COLLOQUIUM MATHEMATICUM

VOL. 110 2008 NO. 1

LARGE SETS OF INTEGERS AND HIERARCHY OF MIXING PROPERTIES OF MEASURE PRESERVING SYSTEMS

BY

VITALY BERGELSON (Columbus, OH) and
TOMASZ DOWNAROWICZ (Wrocław)

Abstract. We consider a hierarchy of notions of largeness for subsets of $\mathbb{Z}$ (such as thick sets, syndetic sets, IP-sets, etc., as well as some new classes) and study them in conjunction with recurrence in topological dynamics and ergodic theory. We use topological dynamics and topological algebra in $\beta\mathbb{Z}$ to establish connections between various notions of largeness and apply those results to the study of the sets $R^{\varepsilon}_{A,B} = \{n \in \mathbb{Z} : \mu(A \cap T^n B) > \mu(A)\mu(B) - \varepsilon\}$ of times of "fat intersection". Among other things we show that the sets $R^{\varepsilon}_{A,B}$ allow one to distinguish between various notions of mixing and introduce an interesting class of weakly but not mildly mixing systems. Some of our results on fat intersections are established in a more general context of unitary $\mathbb{Z}$-actions.

Introduction. Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system. Given $\varepsilon > 0$ and $A, B \in \mathcal{B}$ with $\mu(A) > 0$, $\mu(B) > 0$, let us define the set of times of "fat intersection" by

$$R^{\varepsilon}_{A,B} = \{n \in \mathbb{Z} : \mu(A \cap T^n B) > \mu(A)\mu(B) - \varepsilon\}.$$

When $A = B$, the sets $R^{\varepsilon}_{A,B}$ are intrinsically connected with the various enhancements and applications of the classical Poincaré recurrence theorem and are relatively well understood. For example, the Khinchin recurrence theorem ([Kh]; see also [B1, Section 5]) says that for any, not necessarily ergodic, probability measure preserving system $(X, \mathcal{B}, \mu, T)$, any A with $\mu(A) > 0$ and any $\varepsilon > 0$, the set $R^{\varepsilon}_{A,A}$ is syndetic (i.e., has bounded gaps). This result, in turn, follows from the (stronger) fact that $R^{\varepsilon}_{A,A}$ is a $\triangle^*$-set, i.e. it has nontrivial intersections with any set of the form $\{n_i - n_j\}_{i>j}$, where $(n_i)_{i\in\mathbb{N}}$ is an injective sequence in $\mathbb{Z}$ (see Theorem 3.1 below). Note that while every $\triangle^*$-set is syndetic, not every syndetic set is a $\triangle^*$-set (consider for example the set of all odd numbers).

2000 *Mathematics Subject Classification*: 37A25, 37B20.

Key words and phrases: weak mixing, mild mixing, fat intersections, IP-sets, idempotents, central sets, upper Banach density.

The first author is partially supported by NSF grant DMS-0600042. Research of the second author is supported by grant MENII 1 P03A 021 29, Poland.

Assuming ergodicity, one can show that the sets $R^\varepsilon_{A,B}$ are always syndetic. On the other hand, the natural question whether they are always of the form $E+k$, where E is a $\triangle^*$-set, $k \in \mathbb{Z}$, has, in general, a negative answer (see Theorem 1.7 below). One of the goals of this paper is to introduce and study some new notions of largeness with the intention of better understanding the sets of times of fat intersection and to apply them to the study of mixing properties of dynamical systems.

In order to formulate our main results we have first to introduce and discuss the pertinent notions of largeness. This is done in Section 1, at the end of which the formulations of our main theorems are given. In Section 2 we take a closer look at notions of largeness which are intrinsically related to topological dynamics. In particular, we show that one of the notions playing the decisive role in this paper, namely that of D-sets (see the definition in Section 1), can be naturally viewed as the extension of Furstenberg's notion of central sets (see [F, p. 161]) which proved to be very useful in various applications of ergodic theory to combinatorics (see for example [B1] and [B-M]). In Section 3 we provide the proofs of the characterizations of ergodicity, weak, mild and strong mixing in terms of sets of times of fat intersection. In Section 4 we give an example of a dynamical system which not only proves that two of the classes under study ($\mathcal{IP}^*_+$ and $\mathcal{D}^*_\bullet$) are not contained in one another, but also that one cannot replace $\mathcal{D}^*_\bullet$ by its intersection with $\mathcal{IP}^*_+$ in the characterization of the weak mixing property. Finally, in Section 5 we apply our notions to isolate certain nonempty subclass of weakly mixing but not mildly mixing transformations. The paper is concluded by an appendix containing an explicit example of a topological dynamical system with specific properties. Besides being interesting in its own right, the existence of such a system is important in one of the proofs in Section 2.

Acknowledgements. We are greatful to Sarah Bailey-Frick, Ronnie Pavlov and Neil Hindman for useful comments. We also thank the anonymous referee for numerous pertinent remarks and suggestions.

1. Notions of largeness via duality. Let $\mathcal{F}$ be a family of nonempty subsets of the integers $\mathbb{Z}$. We will denote by $\mathcal{F}^*$ the *dual* family consisting of all sets G such that $G \cap F \neq \emptyset$ for every $F \in \mathcal{F}$. The family $\mathcal{F}$ is *partition regular* if, whenever $F \in \mathcal{F}$ is represented as a union of finitely many sets, then at least one of them belongs to $\mathcal{F}$. It is not hard to verify that if $\mathcal{F}$ is partition regular then its dual $\mathcal{F}^*$ is a filter: the intersection of two elements of $\mathcal{F}^*$ belongs to $\mathcal{F}^*$. (The other requirement for a filter, the property of being closed under taking supersets, is obvious for $\mathcal{F}^*$.) Two elementary examples of this kind are

1. Fix some $n_0 \in \mathbb{Z}$ and let $\mathcal{F} = \{F \subset \mathbb{Z} : n_0 \in F\}$. Then $\mathcal{F}^* = \mathcal{F}$.
2. Let $\mathcal{F} = \mathcal{I} = \{F \subset \mathbb{Z} : |F| = \infty\}$ (infinite sets). Then $\mathcal{F}^* = \mathcal{I}^* = \{F \subset \mathbb{Z} : |\mathbb{Z} \setminus F| < \infty\}$ (cofinite sets).

Let us now mention a more subtle example.

3. A set $F \subset \mathbb{Z}$ is called an *IP-set* if it contains the set $\mathrm{FS}(S)$ of finite sums of some sequence $S = (s_n)_{n \geq 1}$ of nonzero integers:
$$\mathrm{FS}(S) = \{s_{n_1} + \cdots + s_{n_k} : n_1 < \cdots < n_k,\ k \in \mathbb{N}\}.$$
Let $\mathcal{IP}$ be the family of all IP-sets. One can show that both IP-sets and IP*-sets (members of the dual family $\mathcal{IP}^*$) can be characterized (with the help of Hindman's theorem) in terms of idempotents in $\beta\mathbb{Z}$ (see Definition 1.2 below and Theorems 1.2 and 1.5 in [B2]).

Recall that a family of subsets of $\mathbb{Z}$ which is both partition regular and a filter is called an *ultrafilter* (or a maximal filter). Note the obvious fact that the union of any collection of ultrafilters is partition regular. Also, while an intersection of ultrafilters need not be an ultrafilter, it is always a filter. The collection of all ultrafilters is denoted by $\beta\mathbb{Z}$ and, endowed with an appropriate topology, becomes the Stone–Čech compactification of $\mathbb{Z}$. There is a natural semigroup structure in $\beta\mathbb{Z}$ extending the addition operation of $\mathbb{Z}$ (for more details see [H-S]).

The above examples have the following interpretation in terms of ultrafilters: In the first example, $\mathcal{F}$ is nothing but a so-called *principal* ultrafilter, i.e., the ultrafilter representing n_0 in $\beta\mathbb{Z}$ (and so also is $\mathcal{F}^*$). In the second, $\mathcal{F}$ is the union of all nonprincipal ultrafilters and $\mathcal{F}^*$ is the intersection of all such ultrafilters. Finally, in the third example $\mathcal{F}$ is the union of all nonprincipal ultrafilters which are idempotents for the natural semigroup structure of $\beta\mathbb{Z}$ (that is, $\mathcal{F}$ is the union of all idempotents except zero) and $\mathcal{F}^*$ is the intersection of the nonzero idempotents (cf. [B2, Theorem 2.15(i), p. 20]). The above facts are special cases of the following more general statement:

LEMMA 1.1.

(1) *If $\mathcal{F}$ is an ultrafilter then $\mathcal{F}^* = \mathcal{F}$.*
(2) *If $\mathcal{F} = \bigcup_\alpha \mathcal{F}_\alpha$ then $\mathcal{F}^* = \bigcap_\alpha \mathcal{F}^*_\alpha$.*

In particular, whenever $\mathcal{F}$ is a union of some collection of ultrafilters, then $\mathcal{F}^*$ is the intersection of that collection.

Intuitively, if we have the union of a rich collection of families, its dual contains relatively few "very large" sets, namely, sets which intersect nontrivially every member of every family in this collection. This approach to "largeness" will be utilized throughout this paper: a set is "large" if it belongs to the dual of a rich family of sets containing a union of many ultrafilters.

For this reason the first example above is not very useful: the family $\mathcal{F}$ is just a single ultrafilter (and so also is $\mathcal{F}^*$), moreover, $\mathcal{F}^*$ contains finite sets, so being a member of $\mathcal{F}^*$ cannot be considered a criterion for largeness. But leaving this exceptional example aside, we will investigate a whole hierarchy of notions of largeness constructed with the help of dual families, of which the property of being a member of $\mathcal{I}^*$ is the strongest. Several important notions of largeness can be introduced with the help of idempotent ultrafilters.

In order to facilitate the discussion we list some of the important families of large sets in the following definition. (Note that the family $\mathcal{IP}$ appearing in item (1) below was already introduced above.)

Definition 1.2.

(1) The collection $\mathcal{IP}$ (of IP-sets) is the union of all nonzero idempotents $0 \neq p \in \beta\mathbb{Z}$. Accordingly, $\mathcal{IP}^*$ is the intersection of all nonzero idempotents.
(2) The collection $\mathcal{D}$ (of D-sets) is the union of all idempotents $p \in \beta\mathbb{Z}$ such that every member of p has positive upper Banach density ([1]). Accordingly, $\mathcal{D}^*$ is the intersection of all such idempotents.
(3) The collection $\mathcal{C}$ (of C-sets or *central sets*) is the union of all minimal idempotents ([2]). Accordingly, $\mathcal{C}^*$ is the intersection of all minimal idempotents.

Since every member of a minimal idempotent has positive upper Banach density ([3]), we have $\mathcal{C} \subset \mathcal{D}$, hence, directly from the definitions, we obtain the following hierarchy:

$$\mathcal{I}^* \subset \mathcal{IP}^* \subset \mathcal{D}^* \subset \mathcal{C}^* \subset \mathcal{C} \subset \mathcal{D} \subset \mathcal{IP} \subset \mathcal{I}.$$

As we will see below, all these inclusions are in fact proper.

We introduce two more notions of largeness defined via duality:

Definition 1.3.

(1) A subset $F \subset \mathbb{Z}$ is called a $\triangle$-*set*, or we say that F belongs to the family Δ, if there exists an injective sequence $S = (s_n)_{n\geq 1}$ of integers such that the difference set $\triangle(S) = \{s_i - s_j : i > j\}$ is contained in F.

([1]) The *upper Banach density* of a set $E \subset \mathbb{Z}$ is defined as $\limsup_{m-n\to\infty} \frac{1}{m-n}|E \cap [n, m-1]|$. If the corresponding limit exists then it is called the *Banach density* of E.

([2]) An idempotent is *minimal* if it belongs to a minimal right ideal in $\beta\mathbb{Z}$ (see [H-S] and [B2] for details). See also the discussion in Section 2 on various equivalent definitions of the notion of central set.

([3]) This follows from the stronger fact that every member of a minimal idempotent is *piecewise syndetic* (see [B2, Theorem 2.4 and Exercise 7]).

(2) A set $F \subset \mathbb{Z}$ is *thick* if it contains arbitrarily long intervals $[a,b] = \{a, a+1, \ldots, b\}$. The collection of all thick sets will be denoted by $\mathcal{T}$. The dual family $\mathcal{T}^*$ is easily seen to coincide with the collection of all syndetic sets (i.e., sets having bounded gaps).

The family $\varDelta$ is the union of a collection of ultrafilters (see [B-H2, Definition 1.6 and Lemma 1.9]), while that of thick sets is not (because it is not partition regular). It is known (and not very hard to see) that every thick set is an IP-set and every IP-set is a $\triangle$-set, but not the other way around. In particular, the collection of ultrafilters whose union is $\varDelta$ contains more than just idempotents. The hierarchy of notions of largeness introduced so far is as follows:

$$\text{cofinite} = \mathcal{I}^* \subset \varDelta^* \subset \mathcal{IP}^* \subset \mathcal{D}^* \subset \mathcal{C}^* \subset \mathcal{T}^* = \text{syndetic}.$$

Given a family $\mathcal{F}$ and $k \in \mathbb{Z}$, the *shifted family* is defined by $\mathcal{F} + k = \{F + k : F \in \mathcal{F}\}$, where $F + k = \{n + k : n \in F\}$. The extreme classes in the above diagram are shift invariant; a shifted cofinite set remains cofinite, a shifted syndetic set remains syndetic. The other classes fail to be shift invariant. This is not surprising for notions involving idempotents due to the simple fact that if p is an idempotent then $p + k$ is not (unless $k = 0$). To see that the family $\varDelta^*$ is not shift invariant note that it contains the set of all even integers while it does not contain the set of all odd integers. When $\mathcal{F}$ is not shift invariant, there are two natural ways of building a shift invariant family from it:

Definition 1.4. For a given family $\mathcal{F}$, $\mathcal{F}_+$ denotes the union $\bigcup_{k\in\mathbb{Z}}(\mathcal{F}+k)$ while $\mathcal{F}_\bullet$ denotes the intersection $\bigcap_{k\in\mathbb{Z}}(\mathcal{F}+k)$.

When applying these operations to a dual family $\mathcal{F}^*$, we will write $\mathcal{F}^*_+$ and $\mathcal{F}^*_\bullet$, skipping the parentheses in what should formally be $(\mathcal{F}^*)_+$ and $(\mathcal{F}^*)_\bullet$. This convention complies with the existing notation e.g. for IP^*_+-sets. We will call $\mathcal{F}^*_+$ the *extended dual* family. Note that, in general, $\mathcal{F}^*_+$ is not a dual family. On the other hand, by Lemma 1.1(2), the family $\mathcal{F}^*_\bullet$ is the dual of $\mathcal{F}_+$ (it could be written as $(\mathcal{F}_+)^*$, but we will not use this confusing symbol). The elements of $\mathcal{F}^*_\bullet$ are much larger than those of $\mathcal{F}^*$ as they must intersect every set in the family $\mathcal{F}_+$ which is much richer than $\mathcal{F}$. If $\mathcal{F}$ is a union of ultrafilters, so is $\mathcal{F}_+$, thus $\mathcal{F}^*_\bullet$ is an intersection of ultrafilters, and hence in particular a filter. It seems that the type $\mathcal{F}^*_\bullet$ of shift invariant families has not been sufficiently recognized in the existing literature. Here is the diagram including all dual and extended dual classes related to the families discussed so far:

$$\begin{array}{ccccccccccc}
\mathcal{I}^*_\bullet & \subset & \varDelta^*_\bullet & \subset & \mathcal{IP}^*_\bullet & \subset & \mathcal{D}^*_\bullet & \subset & \mathcal{C}^*_\bullet & \subset & \mathcal{T}^*_\bullet \\
\| & & \cap & & \cap & & \cap & & \cap & & \| \\
\mathcal{I}^* & \subset & \varDelta^* & \subset & \mathcal{IP}^* & \subset & \mathcal{D}^* & \subset & \mathcal{C}^* & \subset & \mathcal{T}^* \\
\| & & \cap & & \cap & & \cap & & \cap & & \| \\
\mathcal{I}^*_+ & \subset & \varDelta^*_+ & \subset & \mathcal{IP}^*_+ & \subset & \mathcal{D}^*_+ & \subset & \mathcal{C}^*_+ & \subset & \mathcal{T}^*_+
\end{array}$$

Now we will show that in this diagram no other inclusions hold except the ones that are shown and those obtained by composition. First, observe the following property of all $\triangle$-sets F: a certain distance between elements of F appears infinitely many times. Indeed, in any difference set $\triangle(S)$ with $S = (s_n)$ the distance $|s_2 - s_1|$ occurs between all pairs of elements $s_n - s_1$ and $s_n - s_2$ $(n > 2)$. Obviously, the same property holds for shifted $\triangle$-sets. We conclude that the set of powers of 2 does not contain any shift of any $\triangle$-set, which implies that the complement of the powers of 2 is $\triangle^*_\bullet$. Hence the family $\varDelta^*_\bullet$ is larger than the class $\mathcal{I}^*$ of cofinite sets. Further, the set of all odd numbers is a $\triangle^*_+$-set and is not an IP-set, hence in the diagram it escapes any class contained in $\mathcal{C}^*$. Likewise, the set of all even integers is a $\triangle^*$-set and not $\mathrm{C}^*_\bullet$. The construction of an $\mathrm{IP}^*_\bullet$ but not $\triangle^*_+$ is provided in Theorem 2.11(1). The existence of a $\mathrm{D}^*_\bullet$ but not IP^*_+ will follow from Theorem 1.7 below. A $\mathrm{C}^*_\bullet$ but not D^*_+ example is our Theorem 2.11(2). Finally, a syndetic set which is not C^*_+ is provided in [B2, Theorem 2.10]. All other "unwanted" inclusions are now eliminated by superposition.

It is worth noticing that the family $\mathcal{C}_+$ (shifted central sets) coincides with $\mathcal{PS}$, the family of piecewise syndetic or PS-sets (a set is *piecewise syndetic* if it is the intersection of a thick set and a syndetic set). The proof can be found in [H-S, Theorem 4.43(c)]. Thus, $\mathcal{C}^*_\bullet = \mathcal{PS}^*$, the dual to the family of piecewise syndetic sets. Elements of this dual can be easily identified as "syndetically thick", meaning that for every $E \in \mathcal{PS}^*$ and $n \geq 1$, intervals of length n appear in E with bounded gaps (in [D] such sets have been called S-sets).

This paper focuses on the role the notions of largeness of subsets of $\mathbb{Z}$ play in ergodic theory and topological dynamics. Recall that (X, T) is a (*topological*) *dynamical system* if X is a compact Hausdorff space and $T : X \to X$ is a homeomorphism. The families defined as unions of certain idempotents (IP-sets, C-sets and D-sets) have interpretations (and indeed convenient alternative definitions) as families of sets of the form $\{n \in \mathbb{Z} : (T^n x, T^n y) \in U\}$, where y is a recurrent point, the pair (x, y) is proximal ([4]) and U is a neighborhood of (y, y) in $X \times X$.

While the families of IP-, C- and D-sets are useful in topological dynamics, their dual and extended dual families find applications in ergodic theory.

([4]) Two points x, y in a topological dynamical system (X, T) are *proximal* if the set of pairs $(T^n x, T^n y)$ has an accumulation point on the diagonal.

For example we will show how notions of largeness such as D_+^*, $\mathrm{D}_\bullet^*$ and $\mathrm{IP}_\bullet^*$ can be used to characterize the familiar ergodic-theoretic notions of ergodicity, weak mixing and mild mixing. As already mentioned in the introduction, in this paper we study the sets of times of fat intersection,

$$R_{A,B}^\varepsilon = \{n \in \mathbb{Z} : \mu(A \cap T^n B) > \mu(A)\mu(B) - \varepsilon\}.$$

In the spirit of Khinchin's theorem we will locate those sets for specific types of systems in our diagram of "large sets". First of all, the Khinchin theorem can be strengthened: the set $R_{A,A}^\varepsilon$ is always $\triangle^*$ (see Theorem 3.1). It is not very surprising that the sets $R_{A,A}^\varepsilon$ do not form a shift invariant family. However, to capture the fat intersections for arbitrary two sets A and B (this only makes sense in ergodic systems) one needs a shift invariant notion simply because $R_{A,T^kB}^\varepsilon = R_{A,B}^\varepsilon + k$. The most natural candidate, namely the class Δ_+^*, turns out to be too restrictive. The sets of times of fat intersection are in this class only for certain rather special types of systems, e.g. systems with discrete spectrum. The smallest class in our diagram that suffices for all ergodic systems is the extended dual $\mathcal{D}_+^*$. However, curiously enough, we will show that for the notions of mixing under study, the sets $R_{A,B}^\varepsilon$ are "captured" by the more restrictive shift invariant dual of the form $\mathcal{F}_\bullet^*$: for weak mixing this is $\mathcal{D}_\bullet^*$, for mild mixing this is $\mathcal{IP}_\bullet^*$, and for mixing, directly from the definition, this is $\mathcal{I}^*$ (which can also be written as $\mathcal{I}_\bullet^*$).

Let us briefly recall some of the ergodic-theoretic notions:

Definition 1.5.

(1) The system $(X, \mathcal{B}, \mu, T)$ has *discrete spectrum* if $L^2(\mu)$ is spanned by the eigenfunctions of the unitary operator induced by T.
(2) The system $(X, \mathcal{B}, \mu, T)$ is *weakly mixing* if the product system $(X \times X, \mathcal{B} \times \mathcal{B}, \mu \times \mu, T \times T)$ is ergodic.
(3) The system $(X, \mathcal{B}, \mu, T)$ is *mildly mixing* if there are no nontrivial rigid L^2-functions. (A function $f \in L^2(\mu)$ is *rigid* if $T^{n_k} f \to f$ in L^2 for some sequence $n_k \to \infty$.)
(4) The system $(X, \mathcal{B}, \mu, T)$ is *mixing* if for any two sets $A, B \in \mathcal{B}$ one has $\mu(A \cap T^n B) \to \mu(A)\mu(B)$ as $n \to \infty$.

We stress that the appropriate categorization of fat intersections for all pairs of sets is in many cases equivalent to a given ergodic-theoretic notion, which makes the hierarchy of largeness very useful. In the following theorem we collect formulations of various familiar notions of mixing in terms of sets $R_{A,B}^\varepsilon$ (see also *Final remarks* at the end of the paper). Some of the items in Theorem 1.6 below are mere reformulations of well known facts (see for example [F]), others have relatively easy proofs provided in Section 3 (see also Remark 1 below).

Given a system $(X,\mathcal{B},\mu,T)$ we denote by $\mathcal{R}(X,\mathcal{B},\mu,T)$ the family of all sets of times of fat intersection in this system, $\mathcal{R}(X,\mathcal{B},\mu,T)=\{R^\varepsilon_{A,B}:\varepsilon>0, A,B\in\mathcal{B}\}$.

THEOREM 1.6. *Let $(X,\mathcal{B},\mu,T)$ be an invertible probability measure preserving system. Then*:

(1) *For any $A\in\mathcal{B}$ and any $\varepsilon>0$ we have $R^\varepsilon_{A,A}\in\Delta^*$.*
(2) *If $(X,\mathcal{B},\mu,T)$ is ergodic and has discrete spectrum then $\mathcal{R}(X,\mathcal{B},\mu,T)\subset\Delta^*_+$.*
(3) *$(X,\mathcal{B},\mu,T)$ is ergodic $\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{D}^*_+\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{C}^*_+\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{T}^*$.*
(4) *$(X,\mathcal{B},\mu,T)$ is weakly mixing $\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{D}^*\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{D}^*_\bullet\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{C}^*\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{C}^*_\bullet$.*
(5) *$(X,\mathcal{B},\mu,T)$ is mildly mixing $\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{IP}^*\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{IP}^*_\bullet$* (cf. Chapter 9, Section 4 in [F]).
(6) *$(X,\mathcal{B},\mu,T)$ is mixing $\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{I}^*\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\Delta^*\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\Delta^*_\bullet$* (see [K-Y] and Remark 1(f) below).

REMARK 1. Some of the equivalences in Theorem 1.6 are trivial or very easy, some others follow from known results:

(a) It is clear that in (3) only the first equivalence needs a proof, the other two follow from inclusions of the families of sets and from the fact that in nonergodic systems the family $\mathcal{R}(X,\mathcal{B},\mu,T)$ contains the empty set, so $\mathcal{R}(X,\mathcal{B},\mu,T)\not\subset\mathcal{T}^*$.

(b) Since for any system $(X,\mathcal{B},\mu,T)$ the family $\mathcal{R}(X,\mathcal{B},\mu,T)$ is shift invariant, it is obvious that $\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{F}\Leftrightarrow\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{F}_\bullet$ for any family $\mathcal{F}$.

(c) Notice that if $\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{F}$, where $\mathcal{F}$ is a filter, then, intersecting each set $R^\varepsilon_{A,B}$ with the corresponding set R^ε_{A,B^c}, we find that the sets of times of *accurate intersection*

$$Q^\varepsilon_{A,B}=\{n\in\mathbb{Z}:|\mu(A\cap T^nB)-\mu(A)\mu(B)|<\varepsilon\}$$

also belong to $\mathcal{F}$. In other words, $\mathcal{Q}(X,\mathcal{B},\mu,T)=\{Q^\varepsilon_{A,B}:\varepsilon>0,\ A,B\in\mathcal{B}\}\subset\mathcal{F}$. (Clearly, since $Q^\varepsilon_{A,B}\subset R^\varepsilon_{A,B}$, the converse implication also holds.) Thus statements (4)–(6) in Theorem 1.6 are equivalent to analogous statements with $\mathcal{R}(X,\mathcal{B},\mu,T)$ replaced by $\mathcal{Q}(X,\mathcal{B},\mu,T)$.

(d) If the system $(X,\mathcal{B},\mu,T)$ is not weakly mixing then one can find two sets A and B and an $\varepsilon>0$ such that $R^\varepsilon_{A,A}$ and $R^\varepsilon_{A,B}$ are disjoint (cf. Theorem 4.31 in [F]), and so they cannot both be C*-sets. Thus the condition $\mathcal{R}(X,\mathcal{B},\mu,T)\subset\mathcal{C}^*$ implies weak mixing ([5]). Hence, using remark (b) and

([5]) The same fact is proved (by a different method) in [K-Y, Proposition 5.2], in response to a question formulated in the preliminary version of this paper.

obvious inclusions, we conclude that also in (4) only the first equivalence needs a proof. In fact, the first implication $\Rightarrow$ can be deduced (using (c)) from the classical fact that weak mixing is equivalent to the condition

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=0}^{n-1}|\mu(A\cap T^iB)-\mu(A)\mu(B)|=0 \quad \text{for any sets } A,B\in\mathcal{B}.$$

(e) The first equivalence in (5) (in terms of accurate intersections) is Proposition 9.22 of [F], the second follows from (b).

(f) The first equivalence in (6) applied to accurate intersections becomes merely the definition of mixing. The second equivalence in (6) (formulated for accurate intersections) is nontrivial and has recently been proved in [K-Y, Theorem 4.4] (see also [K-Y, Proposition 5.1], formulated in response to a question in the preliminary version of our paper).

To summarize the content of the above remark, only (1), (2) and portions of (3) and (4) require proofs (see Theorems 3.1, 3.2, 3.8 and 3.9 in the next section, respectively). For completeness we will also provide a proof of (5) using the language of idempotents (see Theorem 3.10).

The following two results (which are proved in Sections 4 and 5) isolate a new class of systems defined in terms of fat intersections and situated strictly between weak and mild mixing. A priori it could happen that for weakly mixing systems the sets $R^{\varepsilon}_{A,B}$ always belong to the intersection of $\mathcal{IP}^*_+$ and $\mathcal{D}^*_\bullet$. The following theorem shows that this is not always so. (It also provides a proof that the family $\mathcal{D}^*_\bullet\setminus\mathcal{IP}^*_+$ is nonempty.)

THEOREM 1.7. *There exists a weakly mixing probability measure preserving system* $(X,\mathcal{B},\mu,T)$, *sets* $A,B\in\mathcal{B}$ *and* $\varepsilon>0$ *such that the set* $R^{\varepsilon}_{A,B}$ *is not* IP^*_+.

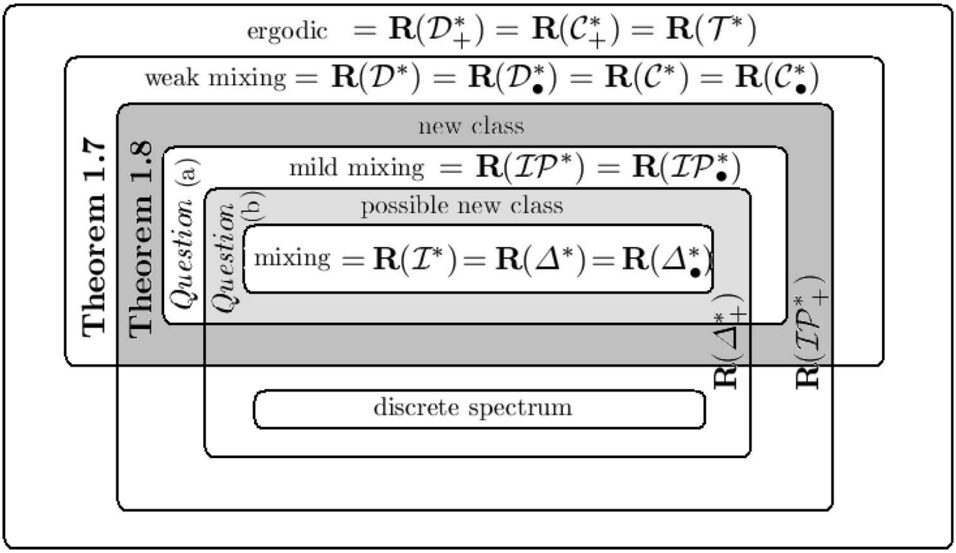
On the other hand, the requirement that all sets $R^{\varepsilon}_{A,B}$ are IP^*_+ is insufficient for mild mixing (in particular $\mathcal{D}^*_\bullet\cap\mathcal{IP}^*_+\setminus\mathcal{IP}^*$ is nonempty):

THEOREM 1.8. *There exists a weakly mixing but not mildly mixing probability measure preserving system* $(X,\mathcal{B},\mu,T)$ *such that all the sets* $R^{\varepsilon}_{A,B}$ *are* IP^*_+ (*but not all of them are* IP^*).

QUESTIONS.

(a) Does there exist a mildly mixing system for which not all sets $R^{\varepsilon}_{A,B}$ belong to $\varDelta^*_+$ (cf. Theorem 2.11(1))?
(b) Does there exist a mildly mixing nonmixing system for which all sets $R^{\varepsilon}_{A,B}$ belong to $\varDelta^*_+$? (Here we do not even know whether the family $\mathcal{IP}^*_\bullet\cap\varDelta^*_+\setminus\varDelta^*$ is nonempty.)
(c) More generally, what is the dynamical condition equivalent to $\mathcal{R}(X,\mathcal{B},\mu,T)\subset\varDelta^*_+$?

The following figure gives an overview of the classes of systems under study and inclusions between them. The symbol $\mathbf{R}(\mathcal{F})$ stands for the class of systems $(X, \mathcal{B}, \mu, T)$ such that $\mathcal{R}(X, \mathcal{B}, \mu, T) \subset \mathcal{F}$.



2. IP-sets, central sets and D-sets in topological dynamics. Recall that $\beta\mathbb{Z}$ is the Stone–Čech compactification of $\mathbb{Z}$ consisting of ultrafilters, which has a natural semigroup structure. On $\beta\mathbb{Z}$ there is also the natural action τ which extends the map $n \mapsto n+1$ on $\mathbb{Z}$.

If $p \in \beta\mathbb{Z}$ is an ultrafilter then the *p-limit* of a sequence x_n of elements of a compact space is defined by the rule

$$p\text{-}\lim x_n = x \iff (\forall \text{ open } U \ni x)\ \{n \in \mathbb{Z} : x_n \in U\} \in p.$$

The following fact will be used repeatedly in our paper: if p is an idempotent and T is a continuous self-map of a compact space then $p\text{-}\lim T^n x = y$ implies $p\text{-}\lim T^n y = y$ (see Proposition 3.2 in [B2]).

Every transitive topological dynamical system (X, T) (with a transitive point x_0) is a topological factor of $(\beta\mathbb{Z}, \tau)$ via the map $p \mapsto p\text{-}\lim T^n(x_0)$ (see e.g. Proposition 7.3 in [E]).

The orbit closure of a point x in a topological dynamical system (X, T) is the set $\overline{O}(x) = \overline{\{T^n x : n \in \mathbb{Z}\}}$. A point x in (X, T) is *recurrent* if for every neighborhood $U_x \ni x$ there exists $n \neq 0$ such that $T^n x \in U_x$.

It is known ([F, Theorem 2.17]) that the set F of return times of a recurrent point x, $F = \{n \in \mathbb{Z} : T^n x \in U_x\}$, is an IP-set. We also have

THEOREM 2.1. *A set $E \subset \mathbb{Z}$ is* IP *if and only if there exist a compact metrizable dynamical system (X, T), a pair of points $x, y \in X$ such that y*

is recurrent and (y,y) *belongs to the orbit closure of* (x,y) *in the product system* $(X\times X, T\times T)$, *and an open neighborhood* $U_{(y,y)}$ *of* (y,y) *such that* $E=\{n\in\mathbb{Z}:(T^nx,T^ny)\in U_{(y,y)}\}$.

REMARK 2. Note that if (y,y) belongs to the orbit closure of (x,y) then x and y are proximal. In general, the conditions that y is recurrent and x is proximal to y do not imply $(y,y)\in\overline{O}(x,y)$. For example, x can be a fixpoint in the orbit closure of a recurrent point $y\neq x$. In order that $(y,y)\in\overline{O}(x,y)$ the recurrence of y and the proximality of x and y must be realized along a common sequence of times.

Proof of Theorem 2.1. Let y and x be such that y is a recurrent point in X with $(y,y)\in\overline{O}(x,y)$ and let $U_{(y,y)}$ be an open neighborhood of (y,y). Consider the set $E'=\{n\in\mathbb{Z}:(T^nx,T^ny)\in U_y\times U_y\}$, where $U_y\times U_y$ is a product neighborhood of (y,y) contained in $U_{(y,y)}$. It is clear that the set E' is infinite, so it contains some $s\neq 0$. Suppose E' contains $\mathrm{FS}(S)$, where S is some finite set not containing zero. Let $V_y\subset U_y$ be an open neighborhood of y such that $T^s(V_y)\subset U_y$ for all $s\in S$. We can find $0\neq s'\notin S$ for which $(T^{s'}x,T^{s'}y)\in V_y\times V_y$. Then $(T^{s'}x,T^{s'}y)\in U_y\times U_y$ and $(T^{s'+s}x,T^{s'+s}y)\in U_y\times U_y$ for every $s\in S$. We have shown that $E'\supset\mathrm{FS}(S')$, where $S'=S\cup\{s'\}$. By induction, we will obtain a set $\mathrm{FS}(S)$ (where S is infinite) contained in E', which proves that E' (as well as E) is an IP-set. To prove the converse, consider an arbitrary IP-set E and let $x=(x(n))_{n\in\mathbb{Z}}$ be the characteristic function of E viewed as an element of the shift system $X=\{0,1\}^{\mathbb{Z}}$. Define $y=p\text{-}\lim T^nx$, where p is an idempotent such that $E\in p$ (see Definition 1.2(1)). Following the proof of Theorem 3.6 in [B2], we claim that the sequence y starts with the symbol 1: $y(0)=1$. By the definition of p-lim, the set $R=\{n\in\mathbb{Z}:(T^nx)(0)=y(0)\}$ belongs to p. So, the intersection $R\cap E$ is nonempty (it belongs to p), which implies that there exists $n\in E$ with $x(n)=y(0)$. But $x(n)=1$ if and only if $n\in E$, so $y(0)=1$. This implies that $E=\{n\in\mathbb{Z}:(T^nx,T^ny)\in U_{(y,y)}\}$, where $U_{(y,y)}$ is defined as $U_1\times X$, where U_1 is the cylinder of elements starting with 1. ■

We will now introduce C-sets and D-sets in a similar way, by imposing additional conditions on the recurrence of y.

A point y contained in a dynamical system (X,T) is *uniformly recurrent* if, for any neighborhood U of y, the set of return times $\{n\in\mathbb{Z}:T^ny\in U\}$ is syndetic. It is well known that y is uniformly recurrent if and only if the orbit closure $\overline{O}(y)$ of y is minimal.

Central sets have been defined by H. Furstenberg ([F, Def. 8.3]) as follows:

DEFINITION 2.2. A set $C\subset\mathbb{Z}$ is *central* if there exists a compact metrizable dynamical system (X,T), a point $x\in X$ proximal to a uniformly recurrent point $y \in X$ and an open neighborhood U_y of y such that

$$C = \{n \in \mathbb{Z} : T^n x \in U_y\}.$$

One can show that C is central if and only if C is a member of a minimal idempotent in $\beta\mathbb{Z}$ (see [B-H1, Corollary 6.12] and [B2, Theorem 3.6]). We have already used this equivalence in Section 1 (Definition 1.2).

Central sets can also be characterized with the help of product systems:

THEOREM 2.3. *A set $C \subset \mathbb{Z}$ is central if and only if there exist a compact metrizable dynamical system (X, T), a pair of points $x, y \in X$ where y is uniformly recurrent and (y, y) belongs to the orbit closure of (x, y) in the product system $(X \times X, T \times T)$, and an open neighborhood $U_{(y,y)}$ of (y, y) such that*

$$C = \{n \in \mathbb{Z} : (T^n x, T^n y) \in U_{(y,y)}\}.$$

Proof. As mentioned in Remark 2, even if y is recurrent and x is proximal to y, (y, y) does not have to belong to the orbit closure of (x, y). Nevertheless, it is easy to see that if y is uniformly recurrent then proximality of x and y does imply that (y, y) belongs to the orbit closure of (x, y). This observation is crucial to the proof. Let C be central, and let x and y be as in Definition 2.2. Then (y, y) belongs to the orbit closure of (x, y), and $C = \{n \in \mathbb{Z} : (T^n x, T^n y) \in U_{(y,y)}\}$, where $U_{(y,y)} = U_y \times X$. Conversely, if $C = \{n \in \mathbb{Z} : (T^n x, T^n y) \in U_{(y,y)}\}$ with assumptions on x and y as in the formulation of the theorem, then C is central directly by Definition 2.2, using (x, y) and (y, y) as a pair of points in the direct product $(X \times X, T \times T)$. Notice that (y, y) is uniformly recurrent in the product system. ∎

Now we focus on D-sets. In the introduction we have defined them analogously to C-sets by replacing minimal idempotents by a wider class of idempotents all of which have positive upper Banach density, so that the class $\mathcal{D}$ of D-sets is (strictly) intermediate between $\mathcal{IP}$ and $\mathcal{C}$. We are interested in obtaining a characterization of D-sets, analogous to those of IP-sets and C-sets (in terms of visits of $(T^n x, T^n y)$ to $U_{(y,y)}$) by imposing on y an appropriate recurrence condition, as defined below.

DEFINITION 2.4. A point y in a (not necessarily metrizable) dynamical system (X, T) is *essentially recurrent* if for any neighborhood U_y of y the set of visits $\{n \in \mathbb{Z} : T^n y \in U_y\}$ has positive upper Banach density.

Obviously, since every syndetic set has positive upper Banach density, every uniformly recurrent point is essentially recurrent. A characterization of essentially recurrent points in terms of the properties of their orbit closures is provided below.

DEFINITION 2.5. A dynamical system (Y, T) will be called *measure saturated* if the union of the topological supports of all invariant probability measures [6] carried by Y is dense in Y. In other words, for every nonempty open set U there exists an invariant measure μ such that $\mu(U) > 0$.

Note that every minimal system is measure saturated.

THEOREM 2.6. *A point y is essentially recurrent if and only if the orbit closure $\overline{O}(y)$ is measure saturated.*

Proof. First let us show that if a point y is essentially recurrent then its orbit closure is measure saturated. Let $U_y \ni y$ be an open set and let $U \ni y$ be open and such that $\overline{U} \subset U_y$. Since y is essentially recurrent, the set $A = \{n \in \mathbb{Z} : T^n y \in U\}$ has positive upper Banach density d. Let I_n be a sequence of intervals in $\mathbb{Z}$ with $|I_n| \to \infty$ (as $n \to \infty$) such that the ratios $|A \cap I_n|/|I_n|$ converge to d. Let μ_n $(n = 1, 2, \dots)$ be the normalized counting measures supported by the sets $\{T^i y : i \in I_n\}$, and let μ be a weak* accumulation point [7] of μ_n. Clearly, μ is T-invariant, supported by $\overline{O}(y)$ and satisfies $\mu(\overline{U}) > 0$, and thus $\mu(U_y) > 0$. We have proved that the closure M of the union of the supports of all invariant measures carried by $\overline{O}(y)$ contains y. Since M is a closed invariant set, it follows that $M = \overline{O}(y)$, i.e., $\overline{O}(y)$ is measure saturated.

Conversely, assume that $\overline{O}(y)$ is measure saturated. Let $U_y \ni y$ be an open set. Then there exists an invariant measure μ supported by $\overline{O}(y)$ such that $\mu(U_y) > 0$. The ergodic theorem ensures that the function

$$f(x) = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} \mathbf{1}_{U_y}(T^i(x))$$

satisfies $\int f\,d\mu = \mu(U_y) > 0$. Thus there exists $y' \in \overline{O}(y)$ with $f(y') = d > 0$. In other words, the set $R = \{n \in \mathbb{Z} : T^n y' \in U_y\}$ has natural density d, i.e., $\lim_n |R \cap [1, n]|/n = d$. Note that for any $m \in \mathbb{N}$ there exists $n \in \mathbb{Z}$ such that for any $i \in [0, m]$, $T^{n+i}(y) \in U_y$ if and only if $T^i(y') \in U_y$. It follows that the set $\{n \in \mathbb{Z} : T^n y \in U_y\}$ has positive upper Banach density (at least d) and hence y is essentially recurrent. ∎

DEFINITION 2.7. Let p be an idempotent in $\beta\mathbb{Z}$. We will call p *essential* if every member of p has positive upper Banach density.

We are in a position to provide a dynamical definition of D-sets, which is completely analogous to the characterizations of IP-sets and central sets.

---

(6) The classical Bogolyubov–Krylov theorem guarantees the existence of at least one invariant probability measure. The *topological support* of a probability measure is the smallest closed set of measure 1.

(7) A sequence of measures μ_n converges to μ *weak** if $\int f\,d\mu_n \to \int f\,d\mu$ for every continuous function f on the space Y.

THEOREM 2.8. *A set $D \subset \mathbb{Z}$ is a D-set if and only there exists a compact metrizable dynamical system (X, T), points $x, y \in X$ with y essentially recurrent for which the orbit closure of (x, y) in the product system $(X \times X, T \times T)$ contains (y, y), and an open neighborhood $U_{(y,y)}$ of (y, y) such that*

$$D = \{n \in \mathbb{Z} : (T^n x, T^n y) \in U_{(y,y)}\}.$$

Before we prove the theorem we need a series of lemmas.

LEMMA 2.9. *An idempotent $q \in \beta\mathbb{Z}$ is an essentially recurrent point in $(\beta\mathbb{Z}, \tau)$ if and only it is essential.*

REMARK 3. Glasner [G] introduces a set Z in $\beta\mathbb{Z}$ defined as the closure of the union of the supports of all invariant measures on $\beta\mathbb{Z}$ and he proves that it is a so-called *kernel* for the family of sets of positive upper Banach density. In fact one implication of the above lemma could be deduced from that result, but we choose to give an independent proof.

Proof of Lemma 2.9. Let q be essentially recurrent and let E be any element of q. The closure $\overline{E}$ of E in $\beta\mathbb{Z}$ can be interpreted as a neighborhood of q. There exists an invariant measure μ such that $\mu(\overline{E}) > 0$. Since μ is supported by the orbit closure of 0, the set of visits of 0 to this neighborhood (which is E) has positive upper Banach density (by the same argument as in the proof of Theorem 2.6). The converse is also true. The map $p \mapsto p + q$ is a factor map from $\beta\mathbb{Z}$ onto $\overline{O}(q)$, and both 0 and q map to q. A neighborhood U_q of q in $\overline{O}(q)$ lifts to a neighborhood V_q of q in $\beta\mathbb{Z}$ and the set R_q of times of visits of q in U_q contains the set R_0 of times of visits of 0 in V_q. But R_0 is a member of q (because its complement is not). Since q is assumed to be an essential idempotent, all members of q have positive upper Banach density (see Definition 2.7). It follows that R_0 has positive upper Banach density and hence, by Definition 2.4, q is essentially recurrent. ∎

It is obvious that if $\pi : X \to Y$ is a topological factor map and $y \in Y$ is uniformly recurrent then there exists a uniformly recurrent π-lift $x \in X$ of y (because the preimage of $\overline{O}(y)$ is invariant and any one of its minimal subsets must map onto $\overline{O}(y)$). The lemma below is an analogous statement for essentially recurrent points.

LEMMA 2.10. *Let $\pi : X \to Y$ be a topological factor map (surjection) between dynamical systems (X, S) and (Y, T). If y is an essentially recurrent point in Y then there exists an essentially recurrent π-lift x of y. Moreover, we can find such an x with $\overline{O}(x)$ containing no proper closed invariant subset mapped by π onto $\overline{O}(y)$.*

Proof. Applying Zorn's lemma to the family of all lifts of $\overline{O}(y)$, i.e., of closed invariant sets mapped by π onto $\overline{O}(y)$, we can find a minimal such lift $X_0 \subset X$. Let x be any lift of y contained in X_0. Since $\overline{O}(x) \subset X_0$ and it maps onto $\overline{O}(y)$, by minimality $\overline{O}(x) = X_0$. On the other hand, since every invariant measure carried by $\overline{O}(y)$ lifts to at least one invariant measure carried by $\overline{O}(x)$, the closure X_1 of the union of the supports of all invariant measures carried by $\overline{O}(x)$ maps onto a closed set containing the union of the supports of all invariant measures carried by $\overline{O}(y)$. Since y was assumed to be essentially recurrent, X_1 maps onto $\overline{O}(y)$ and hence, being a closed invariant subset of X_0, it also equals X_0. This proves that x is essentially recurrent, and that its orbit closure is a minimal lift of $\overline{O}(y)$, as required. ∎

Proof of Theorem 2.8. Let $D = \{n \in \mathbb{Z} : (T^n x, T^n y) \in U_{(y,y)}\}$, where x and y are as in the formulation of the theorem. Consider a factor map $\pi : \beta\mathbb{Z} \to \overline{O}(x,y)$ defined by $p \mapsto \pi(p) := p\text{-}\lim(T^n x, T^n y)$. By assumption, $\overline{O}(y,y) \subset \overline{O}(x,y)$. Since $\overline{O}(y,y)$ is contained in the diagonal, it is topologically conjugate to $\overline{O}(y)$ and hence (y,y) is essentially recurrent. By Lemma 2.10, we can find in $\beta\mathbb{Z}$ an essentially recurrent π-lift p_1 of (y,y) whose orbit closure is a minimal lift of $\overline{O}(y,y)$. We will show that p_1 can be replaced by an idempotent. Consider the set

$$I = \{p \in \overline{O}(p_1) : \pi(p) = (y,y)\}.$$

By an elementary verification, I is a closed semigroup of $\beta\mathbb{Z}$, so it contains an idempotent q. Since q maps to (y,y), its orbit closure maps onto $\overline{O}(y,y)$. By minimality of the lift $\overline{O}(p_1)$, q has the same orbit closure as p_1, and hence is essentially recurrent.

Finally, $D \in q$ follows from two facts: 1) $(T^n x, T^n y)$ does not belong to the neighborhood $U_{(y,y)}$ of (y,y) for all $n \in D^c$; 2) $q\text{-}\lim(T^n x, T^n y) = (y,y)$. This implies $D^c \notin q$, so that D must belong to q. We have completed the proof of one implication.

To prove the converse, let D be a D-set (i.e., a member of an essentially recurrent idempotent). Identify the characteristic function of D with a point x in $\{0,1\}^{\mathbb{Z}}$ and denote the shift transformation by T. Define $y = q(x) := q\text{-}\lim T^n x$. Since q is an idempotent, $q(y) = y$, so $q(x,y) = (y,y)$, i.e., the orbit of (x,y) accumulates at (y,y), as required. Now we repeat the argument used in the proof of Theorem 2.1: The set $R = \{n \in \mathbb{Z} : (T^n x)(0) = y(0)\}$ belongs to q, so $R \cap D \neq \emptyset$. Since $x(n) = 1$ for $n \in D$, we have $y(0) = 1$. As a consequence, $D = \{n \in \mathbb{Z} : (T^n x, T^n y) \in U_{(y,y)}\}$, $U_{(y,y)} = U_1 \times X$, U_1 denotes the cylinder of elements starting with 1, and X denotes the full shift space. The last thing we need to verify is that y is essentially recurrent. But this is immediate, because y is the image of q via the factor map $\pi : \beta\mathbb{Z} \to \overline{O}(x)$ given by $p \mapsto p(x)$, and it is elementary to see that any factor map preserves essentially recurrent points. ∎

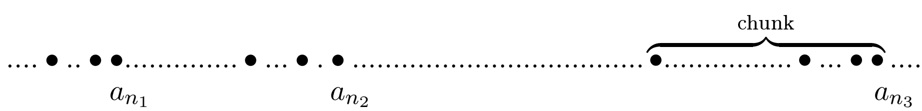
REMARK 4. Note that if y is an essentially recurrent point in the orbit closure of x and x, y are proximal, then the set $\{n \in \mathbb{Z} : T^n x \in U_y\}$ need not be a D-set. For example, let $y = (y(n))$ be a forward transitive point in the full shift on three symbols 0, 1, 2 (such a y is essentially recurrent) with $y(0) = 0$ and let x be as follows: $x(n) = 1$ whenever $y(n) = 1$ (this makes x and y proximal), $x[m, n] = y[0, n - m]$ if $y[m, n] = 2 \dots 2$ and $y(m - 1) \neq 2$ (then x is forward transitive, hence its orbit closure contains y), and $x(n) = 2$ whenever $y(n) = 0$. Then the set $\{n \in \mathbb{Z} : T^n x \in U_y\}$ is not even an IP-set: If $p\text{-}\lim T^n(x) = y$ then $p\text{-}\lim T^n(y) \neq y$ ($p\text{-}\lim T^n(y)$ has the symbol 2 at the zero coordinate), and hence p is not an idempotent.

We now focus on the dual families, more precisely, on proving the "non-containment" claims formulated in the introduction below the main diagram.

THEOREM 2.11.

(1) *There exists an* $\mathrm{IP}^*_\bullet$*-set which is not* $\triangle^*_+$.
(2) *There exists a* $\mathrm{C}^*_\bullet$*-set which is not* D^*_+.

Proof. A set of integers enumerated increasingly as (a_n) (over $n \in \mathbb{Z}$ or $n \in \mathbb{N}$) is said to have *progressive gaps* if it contains a subsequence a_{n_k} (we will call each finite subset $\{a_{n_k+1}, a_{n_k+2}, \dots, a_{n_{k+1}}\}$ a *chunk*) such that for $n_k + 1 < i \leq n_{k+1}$ one has $a_i - a_{i-1} > a_{n_{k+1}} - a_i$ (inside each chunk every gap is larger than the distance to the right end of the chunk) and $a_{n_k+1} - a_{n_k} \to \infty$ (the gaps between the chunks tend to infinity). The structure of a set with progressive gaps is shown below:



A typical example of a set with progressive gaps is the difference set $\triangle(S)$ for a rapidly (for example exponentially) increasing sequence S.

It is not hard to see that in such a set, for any fixed d, the set of elements a_i such that there exists $j > i$ with $a_j - a_i = d$ is either finite or has gaps tending to infinity (because the distance d can eventually occur only inside the chunks and then only once in every chunk).

Notice the following property of all IP-sets F: a certain distance d between elements of F appears along an IP-set. Indeed, if F contains the set of finite sums $\mathrm{FS}(S)$ with $S = (s_n)$ then the distance $|s_1|$ occurs between all pairs b and $s_1 + b$ for every $b \in \mathrm{FS}(A')$, where $A' = (s_n)_{n \geq 2}$. Clearly, an analogous statement holds for shifted IP-sets: a certain distance d occurs along a shifted IP-set. In particular, the gaps between pairs with distance d do not tend to infinity. We conclude that a set with progressive gaps does not contain any shifted IP-set.

Let $(r_k)_{k \geq 1}$ be a sequence containing all integers. Using the above observation we will now describe how to construct a set E as the union over all integers k of $\triangle$-sets E_k shifted by r_k such that E has progressive gaps, hence contains no shifted IP-sets. Clearly, the complement of such a set E is $\mathrm{IP}^*_\bullet$ and not $\triangle^*_+$. Begin with the difference set of a rapidly growing sequence, so that it has progressive gaps. Let E_1 be this difference set shifted by r_1. Inductively, suppose a union of k shifted (by $r_1,\dots,r_k$) difference sets makes a set E_k with progressive gaps. We will now create a new difference set $\triangle(S)$ with progressive gaps, whose chunks "fit into the large gaps" of $E_k - r_{k+1}$ in such a way that E_{k+1} defined as $E_k \cup (\triangle(S) + r_{k+1})$ maintains progressive gaps. Let $s_1 = 1$. Suppose we have defined $s_1,\dots,s_n \in S$. This determines a part of $\triangle(S)$ and the "shape" of the next chunk $\{s_{n+1} - s_n,\dots,s_{n+1} - s_1\}$. The next element s_{n+1} of S determines only the shifting of this new chunk. By an appropriate choice of s_{n+1} we can position this chunk in the central part of some very large gap between the chunks of $E_k - r_{k+1}$. In the union $(E_k - r_{k+1}) \cup \triangle(S)$ this gap splits into two gaps about half the original size with a new chunk in the middle. Similarly we choose s_{n+2}, and so on, until the whole sequence S is defined. It is clear that $(E_k - r_{k+1}) \cup \triangle(S)$ (and hence $E_k \cup (\triangle(S) + r_{k+1})$) maintains progressive gaps. We can pass to step $k+2$ and further steps. If in each step k we split only gaps larger than some increasing (with k) threshold value, the set $E = \bigcup_k E_k$ will maintain progressive gaps, and it is a union of shifted $\triangle$-sets, as needed to complete the proof of statement (1).

We now describe the construction of a $\mathrm{C}^*_\bullet$-set which is not D^*_+. The idea is the same as in the preceding argument, except that we will use different properties of sets. Suppose there exists a non-piecewise syndetic set E such that $E + k$ is a D-set for each $k \in \mathbb{Z}$. Such an E contains no shifted C-sets (recall that $\mathcal{C}_+ = \mathcal{PS}$). Thus the complement of E is a $\mathrm{C}^*_\bullet$-set, and since every shift of E misses a D-set, it is not a D^*_+-set.

It remains to construct a non-piecewise syndetic set E. Consider a topologically weakly mixing ([8]) and measure saturated system (X,T) with the property that the closure of the union of all minimal sets is smaller than X. An explicit construction of such an example is provided in the appendix (the example is in fact topologically mixing, with an invariant measure having full support, and with a fixpoint as the unique minimal set). Another example with the same properties was indicated by F. Blanchard: it is the substitution $0 \mapsto 001$, $1 \mapsto 1$ (see [B-H-S, Proposition 55]). Let U be an open set disjoint from another open set V containing the union of all minimal sets. Notice that the orbit closure of y is conjugate to that of (y,y) in the product system. If y is a transitive point then it is essentially recurrent, and hence so is (y,y). There exists a pair (x,y) transitive in $X \times X$ with both x and y

([8]) A topological dynamical system (X,T) is said to be *topologically weakly mixing* (resp. mixing) if for any nonempty open sets $A, B \subset X$ the set $\{n \in \mathbb{Z} : T^n A \cap B \neq \emptyset\}$ is thick (resp. cofinite).

contained in U. Then, for any integer k, the pair $(T^k x, y)$ is also transitive, hence its orbit closure contains (y, y). Thus the set $\{n-k : T^n x \in U\}$ is a D-set (write it as $\{j : (T^j T^k x, T^j y) \in U \times X\}$). This implies that any shift of the set $E = \{n \in \mathbb{Z} : T^n x \in U\}$ is a D-set, as required. This set E is not piecewise syndetic; if it were we could easily construct a uniformly recurrent point in the closure of U, which is impossible, since all such points are in V. ■

3. Applications of the dual families to unitary and measure preserving actions. This section contains proofs of the nontrivial implications in Theorem 1.6. We begin with the role of the $\triangle^*$- and $\triangle^*_+$-sets.

THEOREM 3.1 (see Theorem 1.6(1)). *In every measure preserving system the set $R^{\varepsilon}_{A,A}$ of times of fat intersection for one set A is $\triangle^*$.*

Proof (cf. [B1, p. 49]; see also [K-Y, Proposition 4.1]). First observe that if A_n is any sequence of sets of equal measure α in a probability space, then for every $\varepsilon > 0$, the inequality $\mu(A_i \cap A_j) > \alpha^2 - \varepsilon$ holds for at least one pair of indices $i < j$. Indeed, suppose otherwise and consider the function $\sum_{i=1}^n \mathbf{1}_{A_i}$. Its inner product with $\mathbf{1}$ equals $n\alpha$, while the square of its L^2-norm is easily seen to be at most $n^2(\alpha^2 - \varepsilon) + n$. For large n this contradicts the Cauchy–Schwarz inequality.

Once this is established, take any injective sequence $S = (s_n)$ and let $A_n = T^{s_n} A$. Then $\mu^2(A) - \varepsilon < \mu(T^{s_i} A \cap T^{s_j} A) = \mu(A \cap T^{s_j - s_i} A)$ for at least one pair of indices $i < j$, proving that $R^{\varepsilon}_{A,A}$ intersects $\triangle(S)$. ■

REMARK 5. We remark that the above proof actually shows that $R^{\varepsilon}_{A,A}$ has nonempty intersection with every large enough finite difference set.

THEOREM 3.2 (see Theorem 1.6(2)). *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic rotation of a compact abelian group (where μ is the Haar measure). Then for any $A, B \in \mathcal{B}$ and $\varepsilon > 0$ the set $R^{\varepsilon}_{A,B}$ is $\triangle^*_+$.*

Proof. The proof is based on a simple observation that for group rotations Khinchin's theorem takes on a stronger form. Namely, if $(X, \mathcal{B}, \mu, T)$ is a (not necessarily ergodic) compact abelian group rotation, then for any $C \in \mathcal{B}$ and $\varepsilon > 0$, one actually sees that the set

$$R^{\varepsilon}_{C} = \{n \in \mathbb{Z} : \mu(C \cap T^{-n} C) \ge \mu(C) - \varepsilon\}$$

is $\triangle^*$ (note that in the displayed formula we have $\mu(C)$ rather than $\mu(C)^2$). Indeed, let $\triangle(S) = \{s_i - s_j\}$ where $S = (s_i)$. Finding a subsequence $s_{i_k} \to \infty$ such that $T^{s_{i_k}}(e)$ converges we obtain a uniformly convergent sequence of maps $T^{s_{i_k}}$. Thus $T^{s_{i_k} - s_{i_l}}$ converges to the identity uniformly (hence strongly in $L^1(\mu)$), which implies that $\mu(C \cap T^n C) \ge \mu(C) - \varepsilon$ for some n of the form $s_{i_k} - s_{i_l}$ (belonging to $\triangle(S)$). Returning to the ergodic case and two sets $A, B \in \mathcal{B}$, let us first find (by ergodicity) an integer n_0 such that

$\mu(A \cap T^{-n_0}B) > \mu(A)\mu(B) - \varepsilon/2$. Setting $C = A \cap T^{-n_0}B$ one easily sees that $R^{\varepsilon}_{A,B} \supset R^{\varepsilon/2}_{C} + n_0$, which implies the assertion. ∎

We will now discuss the connections between essential idempotents and unitary actions. Consider a unitary operator U on a separable Hilbert space H. We will use the orthogonal decomposition $H = H_c \oplus H_{wm}$, where

$$H_c = \{x \in H : \overline{\{U^n x\}}_{n\in\mathbb{Z}} \text{ is compact in the norm topology}\},$$

$$H_{wm} = \left\{x \in H : \frac{1}{N}\sum_{n=0}^{N-1} |\langle U^n x, x\rangle| \to 0\right\}$$

(see [Kr, Section 2.4] and [B2, Theorem 4.5]). Recall that in a Hilbert space the norm convergence $\lim x_n = y$ is equivalent to the conjunction of the weak convergence of x_n to y and the convergence of norms $\lim \|x_n\| = \|y\|$. Since any unitary operator U is an isometry, the relation $p\text{-}\lim U^n x = x$ for some $p \in \beta\mathbb{Z}$ holds in the weak topology if and only if it holds in the strong topology.

LEMMA 3.3. *If $p \in \beta\mathbb{Z}$ is an idempotent then $p\text{-}\lim U^n x = x$ for any $x \in H_c$.*

Proof. By definition of H_c, U acts on the compact metric space $\overline{\{U^n x\}}_{n\in\mathbb{Z}}$ where it is distal (it is actually an isometry). In distal systems one has $p\text{-}\lim U^n x = x$ for any idempotent (if $p\text{-}\lim U^n x = y$ for an idempotent p then also $p\text{-}\lim U^n y = y$, hence x and y are proximal, and so, by distality, $x = y$). ∎

The above statement can be reversed for essential idempotents:

LEMMA 3.4. *If $p \in \beta\mathbb{Z}$ is an essential idempotent and $p\text{-}\lim U^n x = x$ for some $x \in H$ then $x \in H_c$.*

Proof. For $\varepsilon > 0$ consider the set $E = \{n \in \mathbb{Z} : \|U^n x - x\| < \varepsilon/2\}$. Clearly $E \in p$. Note that for any $n_1, n_2 \in E$ one has

$$\|T^{n_1-n_2}x - x\| = \|T^{n_1}x - T^{n_2}x\| \le \|T^{n_1}x - x\| + \|T^{n_2}x - x\| < \varepsilon.$$

Since $E \in p$, it has positive upper Banach density, which implies that $E - E$ is syndetic (see [F, Prop. 3.19(a)] or [B1, p. 8]), i.e., finitely many shifted copies of $E - E$ cover $\mathbb{Z}$. This in turn implies that finitely many preimages of the ε-ball around x cover the orbit of x. Since U is an isometry we have covered the orbit by finitely many ε-balls, hence the orbit of x is precompact, i.e., $x \in H_c$. ∎

LEMMA 3.5. *If $p \in \beta\mathbb{Z}$ is an essential idempotent then $p\text{-}\lim U^n x = 0$ weakly for any $x \in H_{wm}$.*

Proof. By compactness of the ball of radius $\|x\|$ around zero in the weak topology, there exists some y such that $p\text{-}\lim U^n x = y$ weakly. Since H_{wm} is invariant and closed, $y \in H_{wm}$. On the other hand, p is an idempotent, so $p\text{-}\lim U^n y = y$. By Lemma 3.4, $y \in H_c$. This implies $y = 0$. ■

Recall that a unitary operator U acting on a Hilbert space H is called *weakly mixing* if it has no nontrivial eigenvectors. One can show that U is weakly mixing if and only if in the decomposition $H = H_c \oplus H_{wm}$ one has $H_c = \{0\}$ (see [Kr, Thms. 3.4 and 4.4]). Let now $(X, \mathcal{B}, \mu, T)$ be an invertible weakly mixing system. It is not hard to check that in this case the unitary operator induced by T on $L^2(\mu)$ is weakly mixing in the above sense on the orthocomplement of the space of constant functions. This leads to the following corollary of Lemmas 3.4 and 3.5:

COROLLARY 3.6. *An invertible probability measure preserving system $(X, \mathcal{B}, \mu, T)$ is weakly mixing if and only if for every $f \in L^2(X)$ and any essential idempotent $p \in \beta\mathbb{Z}$,*

$$p\text{-}\lim T^n f = \int f \, d\mu \qquad (\textit{in the weak topology}).$$

Equivalently, $(X, \mathcal{B}, \mu, T)$ is weakly mixing if and only if for any $A, B \in \mathcal{B}$ and any essentially recurrent idempotent p, $p\text{-}\lim \mu(A \cap T^n B) = \mu(A)\mu(B)$.

We now turn our attention to the D*-sets. It was proved in [B2, Theorem 4.4] that a unitary operator U acting on a Hilbert space H is weakly mixing if and only if for any $\varepsilon > 0$ and any pair $x, y \in H$ the set $R^{\varepsilon}_{x,y} = \{n \in \mathbb{Z} : \langle U^n x, y\rangle > -\varepsilon\}$ is C*. We will show that a slight modification of that proof provides a somewhat stronger result.

THEOREM 3.7. *A unitary operator U acting on a Hilbert space H is weakly mixing if and only if for any $\varepsilon > 0$ and any pair $x, y \in H$ the set $R^{\varepsilon}_{x,y}$ is* D*.

Proof. If U is weakly mixing then $H_c = \{0\}$ and the result follows from Lemma 3.5. Assume now that for any ε and $x, y \in H$ the set $R^{\varepsilon}_{x,y}$ is D*. If U is not weakly mixing then there exists $x \in H_c$, $x \neq 0$. By Lemma 3.3, $p\text{-}\lim U^n x = x$ for any essential idempotent p. Then $p\text{-}\lim \langle U^n x, x\rangle = \|x\|^2$, which implies that $R^{\varepsilon}_{x,-x}$ is not D* for $\varepsilon > 0$ small enough. ■

We can now continue with proving the statements of Theorem 1.6:

THEOREM 3.8 (see Theorem 1.6(3)). *An invertible probability measure preserving system $(X, \mathcal{B}, \mu, T)$ is ergodic if and only if for any $A, B \in \mathcal{B}$ and $\varepsilon > 0$ the set $R^{\varepsilon}_{A,B}$ belongs to $\mathcal{D}^*_+$.*

Proof (cf. [B2, Theorem 4.11]). Assume that $(X, \mathcal{B}, \mu, T)$ is ergodic. Set $f = 1_A$ and $g = 1_B$. Decompose $g = g_1 + g_2$, $g_1 \in H_c$, $g_2 \in H_{wm}$. Note that

$\int g_1\,d\mu = \mu(B)$. By the von Neumann ergodic theorem,

$$\frac{1}{N}\sum_{n=0}^{N-1}\int f(T^n x)g_1(x)\,d\mu(x) \to \int f\,d\mu \int g_1\,d\mu = \mu(A)\mu(B),$$

hence there exists n_0 satisfying $\int f(T^{n_0}x)g_1(x)\,d\mu(x) > \mu(A)\mu(B) - \varepsilon$. Let p be an essential idempotent. Applying our Lemmas 3.3 and 3.5, we can write

$$\begin{aligned} p\text{-}\lim \mu(T^{n_0}A \cap T^n B) &= p\text{-}\lim \int f(T^{n_0}x)g(T^n x)\,d\mu(x) \\ &= p\text{-}\lim \int f(T^{n_0}x)g_1(T^n x)\,d\mu(x) + p\text{-}\lim \int f(T^{n_0}x)g_2(T^n x)\,d\mu(x) \\ &= \int f(T^{n_0}x)g_1(x)\,d\mu(x) + 0 > \mu(A)\mu(B) - \varepsilon. \end{aligned}$$

This implies that $R^{\varepsilon}_{A,B} - n_0 \in p$, which proves that $R^{\varepsilon}_{A,B}$ is D^*_+.

The converse is obvious: if the sets $R^{\varepsilon}_{A,B}$ are all D^*_+ then they are nonempty, which implies ergodicity. ■

THEOREM 3.9 (see Theorem 1.6(4)). *The system $(X, \mathcal{B}, \mu, T)$ is weakly mixing if and only if for any $A, B \in \mathcal{B}$ and $\varepsilon > 0$ the set $R^{\varepsilon}_{A,B}$ is D^*. Moreover, if $(X, \mathcal{B}, \mu, T)$ is weakly mixing then $R^{\varepsilon}_{A,B}$ has Banach density* 1.

Proof. Assume that $(X, \mathcal{B}, \mu, T)$ is weakly mixing. Then, by Corollary 3.6, for any $A, B \in \mathcal{B}$ and any essential idempotent p we have the equality $p\text{-}\lim \mu(A \cap T^n B) = \mu(A)\mu(B)$, and hence $R^{\varepsilon}_{A,B}$ is a D^*-set. Recalling that weak mixing can be characterized by the relation

$$\lim_{N-M\to\infty} \frac{1}{N-M}\sum_{n=M}^{N-1} |\mu(A \cap T^n B) - \mu(A)\mu(B)| = 0,$$

we see that the set $R^{\varepsilon}_{A,B}$ has Banach density 1.

To prove the converse, assume that $(X, \mathcal{B}, \mu, T)$ is not weakly mixing. If μ is ergodic then there exists an eigenfunction f which takes values in a nontrivial subgroup G of the unit circle and sends the measure μ (via the conjugate map $f^*(\mu)(A) = \mu(f^{-1}A)$) to the Haar measure λ on G. There exists a sequence of trigonometric polynomials W_k defined on the unit circle and converging in $L^2(\lambda)$ to the characteristic function of, say, the upper semicircle $\{z : |z| = 1,\ 0 \le \arg(z) < \pi\}$. Clearly, f assumes values in the upper (half-closed) semicircle with probability $\alpha \in [1/2, 2/3]$. The functions $W_k \circ f$ converge in $L^2(\mu)$ to the characteristic function of a set A of measure α. Since the powers f^k are also eigenfunctions, all eigenfunctions belong to H_c, and since H_c is a closed linear space, $\mathbf{1}_A \in H_c$ ([9]). If μ is not ergodic, the fact that H_c contains a nontrivial characteristic function is immediate.

([9]) The existence of a nontrivial characteristic function in H_c can also be deduced using the classical fact that an ergodic non-weakly mixing system has a nontrivial *Kronecker factor* isomorphic to an ergodic rotation on a compact abelian group.

Now, by Lemma 3.3 one has, for any essential idempotent, $p\text{-}\lim \mathbf{1}_{T^{-n}A} = p\text{-}\lim(T^n\mathbf{1}_A) = \mathbf{1}_A$, and hence $p\text{-}\lim \mu(A\cap T^nA^c) = p\text{-}\lim \mu(T^{-n}A\cap A^c) = 0$, so that R^{ε}_{A,A^c} is not D*. ∎

Theorem 3.10 (see Theorem 1.6(5), see also Proposition 9.22 in [F]). *The system $(X,\mathcal{B},\mu,T)$ is mildly mixing if and only if for any $A,B\in\mathcal{B}$ and $\varepsilon>0$ the set $R^{\varepsilon}_{A,B}$ is* IP*.

Proof. Assume that $(X,\mathcal{B},\mu,T)$ is mildly mixing. Then for every nonzero idempotent $p\in\beta\mathbb{Z}$ one has $p\text{-}\lim \mu(A\cap T^{-n}B) = \mu(A)\mu(B)$. To see this it suffices to verify that $p\text{-}\lim T^nf = 0$ in L^2 for every f with zero integral (and then apply this to $f = \mathbf{1}_B - \mu(B)$). Indeed, if $p\text{-}\lim T^nf = g\neq 0$ then $p\text{-}\lim T^ng = g$ (because p is an idempotent), hence g is a rigid function. This implies $R^{\varepsilon}_{A,B}\in p$, hence $R^{\varepsilon}_{A,B}$ is IP*.

To prove the converse, assume that the system is not mildly mixing. Let $f\in L^2(\mu)$ be a nonconstant real rigid function. For some $t\in\mathbb{R}$ and $\varepsilon>0$ both $A=\{x : f(x)<t\}$ and $B=\{x : f(x)\ge t+\varepsilon\}$ have positive measure. It is easy to see that the set $\{n\in\mathbb{Z} : \|T^nf-f\|<\varepsilon^2\}$ is an IP-set, and on the other hand it is disjoint from $R^{\varepsilon}_{A,B}$. Thus $R^{\varepsilon}_{A,B}$ is not an IP*-set. ∎

4. An example of a weakly mixing system for which $R^{\varepsilon}_{A,B}$ is not IP$^*_+$. Let U be a unitary operator on a separable Hilbert space H. Let $x\in H$. It is known that the sequence $a_n=\langle U^nx,x\rangle$ is positive definite, which implies $a_n=\int z^n\,d\nu$ for some probability measure ν (depending on x) supported by the unit circle $\mathbb{T}=\{z : |z|=1\}$. The action of U on the closed cyclic subspace $\overline{\mathrm{Span}\{U^nx : n\in\mathbb{Z}\}}$ is unitarily isomorphic to the multiplication by the identity function z on $L^2(\nu)$. Temporarily we restrict our attention to such actions only, i.e., H will denote $L^2(\nu)$ and U will stand for the multiplication by the element z. Recall that the Banach–Alaoglu theorem asserts that the unit ball B of $L^2(\nu)$ is weakly compact.

Clearly, U is a self-homeomorphism of B in the weak topology, hence we obtain a topological dynamical system (B,U).

Let C be a subset of $\mathbb{T}$ of positive measure ν. Suppose that $p\text{-}\lim z^n=\mathbf{1}_C$ (in the weak topology) for some ultrafilter $p\in\beta\mathbb{Z}$. We will now show that there exists an idempotent with the same property. First of all, notice that then $p\text{-}\lim z^n\mathbf{1}_C=\mathbf{1}_C$ because the weak convergence holds when restricted to C and outside of C we have changed all functions to zero. This easily implies that the set of ultrafilters p for which $p\text{-}\lim z^n=\mathbf{1}_C$ is a semigroup. It is also closed, so it does contain an idempotent. Actually one easily shows that the converse also holds: $p\text{-}\lim z^n$ is the characteristic function of a set for any idempotent p, but we will not need this.

Now assume that ν_0 is a nonatomic measure supported by a Kronecker set $\varLambda\subset\mathbb{T}$ (see [C-F-S, Appendix 4]; in particular, $\varLambda$ is a topological Cantor set). By definition, the sequence of functions (z^n) restricted to Λ is uniformly dense in the set of all continuous unimodular functions on Λ, which easily implies that this sequence is also weakly dense in the (weakly compact) set $B_0(\nu_0) \subset L^2(\nu_0)$ defined as the set of all functions f satisfying $|f| \le 1$. The system $(B_0(\nu_0), U)$ is now topologically transitive (with the constant function $\mathbf{1}$ as a transitive point), and every measurable subset C of Λ (modulo the measure ν_0) corresponds to at least one idempotent p via the relation $p\text{-}\lim U^n\mathbf{1} = \mathbf{1}_C$ in this system.

For some of the constructions below we will need a symmetric measure ν, i.e., a measure satisfying $\nu(C) = \nu(C^*)$, where $C^* = \{\bar z : z \in C\}$. Recall that for Kronecker sets $\Lambda \cap \Lambda^* = \emptyset$. Let $K = \Lambda \cup \Lambda^*$ and let $\nu = \frac{1}{2}(\nu_0 + \nu_0^*)$, where ν_0^* is a measure on Λ^* symmetric to ν_0. For $f_0 \in B_0(\nu_0)$ define $f \in B_0(\nu)$ by the rule $f(z) = f_0(z)$ for $z \in \Lambda$ and $f(z) = \overline{f_0(\bar z)}$ for $\bar z \in \Lambda^*$. The map $f_0 \mapsto f$ establishes a topological conjugacy between $(B_0(\nu_0), U)$ and $(\widetilde B_0(\nu), U)$, where $\widetilde B_0(\nu)$ now denotes the intersection of $B_0(\nu)$ with the collection of all functions satisfying the symmetry condition $f(\bar z) = \overline{f(z)}$ (in either space, U is the operator of multiplication by z). It is essential that the function z itself satisfies the above symmetry condition, so U is well defined on $\widetilde B_0(\nu)$.

We now proceed with further details of the construction of the example.

Consider $k \in \mathbb{Z}$. There are two possible cases:

$$\nu\{z : |\mathrm{Re}\, z^k| > 1/2\} > 1/3 \tag{1}$$

and (2), the opposite. It follows immediately from the definition of a Kronecker set that both cases are represented by nonempty sets of k's. An elementary (but key) observation is that if k and k_0 satisfy (2) then $k + k_0$ necessarily satisfies (1). We now fix one representative k_0 satisfying (2). If k satisfies (1) then either

$$\nu\{z : \mathrm{Re}\, z^k > 1/2\} \ge 1/6 \quad \text{or} \quad \nu\{z : \mathrm{Re}\, z^k < -1/2\} \ge 1/6.$$

For k satisfying (1) let C_k denote the larger of the above two sets (choose any one if their measures are equal). For k satisfying (2), C_k is defined as the larger of the sets $\{z : \mathrm{Re}\, z^{k+k_0} > 1/2\}$ or $\{z : \mathrm{Re}\, z^{k+k_0} < -1/2\}$. The following facts are obvious for each k:

$$C_k = C_k^*,$$
$$\Big|\int_{C_k} z^k \, d\nu\Big| \ge \frac{1}{12} \quad \text{(in case (1))},$$
$$\Big|\int_{C_k} z^{k+k_0} \, d\nu\Big| \ge \frac{1}{12} \quad \text{(in case (2))}.$$

For unified notation, define $r(k) = 0$ if k satisfies (1) and $r(k) = k_0$ if k

satisfies (2). We can now write

$$\Big|\int\limits_{C_k} z^{k+r(k)}\,d\nu\Big| \ge \frac{1}{12}.$$

Clearly, by symmetry, all the above integrals are real.

Let p_k be an idempotent corresponding to the set C_k, i.e., such that $p_k\text{-}\lim z^n = \mathbf{1}_{C_k}$ (weakly). Then

$$p_k\text{-}\lim\Big|\int z^{n+k+r(k)}\,d\nu\Big| = \Big|\int p\text{-}\lim z^{n+k+r(k)}\,d\nu\Big| = \Big|\int\limits_{C_k} z^{k+r(k)}\,d\nu\Big| \ge \frac{1}{12}.$$

Obviously, because the inequality $|\int g\,d\nu| > 1/13$ holds on a weakly open set of functions, the set of n's for which $|\int z^{n+k+r(k)}\,d\nu| > 1/13$ belongs to the idempotent ultrafilter p_k, hence contains an IP-set M_k. We have proved the following statement:

LEMMA 4.1. *Let U be a unitary operator on a Hilbert space H. If $x \in H$ has spectral measure ν symmetric and concentrated on the union K of a Kronecker set Λ and its complex conjugate reflection Λ^* then for every k there exists an IP-set M_k such that for every $n \in M_k$,*

$$|\langle U^{n+k+r(k)}x, x\rangle| > \frac{1}{13},$$

where $r(k)$ assumes only two values: 0 and some $k_0 \in \mathbb{Z}$. This implies that for $E = \{n \in \mathbb{Z} : |\langle U^n x, x\rangle| < 1/13\}$, the intersection $E\cap(E+k_0)$ is not IP_+^*.

The above construction can be applied to weakly mixing measure preserving transformations, with an interpretation in terms of fat intersections (announced in the introduction as Theorem 1.7):

THEOREM 4.2. *There exists a weakly mixing invertible measure preserving transformation $(X', \mathcal{B}', \mu', T')$, two sets $A', B' \in \mathcal{B}'$ and $\varepsilon > 0$ such that the set $R^\varepsilon_{A',B'}$ of times of ε-fat intersection is not IP_+^*. In particular, the set $R^\varepsilon_{x,y}$ discussed in Theorem 3.7 need not be IP_+^*.*

The construction will involve spectral theory of Gauss–Kronecker systems, namely the fact that there exists a weakly mixing measure preserving transformation $(X, \mathcal{B}, \mu, T)$ and a function $f \in L^2(\mu)$ with zero integral (we will write $f \in L^2_0(\mu)$) such that the spectral measure ν of f with respect to the unitary operator U_T induced by T is supported by a set $K \subset \mathbb{T}$ as described in Lemma 4.1 (see e.g. [C-F-S, Chapter 8, Section 2 and Chapter 14, Section 4).

Define $J = \{n \in \mathbb{Z} : |\langle U_T^n f, f\rangle| > 1/13\}$. By Lemma 4.1, this set contains for each k the shifted IP-set $M_k + k + r(k)$. Outside a small set $A_0 \subset X$ of measure $1/p$ ($p \in \mathbb{N}$) the function f can be uniformly, up to some $1/q$, approximated by a simple zero integral function g constant on elements of some partition $\mathcal{A} = \{A_i : i = 1, \ldots, p\}$ of X into sets of equal measure $1/p$. By choosing q and p large enough we can thus ensure that $|\langle U_T^n g, g\rangle| > 1/14$ for all $n \in J$. Denote by $\mathbf{G} = (g_i)_{i=1,\ldots,p}$ the vector with $g_i = g(A_i)$. The formula

$$F(\mathbf{A}) = \mathbf{G}\mathbf{A}\mathbf{G}^{\mathrm{T}}$$

defines a continuous function on $p \times p$ matrices $\mathbf{A} = [a_{i,j}]$, assuming value 0 at the matrix with all entries equal to $1/p^2$. Thus there exists a positive constant δ such that $|\mathbf{G}\mathbf{A}\mathbf{G}^{\mathrm{T}}| > 1/10$ implies $|a_{i,j} - 1/p^2| > \delta$ for at least one pair of indices (i, j). For given n let $\mathbf{A}_n$ denote the matrix with entries $a_{i,j} = \mu(T^n A_i \cap A_j)$. As is easily verified,

$$\langle U_T^n g, g\rangle = F(\mathbf{A}_n),$$

so for $n \in J$ we deduce that

(∗) $\mu(T^n A_i \cap A_j)$ differs from $\mu(A_i)\mu(A_j) = 1/p^2$ by at least δ for at least one pair of sets A_i, A_j (depending on n).

The final step is a construction of a pair of sets which satisfies a similar "fault of independence" (perhaps with a smaller constant) for all n in the union of $M_k + k$. These sets will be found in the direct $2p^2$-fold Cartesian product $(X', \mu', T') = (X^{\times 2p^2}, \mu^{\times 2p^2}, T^{\times 2p^2})$ as described below. Note that (X', μ', T') remains a weakly mixing system. The desired sets are:

$$A = (A_1 \times \cdots \times A_1) \times (A_2 \times \cdots \times A_2) \times \cdots \times (A_p \times \cdots \times A_p) \times$$
$$\times (T^{k_0} A_1 \times \cdots \times T^{k_0} A_1) \times (T^{k_0} A_2 \times \cdots \times T^{k_0} A_2) \times \cdots \times (T^{k_0} A_p \times \cdots \times T^{k_0} A_p),$$

with p repetitions in each parenthesis, and

$$B = (A_1 \times \cdots \times A_p) \times \cdots \times (A_1 \times \cdots \times A_p),$$

with $2p$ repetitions of $A_1 \times \cdots \times A_p$. Now observe that

$$\mu'(T'^{n+k} A' \cap B') = \prod_{i,j} \mu(T^{n+k} A_i \cap A_j) \cdot \prod_{i,j} \mu(T^{n+k+k_0} A_i \cap A_j).$$

Both products are of p^2 nonnegative numbers whose sum is 1. It is an elementary exercise that among such products the largest is $(1/p^2)^{p^2}$ achieved only if all terms are equal to $1/p^2$. Otherwise it is strictly smaller. So, by continuity, whenever at least one term of this product differs from $1/p^2$ by δ (in either direction), then the whole product is smaller than $(1/p^2)^{p^2} - \gamma$, where γ is some fixed positive number (depending only on δ).

Now let $n \in M_k$. Then either $n+k$ or $n+k+k_0$ belongs to J. So, by (∗), at least one term in at least one of the above products differs from $1/p^2$ by δ, and, as a consequence, one of the products is smaller than $(1/p^2)^{p^2} - \gamma$. Since the other product is still at most $(1/p^2)^{p^2}$, the discussed measure of intersection does not exceed

$$\left(\frac{1}{p^2}\right)^{2p^2} - \left(\frac{1}{p^2}\right)^{p^2}\gamma.$$

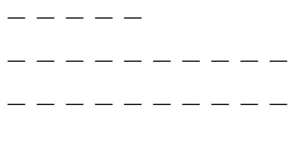
The first term coincides with $\mu'(A')\mu'(B')$. The second term is a positive constant ε. We have proved that the set $R^{\varepsilon}_{A',B'}$ misses all the shifted IP-sets $M_k + k$, so is not IP^*_+. ∎

5. An intermediate class of weakly mixing transformations. This section contains the construction announced in Theorem 1.8.

THEOREM 5.1. *There exists a nonempty class of weakly mixing rank-one rigid transformations* $(X, \mathcal{B}, \mu, T)$ *such that the set* $R^{\varepsilon}_{A,B}$ *of times of* ε*-fat intersection is* IP^*_+ *for every* $\varepsilon > 0$ *and any measurable sets* A, B, *but it is not always* IP^*.

Proof. In the argument below we will skip the tedious but relatively obvious specification of "epsilons" and "deltas".

The construction of $(X, \mathcal{B}, \mu, T)$ follows the standard scheme of "cutting and stacking with spacers" (see e.g. [P, Section 4.5]). We start with the interval $[0, 1]$ which we call tower Δ_1 of height $h_1 = 1$. Having constructed a tower Δ_{2m-1} (with an odd index) of height h_{2m-1} we choose an integer q_{2m-1} such that h_{2m-1}/q_{2m-1} is small, cut the tower into $2q_{2m-1}$ equal width columns and add single spacers above the left q_{2m-1} columns (see figure below).



Then we stack, creating the tower Δ_{2m} whose height equals

$$h_{2m} = 2q_{2m-1}h_{2m-1} + q_{2m-1}.$$

Next, we cut this tower into q_{2m} (which is larger than h_{2m}) columns and we stack them, this time without adding any spacers. This gives us a tower Δ_{2m+1} of height $h_{2m+1} = q_{2m}h_{2m}$. Continuing in this manner (note that we insert spacers only when constructing towers with even indices) we arrive at a space with a bounded measure and a measure preserving transformation. After normalizing we obtain a probability measure preserving rank-one system $(X, \mathcal{B}, \mu, T)$.

Let $L^2_0(\mu)$ denote the subspace of $L^2(\mu)$ consisting of functions with zero integral. Let $f \in L^2_0(\mu)$ be a complex-valued function of norm 1, which is constant on levels of the tower Δ_{2m_0-1} for some $m_0 \in \mathbb{N}$ and zero on the spacers added in the later steps. We are interested in the sequence $\hat{\mu}_f(n) =$

$\langle T^n f, f\rangle$. Fix some $n \in \mathbb{N}$. Choose $m > m_0$. Let x be a "typical" point in X. We are going to observe how the orbits of x and $T^n x$ pass through the tower Δ_{2m}. Let n_1 be the smallest $k \geq 0$ such that $T^{n+k}x$ belongs to the base of the tower, and n_2 be the smallest $k \geq n_1$ such that $T^k x$ belongs to the base of the tower. Define $n_0 = n_2 - n_1$. Clearly, independently of our choice of n, $0 \leq n_0 < h_{2m}$. Consider first the case when $n_0 < h_{2m}/2$. We continue our discussion with the help of the figure below. The top and bottom lines represent the orbits of $T^n x$ and x, respectively; any three dashes correspond to a passage through the tower Δ_{2m-1}; zeros correspond to the visits in the spacers; vertical lines separate the passages through the tower Δ_{2m-1} not separated by spacers; and the question marks indicate possible spacers added at later stages of our construction.

$$\cdots\ 0---0---0---0---|---|---|---|---?\overbrace{---0---0}^{\text{next passage}}\ \cdots$$

$$---0---0---0---0---0---|---|---|---|---?$$

| shift 1 | mixing | shift 2 | mixing |

We distinguish four consecutive intervals on the time axis appearing in the figure:

- The first one, denoted in the figure as "shift 1", roughly of length $h_{2m}/2 - n_0$, where spacers appear in both orbits, so that the pairs of "simultaneous" passages through the tower Δ_{2m-1} for x and $T^n x$ are all shifted in time by the same amount.
- The second one, denoted as "mixing", roughly of length n_0, with spacers in the orbit of x and without spacers in the orbit of $T^n x$, so that the shifts of "simultaneous" passages through Δ_{2m-1} change progressively by a unit.
- The third one, denoted as "shift 2", roughly of the same length as the first one, without spacers in both orbits, with all shifts the same but perhaps different from shift 1.
- The fourth one, which is again of the "mixing" type (in this interval $T^n x$ starts its next passage through Δ_{2m}); the possible spacers appearing at the question marks will not change the mixing type of this last interval.

If $h_{2m}/2 \leq n_0 < h_{2m}$ then one has to interpret the top line as the orbit of x and the bottom line as the orbit of $T^n x$.

Assuming that the mixing intervals are not too short they can be divided into some number of intervals of length $(h_{2m-1} + 1)h_{2m-1}$ which we call "cycles", and short "remainders" at both ends. In every cycle the orbit of one of the points $x, T^n x$ passes $h_{2m-1}+1$ times through the tower Δ_{2m-1} without "hitting" the spacers, while the other orbit passes through this tower h_{2m-1} times "hitting" the spacers (in the figure we have roughly one complete cycle;

the picture is too small to show real proportions). Notice that the average value of $f(T^{i+n}x)\overline{f(T^i x)}$ along a complete cycle equals 0 (since each fixed level of the tower Δ_{2m-1} for x "meets" all levels of the same tower for $T^n x$ the same number of times, f is constant on such levels with average value 0).

Every time the orbit of x passes through Δ_{2m} we observe a "pattern" of four intervals: shift 1/mixing/shift 2/mixing. Such patterns will be repeated throughout the orbit, each with its own parameter n_0. (This parameter will change from one pattern to another only when a higher order spacer appears at a place indicated by a question mark either in the orbit of x or in the orbit of $T^n x$ but not in both.)

These observations lead us to the following conclusions:

(a) If, in a significant fraction of all patterns, the mixing intervals are not too short (i.e., when the parameters n_0 are not too close to 0 or to h_{2m}), then the contribution of the complete cycles causes the value of $\langle T^n f, f\rangle$ to be of modulus essentially smaller than 1.

(b) If the mixing intervals "dominate" (i.e., in most patterns, n_0 is close to $h_{2m}/2$), then the value of $\langle T^n f, f\rangle$ is close to zero. (We assume that q_{2m-1} is so large in comparison with h_{2m-1} that in a pattern dominated by its two mixing intervals, the length of the mixing interval contains so many complete cycles that we can safely ignore the contribution of the "remainders").

Now suppose the value of $\langle T^n f, f\rangle$ is close to 1. By (a), this implies that n_0 is either small or close to h_{2m} in most of the patterns shift 1/mixing/shift 2/mixing. In this case we replace n by $n + h_{2m}/2$, and we will have the domination of mixing intervals, as described in case (b). Then, not only for f but also for any other normalized function $f' \in L^2_0(\mu)$ which is constant on the levels of Δ_{2m_0-1}, the following holds:

(c) $\langle T^{n+h_{2m}/2} f', f'\rangle$ is close to zero.

This is true for every $m > m_0$. For fixed m and any $k \in \mathbb{Z}$ with $|k|$ relatively small compared to h_{2m} (still very large if m is large), $n + k$ is not much different from n in the above arguments, hence the condition that $\langle T^n f, f\rangle$ is close to 1 implies that

(d) $\langle T^{n+k+h_{2m}/2} f', f'\rangle$ is close to zero.

In particular, this proves that T is weakly mixing, since, for f' approximating an eigenfunction, the values of $\langle T^n f', f'\rangle$ which are close to 1 appear with bounded gaps, while the parameter k in (d) can range through arbitrarily long intervals of integers.

Our construction produces a *rank-one* system and it is known that a rank-one transformation has simple spectrum (see [C-N] for more details on rank-one systems), so there exists a cyclic vector f_c in $L_0^2(\mu)$. Fix a pair of functions ϕ and ψ in $L_0^2(\mu)$. These functions can be approximated by finite combinations of the functions of the form $T^k f_c$.

Now consider a nonzero idempotent $p \in \beta\mathbb{Z}$. The weak limit p-$\lim(T^n g)$ exists for every $g \in L_0^2(\mu)$ and equals some $g' \in L_0^2(\mu)$ such that p-$\lim(T^n g') = g'$. Suppose $g' \neq 0$ for some g. Then we can normalize g' and denote it g''. We can now approximate g'' by $f \in L_0^2(\mu)$ constant on the levels of some tower Δ_{2m_0-1} and zero on spacers added in later stages of the construction. Clearly, p-$\lim\langle T^n f, f\rangle$ is a number close to 1. This implies that every IP-set M belonging to p contains a sequence M' along which n does not satisfy (a) (i.e., n_0 is small or close to h_{2m} in most patterns), and hence satisfies (c) and (d).

If m is large enough, the hypotheses (c) and (d) hold (with slightly worse error terms) also for $f' = f_c$. Since every term $\langle T^{n+h_{2m}/2}\phi, \psi\rangle$ splits into a finite combination of terms of the form $\langle T^{n+k+h_{2m}/2} f_c, f_c\rangle$ (with coefficients and k's not depending on n), for sufficiently large m_1 every such term with $n \in M'$ is close to zero. This proves that $R^{\varepsilon}_{\phi,\psi}$ intersects $M + h_{2m_1}/2$. The choice of m_1 is independent of the idempotent p satisfying $g' \neq 0$ for some g (it only depends on ϕ and ψ).

Now assume that p is such that $g' = 0$ for all $g \in L_0^2(\mu)$. In particular this is true for $g = T^{n+h_{2m_1}/2}\phi$ so p-$\lim\langle T^{n+h_{2m_1}/2}\phi, \psi\rangle = 0$, hence again $R^{\varepsilon}_{\phi,\psi}$ intersects $M + h_{2m_1}/2$. We have proved that $R^{\varepsilon}_{\phi,\psi}$ is IP^*_+. This immediately implies an analogous statement for sets A, B.

Finally, observe that the system is rigid along the sequence h_{2m} (because of the many consecutive passages through Δ_{2m} without spacers in the next tower). Thus it is not mildly mixing, hence at least one set $R^{\varepsilon}_{A,B}$ is not IP^*.

This concludes the proof. ■

Appendix

THEOREM A.1. *There exists an invertible topologically mixing symbolic dynamical system (X', T') with a fixpoint as a unique minimal set and having an invariant measure with full topological support.*

Sketch of proof. (The construction is an adaptation of one appearing in Theorem 1 of [D-Y].) Start with an aperiodic strictly ergodic (minimal with unique invariant measure μ) subshift (X, T) on two symbols $\{a, b\}$. From each point $x \in X$ we will create uncountably many points (sequences) x' over three symbols $\{a, b, c\}$, which will constitute our new desired subshift (X', T'). Namely, fix a sequence of closed and open sets (e.g., cylinders) $U_k \subset X$ shrinking to a point x^* so fast that $\sum_{k=1}^{\infty} \mu(U_k) < \infty$. For $x = \ldots, x_{-1}, x_0, x_1, \ldots$ let $(n_i)_{i\in\mathbb{Z}}$ be the times of the visits of x in U_1, and let k_i denote the *depth* of each visit, i.e., the maximal k such that $T^{n_i} x \in U_k$.

Let c be a new (third) symbol and let $c^k = [c, \dots, c]$ stand for the block of k symbols c. Now, from x we create the sequences x' by inserting into x, between x_{n_i-1} and x_{n_i}, either the block c^{k_i} or c^{k_i+1} (all such possible choices lead to uncountably many sequences x' made from one x). For example, one of the points x' will be

$$\dots, x_{n_{-1}-1}, c^{k_{-1}+1}, x_{n_{-1}}, \dots, x_{n_0-1}, c^{k_0+1}, x_{n_0}, \dots, x_{n_1-1}, c^{k_1}, x_{n_1}, \dots$$

The points in the orbit of x^* will produce exceptional sequences x'—either ending or beginning with infinitely many symbols c. Let X' be the closure of the set of all sequences x' so constructed from all $x \in X$. To verify the properties claimed in the formulation of the theorem notice the following:

(1) In each x' and for each k, the blocks c^k appear with bounded gaps. This implies that the fixpoint $c^\infty = \dots ccc \dots$ is the only minimal set in X'.

(2) We now prove that there exists a finite invariant measure whose support is X'. Viewing the symbols c as "spacers", the system (X', T') can be thought of as a "skyscraper": The base is the set $\{x' : x'_0 \neq c\}$, the levels (for $k > 0$) are $\{x' : x'_{-k} \neq c,\ [x'_{-k+1}, \dots, x'_0] = c^k\}$. We do not include in this skyscraper the points x' obtained from points x belonging to the orbit of x^*, but as we will explain, such points form a set of measure zero. The first return time map induced on the base consists in shifting each x' by the distance to the nearest symbol different from c, so that (at coordinate zero) it merely reads the consecutive entries of the original sequence $x \in X$. Note that each point x' is determined by two sequences: x and a $\{0,1\}$-valued sequence $y = (y_i)$ governing the (binary) decisions made while inserting either c^{k_i} or c^{k_i+1}. All (uncountably many) different points obtained from one x remain different in the system induced on the base of the skyscraper, hence this induced system is not isomorphic to (X, T). It is however an extension of (X, T) and it is not hard to see that this extension has the form of a skew product T_S of (X, T) (minus the orbit of x^*) with the full shift (Y, S) on two symbols $\{0, 1\}$ defined by

$$T_S(x, y) = (Tx, S^{\mathbf{1}_{U_1}(x)} y),$$

i.e., we apply the shift on the second coordinate if $x \in U_1$, otherwise the entry on the second coordinate remains unchanged. Clearly, the product measure $\mu \times \lambda$ is T_S-invariant (where λ denotes the homogeneous Bernoulli measure on the two-shift Y), and has full topological support in the product space. Also, we note that the exceptional points created from the orbit of x^* form a set of measure zero for the product measure (this set is the lift of a countable set and μ is nonatomic). Observe that the first level of the skyscraper extends above a dense subset of $U_1 \times Y$ and for $k \geq 2$ the kth level extends above a dense subset of $U_{k-1} \times Y$. Since $\sum_k \mu(U_k) < \infty$, the product measure $\mu \times \lambda$ on the base "lifts" to a finite invariant measure on the whole skyscraper with full topological support in the skyscraper. By an obvious approximation argument, this measure has full support also in X'. The desired probability measure is obtained by normalization.

(3) We will show that under additional assumptions, (X', T') can be made topologically mixing. Let us impose a stronger requirement on the speed of decay of the sets U_k: the smallest gap between visits in U_k ($k \geq 2$) is larger than $2k$ times the largest gap between visits in U_1. This implies that between any two visits in U_k each point visits U_1 at least $2k$ times (of course this can be done by choosing U_k to be contained in balls around x^* of rapidly decreasing radii).

Let $x' \in X'$ be created from a point $x \in X$ not belonging to the orbit of x^*, and let B' be the finite block $x'[-m', m']$. (Note that every block appearing in X' can be obtained this way.) Let $B = x[-m, m]$ be a block (possibly much longer than B') whose appearance at any element $z \in X$ (with coordinate zero at the center) ensures that for a sufficiently long time (forward and backward) the orbit of z visits the sets U_k at exactly the same times as x does, so that among the points z' created from z there exists one with $z'[-m', m'] = B'$. By minimality, B appears at a positive coordinate in x^*, say $B = x^*[r - m, r + m]$. Since, for each k, x^* belongs to U_k, its return to U_k is preceded by at least $2k$ visits in U_1. Begin creating the sequence $x^{*\prime}$ from x^* by insertions. Its negative part is filled with $\ldots, c, c, c$, and positive with the positive part of x^* with appropriate insertions. The insertions into $x^*[r-m, r+m]$ may be arranged so that $x^{*\prime}[r'-m', r' + m'] = B'$. In order to prove the mixing property we need to show that the construction of $x^{*\prime}$ can then be continued to the right in so many ways that any block C' possible in X' will appear in these continuations at all distances larger than some constant. Fix one such C' and let C be a block appearing in X making the creation of C' possible (just like B was chosen for B'). Notice that C appears in x^* with bounded gaps. Let $x_1^{*\prime}, x_2^{*\prime}, \ldots$ denote the sequence of continuations of $x^{*\prime}$ such that in $x_n^{*\prime}$ all insertions to the right of $r + m$ are of the smaller type (i.e., c^{k_i}) except inside one selected (nth after position $r + m$) occurrence of C, where the insertions are adjusted to create the block C'. Let d_n be the distance between the block B' (made from the copy of B centered at position r) and C' (the one made from the nth copy of C) in $x_n^{*\prime}$. We can now enlarge each distance d_n by one, two, or more units, replacing one, two, or more insertions between position $r + m$ and the nth copy of C considered, by insertions of the larger type c^{k_i+1}. The last thing to show is that for n large enough there are at least $d_{n+1} - d_n$ such "regulating insertions" available, so that enlarging the distance d_n we can reach d_{n+1}. This will prove that it is possible to obtain C' at any sufficiently large distance following B'. This is the essence of topological mixing.

Let g denote the maximal gap between the occurrences of C in x^*. Let k_0 be such that the distance between two visits of the orbit of x^* in U_{k_0+1} exceeds g, so that at most one visit in U_{k_0+1} is possible between two blocks C. If n is such that between the nth and $(n+1)$st copy of C (counting from the right end $r+m$ of B) the orbit of x^* visits U_{k_0+1} with some depth $k > gk_0 + r + m + g$, then the distance $d_{n+1} - d_n$ does not exceed $gk_0 + k$ (there are at most g insertions of size k_0, and k is the size of the unique larger insertion). In that case this unique visit to U_k is preceded by at least $2k > k + gk_0 + r + m + g$ visits in U_1, of which at least $k + gk_0 > d_{n+1} - d_n$ fall between B and the nth copy of C, allowing equally many "regulating insertions", as required. If n is such that between the nth and $(n+1)$st copies of C there is no visit of depth larger than $gk_0 + r + m + g$ then $d_{n+1} - d_n$ is bounded (for instance, by $g(gk_0 + r + m + g)$). So, in either case, if n is large enough, the nth copy of C is preceded by sufficiently many visits in U_1 allowing sufficiently many "regulating insertions". ∎

Final remarks. We would like to indicate one natural way of extending statements (4) and (5) of Theorem 1.6. Let $k \in \mathbb{N}$. For $i = 1, \ldots, k$ let $P_i(n)$ be nonconstant polynomials satisfying $P_i(\mathbb{Z}) \subset \mathbb{Z}$. Given a measure preserving system $(X, \mathcal{B}, \mu, T)$, sets $A_i \in \mathcal{B}$ ($i \in [0,k]$) and $\varepsilon > 0$, define

$$R^{\varepsilon}_{A_0,A_1,\ldots,A_k} = \{n \in \mathbb{Z} : \mu(A_0 \cap T^{P_1(n)} A_1 \cap \cdots \cap T^{P_k(n)} A_k) > \mu(A_0)\mu(A_1)\cdots\mu(A_k) - \varepsilon\},$$

$$Q^{\varepsilon}_{A_0,A_1,\ldots,A_k} = \{n \in \mathbb{Z} : |\mu(A_0 \cap T^{P_1(n)} A_1 \cap \cdots \cap T^{P_k(n)} A_k) - \mu(A_0)\mu(A_1)\cdots\mu(A_k)| < \varepsilon\}.$$

Denote by $\mathcal{R}_k(X, \mathcal{B}, \mu, T)$ and $\mathcal{Q}_k(X, \mathcal{B}, \mu, T)$ the family of all sets of the form $R^{\varepsilon}_{A_0,A_1,\ldots,A_k}$ and the family of all sets of the form $Q^{\varepsilon}_{A_0,A_1,\ldots,A_k}$, respectively (note that both $\mathcal{R}_k(X, \mathcal{B}, \mu, T)$ and $\mathcal{Q}_k(X, \mathcal{B}, \mu, T)$ depend on the choice of the polynomials $P_i(n)$). Then one can show that:

(i) $(X, \mathcal{B}, \mu, T)$ is weakly mixing iff $\mathcal{R}_k(X, \mathcal{B}, \mu, T) \in \mathcal{D}^*$ for any $k \geq 1$ and any fixed system of integer-valued polynomials $P_1(n), \ldots, P_k(n)$.

(ii) $(X, \mathcal{B}, \mu, T)$ is mildly mixing iff $\mathcal{R}_k(X, \mathcal{B}, \mu, T) \in \mathcal{IP}^*$ for any $k \geq 1$ and any fixed system of integer-valued polynomials $P_1(n), \ldots, P_k(n)$.

Also, it is easy to see that in (i) the family $\mathcal{D}^*$ can be equivalently replaced by $\mathcal{D}^*_\bullet$, $\mathcal{C}^*$ or $\mathcal{C}^*_\bullet$, while in (ii) the family $\mathcal{IP}^*$ can be equivalently replaced by $\mathcal{IP}^*_\bullet$. Additionally, all resulting statements hold if $\mathcal{Q}_k(X, \mathcal{B}, \mu, T)$ is used in place of $\mathcal{R}_k(X, \mathcal{B}, \mu, T)$ (cf. Remark 1).

The main ingredient in proving the statements (i) and (ii) is provided by multiple recurrence theorems along ultrafilters (see Theorems 4.8 and 5.1(v) in [B2]).

With regard to mixing, it is proved in [K-Y] that

(iii) $(X,\mathcal{B},\mu,T)$ is mixing iff for any three sets $A,B,C\in\mathcal{B}$ all sets of the form

$$Q^{\varepsilon}_{A,B,C}=\{n\in\mathbb{Z}:|\mu(A\cap T^{\alpha_1 n}B\cap T^{\alpha_2 n}C)-\mu(A)\mu(B)\mu(C)|<\varepsilon\}$$

belong to $\varDelta^*$.

Obviously, an analogous statement involving the sets

$$R^{\varepsilon}_{A,B,C}=\{n\in\mathbb{Z}:\mu(A\cap T^{\alpha_1 n}B\cap T^{\alpha_2 n}C)>\mu(A)\mu(B)\mu(C)-\varepsilon\}$$

is also true. As before, the family $\varDelta^*$ can be replaced by $\varDelta^*_\bullet$. No extension of this result to more general sets of polynomials is known.

REFERENCES

[B1] V. Bergelson, *Ergodic Ramsey theory—an update*, in: Ergodic Theory of $\mathbb{Z}^d$ Actions, London Math. Soc. Lecture Note Ser. 228, Cambridge Univ. Press, Cambridge, 1996, 1–61.

[B2] — *Minimal idempotents and ergodic Ramsey theory*, in: Topics in Dynamics and Ergodic Theory, London Math. Soc. Lecture Note Ser. 310, Cambridge Univ. Press, Cambridge, 2003, 8–39.

[B-H1] V. Bergelson and N. Hindman, *Nonmetrizable topological dynamics and Ramsey theory*, Trans. Amer. Math. Soc. 320 (1990), 293–320.

[B-H2] —, —, *Partition regular structures contained in large sets are abundant*, J. Combin. Theory Ser. A 93 (2001), 18–36.

[B-M] V. Bergelson and R. McCutcheon, *Central sets and a non-commutative Roth theorem*, Amer. J. Math., to appear.

[B-H-S] F. Blanchard, W. Huang and Ľ. Snoha, *Topological size of scrambled sets*, this volume, 293–361.

[C-F-S] I. P. Cornfeld, S. V. Fomin and Y. G. Sinai, *Ergodic Theory*, transl. from the Russian by A. B. Sossinskii, Grundlehren Math. Wiss. 245, Springer, New York, 1982.

[C-N] J. R. Choksi and M. G. Nadkarni, *The maximal spectral type of a rank one transformation*, Canad. Math. Bull. 37 (1994), 29–36.

[D] T. Downarowicz, *Strictly nonpointwise Markov operators and weak mixing*, in: Ergodic Theory and its Connections with Harmonic Analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser. 205, Cambridge Univ. Press, 1995, 259–272.

[D-Y] T. Downarowicz and X. D. Ye, *When every point is either transitive or periodic*, Colloq. Math. 93 (2002), 137–150.

[E] R. Ellis, *Lectures on Topological Dynamics*, W. A. Benjamin, New York, 1969.

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

[G] E. Glasner, *Divisible properties and the Stone–Čech compactification*, Canad. J. Math. 32 (1980), 993–1007.

[H-S] N. Hindman and D. Strauss, *Algebra in the Stone–Čech Compactification. Theory and Applications*, de Gruyter Expo. Math. 27, de Gruyter, Berlin, 1998.

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

[Kr] U. Krengel, *Ergodic Theorems*, de Gruyter Stud. Math. 6, de Gruyter, Berlin, 1985.

[K-Y] R. Kuang and X. D. Ye, *Mixing via families for measure preserving transformations*, this volume, 151–165.

[P] K. Petersen, *Ergodic Theory*, Cambridge Univ. Press, Cambridge, 1983.

Department of Mathematics
The Ohio State University
100 Math Tower, 231 West 18th Avenue
Columbus, OH 43210-1174, U.S.A.
E-mail: vitaly@math.ohio-state.edu

Institute of Mathematics
Technical University
Wybrzeże Wyspiańskiego 27
50-370 Wrocław, Poland
E-mail: downar@im.pwr.wroc.pl

Received 17 January 2007 (4859)