertheless, it should be clear that the Henson theory \( \text{Th}_{\text{PET}} \) is uniform in the sense that it imposes
bounds on constants as well as local bounds and local moduli of uniform continuity on distinguished functions. Moreover, $\text{Th}_{\text{PET}}$ obviously is identical to the theory $\text{Th}_{\text{PET}}$ of all abstract PET structures.

• The Følner map $\sigma : \mathbb{N} \to \mathcal{M}$ per Notation 1.1 implies a particular choice of a “notion of averaging” over $\mathbb{Z}$ that is built into $\text{Th}_{\text{PET}}$. Nonequivalent definitions of the PET class over $\mathbb{Z}$ and of $\text{PET}$ are obtained by changing this choice (e.g., letting $\sigma_n = 1/(2n + 1) \sum_{-n \leq i \leq n} \delta_i$ in classical structures), but Theorems 2 and 3 on PET structures over $\mathbb{Z}$ remain true under such alternate choice (in fact; they are special cases of the more general Theorem 4).

• If $\mathcal{M}$ is a PET structure, then the $\mathbb{R}$-named sort $\mathbb{R}^\mathcal{M}$ of $\mathcal{M}$ is (isomorphic to) the set of standard real numbers. Correspondingly, the “Hilbert sort” $\mathcal{H}^\mathcal{M}$ of $\mathcal{M}$ is a classical real Hilbert space. Typically, the $\mathbb{N}$-named sort $\mathbb{N}^\mathcal{M}$ of $\mathcal{M}$ is a proper extension of the set $\mathbb{N}$ of standard natural numbers (when the latter is identified with the set of interpretations $\mathfrak{m}^\mathcal{M}$ of the constant symbols $\mathfrak{m}$ of $\mathcal{L}$, one for each standard natural $m$), and similarly $\mathbb{Z} = \mathbb{Z}^\mathcal{M}$ extends $\mathbb{Z}$ in general. While $\mathcal{B}^\mathcal{M}$ may be identified (via the evaluation map $\mathcal{B} \times \mathcal{H} \to \mathcal{H}$) with an algebra of bounded operators on $\mathcal{H}^\mathcal{M}$, it need not contain all bounded operators. The sort $\mathcal{A}_Z^\mathcal{M}$ may be identified (via $\llbracket \cdot \in \cdot \rrbracket$) with a Boolean algebra of some, but not necessarily all subsets of $\mathbb{Z}^\mathcal{M}$, while $(\mathcal{L}^\infty_{Z,\mathbb{R}})^\mathcal{M}$ may be identified (via the evaluation map) with a space of (not necessarily all) bounded functions $\mathbb{Z} \to \mathbb{R}$.

One of the subtler differences between classical and abstract PET structures is the fact that $\mathcal{M}^\mathcal{M}$ typically consists of measures that are finitely but not countably additive on $\mathbb{Z}^\mathcal{M}$ (in particular, such measures need not have atomic decompositions as in the classical case). Fortunately, this difference turns out not to be critical, at least if one works in saturated PET structures: In this setting, the interplay between sorts $\mathbb{Z}^\mathcal{M}$, $\mathcal{A}_Z^\mathcal{M}$ and $(\mathcal{L}^\infty_{Z,\mathbb{R}})^\mathcal{M}$ comes to the rescue via analogues of Loeb measure and Loeb integration [DnI17]. Appendix A.2 contains a basic discussion of Loeb structures.

2. Ergodic averages and Leibman sequences in PET structures

Throughout the end of section 3, $\mathcal{L}$ will be the Henson language for PET structures over $\mathbb{Z}$. All structures will be in the class $\text{PET}$ of abstract PET structures over $\mathbb{Z}$.

2.1. The sequence of ergodic averages.

Convention 2.1. Henceforth, the standalone symbols $\mathbb{R}, \mathbb{Z}, \mathbb{N}$ shall denote the usual sets of real, integer and natural numbers. If $\mathcal{M}$ is a PET structure, we shall use interpretation of constants (and the density of $\mathbb{Q}$ in $\mathbb{R}$) to identify $\mathbb{R}$ with the sort $\mathbb{R}^\mathcal{M}$, and also $\mathbb{Z}$ and $\mathbb{N}$ with subsets of $\mathbb{Z} = \mathbb{Z}^\mathcal{M}$ and $\mathbb{N} = \mathbb{N}^\mathcal{M}$, respectively. (However, there is no need to identify $\mathbb{N}$ with a subset of $\mathbb{Z}$—it is actually best not to do so.) By an abuse of notation, when the structure $\mathcal{M}$ is clear from context, we may omit the superscript and write $\mathcal{H}$, $\mathcal{B}$, $\mathcal{A}_Z$, $\mathcal{L}^\infty_{Z,\mathbb{R}}$, … to denote the sorts $\mathcal{H}^\mathcal{M}$, $\mathcal{B}^\mathcal{M}$, $\mathcal{A}_Z^\mathcal{M}$, $(\mathcal{L}^\infty_{Z,\mathbb{R}})^\mathcal{M}$, … of $\mathcal{M}$.

Definition 2.2 (Ergodic averages). Let $\mathcal{M}$ be a PET structure over $\mathbb{Z}$ and let $T \in \mathcal{L}^\infty_{Z,\mathbb{B}}$. Via the evaluation $\mathcal{L}^\infty_{Z,\mathbb{B}} \times \mathbb{Z} \to \mathcal{B}$, one may regard $T$ as a function $i \mapsto T_i$. For $n \in \mathbb{N}$, the
n-th average of $T$ is
\[ \text{AV}_n T = \langle T, \sigma_n \rangle. \]
The sequence of averages of $T$ is $\text{AV}_n T = (\text{AV}_n T : n \in \mathbb{N})$.

Similarly, for $x \in \mathcal{H}$, the n-th average of $x$ under $T$ is $\text{AV}_n T(x)$. The sequence of averages of $x$ under $T$ is $\text{AV}_n T(x) = (\text{AV}_n T(x) : n \in \mathbb{N})$.

We remark that $\text{Th}_{\text{PET}}$ ensures the validity of the identities
\[ \text{AV}_n T = \frac{1}{n+1} \sum_{0 \leq i \leq n} T_i \quad \text{and} \quad \text{AV}_n T(x) = \frac{1}{n+1} \sum_{0 \leq i \leq n} T_i(x) \]
for (standard) $n \in \mathbb{N}$; however, averages as defined above are non-classical if $n \in \mathbb{N} \setminus \mathbb{N}$. On the other hand, the terms of sequences of averages above are all classical.

### 2.2. Leibman sequences of unitary operators.

Leibman introduced the notion of polynomial sequences in a group [Lei98]. Relevant to our purposes is the group $\mathbb{Z}$ of Leibman sequences of any degree $d$. Let $\mathcal{M}$ be a PET structure over $\mathbb{Z}$. The discrete-difference operator is the function $\Delta^\bullet : \mathcal{L}_{\mathcal{Z},}\mathcal{B} \to \mathcal{L}_{\mathcal{Z},}\mathcal{B}$ uniquely characterized by the identity
\[ \Delta^\bullet(T)_{i,j} = T_{i+j} \circ T_j^* \quad \text{for all } i, j \in \mathbb{Z}. \]
(Note that $(T_j^*)_j = (T_j)^*$ so parentheses may be omitted without ambiguity.) Alternatively, for $i \in \mathbb{Z}$, the left evaluation at $i$ of $\Delta^\bullet T$ is $\Delta^\bullet T = i^\bullet T \circ T^*$. Let $\mathbb{N} = I(\bullet)$ denote the constant family $i \mapsto I$ in $\mathcal{L}_{\mathcal{Z},}\mathcal{B}$. Given $d \in \mathbb{N}$, a unitary Leibman sequence of degree at most $d$ is any $T \in \mathcal{L}_{\mathcal{Z},}\mathcal{B}$ satisfying
\[ \Delta^d (\ldots (\Delta^{i_0}(\Delta^{i_1}(T)) \ldots)) = 1 \quad \text{for all } i_0, i_1, \ldots, i_d \in \mathbb{Z}, \]
that takes values in $\mathcal{U}_{\mathcal{Z},}$, i.e., also satisfying $T^* \circ T = \mathbb{1} = T \circ T^*$. A Leibman sequence is a Leibman sequence of any degree $d$; its degree $\text{deg}_{\mathcal{L}}(f)$ is the least such $d$. (We define formally $\text{deg}_{\mathcal{L}}(\mathbb{1}) = -\infty$.)

**Remarks 2.4.**

- We have $\Delta^\bullet T = \bullet T \circ (T^*_\bullet)$; hence, discrete differentiation is obtained from the $\mathbb{Z}$-action on $\mathcal{L}_{\mathcal{Z},}\mathcal{B}$, the right inclusion $\mathcal{L}_{\mathcal{Z},}\mathcal{B} \hookrightarrow \mathcal{L}_{\mathcal{Z},}\mathcal{B}$, plus the pointwise operations of composition and taking adjoint; thus, $\Delta^\bullet$ may as well be regarded as a distinguished function of any PET structure $\mathcal{M}$. 

**Definition 2.3 (Discrete difference and Leibman sequence).** Let $\mathcal{M}$ be a PET structure over $\mathbb{Z}$. The discrete-difference operator is the function $\Delta^\bullet : \mathcal{L}_{\mathcal{Z},}\mathcal{B} \to \mathcal{L}_{\mathcal{Z},}\mathcal{B}$ uniquely characterized by the identity
\[ (\Delta^\bullet T)_{i,j} = T_{i+j} \circ T_j^* \quad \text{for all } i, j \in \mathbb{Z}. \]
• Note that Leibman sequences as defined above are “internal”, i.e., obtained from elements of \( L^\infty_{Z,A} \)—that are otherwise only incidentally regarded as functions \( Z \to A \) via evaluation. Accordingly, the defining property of a Leibman sequence amounts to the requirement that \( d + 1 \) discrete-difference operations, possibly involving nonstandard elements \( j \in Z \), always transform \( T \) into \( \mathbb{1} \), i.e., the constant function \( i \mapsto I \) for all \( i \in Z \), not merely for all \( i \in Z \). The proofs of Theorems 2 and 3 below crucially depend on the richer structure of \( Z \) in saturated PET structures—even if ultimately the results are valid in all PET structures, including classical ones whose Leibman sequences are bona fide functions \( Z \to U_{\mathcal{E}} \subset A \).

• Translations commute with adjoints and with discrete differences, i.e., \((j^*)^* = j(T^*)^* \) and \( j(\Delta^iT) = \Delta^i(j^*)^* \). (The latter equality depends on the commutativity of addition on \( Z \).) In particular, Leibman degree is invariant under translation. However, the discrete difference operators do not commute with adjoints, so Leibman degree is not invariant under taking adjoints. Correspondingly, \( T^* \) need not be a Leibman polynomial if \( T \) is.

• The predicate “\( T \) is a unitary Leibman sequence of degree at most \( d \)” is captured by a single Henson formula \( \lambda_d(T) \), namely

\[
(T \circ T^* = 1 = T^* \circ T) \land \forall i_d \ldots \forall i_1 \forall i_0 [\Delta^i_d(\ldots (\Delta^i_1(\Delta^i_0T))\ldots) = \mathbb{1}].
\]

• In classical structures with \( Z^\mathbb{M} = Z \), it can be shown (essentially by the familiar method of finite differences) that if \((T_k)\) is a unitary Leibman sequence of degree at most \( d \) in an abelian subgroup of \( U_{3\mathcal{E}} \), then there exist pairwise-commuting unitary transformations \( U_0, U_1, \ldots, U_d \) such that

\[
T_k = U_0 \circ U_1^k \circ U_2^{k^2} \circ \cdots \circ U_d^{k^d}
\]

for all \( k \in Z \), where \( \binom{k}{j} = k(k-1) \cdots (k-j+1)/k! \) is the \( j \)-th binomial coefficient. One may regard \( U_0, U_1, \ldots, U_d \) as the “coefficients” of the (classical) abelian unitary Leibman polynomial \( T \). In particular, this abelian case comprises all families \((U^{p(k)})\) where \( p \) is a polynomial \( Z \to Z \) and \( U \in U_{3\mathcal{E}} \) is fixed. Theorem 1 states the convergence of ergodic averages in this particular case; nevertheless, Theorems 2, 3 and 4 only assume that \( T \) is a unitary Leibman sequence per Definition 2.3 but no explicit commutativity hypotheses otherwise.

• A natural generalization of Theorem 2 (namely, Theorem 4) holds with an arbitrary abelian group \( G \) replacing \( Z \)—the only properties of \( Z \) used in the proof of Theorem 2 in an essential fashion are the commutativity of the group operation and the existence of a countable Følner net.

Example 2.5. We show that the range of a Leibman sequence (on \( Z \)) need not generate a nilpotent subgroup of \( U_{3\mathcal{E}} \).

Let \( K \) be a multiplicative abelian group freely generated by a collection \((b_i : i \in Z)\) of distinct formal elements \( b_i \). Let \( a \) be yet another formal element generating an infinite cyclic group \( H = \langle a \rangle \), and let \( G = K \rtimes H \) be the semidirect product with respect to the formal conjugation action \( a^{-i} b_j a^i = b_{i+j} \) of \( H \) on \( K \rtimes G \). (Succinctly, the resulting group \( G \) is the restricted wreath product \( Z \wr Z_i \).) Observe that \( G \) is generated by \( a \) and \( b = b_0 \). It is easy

3Expressions of the type \( x = y \) such as those in (2.1) are not Henson formulas sensu stricti, but may be regarded as abbreviations of formulas \( d(x, y) \leq 0 \) (or \( \|y - x\| \leq 0 \) in Banach sorts).
to see that each of the groups $H, K$ is its own centralizer in $G$, so $G$ has trivial center. In particular, $G$ is not nilpotent (although it is certainly solvable, being the semidirect product of two abelian groups).

Let $G$ be realized as a subgroup of a suitable unitary group $U_{\mathfrak{g}}$ It can be shown that the relations satisfied by $a$ and $(b_j)$ imply that a unique quadratic Leibman sequence $(T_k)$ satisfies $T_0 = I$, $\Delta^1 T_0 = a$ and $\Delta^1 \Delta^1 T_0 = b$, namely the sequence

$$T_k = \begin{cases} I, & k = 0; \\ b_{k-2}b_{k-3}^2 \ldots b_{1}^{k-2}b_{0}^{k-1}a^k, & k > 0; \\ b_{k}b_{k+1}^2 \ldots b_{-2}^{k-1}b_{-1}^{k}a^k, & k < 0. \end{cases}$$

Clearly, the range of $(T_k)$ is a subset of $G$. A moment’s reflection shows that $(T_k)$ generates $G$. To our knowledge, this is the first example of a sequence generating a non-nilpotent group whose pointwise averages converge as ensured by Theorem 2 below.

3. An ergodic theorem for unitary polynomial actions of $\mathbb{Z}$

**Theorem 2** (Poly-MET/$\mathbb{Z}$: Mean Ergodic Theorem for unitary polynomial actions of $\mathbb{Z}$). Let $\mathcal{M}$ be a PET structure over $\mathbb{Z}$, and let $T \in (\mathcal{L}_{\mathbb{Z}, \mathbb{R}}^\infty)^{\mathcal{M}}$ be a Leibman sequence of unitary operators on the Hilbert space $\mathcal{H} = \mathcal{H}^\mathcal{M}$. For every $x \in \mathcal{H}$, the sequence $AV_n(T)(x) = (AV_n T(x) : n \in \mathbb{N})$ of averages of $x$ under $T$ converges in the norm topology of $\mathcal{H}$.

**Theorem 2** admits the following uniformly metastable strengthening.

**Theorem 3** (Metastable Poly-MET/$\mathbb{Z}$). Fix $d \in \mathbb{N}$. There exists a universal metastability rate $E^d_\bullet$, depending only on $d$, that applies uniformly to all sequences $AV_n T(x)$ of averages of arbitrary $x$ in the unit ball of the Hilbert-space sort $\mathcal{H}$ under any Leibman sequence $T$ in $U_{\mathfrak{g}}$ of degree at most $d$ in any PET structure $\mathcal{M}$ over $\mathbb{Z}$.

The rest of Section 3 is devoted to proving Theorems 2 and 3

3.1. Preliminaries.

**Lemma 3.1.** Let $\mathcal{L}$ be the language of PET structures. Let $\mathcal{M}$ be any saturated PET structure. Let $\varphi_\bullet = (\varphi_n : n \in \mathbb{N})$ be a bounded sequence in $\mathcal{L}_{\mathbb{Z}, \mathbb{R}}^\infty$. For all $x \in \mathcal{H}^\mathcal{M}$ assume that the sequence $\varphi_\bullet(x) = (\varphi_n(x) : n \in \mathbb{N})$ in $\mathcal{H}^\mathcal{M}$ is convergent. Then, for arbitrary $\mu \in \mathcal{M}^\mathcal{M}$, the sequence $\langle \varphi_\bullet, \mu \rangle = (\langle \varphi_n, \mu \rangle : n \in \mathbb{N})$ in $\mathcal{H}^\mathcal{M}$ is convergent.

(We will only require the special case of Lemma 3.1 in which $\varphi$ is of the form $n \mapsto AV_n T$ with $T \in \mathcal{L}_{\mathbb{Z}, \mathbb{R}}^\infty$.)

**Proof.** A saturated PET structure $\mathcal{M}$ is a Banach integration framework (with Banach sort $\mathcal{H}^\mathcal{M}$ and measure-space sort $\mathbb{Z}^\mathcal{M}$) as defined in Appendix A.4. Thus, Lemma 3.1 follows from Theorem 6 whose statement and proof are in Appendix A.5. \qed

To state the next lemma we need a definition. Let the reverse difference operator $\nabla^\bullet : \mathcal{L}_{\mathbb{Z}, \mathbb{R}}^\infty \to \mathcal{L}_{\mathbb{Z}, \mathbb{R}}^\infty$ be the mapping $T \mapsto \nabla^\bullet T$ characterized by the property that $\nabla^\bullet T$ evaluates to the function $(i,j) \mapsto T_j \circ T_{i+j}^*$. We write $\nabla^i T$ to denote the left evaluation of $\nabla^\bullet T$, i.e., $\nabla^i T$ evaluates to the mapping $j \mapsto T_j \circ T_{i+j}^*$. Note that $\nabla^i T$ is the translate by $i$ of the

---

4One may identify $G$ with its faithful homomorphic image under the translation action $G \subset \mathcal{L}^2(G)$, which realizes $G$ as a group of unitary transformations of $\mathcal{H} = \mathcal{L}^2(G)$. 

---

...
(forward) difference of $T$ with step $-i$, i.e., $\nabla^i T = i^{\Delta^{-i}} T$ holds for all $i \in \mathbb{Z}$. Just like the forward difference operator $\Delta^*$, the reverse difference operator $\nabla^*$ is explicitly definable in any PET structure since it is obtained by composing functions of the structure (the $\mathbb{Z}$-action $L^\infty_{\mathbb{Z},\mathcal{B}} \to L^\infty_{\mathbb{Z},\mathcal{B}}$, the shear map on $L^\infty_{\mathbb{Z},\mathcal{B}}$, the pointwise adjoint operation $L^\infty_{\mathbb{Z},\mathcal{B}} \to L^\infty_{\mathbb{Z},\mathcal{B}}$, the left inclusion $L^\infty_{\mathbb{Z},\mathcal{B}} \hookrightarrow L^\infty_{\mathbb{Z},\mathcal{B}}$, and the pointwise composition $L^\infty_{\mathbb{Z},\mathcal{B}} \times L^\infty_{\mathbb{Z},\mathcal{B}} \to L^\infty_{\mathbb{Z},\mathcal{B}}$); thus, $\nabla^*$ may as well be considered a distinguished function of any PET structure.

**Lemma 3.2.** Let $\mathcal{M}$ be a PET structure such that $\mathcal{N} = \mathbb{N}^\mathbb{N}$ contains a nonstandard natural number $M \in \mathcal{N} \setminus \mathbb{N}$. For every standard natural $n \in \mathbb{N}$ and $T \in L^\infty_{\mathbb{Z},\mathcal{B}}$:

$$(AV_n T) \circ (AV_M T)^* = \langle \langle \nabla^* T, \sigma_n \rangle, \sigma_M \rangle.$$ 

In less cryptic notation, the equation above reads:

$$(AV_n T) \circ (AV_M T)^* = \int AV_n (\nabla^i T) d\sigma_M(i).$$

Every step of the proof below is justified by an axiom of $\text{Th}_{\text{PET}}$. We prefer to use informal integral notation to make the argument transparent.

**Proof.** Note that $(AV_M T)^* = AV_M (T^*)$ (in fact, $\langle T, \mu \rangle^* = \langle T^*, \mu \rangle$ for all $\mu \in \mathfrak{M}$). For $m, n \in \mathbb{N}$ we have:

$$(AV_n T) \circ (AV_m T)^* = \int T_j d\sigma_n(j) \circ \int T_i^* d\sigma_m(i) = \int \int T_j \circ T_i^* d\sigma_m(i) d\sigma_n(j)$$

$$= \int \int T_j \circ T_i^* d\sigma_m(i + j) d\sigma_n(j)$$

(letting $i = i + j$ in the inner integral)

$$= \int \int \nabla^i T_j d\sigma_m(i + j) d\sigma_n(j)$$

$$= \int \int \nabla^i T_j d\sigma_n(j) d\sigma_m(i)$$

$$- \int \int \nabla^i T_j d[\sigma_m - j(\sigma_m)](i) d\sigma_n(j),$$

where $j(\sigma_m)$ is the translation of $\sigma_m$ by $j$. Since $\|i \Delta^k T_j\| \leq \|T\|^2$ for all $i, j, k$:

$$\left\| (AV_n T) \circ (AV_m T)^* - \int AV_n (\nabla^i T) d\sigma_m(i) \right\| \leq \|T\|^2 \cdot \max_{0 \leq j \leq m} \|\sigma_m - j(\sigma_m)\|.$$ 

Given fixed $\epsilon > 0$ and $n \in \mathbb{N}$, let $m_{\epsilon,n}$ be the smallest natural number satisfying $m_{\epsilon,n} \geq 2n/\epsilon$. Clearly, $\|\sigma_m - j(\sigma_m)\| \leq 2n/(m + 1) \leq \epsilon$ for $j \leq n \leq m + 1$. (For $m$ large and $n$ small, this inequality captures the “approximate invariance” of the long interval $\{0, 1, \ldots, m\}$ of $\mathbb{Z}$ under small translations, i.e., the Fölner property of the collection of such intervals.) Then we have

$$\left\| (AV_n T) \circ (AV_m T)^* - \int AV_n (\nabla^i T) d\sigma_m(i) \right\| \leq \epsilon \|T\|^2 \quad \text{whenever } m \geq m_{\epsilon,n}. \tag{3.1}$$

Since $M \in \mathcal{N} \setminus \mathbb{N}$ satisfies $M \geq m_{\epsilon,n}$ for all $n \in \mathbb{N}$ and $\epsilon > 0$, the assertion in Lemma 3.2 follows. \qed
We offer some remarks on the proof of Lemma 3.2 above, which is the crux of our approach to proving Theorem 2. Despite its rather short length, it sheds light on the various sorts and distinguished functions in PET structures. There are no double integrals as such but rather iterated integrals $L^\infty_{\mathbb{Z},\mathcal{B}} \to L^\infty_{\mathbb{Z},\mathcal{B}}$ and $L^\infty_{\mathbb{Z},\mathcal{B}} \to \mathcal{B}$—the order of the integration (first on the left and then on the right variable, or vice versa) is immaterial as ensured by the PET axiom

$$(\forall \mathcal{I} \in L^\infty_{\mathbb{Z},\mathcal{B}})(\forall \mu, \nu \in \mathcal{M})[\langle \langle \mu, \mathcal{I} \rangle, \nu \rangle = \langle \langle \mathcal{I}, \nu \rangle, \mu \rangle].$$

The validity of the substitution $i = i + j$ in the inner integral is justified by the compatibility of the shear transformation on $L^\infty_{\mathbb{Z},\mathcal{B}}$ and the action of $\mathbb{Z}$ on $\mathcal{M}$:

$$(\forall \mathcal{I} \in L^\infty_{\mathbb{Z},\mathcal{B}})(\forall \mu \in \mathcal{M})[\langle \mathcal{I}, \mu \rangle = \langle \tilde{\mathcal{I}}, \mu \rangle].$$

Other steps in the proof admit similar formal justifications by axioms of PET structures.

**Lemma 3.3.** Let $\mathcal{M}$ be a saturated Henson structure with an ordered sort $(\mathbb{N}, \leq)$ extending $(\mathbb{N}, \leq)$, and let $\varphi_i$ be a sequence explicitly defined by a $\mathcal{L}[S]$-term $\varphi(n)$ of sort $s$ (i.e., $\varphi_i$ is the sequence $(\varphi(n) : n \in \mathbb{N})$ in $\mathcal{M}$, where $n$ is a variable of sort $\mathbb{N}$ and $S$ is some set of parameters of the universe of $\mathcal{M}$). Then every sub-sequential limit of $\varphi$ is of the form $\varphi(M)$ for some $M \in \mathbb{N} \setminus \mathbb{N}$. If $\varphi(M) = \varphi(N)$ for all $M \in \mathbb{N} \setminus \mathbb{N}$, then the sequence $\varphi_i = (\varphi(n) : n \in \mathbb{N})$ converges. In such case, the common value $\varphi(M)$ is the limit $\lim_{n \to \infty} \varphi(n)$.

**Proof.** This is a routine application of saturation. □

**Lemma 3.4.** Let $X$, $Y$ be metric spaces, and let $\varphi : (x, n) \mapsto \varphi_n(x)$ be a function from $X \times \mathbb{N}$ to $Y$ such that $\varphi_n(\cdot) : X \to Y$ is 1-Lipschitz for each $n \in \mathbb{N}$ (i.e., $d(\varphi_n(x), \varphi_n(y)) \leq d(x, y)$ for $x, y \in X$). Let $S$ be a dense subset of $X$ such that $\varphi_n(x)$ converges for all $x \in S$. Then $\varphi_i(x)$ converges for all $x \in X$.

We omit the straightforward proof.

### 3.2. Proof of Theorem 2

For each fixed Leibman degree $d \in \mathbb{N}$, we first prove Theorem 2 for unitary Leibman sequences of degree at most $d$ in any saturated PET structure $\mathcal{M}$. The descent argument on the degree $d$ is characteristic of Bergelson’s PET induction [Ber87].

The assertion is trivial for $T = 1$. If $T$ (is pointwise unitary and) has Leibman degree $\deg_{\mathcal{L}} T = 0$, we have $T_i \circ T^n = \Delta^i T_0 = I$, hence $T_i = T_0$ for all $i \in \mathbb{Z}$, so $T$ is constant. Thus, the sequences AV$_i T$ and AV$_\bullet T(x)$ are also constant (all terms are equal to $T_0$ and $T_0(x)$, respectively), so Theorem 2 follows for Leibman polynomials of degree 0.

Assume now that the assertion in Theorem 2 is proved for all $T \in L^\infty_{\mathbb{Z},\mathcal{B}}$ having Leibman degree less than some positive integer $d$. Fix $T \in L^\infty_{\mathbb{Z},\mathcal{B}}$ with $\deg_{\mathcal{L}}(T) = d$.

**Lemma 3.5.** If $x = (AV_M T)^*(y)$ for some $M \in \mathbb{N} \setminus \mathbb{N}$ and $y \in \mathcal{H}$, then AV$_\bullet T(x)$ converges.

**Proof.** Note that AV$_\bullet T(x)$ is bounded by $\|T\| \|x\|$. By Lemma 3.2, we have AV$_n T(x) = AV_n T \circ (AV_M T)^*(y) = \langle \langle \nabla^i T(y), \sigma_M \rangle, \sigma_M \rangle$ for $n \in \mathbb{N}$. For $i \in \mathbb{Z}$, we have $\nabla^i T = [\Delta^{i-d}] T$. By the invariance of Leibman degree under translation and the assumption $\deg_{\mathcal{L}}(T) \leq d$, we have $\deg_{\mathcal{L}}(\nabla^i T) = \deg_{\mathcal{L}}([\Delta^{i-d}] T) < d$, and hence AV$_\bullet (\nabla^i T)(y)$ is convergent for all $i \in \mathbb{Z}$ by the inductive hypothesis. An application of Lemma 3.1 concludes the proof. □

The space Struct of structured elements of $\mathcal{H}$ (relative to $T$) is the closure of the linear span of all elements of the form $(AV_M T)^*(y)$ for $M \in \mathbb{N} \setminus \mathbb{N}$ and $y \in \mathcal{H}$. By linearity and Lemma 3.4, AV$_\bullet T(x)$ converges for all structured elements $x$. (The 1-Lipschitz condition
follows from the inequalities \( \|AV_nT(x) - AV_nT(y)\| \leq \|AV_nT\|\|y - x\| \) and \( \|AV_nT\| = \|T, \sigma_n\| \leq \|T\| \cdot \|\sigma_n\| = 1 \cdot 1 = 1 \).

The space \( \psi \text{Rand of pseudorandom elements} \) of \( \mathcal{H} \) (relative to \( T \)) is the orthogonal complement of Struct in \( \mathcal{H} \). By the fundamental theorem of linear algebra and the definition of structured elements, an element \( x \in \mathcal{H} \) is pseudorandom precisely when \( AV_M T(x) = 0 \) for all \( M \in \mathbb{N} \setminus \mathbb{N} \). By Lemma 3.3, \( AV_\bullet T(x) \) converges to zero in this case.

Combining the pseudorandom and structured cases using linearity, we deduce that all averages \( AV_\bullet T(x) \) converge. This concludes the inductive step of the proof of Theorem 3 in any saturated PET structure \( \mathcal{M} \).

To conclude the proof for any PET structure over \( \mathbb{Z} \), let \( \mathcal{M} \) be any (not necessarily saturated) PET structure, and let \( \tilde{\mathcal{M}} \) be a saturated elementary \( \mathcal{L} \)-extension of \( \mathcal{M} \) in Henson’s logic. For fixed degree \( d \in \mathbb{N} \) and \( T \in (\mathcal{L}_\mathbb{Z})^d \), the property that \( T \) is a unitary Leibman sequence of degree at most \( d \) is \( \mathcal{L} \)-axiomatizable, hence it is true in \( \tilde{\mathcal{M}} \) when \( T \) is regarded as an element of \( (\mathcal{L}_\mathbb{Z})^d \). For \( x \in \mathcal{H}^d \) we have proved that \( AV_\bullet T(x) \) converges since \( \deg L(T) \leq d \). A fortiori, \( AV_\bullet T(x) \) converges for \( x \in \mathcal{H}^d \). This concludes the proof of Theorem 3 in full generality.

### 3.3. Proof of Theorem 3

Let \( \tilde{\mathcal{L}} \) expand the language \( \mathcal{L} \) of PET structures with new constants \( (T, x, y_n : n < \omega) \), with \( T \) of sort \( \mathcal{L}_\mathbb{Z} \), and \( x \) and all \( y_n \) of sort \( \mathcal{H} \). For fixed \( d \in \mathbb{N} \), consider the \( \mathcal{L} \)-theory

\[
\Lambda_d = \text{Th}_{\text{PET}} \cup \{ \lambda_d(T), \|x\| \leq 1, y_n = AV_n T(x) : n < \omega \}
\]

where \( \lambda_d(T) \) is formula \( 2.1 \) stating that the interpretation of \( T \) is Leibman of degree at most \( d \). Note that \( \Lambda_d \) is a uniform theory: \( \lambda_d(T) \) implies \( \|T\| \leq 1 \), hence also \( \|y_n\| \leq 1 \). Every model \( \tilde{\mathcal{M}} \) of \( \Lambda_d \) is an expansion of a PET structure \( \mathcal{M} \) having the form \( \mathcal{M} = (\mathcal{M}, T, x, AV_\bullet T(x)) \). By Theorem 2 all sequences \( (c_n^d) = AV_\bullet T(x) \) are convergent. An application of Proposition A.10 finishes the proof of Theorem 3.

### 3.4. Proof of Theorem 1

A classical Leibman sequence \( T_\bullet = (T_k : k \in \mathbb{Z}) \) with \( \deg L(T) \leq d \) in an abelian subgroup \( \mathcal{K} \) of the group \( U_{\mathcal{K}} \) of unitary operators on a Hilbert space \( \mathcal{H} \) is easily shown to have the form \( T_k = U_0 \circ U_1^k \circ U_2^k \circ \cdots \circ U_d^k \) where \( U_j = \Delta^j T_0 \) and \( \binom{k}{j} = k(k-1)\ldots(k-j+1)/j! \) are binomial coefficients for \( j = 0, 1, \ldots, d \). Since the functions \( k \mapsto \binom{k}{j} \) are a \( \mathbb{Z} \)-basis for polynomial mappings \( p : \mathbb{Z} \to \mathbb{Z} \) of degree at most \( d \), Theorem 1 is an immediate corollary of Theorems 2 and 3.

### 4. A Mean Ergodic Theorem for Unitary Polynomial Actions of Abelian Groups

To formulate our most general result on convergence of averages, we replace \( \mathbb{N} \) with an arbitrary countable directed set \( (\mathbb{D}, \leq) \) and \( \mathbb{Z} \) with an arbitrary abelian group \( (\mathcal{G}, +) \) endowed with a countable Følner \( \mathbb{D} \)-net \( \mathcal{F}_\bullet = (\mathcal{F}_j : j \in \mathbb{D}) \) of nonempty finite subsets of \( \mathcal{G} \). These assumptions are sufficient to ensure that the proofs of natural generalizations of Theorems 2 and 3 carry through in this more general context, mutatis mutandis, from those given in Section 3.

**Theorem 4** (Poly-MET: Mean Ergodic Theorem for unitary polynomial actions of an abelian group). Fix an abelian group \( (\mathcal{G}, +) \) and a Følner net \( \mathcal{F}_\bullet = (\mathcal{F}_j : j \in \mathbb{D}) \) of subsets
of $G$, indexed by a countable directed set $(D, \leq)$. Let $H$ be a Hilbert space. Let $T : G \to U_H$ be a polynomial mapping, in Leibman’s sense, into the group $U_H$ of unitary transformations of $H$. For every $x \in H$, the $D$-net $AV_i T(x) = (AV_i T(x) : i \in D)$ in $H$, of averages relative to $F_i$:

$$AV_i T(x) = \frac{1}{\#F_i} \sum_{g \in F_i} T_g(x)$$

of $x$ under $T$, converges in the norm topology of $H$.

In fact, given fixed choices of $G$, $D$, $F_i$, and $d \in \mathbb{N}$, there exists a rate of metastability

$$E_{\epsilon, \eta} = \left( E_{\epsilon, \eta} : \epsilon > 0, \eta \in \prod_{i \in D} P_{\text{fin}}(D_{\geq i}) \right)$$

(with $E_{\epsilon, \eta} \in P_{\text{fin}}(D)$ for each $\epsilon, \eta$) that applies universally to all sequences $AV_i T(x)$ for any element $x$ in the unit ball of any Hilbert space $H$ and any Leibman polynomial $T : D \to U_H$ of degree at most $d$.

**Remark 4.1.** The definition of Leibman polynomial mapping $T : G \to U_H$ is a straightforward generalization of that of Leibman sequence $Z \to U_H$ [Lei02]. The discrete difference $\Delta^g T$ of $T$ with step $g \in G$ is the mapping $h \mapsto T_{g+h} \circ T^{-1}_g$. Define $\deg(T) \leq 0$ if $\Delta^g T = 1$ (where $1$ is the constant mapping $g \mapsto I$). Recursively, let $\deg(T) \leq d+1$ mean that $\deg(\Delta^g T) \leq d$ for all $g \in G$. Then $T$ is a Leibman mapping if $\deg(T) \leq d$ for some $d$; the least such $d$ is $\deg(T)$, although we adopt the convention $\deg(1) = -\infty$.

We only provide an outline of the proof of Theorem 4 since it is formally identical to the arguments in Sections 3.1–3.3. The definition of classical PET structure over $G$ is completely analogous to that of PET structure over $Z$ in section 1.1—simply replace all instances of $Z$ by $G$ and those of $N$ by $D$. The Følner sequence $F_i$ is captured indirectly via the Følner measure map $\sigma : D \to \mathcal{M}$, where

$$\sigma_i = \frac{1}{\#F_i} \sum_{g \in F_i} \delta_g \quad \text{for all } i \in D.$$
The arguments in Sections 3.2 and 3.3 apply verbatim once the lemmas in Section 3.1 have been adapted, completing the proof of Theorem 4.

Appendix A. A Dominated Convergence Theorem for notions of integration in Banach spaces

This appendix bears a close connection to our prior manuscript on measure, integration and metastable convergence in Henson structures [DN17]. Our main goal is proving Lemma 3.1. Rather than doing so in the specific context of PET structures, we prove a more general result (Theorem 5) about sequences of integrals of functions on a finite measure space taking values in a Banach space. This requires a number of preliminary steps.

A.1. Integration structures. We recall the class of integration structures (with underlying finite positive measure) introduced in our earlier manuscript, to which we refer the reader for details [DN17]. These are saturated models of the Henson theory Th$_f$ of integration with respect to a positive finite measure on structures with (classical) sorts $\mathbb{R}, \Omega, \mathcal{A}_\Omega, \mathcal{L}^\infty_{\Omega,\mathbb{R}}$ where $\mathbb{R}$ is the set of real numbers, $\mathcal{A}_\Omega$ is a $\sigma$-algebra of subsets of $\Omega$, and $\mathcal{L}^\infty_{\Omega,\mathbb{R}}$ is the set of bounded $\mathcal{A}_\Omega$-measurable (everywhere-defined) real functions on $\Omega$. (Here $\Omega, \mathcal{A}_\Omega$ are discrete while $\mathbb{R}, \mathcal{L}^\infty_{\Omega,\mathbb{R}}$ are real Banach spaces.) This theory contains all Henson formulas involving the functions and distinguished constants below that are valid in such structures:

- **Constants:** Rational numbers $r \in \mathbb{Q}$, zero vector in the Banach sort $\mathcal{L}^\infty_{\Omega,\mathbb{R}}$, an arbitrary point (“anchor”) $\omega_0 \in \Omega$, the empty set $\emptyset \in \mathcal{A}_\Omega$, and the improper subset $\Omega \in \mathcal{A}_\Omega$.
- **Functions:**
  - Arithmetic operations (addition and multiplication), absolute value and lattice operations (binary min and max) on $\mathbb{R}$;
  - The characteristic function $[\cdot \in \cdot] : \Omega \times \mathcal{A}_\Omega \to \{0, 1\} \subseteq \mathbb{R}$ of the membership relation $\in$ on $\Omega \times \mathcal{A}_\Omega$;
  - Banach operations (addition, scalar multiplication) and norm on $\mathcal{L}^\infty_{\Omega,\mathbb{R}}$ (namely, $\|f\| = \sup_{x \in \Omega} |f(x)|$ for $f \in \mathcal{L}^\infty_{\Omega,\mathbb{R}}$—note that an almost-everywhere null function $f$ has positive norm per this definition unless $f = 0$ everywhere);
  - The evaluation map $\mathcal{L}^\infty_{\Omega,\mathbb{R}} \times \Omega \to \mathbb{R} : (f, x) \mapsto f(x)$;
  - The Banach lattice operations (binary min and max) on $\mathcal{L}^\infty_{\Omega,\mathbb{R}}$;
  - The unary operation of pointwise absolute value $f \mapsto |f|$ on $\mathcal{L}^\infty_{\Omega,\mathbb{R}}$ where $|f| \in \mathcal{L}^\infty_{\Omega,\mathbb{R}}$ is the function $x \mapsto |f(x)|$;
  - The Boolean algebra operations of union, intersection and relative complement $S \mapsto S^c = \Omega \setminus S$ on $\mathcal{A}_\Omega$;
  - The characteristic-function map $\chi : \mathcal{A}_\Omega \to \mathcal{L}^\infty_{\Omega,\mathbb{R}} : S \mapsto \chi_S$;
  - A positive finite measure $\mu$ on $\Omega$;
  - The integration operator $I : \mathcal{L}^\infty_{\Omega,\mathbb{R}} \to \mathbb{R} : f \mapsto \int_\Omega f \, d\mu$.

Let $\mathcal{L}$ be any Henson language including sort symbols $\mathbb{R}, \Omega, \mathcal{A}_\Omega, \mathcal{L}^\infty_{\Omega,\mathbb{R}}$ as well as constant and function symbols matching the lists above, and let Th$_f$ be the $\mathcal{L}$-theory of such structures $\mathcal{M} = (S, F, C)$, where $S$ is the list of sorts, $F$ the collection of distinguished functions, and $C$ the set of distinguished elements of $\mathcal{M}$. An (abstract) pre-integration structure is a model of Th$_f$. An integration structure is a saturated model of Th$_f$. If $\mathcal{M}$ is any pre-integration structure (whether saturated or not), then via interpretation of constants, the membership relation $[\cdot \in \cdot]$, and the evaluation map, we may identify $\mathbb{R}^\mathcal{M}$ with $\mathbb{R}, \mathcal{A}_\Omega^\mathcal{M}$ with a Boolean algebra of subsets of $\Omega^\mathcal{M}$, and $\mathcal{L}^\infty_{\Omega,\mathbb{R}}^\mathcal{M}$ with a set of functions $\Omega \to \mathbb{R}$. However,
A_\Omega need not be a \sigma\text{-algebra. Accordingly, } \mu^{\|} \text{ is typically just a finitely (not countably) additive measure on } (\Omega^{\|}, \mathcal{A}_\Omega^{\|}) \text{, while elements } f \in \mathcal{L}_\Omega^{\infty} \text{ are identified with uniformly bounded functions on } \Omega \text{ that may only be approximately } A_\Omega\text{-measurable}^5 \text{. Nevertheless, in earlier work we have shown how the classical (i.e., } \sigma\text{-additive) theory of integration of bounded measurable functions over a finite measure space and the corresponding version of the Dominated Convergence Theorem are recovered essentially verbatim in saturated Henson integration structures via an analogous construction to that of Loeb measure in nonstandard analysis} \text{ [DnI17].}

A.2. Loeb structures.

**Definition A.1** (Loeb structure). Let \text{Th}_{\text{Loeb}} be the reduct of the \mathcal{L}\text{-theory } \text{Th}_f \text{ of integration structures with a positive measure to the language } \mathcal{L}' \text{ obtained by removing from } \mathcal{L} \text{ the symbol for sort } \mathcal{L}_\Omega^{\infty} \text{ as well as all functions and constants involving } \mathcal{L}_\Omega^{\infty} \text{ (such as the symbol } I \text{ for the integral). A model of } \text{Th}_{\text{Loeb}} \text{ is a } \text{pre-Loeb structure. (Note that } M \text{ may be an } \mathcal{L}'\text{-structure for a language } \mathcal{L}' \text{ properly extending the language } \mathcal{L}' \text{ of Loeb structures, and thus have other sorts, functions and constants prescribed by } \mathcal{L} \text{ but not by } \mathcal{L}' \text{.)}

A Loeb structure is a saturated pre-Loeb structure.

Note that, for the present discussion, we are requiring the measure } \mu^{\|} \text{ in a pre-Loeb structure } M \text{ to be positive.}

If } M \text{ is any pre-Loeb structure, the set underlying a given } A \in \mathcal{A}_\Omega^{\|} \text{ is}

\[ [A] = \{ x \in \Omega^{\|} : \llbracket x \in A \rrbracket = 1 \}. \]

We may (externally) identify } A \text{ with } [A] \text{ since}

\[ (\forall A)(\forall B)(A = B \leftrightarrow (\forall x)(\llbracket x \in A \rrbracket = \llbracket x \in B \rrbracket)) \]

is a sentence in } \text{Th}_{\text{Loeb}} \text{.}

**Definition A.2** (Loeb measure and Loeb-measurable sets). Let } M \text{ be a pre-Loeb structure with positive measure } \mu = \mu^{\|} \text{.}

A set } S \subseteq \Omega^{\|} \text{ is } A_\Omega\text{-measurable (or just measurable) if } S = [A] \text{ for some } A \in \mathcal{A}_\Omega^{\|} \text{ (i.e., if } "S \in \mathcal{A}_\Omega^{\|}" \text{—modulo the identification of } S = [A] \text{ with } A \text{ itself).}

A set } S \subseteq \Omega^{\|} \text{ is } \mu\text{-measurable (or Loeb-measurable (modulo } \mu) \text{) if for every } \epsilon > 0 \text{ there exist measurable } A, B \in \mathcal{A}_\Omega^{\|} \text{ such that } [A] \subseteq S \subseteq [B] \text{ and } \mu(B - A) \leq \epsilon.

The Loeb measure of a Loeb-measurable set } S \text{ is}

\[ \mu_L(S) = \sup\{ \mu(A) : A \in \mathcal{A}_\Omega^{\|}, [A] \subseteq S \} = \inf\{ \mu(B) : B \in \mathcal{A}_\Omega^{\|}, [B] \supseteq S \}. \]

The Loeb algebra of } M \text{ is the collection } \llbracket A \rrbracket_\mu \text{ of all Loeb-measurable subsets of } \Omega^{\|.}

\text{A function } f \text{ on } \Omega \text{ is approximately } A_\Omega\text{-measurable if for all rational } r < s \text{ there exists } A \in \mathcal{A}_\Omega \text{ such that } f^{-1}((-\infty, r)) \subseteq A \subseteq f^{-1}((-\infty, s)) \text{—a property axiomatizable by countably many Henson formulas in the logic of approximate satisfaction} \text{ [DnI17, Proposition 4.4].}

\text{Although Henson’s languages have no conditional connective } "\rightarrow", \text{ when } P \text{ is a discrete predicate (i.e., a term taking only the values } 0, 1 \text{,} \text{ a non-Henson formula such as } P(x) \rightarrow \varphi(x) \text{ can be semantically identified with the Henson formula } (P(x) \leq 1/2) \vee \varphi(x). \text{ (By contrast, the converse } \varphi(x) \rightarrow P(x) \text{ is not semantically equivalent to a Henson formula in general.) When both } P, Q \text{ are discrete, a biconditional } P(x) \leftrightarrow Q(x) \text{ can similarly be rewritten as a Henson formula. The assertion } "R \text{ is discrete}" \text{ is captured by the Henson formula } (\forall x)(R(x) = 0 \vee R(x) = 1), \text{ where } "R(x) = r" \text{ is itself an abbreviation for } "(R(x) \leq r) \land (R(x) \geq r)".}
Note that $[A]_\mu$ is an external collection of subsets of $\Omega^\mathbb{M}$. It depends on $\mathbb{M}$ and has no intrinsic definition otherwise. It is easy to check that $[A]_\mu$ is an algebra of sets (i.e., closed under finite unions and intersections as well as complements). In fact, as soon as $\mathbb{M}$ is at least $\omega$-saturated (i.e., types over a countable set of parameters are realized), $[A]_\mu$ is a $\sigma$-algebra that is complete for $\mu_L$ in the sense that any subset of a $\mu_L$-null set is itself $\mu_L$-null ([Dml17], Proposition 3.4). On the other hand, no degree of saturation ensures that $[A]_\mu$ is closed under unions of subfamilies of size $\omega_1$ or more.

A.3. Integration frameworks. We need to introduce the notion of (real) integration framework, which generalizes integration structures as presented in section A.1. Roughly speaking, an integration framework is a saturated model of the theory of the operations of integration ([Dml17], Proposition 3.4). On the other hand, no degree of saturation ensures that $[A]_\mu$ is closed under unions of subfamilies of size $\omega_1$ or more.

Consider the reduct $\tilde{\mathbb{M}}$ of a classical pre-integration structure $\mathbb{M}$, obtained by removing from $\mathbb{M}$ the distinguished measure $\mu$ and all the functions involving $\mu$ (including the integration operator $I$). Now expand $\mathbb{M}$ to a structure $\mathbb{M}'$ with a new Banach sort $\mathbb{M}_\Omega$ containing all finite (signed, real-valued) measures $\mu$ on $\Omega$ plus the following distinguished functions and constants:

- Constants: The zero measure $0 \in \mathbb{M}_\Omega$;
- Functions:
  - Vector space operations of addition and scalar multiplication on $\mathbb{M}_\Omega$;
  - Banach norm of total variation on $\mathbb{M}_\Omega$: $\|\mu\| = \sup\{A \in \mathcal{A}_\Omega : |\mu(A)| + |\mu(A^c)|\}$;
  - The inclusion maps:
    - $\Omega \hookrightarrow \mathcal{A}_\Omega : x \mapsto \{x\}$,
    - $\Omega \hookrightarrow \mathbb{M}_\Omega : x \mapsto \delta_x$ (the unit point mass at $x$);
  - The evaluation map $\mathbb{M}_\Omega \times \mathcal{A}_\Omega \to \mathbb{R} : (\mu, A) \mapsto \mu(A)$ (which is 1-Lipschitz by definition of the norm $\|\cdot\|$ on $\mathbb{M}_\Omega$);
  - The total variation map $\mathcal{L}^\infty_{\mathbb{M}_\Omega} \to \mathcal{L}^\infty_{\mathcal{M}_\Omega}$: $\mu \mapsto |\mu|$ where $|\mu|$ is the (positive) measure of total variation of $\mu$: $|\mu|(A) = \sup\{|\mu(A \cap B)| + |\mu(A \cap B^c)| : B \in \mathcal{A}_\Omega\}$;
  - The integration operator $\langle \cdot, \cdot \rangle : \mathcal{L}^\infty_{\mathbb{M}_\Omega} \times \mathcal{M}_\Omega \to \mathbb{R} : f \mapsto \langle f, \mu \rangle = \int f \, d\mu$.

Given a language $\mathcal{L}$ for pre-integration structures, let $\mathcal{L}'$ be obtained from $\mathcal{L}$ by removing the symbols $\mu, I$ for the distinguished measure and integral operator, and adding a new sort symbol $\mathbb{M}_\Omega$ as well as new constant and function symbols per the list above.

**Definition A.3 (Real integration framework).** A classical (real) pre-integration framework is any $\mathcal{L}'$-structure $\mathbb{M}'$ as described above. An (abstract) real pre-integration framework is any model of the Henson theory $\text{Th}_{f,\mathbb{R}}$ of classical pre-integration frameworks\(^7\)

More generally, any structure $\mathbb{M}$ in a language expanding $\mathcal{L}'$ such that the $\mathcal{L}'$-reduct of $\mathbb{M}$ is a pre-integration framework in the above sense will be called a pre-integration framework.

A.4. Banach integration frameworks. There is no completely general notion of integration of functions taking values in an arbitrary Banach space $\mathcal{B}$—not even for bounded functions $F : \Omega \to \mathcal{B}$ on a finite measure space $(\Omega, \mathcal{A}_\Omega)$. However, it is very natural to require that any such notion of Banach integration should build upon the classical integral of real-valued functions. Our viewpoint is that any reasonable notion of Banach integration

\(^7\)It is straightforward to verify that $\text{Th}_{f,\mathbb{R}}$ is a uniform theory.
must expand a (pre-)integration framework to a Banach (pre-)integration framework per Definition A.4 below.

Consider expansions \( \mathcal{M}' = (S', F', C') \) of real integration frameworks \( \mathcal{M} = (S, F, C) \) where \( S' \supset S \) contains two new sorts \( B \) and \( L^\infty_{\Omega, B} \), while \( F' \supset F \) and \( C' \supset C \) contain new functions and symbols as follows:

- Addition, scalar product and norm on \( B \) making it a real Banach space.
- Addition, scalar product and norm on \( L^\infty_{\Omega, B} \) making it a Banach space.
- The zero elements of \( B \) and \( L^\infty_{\Omega, B} \).
- An evaluation map \( L^\infty_{\Omega, B} \times \Omega \rightarrow B : (F, x) \mapsto F(x) \) such that \|F\| = \sup\{\|F(x)\| : x \in \Omega\}. (Thus, elements of \( L^\infty_{\Omega, B} \) may be identified with functions \( \Omega \rightarrow B \)).
- The operation of multiplication \( L^\infty_{\Omega, B} \times L^\infty_{\Omega, B} \rightarrow L^\infty_{\Omega, B} : (f, F) \mapsto fF \) (such that \((fF)(x) = f(x)F(x)\) for all \( x \in \Omega \)) under which \( L^\infty_{\Omega, B} \) is a \( L^\infty_{\Omega, B} \)-module.
- The inclusion map \( B \rightarrow L^\infty_{\Omega, B} : T \mapsto T(\cdot) \) where \( T(\cdot) \in L^\infty_{\Omega, B} \) is identified via evaluation with the constant function \( \Omega \rightarrow B : x \mapsto T \).
- The pointwise-norm map \( |\cdot| : L^\infty_{\Omega, B} \rightarrow L^\infty_{\Omega, B} \) satisfying \( |F|(x) = \|F(x)\| \) for all \( F \in L^\infty_{\Omega, B}, x \in \Omega \).
- An operation of Banach integration, namely a pairing \( \langle \cdot, \cdot \rangle : L^\infty_{\Omega, B} \times M_\Omega \rightarrow B \) satisfying the following properties:
  1. \( \langle \cdot, \cdot \rangle \) is bilinear;
  2. \( \langle \cdot, \cdot \rangle \) is compatible with the integration \( \langle \cdot, \cdot \rangle \) of real functions:
     a. For all \( T \in B \) and \( f \in L^\infty_{\Omega, B} \): \( \langle fT, \mu \rangle = (f, \mu)T \).
     b. For all \( F \in L^\infty_{\Omega, B} \): \( \|\langle F, \mu \rangle\| \leq \|\langle |F|, |\mu| \rangle\| \).

**Definition A.4** (Banach integration framework). A language \( L' \) expanding the language \( L \) of real pre-integration frameworks with the new sort symbols plus symbols for the functions and constants above is called a language for Banach integration frameworks.

Let \( \text{Th}_{f, B} \) be the Henson \( L \)-theory of real pre-integration frameworks, and let \( \text{Th}_{f, B} \) extend \( \text{Th}_{\tilde{L}} \) with further Henson \( L' \)-axioms capturing the properties of new sorts, functions and constants stated above (in semantically equivalent terms, let \( \text{Th}_{f, B} \) be the Henson \( L' \)-theory of those expansions \( \mathcal{M}' \) of real pre-integration frameworks having the properties above).⁸

A Banach pre-integration framework is a model of \( \text{Th}_{f, B} \). More generally, if \( \tilde{L} \) is a language extending \( L' \) and \( \mathcal{M} \) is an \( \tilde{L} \)-structure whose reduct \( \mathcal{M} | L' \) is a model of \( \text{Th}_{f, B} \), we shall still call \( \mathcal{M} \) a Banach pre-integration framework.

A Banach integration framework is a saturated model of \( \text{Th}_{f, B} \).

**Remark A.5**. The question whether a real pre-integration framework \( \mathcal{M} \) admits an expansion to a Banach pre-integration framework \( \mathcal{M}' \) is very delicate. In general, the answer may be negative. However, when \( \Omega^\# \) is a finite set the answer is affirmative: It suffices to let \( (L^\infty_{\Omega, B})^\# \) be the set of all functions \( F : \Omega^\# \rightarrow \mathbb{B}^\# \), and also let \( \langle F, \mu \rangle = \sum_{x \in \Omega^\#} F(x)\mu(\{x\}) \). The remaining ingredients of the expansion are defined in the obvious manner. Similarly, an expansion \( \mathcal{M}' \) also exists if the Banach sort \( \mathbb{B}^\# \) has finite dimension (using a basis for \( \mathbb{B}^\# \), real-valued integration extends to \( \mathbb{B}^\# \)-valued integration in the straightforward classical fashion).

⁸The verification that \( \text{Th}_{f, B} \) is a uniform theory is routine.
Averages of Unitary Polynomial Actions

20

A.5. A Dominated Convergence Theorem for nets of functions in Banach integration frameworks. In order to formulate a version of the Dominated Convergence Theorem for integration frameworks below, we fix a directed set \((\mathbb{D}, \preceq)\) so we can eventually discuss convergence of nets on it. Classical sequences indexed by the directed set \((\mathbb{N}, \leq)\) of natural numbers are of particular interest. Only infinite directed sets are useful as tools to define and study notions of convergence in analysis and topology; on the other hand, critical results such as Theorem 5 depend on the countability of the directed set, so we may as well fix an infinite countable directed set \((\mathbb{D}, \preceq)\) for the reminder of the manuscript (this hypothesis will be made explicit whenever needed).

Definition A.6. Fix a directed set \((\mathbb{D}, \preceq)\). For \(i \in \mathbb{D}\), the final segment of \(\mathbb{D}\) starting at \(i\) is \(\mathbb{D}_{\geq i} = \{j \in \mathbb{D} : j \succeq i\}\) (i.e., the set of elements equal to or greater than \(i\) in \(\mathbb{D}\)). A \(\mathbb{D}\)-net \(a_\bullet\) in a metric space \((X, d)\) is any function \(\mathbb{D} \to X : i \mapsto a_i\). The spread of \(a_\bullet\) from \(i\) is

\[
\text{spr}_{\geq i}(a_\bullet) = \sup_{j,k \succeq i} d(a_j, a_k).
\]

The oscillation of \(a_\bullet\) is

\[
\text{osc}(a_\bullet) = \inf_{i \in \mathbb{D}} \text{spr}_{\geq i}(a_\bullet).
\]

The net \(a_\bullet\) converges if \(\text{osc}(a_\bullet) = 0\).

Theorem 5. (Dominated Convergence Theorem in Banach integration frameworks) Fix a countable directed set \(\mathbb{D}\). Let \(\mathcal{M}\) be any (saturated) Banach integration framework. Let \(\varphi_\bullet\) be a bounded \(\mathbb{D}\)-net in \((L^\infty_{\Omega})^\#\). For every \(x \in \Omega^\#\) and \(\mu \in \mathcal{M}_\Omega\), let \(\varphi_\bullet(x)\) denote the net \((\varphi_j(x) : j \in \mathbb{D})\) and \(\{\varphi_\bullet, \mu\}\) the net \((\{\varphi_j, \mu\} : j \in \mathbb{D})\) in \(\mathcal{B}^\#\). Then we have

\[
\text{osc}(\{\varphi_\bullet, \mu\}) \leq \|\mu\| \sup_{x \in \Omega^\#} \text{osc}(\varphi_\bullet(x)).
\]

In particular, if the net \(\varphi_\bullet(x)\) is convergent for all \(x \in \Omega^\#\), then \(\{\varphi_\bullet, \mu\}\) is convergent.

The proof of Theorem 5 below is an adaptation of our earlier one for real-valued functions ([Dn17], Proposition 5.3).

Definition A.7. If \(X\) is any nonempty set, let \(P^*_\text{fin}(X)\) be the family of finite nonempty subsets of \(X\). We call a filter (or prefilter) \(\mathcal{F}\) on \(P^*_\text{fin}(X)\) greedy if it extends the prefilter consisting of all sets \(X_{\geq S} = \{T \in P^*_\text{fin}(X) : T \supseteq S\}\) for \(S \in P^*_\text{fin}(X)\).

Note that \(\{X \supseteq S : S \in P^*_\text{fin}(X)\}\) is a prefilter on \(P^*_\text{fin}(S)\) since \(X_{\geq S} \cap X_{\geq T} = X_{\geq S \cup T}\). By a routine application of the axiom of choice, greedy ultrafilters on \(P^*_\text{fin}(X)\) exist whenever \(X\) is nonempty. Observe that the principal ultrafilter generated by a fixed \(S \in P^*_\text{fin}(X)\) is greedy precisely when \(S = X\); thus, if \(X\) is infinite, greedy ultrafilters on \(P^*_\text{fin}(X)\) are nonprincipal.

Lemma A.8. Let \(f : \Omega^\# \to \mathbb{R}\) be bounded and (externally) \(\mu_\text{L}\)-measurable. Then there exists \(\bar{f} \in (L^\infty_{\Omega, \bar{\mu}})^\#\) such that \(f(x) = \bar{f}(x)\) for \(\mu_\text{L}\)-almost all \(x \in \Omega^\#\) and \(\inf_x f(x) \leq \bar{f} \leq \sup_x f(x)\).

Proof. Let \(f\) be a \(\mu_\text{L}\)-measurable bounded external function on \(\Omega^\#\), and let \(a = \inf_x f\), \(b = \sup_x f\). The assertion is trivial if \(a = b\) or if \(\|\mu\| = 0\)—just take a constant \(\bar{f}\) in \([a, b]\).

Otherwise, we have \(a < b\) and, replacing \(\mu\) with \(|\mu|\), we may assume \(\mu\) to be a positive measure without loss of generality. By definition of Loeb measurability, for rational \(r \in [a, b]\) and integer \(n \geq 1\) there exist \(A^*_n, B^*_n \in \mathcal{A}^\#_\Omega\) such that \([A^*_n] \subseteq \{f \leq r\}, [B^*_n] \subseteq \{f \geq r\}\), and

\[\text{i.e., } \preceq \text{ is a nonstrict partial order on } \mathbb{D} \text{ such that any two } i, j \in \mathbb{D} \text{ have an upper bound } k.\]
functions such that \( \sigma \) the first assertion follows from Lemma A.8.

Let \( f_n^r = a \cdot (1 - \chi_{B_n^r}) + r \cdot \chi_{B_n^r} \) and \( g_n^r = b \cdot (1 - \chi_{A_n^r}) + r \cdot \chi_{A_n^r} \). Let \( Q \) be the set of rational numbers in \([a, b]\). The construction of \((A_n^r)\) and \((B_n^r)\) implies that \( f_n^r \leq f_n^r \leq g_n^r \leq g_n^r \) for all \( r \in Q \) and \( m \leq n \). For \( I \in \mathcal{P}_\text{fin}^r(Q) \) of cardinality \( n \), let \( f^I = \max\{f_n^r : r \in I\} \), \( g^I = \min\{g_n^r : r \in I\} \). Observe that \( f_n^r \leq f^I \leq f^I \leq g^I \leq g^I \leq g_n^r \) if \( I \subseteq J \), \( r \in I \) and \( \text{card}(I) \geq n \). Since \( a < b \) by assumption, \( Q \) is infinite countable. Let \( \mathcal{U} \) be a greedy ultrafilter on \( \mathcal{P}_\text{fin}^r(Q) \). By saturation, there are \( \tilde{f}, \tilde{g} \in (\mathcal{L}_\Omega^∞)^{\mathcal{U}} \) realizing the \( \mathcal{U} \)-ultralimit of the types \( tp_S(f^I, g^I) \) over the set of parameters \( S = \{\mu, a, b\} \cup \{A_n^r, B_n^r\}_{r \in Q, n \in \mathbb{N}^*} \). From the construction of \( \mathcal{U} \) as a greedy ultrafilter, the definition of ultralimit, and the meaning of realization of a type, it is easy to verify that

\[
a \leq f_n^r \leq \tilde{f} \leq \tilde{g} \leq g_n^r \leq b \quad \text{for all } r \in Q \text{ and } n \geq 1.
\]

It follows that for fixed \( r \in Q \) we have \( S_r := \bigcup_n A_n^r = \bigcup_n \{g_n^r \leq r\} \subseteq \{\tilde{g} \leq r\} \). On the other hand, by construction of \( A_n^r \) we have \( S_r \subseteq \{f \leq r\} \) and \( \mu_L(S_r) = \sup_n \mu(A_n^r) = \mu_L\{f \leq r\} \). Thus, \( S_r \subseteq \{f \leq r\} \cap \{\tilde{g} \leq r\} \) and \( \mu_L(S_r) = \mu_L\{f \leq r\} \); hence, \( \{f \leq r\} \) is \( \mu_L \)-almost included in \( \{\tilde{g} \leq r\} \). By a completely analogous argument, \( \{\tilde{f} \geq r\} \mu_L \)-almost includes \( \{f \geq r\} \). These almost-inclusions for every (rational) \( r \in Q \) are easily shown to imply the \( \mu_L \)-a.e. inequalities \( \tilde{g} \leq f \leq \tilde{f} \). However, \( \tilde{f} \leq \tilde{g} \), so in fact \( \tilde{f} = f = \tilde{g} \) (\( \mu_L \)-a.e.) \( \square \)

**Lemma A.9.** Fix \( \mu \in \mathcal{M}_\Omega^\mathcal{U} \) and let \( a_\bullet = (a_i : i < \omega) \) be a sequence of external \( \mu_L \)-measurable functions \( \Omega^\mathcal{U} \to \mathbb{R} \). Then there exist \( \sigma, \iota \in (\mathcal{L}_\Omega^\infty)^{\mathcal{U}} \) such that \( \sigma(x) = \sup_{i<\omega} a_i(x) \) and \( \iota(x) = \inf_{i<\omega} a_i(x) \) for \( \mu_L \)-a.e. all \( x \in \Omega^\mathcal{U} \).

If the (external) sequence \( a_\bullet \) consists of internal functions, i.e., it is a sequence in \( (\mathcal{L}_{\Omega^\infty})^{\mathcal{U}} \), then \( \sigma, \iota \) may be chosen so \( \sigma \geq \sup_{i<\omega} a_i \) and \( \iota \leq \inf_{i<\omega} a_i \).

**Proof.** It is routine to show that \( f = \sup_{i<\omega} a_i \) and \( g = \inf_{i<\omega} a_i \) are \( \mu \)-measurable, so the first assertion follows from Lemma A.8.

When \( a_\bullet \) is a sequence of internal functions, let \( \mathcal{U} \) be any nonprincipal ultrafilter on \( \omega \) and let \( \sigma \) realize the \( \mathcal{U} \)-ultralimit of the types \( tp_S(a^k) \) over the set of parameters \( S = \{\mu\} \cup \{a_i\}_{i \in \omega} \), where \( a^k = \max\{a_i : i \leq k\} \). (Recall that \( \mathcal{L}_{\Omega^\infty} \) is endowed with the binary lattice operation \( \max\{a, b\} \), which trivially defines \( n \)-ary maximum operations for all \( n \geq 1 \).) The verification that \( \sigma \) has the required properties is routine. The construction of \( \iota \) is identical upon replacing “max” by “min”. \( \square \)

**Proof of Theorem 5.** The asserted inequality evidently holds if \( \mu = 0 \). Otherwise, using a Jordan decomposition \( \mu = \mu_+ - \mu_- \) where \( \mu_+ = (|\mu| + \mu)/2 \) and \( \mu_- = (|\mu| - \mu)/2 \) are positive, the proof is easily reduced to the case in which \( \mu \) is a probability measure, which we assume henceforth.

Choose \( C \) such that \( ||\varphi_i|| \leq C \) for all \( i \). For \( j, k \in \mathbb{D} \) let \( \varphi_{j}^{k} = |\varphi_k - \varphi_j| \in (\mathcal{L}_{\Omega^\infty})^{\mathcal{U}} \). Since \( \mathbb{D} \) is countable, Lemma A.9 implies that for each \( i \in \mathbb{D} \) there is \( \sigma^i \in (\mathcal{L}_{\Omega^\infty})^{\mathcal{U}} \) with \( ||\sigma^i|| \leq 2C \) such that \( \sigma^i \) is \( \mu_L \)-a.e. equal to \( \text{spr}_{\leq i} \varphi_\bullet = \sup_{i, k \leq i} \varphi_{j}^{k} \). Similarly, \( \text{osc} \varphi_\bullet = \inf_{i} \text{spr}_{\leq i} \varphi_\bullet \) is \( \mu_L \)-a.e. equal to \( \inf_{i} \sigma^i \), hence to some \( \omega \in (\mathcal{L}_{\Omega^\infty})^{\mathcal{U}} \) with \( ||\omega|| \leq 2C \).

Let \( s = \sup_{i} \text{osc}(\varphi_i(x)) \) and fix \( t > s \). Since \( \{x : \text{osc} \varphi_i(x) \geq t\} \) is empty (by choice of \( s \) and \( t \)) and \( \omega(x) = \text{osc}(\varphi_i(x)) \) for \( \mu_L \)-a.e. \( x \), the set \( \{x : \omega(x) \geq t\} \) is \( \mu_L \)-null. Since \( \omega = \inf_{i} \text{spr}_{\leq i} \varphi_\bullet \) (\( \mu_L \)-a.e.), we have \( \inf_{i} \mu_L \{x : \text{spr}_{\leq i} \varphi_\bullet (x) \geq t\} = 0 \). Thus, for arbitrary
fixed $\varepsilon > 0$ we have $\mu_L \{ x : \text{spr}_{\geq i} \varphi_\bullet(x) \geq t \} \leq \varepsilon$ for some $i = i_\varepsilon \in \mathbb{D}$. (This depends crucially on the hypothesis that $\mathbb{D}$ is countable.) It follows that for $j, k \geq i$:

$$\|\varphi_k - \varphi_j, \mu\| \leq \|\varphi_k - \varphi_j, \mu\| = \int |\varphi_k(x) - \varphi_j(x)| \, d\mu_L(x)$$

$$= \left( \int \{ x : \text{spr}_{\geq i} \varphi_\bullet(x) < t \} + \int \{ x : \text{spr}_{\geq i} \varphi_\bullet(x) \geq t \} \right) |\varphi_k(x) - \varphi_j(x)| \, d\mu_L(x)$$

$$\leq t \cdot \mu_L \{ x : \text{spr}_{\geq i} \varphi_\bullet(x) < t \} + 2C \varepsilon \leq t + 2C \varepsilon.$$ 

This proves that $\text{spr}_{\geq i} \langle \varphi_\bullet, \mu \rangle \leq t + 2C \varepsilon$. As $t > s$ and $\varepsilon > 0$ are arbitrary, $\text{osc} \langle \varphi_\bullet, \mu \rangle \leq s$. □

A.6. A Uniform Metastability Principle for nets in Henson structures.

**Proposition A.10** (Uniform Metastability Principle (UMP)). Fix a directed set $(\mathbb{D}, \preceq)$. Fix a Henson language $\mathcal{L}$ including constants $(a_j : j \in \mathbb{D})$ of all a common sort $\mathbb{S}$. Let $\mathcal{T}$ be a uniform $\mathcal{L}$-theory such that for every model $\mathcal{M}$ of $\mathcal{T}$ the net $a_\bullet = (a_j^\bullet : j \in \mathbb{D})$ is convergent. Then there exists a metastability rate $E_\bullet = E_\mathcal{T}$ depending only on $\mathcal{T}$ that applies uniformly to all sequences $a_\bullet^\#$ in all models $\mathcal{M}$ of $\mathcal{T}$.

**Proof.** ([DnII17], Proposition 2.4.) Assume no such rate of metastability exists. Then there exist $\varepsilon > 0$ and a sampling $\eta \in \prod_{i \in \mathbb{D}} \mathcal{P}^*_{\text{fin}}(\mathbb{D}_{\geq i})$ such that for every $S \in \mathcal{P}^*_{\text{fin}}(\mathbb{D})$ there is a model $\mathcal{M} = \mathcal{M}^S_{\eta}$ of $\mathcal{T}$ such that $a_\bullet = a_\bullet^\#$ satisfies $\varepsilon \leq \text{spr}_S(a_\bullet) = \max \{ d(a_j, a_k) : j, k \in \eta_i \}$ for all $i \in S$. By the compactness theorem for Henson logic, there is a model $\mathcal{M}$ of $\mathcal{T}$ such that $a_\bullet = a_\bullet^\#$ satisfies $\text{spr}_S(a_\bullet) \geq \varepsilon$ for all $i \in \mathbb{D}$, and hence $\text{osc}(a_\bullet) \geq \varepsilon$, contradicting the hypothesis that $a_\bullet^\#$ converges. □

**References**

[Aus15a] Tim Austin, *Pleasant extensions retaining algebraic structure, I*, J. Anal. Math. **125** (2015), 1–36. MR 3317896

[Aus15b] Tim Austin, *Pleasant extensions retaining algebraic structure, II*, J. Anal. Math. **126** (2015), 1–111. MR 3358029

[Aus16] Tim Austin, *Non-conventional ergodic averages for several commuting actions of an amenable group*, J. Anal. Math. **130** (2016), 243–274. MR 3574655

[Ber87] V. Bergelson, *Weakly mixing PET*, Ergodic Theory Dynam. Systems **7** (1987), no. 3, 337–349. MR 912373 (89g:28022)

[Bir31] G. D. Birkhoff, *Proof of the ergodic theorem*, Proc. Nat. Acad. Sci. U. S. A. **17** (1931), 656–660.

[DnII17] Eduardo Dueñez and José Iovino, *Model theory and metric convergence I: Metastability and dominated convergence*, Beyond first order model theory, CRC Press, Boca Raton, FL, 2017, pp. 131–187. MR 3729326

[Fur77] Harry Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204–256. MR 0498471 (58 #16583)

[HI02] C. Ward Henson and José Iovino, *Ultraproducts in analysis*, Analysis and logic (Mons, 1997), Analysis and logic (Mons, 1997), vol. 262, Cambridge Univ. Press, Cambridge, 2002, pp. 1–110. MR 1 967 834

[HK05] Bernard Host and Bryna Kra, *Nonconventional ergodic averages and nilmanifolds*, Ann. of Math. (2) **161** (2005), no. 1, 397–488. MR 2150389 (2007b:37004)

[Iov14] José Iovino, *Applications of model theory to functional analysis*, Dover Publications, Inc., Mineola, NY, 2014, Revised reprint of the 2002 original, With a new preface, notes, and an updated bibliography. MR 3362124

[Lei98] A. Leibman, *Polynomial sequences in groups*, J. Algebra **201** (1998), no. 1, 189–206. MR 1608723 (99c:20044)
[Lei02] ______, Polynomial mappings of groups, Israel J. Math. 129 (2002), 29–60. MR 1910931 (2003g:20060)

[Lei05] ______, Convergence of multiple ergodic averages along polynomials of several variables, Israel J. Math. 146 (2005), 303–315. MR 2151605

[Sze75] E. Szemerédi, On sets of integers containing no k elements in arithmetic progression, Acta Arith. 27 (1975), 199–245, Collection of articles in memory of Juriı̆ Vladimirovič Linnik. MR 0369312 (51 #5547)

[Tao08] Terence Tao, Norm convergence of multiple ergodic averages for commuting transformations, Ergodic Theory Dynam. Systems 28 (2008), no. 2, 657–688. MR 2408398 (2009k:37012)

[Tao12] ______, Walsh’s ergodic theorem, metastability, and external Cauchy convergence, http://terrytao.wordpress.com/2012/10/25/ Oct 2012.

[vN32] John von Neumann, Proof of the quasi-ergodic hypothesis, Proc. Nat. Acad. Sci. U. S. A. 18 (1932), no. 1, 70–82.

[Wal12] Miguel N. Walsh, Norm convergence of nilpotent ergodic averages, Ann. of Math. (2) 175 (2012), no. 3, 1667–1688. MR 2912715

[ZK16] Pavel Zorin-Kranich, Norm convergence of multiple ergodic averages on amenable groups, J. Anal. Math. 130 (2016), 219–241. MR 3574654

Department of Mathematics, The University of Texas at San Antonio, One UTSA Circle, San Antonio, TX 78249-0664, U.S.A.
E-mail address: eduardo.duenez@utsa.edu
E-mail address: jose.iovino@utsa.edu