Piecewise-Bohr Sets of Integers and Combinatorial Number Theory

Vitaly Bergelson[1*], Hillel Furstenberg[2], and Benjamin Weiss[3]

[1] Department of Mathematics, Ohio State University, Columbus, OH 43210, USA
vitaly@math.ohio-state.edu

[2] Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem, Israel
harry@math.huji.ac.il

[3] Institute of Mathematics, Hebrew University of Jerusalem, Jerusalem, Israel
weiss@math.huji.ac.il

Summary. We use ergodic-theoretical tools to study various notions of "large" sets of integers which naturally arise in theory of almost periodic functions, combinatorial number theory, and dynamics. Call a subset of **N** a Bohr set if it corresponds to an open subset in the Bohr compactification, and a piecewise Bohr set (PWB) if it contains arbitrarily large intervals of a fixed Bohr set. For example, we link the notion of PWB-sets to sets of the form A+B, where A and B are sets of integers having positive upper Banach density and obtain the following sharpening of a recent result of Renling Jin.

Theorem. If A and B are sets of integers having positive upper Banach density, the sum set A+B is PWB-set.

AMS Subject Classification. 37A45.
Keywords. Bohr sets, Banach density, measure preserving system, Kronecker factor.

1 Introduction to Some Large Sets of Integers

In combinatorial number theory, as well as in dynamics, various notions of "large" sets arise. Some familiar notions are those of sets of positive (upper) density, syndetic sets, thick sets (also called "replete"), return-time sets (in dynamics), sets of recurrence (also known as Poincaré sets), (finite or infinite) difference sets, and Bohr sets. We will here introduce the notion of "piecewise-Bohr" sets (or PWB-sets), as well as "piecewise-Bohr$_0$" sets (or PWB$_0$-sets), and we'll show how they arise in some combinatorial number-theoretic questions.

We begin with some basic definitions and elementary considerations. We'll say that a subset $A \subset \mathbb{Z}$ has *positive upper* (Banach) *density*, $d^*(A) > 0$,

* V. Bergelson was partially supported by NSF, grant DMS-0345350.

if for some $\delta > 0$, there exist arbitrarily large intervals of integers $J = \{a, a+1, \dots, a+l-1\}$ with $\frac{|J \cap A|}{|J|} \geq \delta$. (Here $|S|$ is the cardinality of the set S; $d^*(A) = \text{l.u.b.}\{\delta \text{ as above}\}$.) Syndetic sets are special cases of sets with positive upper density. Namely, A is *syndetic* if for some l, every interval J of integers with $|J| \geq l$ intersects A. Clearly $d^*(A) \geq 1/l$ in this case. We'll say a set A is *thick* if it contains arbitrarily long intervals; thus A is syndetic $\Leftrightarrow \mathbb{Z} \backslash A$ is not thick $\Leftrightarrow A \cap B \neq \emptyset$ for any thick set B. For any distinct r integers $\{a_1, a_2, \dots, a_r\}$ the set $\{a_j - a_i | 1 \leq i < j \leq r\}$ is called an r-*difference* set or a Δ_r-set. Every thick set contains some r-difference set for every r. This is obvious for $r = 2$, and inductively, if A is thick and if A contains the $(r-1)$-difference set formed from $\{a_1, \dots, a_{r-1}\}$, by choosing a_r in the middle of a large enough interval in A, we can complete this to an r-difference set. It follows that for any r, a set that meets every r-difference set is syndetic. An example of this is the set of (non-zero) differences $A - A = \{x - y : x, y \in A, x \neq y\}$ when A has positive upper density. For if $d^*(A) > 1/r$ and if the numbers $a_1, a_2, \dots, a_r$ are distinct, the sets $A + a_1, A + a_2, \dots, A + a_r$ cannot be disjoint; so, for some $1 \leq i < j \leq r$, $a_j - a_i \in A - A$. One conclusion which is behind much of our subsequent discussion is that if A has positive upper density, then $A - A$ is syndetic. We shall see in §3 that $d^*(A) > 0$ implies that $A - A$ is a *piecewise-Bohr* set.

Definition 1.1. *$S \subset \mathbb{Z}$ is a* Bohr *set if there exists a trigonometric polynomial $\psi(t) = \sum_{k=1}^{m} c_k e^{i\lambda_k t}$, with the λ_k real numbers, such that the set*

$$S' = \{n \in \mathbb{Z} : \operatorname{Re} \psi(n) > 0\}$$

is non-empty and $S \supset S'$. When $\psi(0) > 0$ we say S is a Bohr$_0$ set. (Compare with [Bilu97]).

The fact that a Bohr set is syndetic is a consequence of the almost periodicity of trigonometric polynomials. It is also a consequence of the "uniform recurrence" of the Kronecker dynamical system on the m-torus

$$(\theta_1, \theta_2, \dots \theta_m) \longrightarrow (\theta_1 + \lambda_1, \theta_2 + \lambda_2, \dots, \theta_m + \lambda_m).$$

Indeed, it is not hard to see that a set $S \subset \mathbb{Z}$ is Bohr if and only if there exist $m \in \mathbb{N}$, $\alpha \in \mathbb{T}^m$ and an open set $U \subset \mathbb{T}^m$ such that $S \supset \left\{n \in \mathbb{Z} : n\alpha \in U\right\}$.

Alternatively we can define Bohr sets and Bohr$_0$ sets in terms of the topology induced on the integers $\mathbb{Z}$ by imbedding $\mathbb{Z}$ in its Bohr compactification. Namely, a set in $\mathbb{Z}$ is Bohr if it contains an open set in the induced topology, and it is Bohr$_0$ if it contains a neighborhood of 0 in this topology.

We can apply the foregoing observations regarding $A - A$ to dynamical systems. We shall be concerned with *measure preserving systems* $(X, \mathcal{B}, \mu, T)$, where $(X, \mathcal{B}, \mu)$ is a probability space, $T: X \to X$ a measurable measure preserving transformation. We assume (for simplicity) that the system is *ergodic*

($T^{-1}A = A$ for $A \in \mathcal{B} \Rightarrow \mu(A)\mu(X\backslash A) = 0$). The ergodic theorem then ensures that for $A \in \mathcal{B}$ with $\mu(A) > 0$, the orbit $\{T^n x\}_{n\in\mathbb{Z}}$ of almost every x visits A along a set of times $V(x, A) = \{n : T^n x \in A\}$ of positive density. If we set $R_1(A) = \{n : A \cap T^{-n}A \neq \emptyset\}$ (the return time set of A), then for any x, $R_1(A) \supset V(x, A) - V(x, A)$. Hence $R_1(A)$ is syndetic. We can define a smaller set $R(A) = \{n : \mu(A \cap T^{-n}A) > 0\} = R(A')$ where $A' = A \setminus \bigcup\{(A \cap T^{-n}A) : \mu(A \cap T^{-n}A) = 0\}$, and it follows that $R(A)$ is also syndetic. This can be seen directly as well (and for arbitrary measure preserving systems), but the present argument illustrates the connection of dynamics to combinatorial properties of sets. We shall call sets containing sets of the form $R(A)$, where $\mu(A) > 0, RT$-sets (for return time). A set meeting every RT-set is called a Poincaré set since Poincaré's recurrence theorem gives content to the property by implying that $R(A)$ is never empty for $\mu(A) > 0$ even if T is not ergodic. These are also known in the literature as *intersective* sets. (See [Ruz82]). Much is known about these (see [Fur81], [B-M86], [BH96], [BFM96]). In particular $\{n^r; n = 1, 2, \ldots\}$ is a Poincaré set for each $r = 1, 2, 3, \ldots$

For a family $\mathcal{F}$ of subsets of $\mathbb{Z}$ it is customary to denote by $\mathcal{F}^*$ the dual family: $\mathcal{F}^* = \{S \subset \mathbb{Z} : \forall S' \in \mathcal{F}, S \cap S' \neq \emptyset\}$. Note that $\{\text{syndetic}\} = \{\text{thick}\}^*$, $\{\text{thick}\} = \{\text{syndetic}\}^*$ and $\{RT\} = \{\text{Poincaré}\}^*, \{\text{Poincaré}\} = \{\text{RT}\}^*$.

We have seen above that a Δ_r^*-set is necessarily syndetic. One of our objectives is to sharpen this statement.

We will need the notion of a "PW-$\mathcal{F}$" set for a family $\mathcal{F}$ of subsets of $\mathbb{Z}$. "PW" stands for "piecewise" and if $S \in \mathcal{F}$ and Q is a thick set then we shall say $S \cap Q$ is PW-$\mathcal{F}$ (or $S \cap Q \in$ PW-$\mathcal{F}$). Clearly this notion is useful only for families of syndetic sets. "PW-syndetic" is itself a useful notion. Van der Waerden's theorem [GRS80] implies that syndetic sets contain arbitrarily long arithmetic progressions. In fact this is true for PW-syndetic sets. Unlike the family of syndetic sets, the latter have the "divisibility" property: if S is PW-syndetic and $S = S_1 \cup S_2 \cup \cdots \cup S_k$ is a finite partition, then some S_i is PW-syndetic, see [Bro71]. A recent result of Renling Jin [Jin02] is the following:

Theorem 1.2. *If $A, B \subset \mathbb{Z}$ and $d^*(A) > 0,\ d^*(B) > 0$, then $A + B$ is PW-syndetic.*

We will sharpen this to

Theorem I. *If $A, B \subset \mathbb{Z}$ and $d^*(A) > 0,\ d^*(B) > 0$, then $A + B$ is a PW-Bohr set (PWB-set).*

In particular $d^*(A) > 0$ will imply that $A - A$ is a PW-Bohr set. More precisely it is a PW-Bohr$_0$ (PWB$_0$)-set. This will also follow from our earlier observation that it is a Δ_r^*-set for sufficiently large r, and from

Theorem II. *For each $r \geq 2$, a Δ_r^*-set is PW-Bohr$_0$.*

It is not hard to see that the prefix "PW" is indispensable in these theorems. For example $A = \bigcup[10^n, 10^n + n]$ has $d^*(A) = 1$ but $A + A$ is not syndetic. Also since $x^3 + y^3 = z^3$ has no solution in non-zero integers, it follows that the set of non-cubes $S = \mathbb{Z}\backslash\{n^3; n = \pm1, \pm2, \pm3, \ldots\}$ is a Δ_3^* set. But by Weyl's equidistribution theorem S is not a Bohr_0-set. (See Theorem 4.1 below for a stronger form of this observation.)

From Theorem I we shall deduce the following result which should be compared with a theorem due to Ruzsa ([Ruz82], Theorem 3) which states that if $d^*(A) > 0$, then $A + A - A$ is a Bohr set. (Both Ruzsa's theorem and our result can be viewed as improvements on a theorem of Bogoliouboff ([Bog39], [Føl54]) which implies that if $d^*(A) > 0$, then $A - A + A - A$ is a Bohr set.)

Corollary 1.3. *If A, B, C are three subsets of $\mathbb{Z}$ with positive upper density and one of them is syndetic, then $A + B + C$ is a Bohr set.*

2 Measure Preserving Systems, Time Series, and Generic Schemes

In this section we introduce a basic tool which will be needed repeatedly: the correspondence between data given on large intervals of time ("time series") and measure preserving dynamical systems. This tool has been used previously under the name "correspondence principle" (see e.g., [Ber96]) and here we present it in a more general form. We repeat the definition of a measure preserving system which was given informally in §1.

Definition 2.1. *A* measure preserving system *is a quadruple $(X, \mathcal{B}, \mu, T)$ where $(X, \mathcal{B}, \mu)$ is a probability space where we assume $\mathcal{B}$ is countably generated, and T is a measurable, invertible, and measure preserving map, $T: X \to X$. The system is* ergodic *if every measurable T-invariant set has measure 0 or 1.*

For a measurable function $f: X \to \mathbb{C}$ we denote by Tf the function $Tf(x) = f(Tx)$. We take note of the ergodic theorem (see, for example, [Kre85]):

Theorem 2.2. *If $(X, \mathcal{B}, \mu, T)$ is a measure preserving system and $f \in L^1(X, \mathcal{B}, \mu)$, then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n f = \overline{f}$$

exists almost everywhere. If $f \in L^p(X, \mathcal{B}, \mu)$, $1 \le p < \infty$, the convergence is in L^p as well. If the system is ergodic then $\overline{f} = \int f d\mu$ a.e., so that the average of the sequence $\{f(T^n x)\}$ equals a.e. the average of f over X.

Sequences of the form $\{f(T^n x)\}_{a\le n\le b}$ are referred to as "time series". In a certain sense the ergodic theorem enables one to reconstruct a dynamical system from "time series data". We shall make this precise in the notion of "generic schemes" which we proceed to define. In the next definitions the indices l and r range over the natural numbers.

Definition 2.3. *An* array *is a sequence* $\{J_l\}$ *of intervals of integers,* $J_l = \{a_l, a_l+1, \dots, b_l\}$ *for which* $|J_l| = b_l - a_l + 1 \to \infty$ *as* $l \to \infty$.

Definition 2.4. *A* scheme $(\{J_l\}, \{\xi_r^l\})$ *is an array* $\{J_l\}$ *together with a doubly indexed set of complex-valued functions* $\{\xi_r^l\}$ *where, for each* r, $\xi_r^l(n)$ *is defined for* $n \in J_l$ *and, for each* r, *the functions* $\{\xi_r^l; l = 1, 2, \dots\}$ *are uniformly bounded. For* $n \notin J_l$ *we take* $\xi_r^l(n) = 0$. *The* $\{\xi_r^l\}$ *will be referred to as* time series. *They are defined on all of* $\mathbb{Z}$ *but only the values on* J_l *have significance. The following notion relates closely to that of a "stationary stochastic process".*

Definition 2.5. *A* process $(X, \mathcal{B}, \mu, T, \Phi)$ *consists of a measure preserving system* $(X, \mathcal{B}, \mu, T)$ *together with an at most countable ordered set* $\Phi = \{\varphi_1, \varphi_2, \dots\}$ *of* L^∞*-functions on* X *such that* $\mathcal{B}$ *is the* σ*-algebra generated by the functions of* Φ *and their translates under* T. *(When the* φ_i *are complex valued we assume* Φ *closed under conjugation). A process is ergodic if the underlying measure preserving system is ergodic.*

Finally we have

Definition 2.6. *A scheme* $(\{J_l\}, \{\xi_r^l\})$ *is* generic *for a process* $(X, \mathcal{B}, \mu, T, \Phi)$ *if for every* m *and for every choice of* $i_1, i_2, \dots, i_m$ *and* $j_1, j_2, \dots, j_m$ *(the indices here need not be distinct):*

$$\lim_{l\to\infty} \frac{1}{|J_l|} \sum_{n\in J_l} \xi_{i_1}^l(n+j_1)\xi_{i_2}^l(n+j_2)\cdots\xi_{i_m}^l(n+j_m) \tag{1}$$
$$= \int_X T^{j_1}\varphi_{i_1} T^{j_2}\varphi_{i_2} \cdots T^{j_m}\varphi_{i_m}\, d\mu$$

It will be convenient to introduce the countable family Φ^* consisting of the products appearing in (1):

$$\Phi^* = \{\psi = T^{j_1}\varphi_{i_1} T^{j_2}\varphi_{i_2} \cdots T^{j_m}\varphi_{i_m}\}$$

The corresponding time series have the form

$$\zeta^l(n) = \xi_{i_1}^l(n+j_1)\xi_{i_2}^l(n+j_2)\cdots\xi_{i_m}^l(n+j_m),$$

and when (1) holds, we say that $\{\zeta^l\}$ *represents* ψ.

It will be convenient in the sequel to regard Φ^* as the increasing union of finite sets, $\Phi^* = \bigcup_{h=1}^{\infty} \Phi_h^*$. The subscript h has no significance other than as an index with $\Phi_1^* \subset \Phi_2^* \subset \cdots \subset \Phi_h^* \subset \cdots$.

We note that the ergodic theorem implies that if $(X, \mathcal{B}, \mu, T)$ is ergodic, then for almost every $x_0 \in X$, the scheme $(\{J_l\}, \{\xi_r^l\}$ is generic for the process $(X, \mathcal{B}, \mu, T, \Phi)$ with $J_l = [1, l]$ and $\xi_r^l(n) = \varphi_r(T^n x_0)$ independently of l.

The main result of this section goes in the opposite direction, and will attach to an arbitrary scheme an ergodic process. First we need the notions of *subarrays* and *subschemes.*

Definition 2.7. *An array $\{H_l\}$ is a* subarray *of $\{J_l\}$ if $l \to L_l$ is a monotone increasing function from $\mathbb{N}$ to $\mathbb{N}$ and H_l is a subinterval of J_{L_l}.*

Definition 2.8. *A scheme $(\{H_l\}, \{\eta_r^l\})$ is a* subscheme *of $(\{J_l\}, \{\xi_r^l\})$ if $\{H_l\}$ is a subarray of $\{J_l\} : H_l \subset J_{L_l}$, and η_r^l is the restriction of $\xi_r^{L_l}$ to H_l.*

Our main result in this section is

Theorem 2.9. *For any scheme $(\{J_l\}, \{\xi_r^l\})$ there exists a subscheme and an ergodic process for which the subscheme is generic.*

Proof. First we will pass to a subscheme which is generic for a process $(X, \mathcal{B}, \mu, T, \Phi)$ which is not necessarily ergodic. For each r, let $\Lambda_r \subset \mathbb{C}$ be a compact set with $\xi_r^l(n) \in \Lambda_r$ for all l and n. Let $\tilde{\Lambda} = \prod \Lambda_r$ and let $X = \tilde{\Lambda}^{\mathbb{Z}}$. We denote by ξ_r^l the point in $\Lambda_r^{\mathbb{Z}}$ with $\xi_r^l = (\ldots, \xi_r^l(-1), \xi_r^l(0), \xi_r^l(1), \ldots)$ and form $\tilde{\xi}^l = (\xi_1^l, \xi_2^l, \ldots) \in \tilde{\Lambda}^{\mathbb{Z}} = X$. X is a compact metrizable space and we form the measures

$$\nu_l = \frac{1}{|J_l|} \sum_{n \in J_l} \delta_{T^n \tilde{\xi}^l} \tag{2}$$

where $T : X \to X$ denotes the shift map $T\omega(n) = \omega(n+1)$. Since $|J_l| \to \infty$, any weak limit of a subsequence of ν_l is T-invariant, and we let ν be some such limit: $\nu = \lim \nu_{L_l}$. It is not hard to see that $(\{J_{L_l}\}, \{\xi_r^{L_l}\})$ is generic for the process $(X, \mathcal{B}, \nu, T, \Phi)$ where $\mathcal{B}$ is the Borel σ-algebra of sets in X and $\Phi = \{\varphi_1, \varphi_2, \ldots\}$ with φ_r the functions on $\tilde{\Lambda}^{\mathbb{Z}}$ given by $\varphi_r(\omega) = \omega(0)(r)$. By ergodic decomposition there will be an ergodic measure μ whose support is a subset of the support of ν. Any point in the support of μ is a limit of points of the form $T^n \tilde{\xi}^l$ with $n \in J_l$ and $l \to \infty$, by (2). Since μ is ergodic, almost every point ω in its support is *generic* for μ, in the sense that averages of a given bounded measurable function along the orbit of ω tend to the integral of the function. In particular for functions in Φ^* we have:

$$\frac{1}{N} \sum_{n=k}^{k+N-1} T^{j_i} \varphi_{i_1} T^{j_2} \varphi_{i_2} \cdots T^{j_m} \varphi_{i_m} (T^n \omega) \longrightarrow \int T^{j_1} \varphi_{i_1} T^{j_2} \varphi_{i_2} \cdots T^{j_m} \varphi_{i_m} d\mu \tag{3}$$

uniformly for $|k| \leq N$.

We can find N sufficiently large that the difference of the two sides in (3) is $< \varepsilon$ for all $T^{j_1} \varphi_{i_1} \cdots T^{j_m} \varphi_{i_m} \in \Phi_h^*$. We then choose $T^n \tilde{\xi}^l$ close enough to ω, $n \in J_l$, so that the difference of the two sides of (3) remains $< \varepsilon$ with ω replaced by $T^n \tilde{\xi}^l$. Since $n \in J_l$, assuming l sufficiently large, we will have

$H_l = [n+k, n+k+N-1] \subset J_l$ for some k with $|k| \leq N$. We now let $\varepsilon \to 0$, $h \nearrow \infty$, and choose an appropriate subsequence of l; rescrambling the information in (3) we find a subscheme $(\{H_l\}, \{\xi_r^l\})$ which is generic for $(X, \mathcal{B}, \mu, T, \Phi)$. □

Scholium to Theorem 2.9. If for some r,

$$\limsup_{l\to\infty} \frac{1}{|J_l|} \left| \sum_{n\in J_l} \xi_r^l(n) \right| > 0,$$

we can add the condition that the corresponding φ_r does not vanish a.e. This follows from the fact that the measure ν satisfies $\int \varphi_r d\nu \neq 0$ and so ν must have an ergodic component with $\int \varphi_r d\mu \neq 0$.

We remark that in the case of ergodic processes, given a generic scheme, "many" subschemes will again be generic. This is made precise in the following: For any process $(X, \mathcal{B}, \mu, T, \Phi)$, Φ^* is countable and we fix an increasing family of finite sets $\Phi_h^* \subset \Phi^*$ increasing to Φ^*. Given a scheme $(\{J_l\}, \{\xi_r^l\})$ and fixing l, and letting $\varepsilon > 0$, we shall say that an interval $H \subset J_l$ is *ε-h-generic* for the process $(X, \mathcal{B}, \mu, T, \Phi)$ if (1) holds approximately; i.e, if for every $\psi \in \Phi_h^*$ and corresponding time series $\zeta^l(n)$.

$$\left| \frac{1}{|H|} \sum_{n\in H} \zeta^l(n) - \int \psi d\mu \right| < \varepsilon. \tag{4}$$

Assume now a process $(X, \mathcal{B}, \mu, T, \Phi)$ given with $\Phi^* = \bigcup \Phi_h^*$ as above, and let $(\{T_l\}, \{\xi_r^l\})$ be a generic scheme for the process.

Proposition 2.10. *If $(X, \mathcal{B}, \mu, T, \Phi)$ is an ergodic process, then for any $\varepsilon > 0$ and $h \in \mathbb{N}$ there exists $p_0 \in \mathbb{N}$ so that for any $p \geq p_0$ there exists a positive number $l_0(\varepsilon, h, p)$ so that for $l > l_0(\varepsilon, h, p)$, at least $(1-\varepsilon)(|J_l| - p + 1)$ of the $(|J_l| - p + 1)$ intervals of length p in J_l are ε-h-generic for the process.*

Letting p and l grow we see, according to the proposition, that the intervals J_l can be replaced by many choices of subintervals, and the scheme will remain generic. It is easy to see that this is not true for non-ergodic processes (where time series have different statistical behavior along different intervals of time).

Proof of Proposition 2.10. It suffices to treat a single function and the corresponding time series. For if for each of the $|\Phi_h^*|$ functions in Φ_h^* we have $(1-\varepsilon_1)(|J_l| - p + 1)$ "ε_1-generic" intervals with $\varepsilon_1 |\Phi_h^*| < \varepsilon$, the number of intervals common to all of these will not be less than $(1-\varepsilon)(|J_l| - p + 1)$, and these intervals are ε_1-h-generic, and so also ε-h-generic. So let $\psi \in \Phi^*$. Ergodicity assures that for p large, $\frac{1}{p} \sum_{q=0}^{p-1} T^q \psi$ is L^2-close to $\int \psi d\mu$, and so

$$\int \left(\frac{1}{p} \sum_{q=0}^{p-1} T^q \psi \right)^2 d\mu - \left(\int \psi d\mu \right)^2$$

is small. Fix p and set $\eta(n) = \frac{1}{p}\sum_{q=0}^{p-1}\zeta(n+q)$. η and ζ have the same long-term averages,

$$\frac{1}{|J_l|}\sum_{n\in J_l}\left(\eta(n) - \int\psi d\mu\right)^2 = \frac{1}{J_l}\sum_{n\in J_l}\eta(n)^2 - 2\left(\frac{1}{|J_l|}\sum_{n\in J_l}\eta(n)\right)\left(\int\psi d\mu\right) + \left(\int\psi d\mu\right)^2$$
$$\longrightarrow \int\left(\frac{1}{p}\sum_{q=0}^{p-1}T^q\psi\right)^2 d\mu - \left(\int\psi d\mu\right)^2$$

which is small for large p. But this implies that most $\eta(n)$ are close to $\int\psi d\mu$ as asserted in the proposition. □

3 Some Examples of PW-Bohr Sets

3.1 Fourier Transforms

Our first example of PW-Bohr sets will lead to three more in the following subsections.

Theorem 3.1. *Let ω be a non-negative measure on $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ with a non-trivial discrete (atomic) component, and let $\hat\omega$ denote its Fourier transform: $\hat\omega(n) = \int_{\mathbb{T}} e^{2\pi itn}d\omega(t)$. If*

$$S = \{n : \mathrm{Re}\,\hat\omega(n) > 0\},$$

then S is a PW-Bohr$_0$ set.

Proof. Let ω_d denote the discrete component of ω: $\omega_d = \sum_{\lambda\in\Lambda}\omega(\{\lambda\})\delta_\lambda$ where Λ consists of all the atoms of ω . Let Λ_0 be a finite subset of Λ so that $\omega_d(\Lambda_0) > \frac{3}{4}\omega_d(\Lambda)$. Set

$$\psi(\tau) = \sum_{\lambda\in\Lambda_0}\omega_d(\lambda)e^{2\pi i\lambda\tau}$$

and let B_0 be the Bohr$_0$ set: $B_0 = \{n : \mathrm{Re}\,\psi(n) > \frac{2}{3}\omega_d(\Lambda_0)\}$. The measure $\omega - \omega_d$ is continuous and so by Wiener's theorem (see [Kre85], p.96)

$$\lim_{N\to\infty}\frac{1}{2N+1}\sum_{n=-N}^{N}\left|\hat\omega(n) - \hat\omega_d(n)\right|^2 = 0$$

It follows that $Q' = \left\{n : \left|\hat\omega(n) - \hat\omega_d(n)\right| \geq \frac{1}{3}\omega_d(\Lambda_0)\right\}$ has density 0 so that $Q = \mathbb{Z}\backslash Q'$ is a thick set.

In $B_0 \cap Q$,

$$\begin{aligned}\operatorname{Re}\hat{\omega}(n) &> \operatorname{Re}\hat{\omega}_d(n) - \tfrac{1}{3}\omega_d(\Lambda_0)\\ &\geq \operatorname{Re}\psi(n) - \omega_d(\Lambda\backslash\Lambda_0) - \tfrac{1}{3}\omega_d(\Lambda_0)\\ &> \operatorname{Re}\psi(n) - \tfrac{1}{4}\omega_d(\Lambda) - \tfrac{1}{3}\omega_d(\Lambda_0)\\ &> \tfrac{2}{3}\omega_d(\Lambda_0) - \tfrac{1}{3}\omega_d(\Lambda_0) - \tfrac{1}{3}\omega_d(\Lambda_0) = 0\end{aligned}$$

so that $B_0 \cap Q \subset S$. It follows that S is PWB_0. ◻

3.2 Positive Definite Sequences

Theorem 3.2. *Let $\{a(n)\}_{n\in\mathbb{Z}}$ be a positive definite sequence of non-negative reals for which* $\lim \frac{1}{2N+1} \sum_{n=-N}^{N} a(n) > 0$. *Then $S = \{n : a(n) > 0\}$ is a PWB_0 set.*

Proof. By Herglotz's theorem $a(n) = \hat{\omega}(n)$ for some non-negative measure ω on $\mathbb{T}$ [Hel83], and the hypothesis of the theorem implies that $\omega\{0\} > 0$. The previous theorem applies and so S is PWB_0. ◻

3.3 Return Time Sets

A consequence of the foregoing is that RT-sets are PW-Bohr_0 sets. Recall a return time set has the form $S \supset R(A) = \{n : \mu(A \cap T^{-n}A) > 0\}$ where $(X, \mathcal{B}, \mu, T)$ is a measure preserving system, $A \in \mathcal{B}$ and $\mu(A) > 0$. If $a(n) = \mu(A \cap T^{-n}A)$ we can write $a(n) = \int fT^n f d\mu$ with $f = 1_A$ and T is a unitary operator. It is easily checked that $\sum_{m,n=1}^{N} a(n-m)x_n\overline{x}_m \geq 0$ for any $x_1, x_2, \ldots, x_N$, and so $\{a(n)\}$ is a positive definite sequence. We also have

$$\frac{1}{2N+1} \sum_{n=-N}^{N} \int fT^n f d\mu \longrightarrow \int fP_T f d\mu,$$

where P_T is the self-adjoint projection of $L^2(X, \mathcal{B}, \mu)$ to the subspace of T-invariant functions. Since $\int P_T f d\mu = \mu(A)$, it follows that $P_T f \neq 0$, and since $\int fP_T f d\mu = \int fP_T^2 f d\mu = \int (P_T f)^2 d\mu > 0$ the hypotheses of Theorem 3.2 are fulfilled. This proves

Theorem 3.3. *RT sets are PW-$Bohr_0$.*

3.4 Difference Sets of Sets of Positive Upper Density

Proposition 3.4. *Let $\{J_l\}$ be an array and, for each l, let $S_l \subset J_l$ with $|S_l| > \delta|J_l|$ for fixed $\delta > 0$. Then $\bigcup(S_l - S_l)$ is PW-$Bohr_0$.*

This leads immediately to

Theorem 3.5. *If $d^*(S) > 0$, then $S - S$ is PWB_0 for $S \subset \mathbb{Z}$*

Proof of Proposition 3.4. We form the scheme $(\{J_l\}, \{\xi^l\})$, where the usual index r is suppressed since it takes only one value, and we define $\xi^l(n) = 1_{S_l}(n)$. We pass to a subscheme which is generic for a process $(X, \mathcal{B}, \mu, T, \{\varphi\})$ where, according to the scholium following Theorem 2.2, φ is not almost everywhere 0. By the construction ($\Lambda = \{0,1\}$), φ takes on the values 0, 1 and so $\varphi = 1_A$ for $A \in \mathcal{B}$, $\mu(A) > 0$. By definition of a generic scheme

$$\mu(A \cap T^{-k}A) = \int \varphi T^k \varphi d\mu = \lim \frac{1}{|H_l|} \sum_{n \in H_l} \xi^l(n)\xi^l(n+k)$$

which will be > 0 only if $k \in \bigcup(S_l - S_l)$. This proves the proposition. □

In the sequel we will use a stronger version of Proposition 3.4. Let us say that a set Q is *uniformly thick* if for every $l \in \mathbb{N}$, $\exists l' \in \mathbb{N}$ so that every interval J of length l' meets Q in a set containing an interval of length l. This will happen if $\frac{1}{N} \sum_{j=n+1}^{n+N} 1_Q(j) \to 1$ uniformly in n. If ω is a continuous measure on $\mathbb{T}$ then Wiener's Theorem can be sharpened to

$$\frac{1}{N} \sum_{j=n+1}^{n+N} |\hat{\omega}(j)|^2 \longrightarrow 0$$

uniformly in n. Using this in the proof of Theorem 3.1 we find that the set S of that theorem is the intersection of a Bohr_0-set and a uniformly thick set. If we call a set a UPW-Bohr_0 set if it contains intersection of a Bohr set and a uniformly thick set, we can replace PW-Bohr_0 throughout this section by UPW-Bohr_0. For later reference we re-write Proposition 3.4 in its strengthened form as

Proposition 3.6. *Let $\{J_l\}$ be an array and for each l, let $S_l \subset J_l$ with $|S_l| > \delta|J_l|$ for fixed $\delta > 0$. Then $\bigcup(S_l - S_l)$ is a UPW-$Bohr_0$ set.*

4 The Hierarchy of Families of Large Sets

We consider the following families of "large sets":

(a) B_0 = Bohr_0 sets
(b) RT = return time sets
(c) $\bigcup \Delta_r^*$ = sets which for some r meet every $(S - S)\backslash\{0\}$ provided $|S| \geq r$
(d) PWB_0 = piecewise Bohr_0 sets
(e) PWB = piecewise Bohr sets
(f) PW Syn = piecewise syndetic sets
(g) PD = sets of positive upper Banach density = $\{S : d^*(S) > 0\}$

It is easily seen that the first of these families is contained in the second, the second in the third, the fourth in the fifth and the fifth in the sixth. That $\bigcup \Delta_r^* \subset \mathrm{PWB}_0$ is the content of our Theorem II to be proved in §9. In fact all these inclusions are proper, and in this section we shall show that (b) $\neq$ (c), (c) $\neq$ (d) and (e) $\neq$ (f). The fact that (a) $\neq$ (b) follows from work of I. Kříž [Kříž87] and that (f) $\neq$ (g) is an exercise.

Theorem 4.1. *There are Δ_3^*-sets which do not contain RT-sets. So (b) $\neq$ (c).*

Proof. We use the fact ([Fur81], [Sár78]) that for every $r = 1, 2, \ldots$ the set $P_r = \{n^r\}_{n\in\mathbb{Z}}$ is a Poincaré set; i.e., it meets every return time set. Hence $\mathbb{Z}\backslash P_r$ does not contain any RT-set. On the other hand, when $r \geq 3, \mathbb{Z}\backslash P_r$ is a Δ_3^*-set. For, by Fermat's theorem, for any distinct a, b, c, we cannot have $b - a$, and $c - b$ as well as $c - a = (b - a) + (c - b)$ all in P_r. ◻

To prove that (c) $\neq$ (d) we produce a set of density 0 in $\mathbb{Z}$ that contains a Δ_r-set for every r. The complement of this set cannot belong to any Δ_r^*. On the other hand, the complement of a set of density 0 contains arbitrarily long intervals, and so is thick, and in particular it is PWB_0. So we take as a Δ_r-set a set of the form

$$D_r = \{-rq_r, -(r-1)q_r, \ldots, -q_r, 0, q_r, \ldots, (r-1)q_r, rq_r\}$$

Choosing $q_r = r^3$ we can check that the density of $\bigcup D_r$ is 0. This proves

Theorem 4.2.

$$\bigcup \Delta_r^* \neq PWB_0$$

Finally we have (e) $\neq$ (f) by the following:

Theorem 4.3. *There are syndetic sets that are not PWB.*

Proof. We use considerations from topological dynamics. Let $\Omega = \{0,1\}^{\mathbb{Z}}$ and define the shift T on Ω by $T\omega(n) = \omega(n+1)$. If $M \subset \Omega$ is a minimal closed T-invariant subset, $M \neq \{0\}$, then for any $\omega \in M$, $\{n : \omega(n) = 1\}$ is syndetic. We can choose M so that the system (M, T) is weakly mixing ([Fur81]). Let $\xi \in M$ and set $S = \{n : \xi(n) = 1\}$. Assume S is PWB; then $S = Q \cap P$ where Q is thick and P is a Bohr set. If $\eta = 1_P$ then ξ and η agree on arbitrarily long intervals and for some $\{n_k\}, \lim T^{n_k}\xi = \lim T^{n_k}\eta$. Let $L = \overline{\{T^n\eta\}_{n\in\mathbb{Z}}}$ be the closed invariant set generated by η (so that $M \cap L \neq \emptyset$). Since M is minimal, $M \subset L$. By definition of a Bohr set there is a torus $\mathbb{T}^m$, a rotation $R : \mathbb{T}^m \to \mathbb{T}^m$, $R(\theta) = \theta + \alpha$, and an open set $U \subset \mathbb{T}^m$ so that $R^n(0) \in U \Rightarrow \eta(n) = 1$. Let $A = \{\omega : \omega(0) = 1\}$; then $R^n(0) \in U \Rightarrow T^n\eta \in A$. Let $Z \subset \mathbb{T}^m$ be the closed subgroup of $\mathbb{T}^m$ generated by α. By [Fur67] (Z, R) and (M, T) are disjoint, and since both are minimal, $Z \times M$ is minimal for $R \times T$. This implies that $\{(R^n(0), T^n\eta)\}$ is dense in $Z \times M$. But from the foregoing, when the first coordinate is in U the other is in A. It follows that $U \times A$ is dense in $U \times M$; hence $M = A$ and $\xi \equiv 1$. Choosing M non-degenerate gives us the example we seek. ◻

5 The Sum Set of Positive Density Sets

In this section we will prove Theorem I which asserts that the sum set of two sets A, B with positive upper density is a PW-Bohr set.

We begin with an elementary lemma.

Lemma 5.1. *Let $J, J' \subset \mathbb{Z}$ be intervals of length l, l' respectively. Let $S \subset J$, $S' \subset J'$ be subsets satisfying $|S| \geq \delta l$, $|S'| \geq \delta' l'$. We can find an interval L and a subset $R \subset L$ so that for some t, $S + S' \supset R - R + t$ and such that $|R| \geq \frac{\delta\delta'}{2}|L|$.*

Proof. Without loss of generality we suppose $l \leq l'$. For each $t \in \mathbb{Z}$, form $R_t = S \cap (t - S')$. $|R_t|$ equals the number of points of $S \times S'$ lying on the line $x + y = t$. The number of such lines meeting $S \times S'$ doesn't exceed $l + l'$, and so for some t,

$$|R_t| \geq \frac{|S \times S'|}{l + l'} \geq \frac{\delta\delta' l l'}{l + l'} \geq \frac{\delta\delta'}{2} l.$$

Take $R = R_t$ so that $R - R \subset S + (S' - t)$, and take $L = J$. □

Theorem I will now follow from

Theorem 5.2. *Let $\{J_k\}$ be an array (Def. 2.3), and let $S_k \subset J_k$, with $|S_k| > \delta|J_k|$ where $\delta > 0$. Let $\{t_k\}$ be an arbitrary set of integers. The set $A = \bigcup_{k=1}^{\infty} (S_k - S_k + t_k)$ is PW-Bohr.*

Our next step is to reduce Theorem 5.2 to a special case in which the sets S_k are related. For two sets of integers S', S'', let us write $S' \prec S''$ if for some $c \in Z$, $S' + c \subset S''$. Clearly $S' \prec S''$ implies that $S' - S' \subset S'' - S''$.

Lemma 5.3. *Theorem 5.2 is true in general if it is true for the case that $S_k \prec S_{k+1}$ for each $k = 1, 2, 3, \ldots$.*

Proof. We consider the general case of an arbitrary array $\{J_k\}$ with subsets $S_k \subset J_k$. We follow the procedure in the proof of Proposition 3.4 based on Theorem 2.2 to obtain a subscheme of $(\{J_k\}, \{1_{S_k}\})$ generic for an ergodic process $(X, \mathcal{B}, \mu, T, 1_A)$ with $\mu(A) > 0$. Reindexing and renaming sets we suppose that $(\{J_k\}, \{1_{S_k}\})$ is generic for the above process. Note that the hypothesis of genericity implies that we will still have $|S_k| > \delta'|J_k|$ for some positive δ'. We now pass to a further subscheme for which $S'_k \prec S'_{k+1}$. This is done as follows. Removing a set of measure 0 from A we can assume that any non-empty intersection $A \cap T^{-\tau_1} A \cap T^{-\tau_2} A \cap \cdots \cap T^{-\tau_r} A$ has positive measure. It follows from the ergodic theorem that there exist points x with $T^{\tau_n} x \in A$ for a sequence $\tau_1 < \tau_2 < \cdots < \tau_n < \cdots$ (depending on x) with $\lim \frac{\tau_n}{n} < \infty$. Thus $A \cap T^{-\tau_1} A \cap T^{-\tau_2} A \cap \cdots \cap T^{-\tau_r} A$ is non-empty for each r and by our assumption $\mu(A \cap T^{-\tau_1} A \cap T^{-\tau_2} A \cap \cdots \cap T^{-\tau_r} A) > 0$ for each r. By genericity of $(\{J_k\}, \{1_{S_k}\})$ this implies that translating $\{0, \tau_1, \tau_2, \cdots, \tau_r\}$ by some c_r we will obtain a subset of some $S_k : \{c_r, c_r + \tau_1, c_r + \tau_2, \cdots, c_r + \tau_r\} \subset S_{k(r)}$. We now set $J'_r = [c_r, c_r + \tau_r] \subset J_{k(r)}$ and $S'_r = \{c_r, c_1 + \tau_1, c_r + \tau_2, \cdots, c_r + \tau_r\}$. Then $S'_r \prec S'_{r+1}$ and since $S'_r \subset S_{k(r)}$, $\bigcup(S_k - S_k + t_k) \supset \bigcup(S'_r - S'_r + t_{k(r)})$. At the same time $\lim \frac{\tau_n}{n} < \infty$ so that $\exists \alpha > 0$ with $r = |S'_r| \geq \alpha |J'_r|$. ◻

We show now how Theorem 5.2 follows from Proposition 3.6.

Proof of Theorem 5.2. According to the foregoing lemma, we may assume that for each k, $S_k - S_k \subset S_{k+1} - S_{k+1}$. For each $m = 1, 2, 3, \ldots$, let $k(m)$ be chosen so that $(S_k - S_k) \cap [-m, m]$ is a fixed set for $k \geq k(m)$. Write $S'_m = S_{k(m)}$ and $t'_m = t_{k(m)}$; we will show that $\bigcup(S'_m - S'_m + t'_m)$ is PW-Bohr. By Proposition 3.6 $\bigcup(S'_m - S'_m)$ is UPW-Bohr$_0$; i.e., it contains the intersection of a Bohr$_0$ set H and a *uniformly* thick set Q. Thus there is a trigonometric polynomial $\psi(t) = \sum_{j=1}^{N} a_j e^{i\lambda_j t}$ with $\mathrm{Re}\,\psi(0) > 0$ such that for any $n \in Q$, if $\mathrm{Re}\,\psi(n) > 0$ then $n \in \bigcup(S'_m - S'_m)$. Form $\psi_m(t) = \psi(t - t'_m)$ and pass to a subsequence $\{m_p\}$ so that these converge uniformly to a polynomial $\psi'(t)$. Let $0 < \alpha < \mathrm{Re}\,\psi(0)$. By almost periodicity of $\psi(t)$ it follows that $\mathrm{Re}\,\psi'(n) > \alpha$ on a non-empty (and therefore syndetic) set of n. We can suppose that the subsequence $\{m_p\}$ is such that $\mathrm{Re}\,\psi'(n) > \alpha$ implies $\mathrm{Re}\,\psi(n - t'_{m_p}) > 0$ for each p. Form the set $Q' = \bigcup([-m_p, m_p] \cap Q + t'_{m_p})$. Suppose $\mathrm{Re}\,\psi'(n) > \alpha$ with $n \in Q'$. Then for some p, $n - t'_{m_p} \in [-m_p, m_p] \cap Q$ and $\mathrm{Re}\,\psi(n - t'_{m_p}) > 0$. It follows that $n - t'_{m_p} \in (\bigcup(S'_m - S'_m)) \cap [-m_p, m_p]$. By the choice of $\{S'_m\}$ this implies $n \in S'_{m_p} - S'_{m_p} + t'_{m_p}$. Since Q is uniformly thick, for large p, $[-m_p, m_p] \cap Q$ contains large intervals and this implies that Q' is a thick set. This proves that $\bigcup(S'_m - S'_m + t'_m)$ is a PW-Bohr set. ◻

This completes the proof of Theorem I.

Corollary 5.4 (Corollary 1.3 of §1). *If $A, B, C, \subset \mathbb{Z}$ are three sets with positive upper density, one of which is syndetic, then $A + B + C$ is a Bohr set.*

This will follow from the Theorem I together with the following lemma:

Lemma 5.5. *If R is a PW-Bohr set and S is syndetic in $\mathbb{Z}$ then $R + S$ is Bohr.*

Proof. A translate of R will be PW-Bohr$_0$ and the opposite translate of S is syndetic, so we can assume that R is a PW-Bohr$_0$ set. This means that there is a torus $\mathbb{T}^m$, an $\alpha \in \mathbb{T}^m$, a neighborhood U of 0 in $\mathbb{T}^m$ and a thick set Q with $R \supset \{n : n\alpha \in U\} \cap Q$. Let V be a neighborhood of 0 in $\mathbb{T}^m$ with $V - V \subset U$ and let $\beta_1, \beta_2, \ldots \beta_k \in \mathbb{T}^m$ so that $\mathbb{T}^m = \bigcup_{l=1}^{k}(\beta_l + V)$.

We claim that for some $l, 1 \leq l \leq k$, $S + R \supset \{n : n\alpha \in \beta_l + V\}$ which implies that $S + R$ is a Bohr set. Assume this isn't so; then for each l, $\exists x_l$ with $x_l\alpha \in \beta_l + V$ and $x_l \notin S + R$. Let $S_l = S \cap \{n : n\alpha \in \beta_l + V\}$ so that $S = \bigcup S_l$. We have $x_l \notin S + R$ and so $x_l - S_l \cap R = \varnothing$. Since $x_l\alpha \in \beta_l + V$ and $S_l\alpha \subset \beta_l + V$ we have $(x_l - S_l)\alpha \subset U$. Now $R \supset \{n : n\alpha \in U\} \cap Q$ so $(x_l - S_l) \cap R = \varnothing$

implies that $(x_l - S_l) \subset Q^c$, the complement of Q. Equivalently $S_l \subset (x_l - Q)^c$, so $S = \bigcup S_l \subset \left(\bigcap (x_l - Q)\right)^c$. But the intersection of finitely many translates of a thick set is thick whereas S is syndetic. This contradiction proves our assertion. ◻

6 Kronecker-complete Processes

The remaining sections are directed to giving a proof of Theorem II of §1. The crucial step in this proof is a proposition to be proved in §8 which generalizes the fact (Theorem 3.5) that $d^*(A) > 0$ implies that $A - A$ is PWB_0. To achieve this generalization we will use once more the correspondence described in §2 between schemes and processes. Another ingredient that will enter is the point spectrum of an ergodic system, i.e., the eigenvalues of the operator T on the L^2-space of the system. It will be of importance that in a scheme generic for a process for which non-trivial eigenvalues exist, the eigenfunctions are also represented. This leads to the notion dealt with in this section of a "Kronecker-complete process."

We begin by recalling the notion of the "Kronecker factor" of an ergodic system: Let $(X, \mathcal{B}, \mu, T)$ be an ergodic measure preserving system. There is a compact abelian group Z and an element $\alpha \in Z$ whose multiples $\{n\alpha\}$ are dense in Z, and a map $\pi : X \to Z$ which is measurable and measure preserving with respect to Haar measure dz on Z, and such that for a.e. $x \in X$, $\pi(Tx) = \pi(x) + \alpha$. If $\chi \in \hat{Z}$ is a character on Z then $f = \chi \circ \pi$ is an eigenfunction of T: $f(Tx) = \chi(\pi(x) + \alpha) = \chi(\alpha) f(x)$, and *every* eigenfunction of T in $L^2(X, \mathcal{B}, \mu)$ is a multiple of one derived from a character. (Z, α) is unique up to isomorphism and is called the *Kronecker factor* of $(X, \mathcal{B}, \mu, T)$. The eigenvalues of T are $\{\chi(\alpha)\}_{\chi \in \hat{Z}}$, so that $Z \cong$ the dual group to the (discrete) group of eigenvalues of T. The system $(X, \mathcal{B}, \mu, T)$ is *weakly mixing* if and only if there are no eigenvalues other than 1 if and only if Z is the trivial one-element group. The discussion in this section will be vacuous in the case of weakly mixing systems.

We turn to processes. When we speak of an eigenfunction f we will assume $f \not\equiv 0$.

Definition 6.1. *A process* $(X, \mathcal{B}, \mu, T, \Phi)$ *is* Kronecker-complete *if it is ergodic and if every eigenfunction of* T *is proportional to some function in* Φ.

Note that for an ergodic system, if $Tf = \lambda f$ for a measurable f, it is easily seen that $|\lambda| = 1$ and that $|f(x)|$ is constant a.e., so that $f \in L^\infty(X, \mathcal{B}, \mu)$. Also note that $Tf_1 = \lambda f_1, Tf_2 = \lambda f_2$ implies that f_1/f_2 is invariant so that by ergodicity, f_1, f_2 are proportional. Thus a process is Kronecker-complete if Φ contains *some* eigenfunction for each eigenvalue. Under our standing hypothesis that $\mathcal{B}$ is a countably generated σ-algebra, the set of eigenvalues is at most countable. As a result we can always "complete" a non-Kronecker-complete process. The principal result in this section states that if a scheme is generic for a non-Kronecker-complete process, by augmenting the process and the scheme and passing to a subscheme, we will obtain a scheme generic for a Kronecker-complete process.

Theorem 6.2. *Let* $(\{J_l\},\{\xi_r^l\})$ *be generic for an ergodic process* $(X,\mathcal{B},\mu,T,\Phi)$. *Denote by* Λ *the subgroup of the unit circle* S^1 *consisting of eigenvalues of* T *on* $L^2(X,\mathcal{B},\mu)$. *We can find eigenfunctions* ψ_λ *for each* $\lambda \in \Lambda$ *and a subscheme* $(\{H_k\},\{\eta_r^k\})$ *so that setting* $\eta_\lambda^k(n) = \lambda^n$ *independent of* k *and letting* $\Psi = \{\psi_\lambda\}_{\lambda\in\Lambda}$, *the process* $(X,\mathcal{B},\mu,T,\Phi\cup\Psi)$ *will be Kronecker-complete, and the scheme* $(\{H_k\},\{\xi_r^k\}\cup\{\eta_\lambda^k\})$ *will be generic for* $(X,\mathcal{B},\mu,T,\Phi\cup\Psi)$.

In the weak mixing case we merely need to adjoin the function 1 to the process and to the scheme. In the general case we proceed by successively adjoining eigenfunctions, passing to a subarray at each stage. We will thus obtain a sequence of subarrays which is "decreasing" and a sequence $\Phi_n = \Phi \cup \{\Psi_{\lambda_1},\Psi_{\lambda_2},\dots,\Psi_{\lambda_n}\}$ of sets of functions with the corresponding $\{\eta_r^{(k)}\} \cup \{\eta_{\lambda_1},\eta_{\lambda_2},\dots,\eta_{\lambda_n}\}$ of representative time series. Our final scheme is obtained by choosing from successive schemes intervals that are "ε-h-generic" for the final process $(X,\mathcal{B},\mu,T,\Phi\cup\Psi)$ with $\varepsilon \searrow 0$, $h \nearrow \infty$. Such intervals will be found in the array for Φ_n with n sufficiently large.

Adjoining a single eigenfunction will also entail a procedure of successive approximation. We assume given a scheme $(\{J_l\},\{\xi_r^l\})$ generic for $(X,\mathcal{B},\mu,T,\Phi)$ and we wish to adjoin an eigenfunction for the eigenvalue λ. Fix an eigenfunction $f, Tf = \lambda f$, with $|f| = 1$. Since we have fixed the representative time series for the eigenfunction as η_λ where $\eta_\lambda(n) = \lambda^n$, the corresponding φ_λ to be adjoined will be some multiple cf, $|c| = 1$. Our task is to find subintervals of J_l that give better and better representation for the augmented $\Phi\cup\{cf\}$ in a sense analogous to ε-h-genericity (§2). In our procedure of successive approximation we can let c vary, since a subsequence will converge to a fixed value for which the intervals that have been found will still provide good representation. We form Φ^* from Φ as in §2, and express Φ^* as a union $\Phi^* = \bigcup \Phi_h^*$ of increasing finite subsets. Now $\{cf\}$ enters the picture and we say that the interval $J \subset J_l$ is "ε-h-m-generic" for $(X,\mathcal{B},\mu,T,\Phi\cup cf)$ if for every $\varphi \in \Phi_h^*$ and the corresponding time series ξ^l, and for a an integer with $0 \le a \le m$,

$$\left| \frac{1}{|J|} \sum_{n\in J} \xi^l(n)\lambda^{an} - \int_X \varphi \cdot c^a f^a d_\mu \right| < \varepsilon. \tag{5}$$

Note that for $a = 0$ this is ε-h-genericity. What will be shown for the proof of the theorem is the existence of ε-h-m-generic intervals inside J_l for large l for arbitrary ε, h, m, and putting these together we obtain the subscheme that is sought.

In establishing (5) we will use the following lemma.

Lemma 6.3. *Let $\alpha_1, \alpha_2, \ldots, \alpha_N$ be N complex numbers and form, for $a, b = 0, 1, 2, \ldots$, the averages*

$$u(a,b) = \frac{1}{N} \sum_{i=1}^{N} \alpha_i^a \overline{\alpha}_i^b.$$

There is a function $\delta(\varepsilon, p) > 0$ for $\varepsilon > 0$ and $p \in \mathbb{N}$ so that if $|u(a,b) - 1| < \delta(\varepsilon, p)$ for $0 \le a, b \le p$, then $\exists \beta$ so that

$$\frac{1}{N} \sum_{i=1}^{N} |\alpha_i - \beta|^{2p} < \varepsilon$$

Proof. We form the average

$$\begin{aligned} \frac{1}{N^2} \sum_{i,j=1}^{N} |\alpha_i - \alpha_j|^{2p} &= \frac{1}{N^2} \sum_{i,j=1}^{N} (\alpha_i - \alpha_j)^p (\overline{\alpha}_i - \overline{\alpha}_j)^p \\ &= \frac{1}{N^2} \sum_{q=0}^{p} \sum_{q'=0}^{p} \sum_{i,j=1}^{N} (-1)^{2p-q-q'} \tbinom{p}{q} \tbinom{p}{q'} \alpha_i^q \, \overline{\alpha}_i^{q'} \, \alpha_j^{p-q} \, \overline{\alpha}_j^{p-q'} \\ &= \sum_{q,q'=0}^{p} (-1)^{2p-q-q'} \tbinom{p}{q} \tbinom{p}{q'} u(q,q') u(p-q,\, p-q') \end{aligned}$$

The latter expression is continuous in the $(p+1)^2$ expressions $\{u(q,q'), 0 \le q,\, q' \le p\}$ and we can evaluate it for $u(q,q') = 1$ by setting all $\alpha_i = 1$. Since the expression in question vanishes when $\alpha_i = 1$, it follows that we can find $\delta(\varepsilon, p) > 0$ so that the hypothesis of the lemma implies

$$\frac{1}{N} \sum_{j=1}^{N} \left(\frac{1}{N} \sum_{i=1}^{N} |\alpha_i - \alpha_j|^{2p} \right) < \varepsilon.$$

But this implies that for some index j the inside average is $< \varepsilon$, so with $\beta = \alpha_j$ we get the desired result. □

Proof of Theorem 6.2. We have seen that to prove the theorem we have to show the existence of long intervals J inside J_l for sufficiently large l, for which (5) is valid, where φ ranges over Φ_h^*, f is an eigenfunction $Tf = \lambda f$, and the $\xi^l(n)$ are the time series representing φ in the respective J_l, and the exponent "a" ranges from 1 to m.

Our assumption in Definition 2.5 that the functions of Φ generate the σ-algebra $\mathcal{B}$ for the process $(X, \mathcal{B}, \mu, T, \Phi)$ implies that linear combinations of functions in Φ^* will approximate any function in $L^p(X, \mathcal{B}, \mu)$ in the L^p-norm, for any p, $1 \le p < \infty$. We wish to approximate f and for any $\varepsilon_1 > 0$ we can find σ in the linear space spanned by Φ^* with $||\sigma - f||_{L^q} < \varepsilon_1$ where $q = q(m) \ge 8$ will be made explicit further on. Taking appropriate combinations of the time series $\zeta^l(n)$ representing σ in the given scheme, we find that

$$\frac{1}{|J_l|}\sum_{n\in J_l}\left(\zeta^l(n)\right)^r\left(\overline{\zeta^l(n)}\right)^s\left(\zeta^l(n+k)\right)^t\left(\overline{\zeta^l(n+k)}\right)^u \longrightarrow \int_X \sigma^r\overline{\sigma}^s T^k(\sigma^t\overline{\sigma}^u)d\mu. \tag{6}$$

We're going to apply Lemma 6.3 with $p=2$ to the $N=K|J_l|$ numbers:

$$\alpha_{k,n}=\lambda^{-k}\zeta(n+k)\overline{\zeta(n)} \qquad 0\le k\le K-1,\quad n\in J_l$$

where $\zeta=\zeta^l$. K will be arbitrary and l will be large. We have

$$u(a,b)=\frac{1}{|J_l|}\frac{1}{K}\sum_{k=0}^{K-1}\sum_{n\in J_l}\lambda^{(b-a)k}\zeta(n+k)^a\overline{\zeta(n+k)}^b\zeta(n)^b\overline{\zeta(n)}^a$$

When l is large this is close to $\int_X \frac{1}{K}\sum_{k=0}^{K-1}\lambda^{(b-a)k}\sigma^b\overline{\sigma}^aT^k(\sigma^a\overline{\sigma}^b)d\mu$. The latter expression will be within ε_2 of

$$\int_X \frac{1}{K}\sum_{k=0}^{K-1}\lambda^{(b-a)k}f^b\overline{f}^aT^k(f^a\overline{f}^b)d\mu=\int_X \frac{1}{K}\sum_{k=0}^{K-1}\lambda^{(b-a)k}f^{b-a}T^k(f^{a-b})d\mu=1$$

where $\varepsilon_2=\varepsilon_2(\varepsilon_1)\to 0$ as $\varepsilon_1\to 0$, using the fact that σ is close to f in L^8 and the total exponent in the integrals above is $2a+2b\le 8$, and the fact that $T^kf=\lambda^kf$. Having chosen ε_1 sufficiently small, we find by Lemma 6.3 that for l large we can find β_l so that

$$\frac{1}{K}\frac{1}{J_l}\sum_{k,n}\left|\lambda^{-k}\zeta(n+k)\overline{\zeta(n)}-\beta_l\right|^4<\varepsilon_0 \tag{7}$$

where ε_0 is given.

We wish to use (7) to estimate

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\zeta(n+k)\overline{\zeta(n)}\zeta(n)-\lambda^k\beta_l\zeta(n)\right|^2=$$
$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|(\lambda^{-k}\zeta(n+k)\overline{\zeta(n)}-\beta_l\right|^2|\zeta(n)|^2\le\sqrt{\varepsilon_0}\theta_l$$

where $\theta_l^2=\frac{1}{|J_l|}\sum_{n\in J_l}|\zeta(n)|^4=\frac{1}{|J_l|}\sum_{n\in J_l}\zeta(n)^2\,\overline{\zeta(n)}^2$, and by (6), $\theta_l^2\to\int|\sigma|^4d\mu$ as $l\to\infty$. Since $\|\sigma-f\|_4<\varepsilon_1$ the latter expression is $<(1+\varepsilon_1)^4$ and we can assume this ≤ 4. We get for large l

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\zeta(n+k)|\zeta(n)|^2-\lambda^k\beta_l\zeta(n)\right|^2<2\sqrt{\varepsilon_0}. \tag{8}$$

Finally we wish to use this to estimate

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\zeta(n+k)-\lambda^k\beta_l\zeta(n)\right|^2$$

and for this we need an estimate of

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\zeta(n+k)|\zeta(n)|^2-\zeta(n+k)\right|^2. \tag{9}$$

As $l\to\infty$, (9) approaches

$$\int\Big(|\sigma|^4\,T^k|\sigma|^2-2|\sigma|^2\,T^k\,|\sigma|^2\;+T^k|\sigma|^2\Big)\,d\mu. \tag{10}$$

The corresponding expression for f instead of σ vanishes so that for some C, the expression in (10) is bounded by $C\varepsilon_1$, and the same will be true for (9) when l is large. Combining this with (8) gives

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\zeta(n+k)-\lambda^k\beta_l\zeta(n)\right|^2<\varepsilon_3=\varepsilon_3(\varepsilon_1)$$

for large l, where $\varepsilon_3(\varepsilon_1)\to 0$ for $\varepsilon_1\to 0$.

Using the Hilbert space inequality

$$\|u\|^2-\|v\|^2\le\Big(\|u\|+\|v\|\Big)\,\|u-v\|$$

we find for large l

$$\left|\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}|\zeta(n+k)|^2-\frac{1}{K}\frac{1}{|J_l|}\,|\beta_l|^2\sum_{k,n}|\zeta(n)|^2\right|\le C'\sqrt{\varepsilon_3}$$

from which it follows that $|\beta_l|\to 1$. To summarize the foregoing, we have shown that for any $\varepsilon>0$ we can find a function σ with time series $\zeta^l(n)$ and γ_l with $|\gamma_l|=1$ so that for l sufficiently large, and any K,

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\zeta^l(n+k)-\lambda^k\gamma_l\zeta^l(n)\right|^2<\varepsilon.$$

To apply this to (5) we let $1\le a\le m$ and we estimate for a time series $\xi^l(n)$

$$\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}\left|\Big(\zeta^l(n+k)\Big)^a-\lambda^{ak}\gamma_l\Big(\zeta^l(n)\Big)^a\right|\,\left|\xi^l(n+k)\right| \tag{11}$$

Writing $x^a-y^a=(x-y)(x^{a-1}+x^{a-2}y+\cdots+y^{a-1})$ we obtain for large l that the expression in (11) is bounded by $M\sqrt{\varepsilon}$ where

$$M^2=a\left(\sum_{j=0}^{a-1}\frac{1}{K}\frac{1}{|J_l|}\sum_{k,n}|\zeta^l(n+k)|^{2(a-j-1)}|\zeta^l(n)|^{2j}|\xi^l(n+k)|^2\right)$$

If ξ^l represents the function φ, the limit of the foregoing expression, as $l \to \infty$, is

$$\frac{a}{K} \sum_{k=0}^{K-1} \int T^k |\sigma|^{2(a-j-1)} |\sigma|^{2j} T^k |\varphi|^2 d\mu,$$

and provided $q(m) \geq 2m + 2$ with $\|\sigma - f\|_{L^q} < 1$, the expression in (11) will be bounded by $C'\|\varphi\|_{L^q}\sqrt{\varepsilon},\ C' = C'(m)$.

In all the estimates for averages over $0 \leq k < K,\ n \in J_l$, if the overall average is $< \theta$, then for at least half of the $n \in J_l$, the average over k cannot exceed 2θ. For large l, we let $N_l \subset J_l$ consist of the n with $\{n, n+1, \ldots, n + K - 1\} \subset J_l$ and

$$\left| \frac{1}{K} \sum_{k=0}^{K-1} \left(\left(\zeta^l(n+k)\right)^a - \lambda^{ak} \left(\gamma_l \zeta^l(n)\right)^a \right) \xi^l(n+k) \right| < 2C'\|\varphi\|_{L^q}\sqrt{\varepsilon}. \quad (12)$$

We now refer to Theorem 2.9 applied to the functions $\sigma_\varphi^a,\ 1 \leq a \leq m$, $\varphi \in \Phi_h^*$ which are in the linear span of Φ^*. These functions are represented in the given scheme by $\left(\zeta^l(n)\right)^a \xi^l(n)$, and with $\delta > 0$ given, there will be a K so that for sufficiently large l, the inequalities

$$\left| \frac{1}{K} \sum_{k=0}^{K-1} \left(\zeta^l(n+k)\right)^a \xi^l(n+k) - \int \sigma^a \varphi d\mu \right| < \delta \quad (13)$$

hold for most $n \in J_l$ provided $|J_l| \gg K$. This implies that (12) and (13) will hold simultaneously for most $n \in N_l$ for which we will then have

$$\left| \frac{1}{K} \sum_{k=0}^{K-1} \lambda^{ak} \left(\gamma_l \zeta^l(n)\right)^a \xi^l(n+k) - \int \sigma^a \varphi d\mu \right| < \delta + 2C'\|\varphi\|_{L^q}\sqrt{\varepsilon}.$$

Set $c_{l,n} = \lambda^n \gamma_l^{-1} \zeta^l(n)^{-1}$ and we can write

$$\left| \frac{1}{K} \sum_{k=0}^{K-1} \lambda^{a(n+k)} \xi^l(n+k) - c_{l,n}^a \int \sigma^a \varphi d\mu \right| < |c_{l,n}|^a \left(\delta + 2C'\|\varphi\|_{L^q}\sqrt{\varepsilon}\right) \quad (14)$$

We write $n \in N_l'$ if (14) is valid.

$$\left| \frac{1}{K} \sum_{k=0}^{K-1} \lambda^{a(n+k)}\ \xi^{(l)}(n+k) - c_{l,n}^a \int_X f^a \varphi d\mu \right| < \quad (15)$$
$$|c_{l,n}|^a \left(\delta + 2C'\|\varphi\|_{L^q}\sqrt{\varepsilon}\right) + |c_{l,n}|^a c'' \|\varphi\|_{L^{m+1}} \|\sigma - f\|_{L^{m+1}}$$

If J is the interval $\{n, n+1, \ldots, n+k-1\}$ then (15) has the form (5) if the right hand side can be made small and if $|c_{l,n}|$ is close to 1. All this can be achieved by choosing σ with $\|\sigma - f\|_{L^q}$ small, and finding $n_l \in J_l$ for which (15) holds with $|c_{l,n}| = |\gamma_l|^{-1}|\zeta^l(n)|^{-a}$ close to 1. The domain of n is N_l' which depends on ζ^l, but $|N_l'|/|J_l|$ is bounded from below. It suffices to show that by choosing $\|\sigma - \delta\|_{L^q}$ small we will have (for ζ^l representing σ) $\left||\zeta^{(l)}(n)| - 1\right| < \theta$ for a preassigned $\theta > 0$ for most $n \in J_l$. But this follows from the fact that

$$\frac{1}{K}\sum_{n\in J_l}\left(|\zeta^l(n)|^2 - 1\right)^2 \longrightarrow \int \left(|\sigma|^2 - 1\right)^2 d\mu$$

as $l \to \infty$ and the latter expression is small if $\|\sigma - f\|$ is small. With this we have completed the proof of Theorem 6.2. □

Corollary 6.4. *If an ergodic process is Kronecker-complete, it has a generic scheme whereby eigenfunctions are represented by the time series $c_\lambda \lambda^n$ for all intervals of the array $\{J_l\}$.*

Suppose now that we have a generic scheme for a Kronecker-complete process, $(X, \mathcal{B}, \mu, T, \Phi)$ and let $\Lambda \subset S^1$ be the group of eigenvalues of the process. If we identify the Kronecker factor of $(X, \mathcal{B}, \mu, T)$ with $Z = \hat{\Lambda}$ we can define a *canonical map* $\pi: X \to Z$. Namely for $\lambda \in \Lambda$ there is a unique eigenfunction φ_λ on X with $T\varphi_\lambda = \lambda\varphi_\lambda$, and which is represented in the scheme by $\eta_\lambda(n) = \lambda^n$. We set $\alpha \in Z = \hat{\Lambda}$ to correspond to the inclusion map of $\Lambda \to S^1 : \alpha(\lambda) = \lambda$. Notice that since $\eta_{\lambda_1\lambda_2} = \eta_{\lambda_1}\eta_{\lambda_2}$ we will have $\varphi_{\lambda_1\lambda_2} = \varphi_{\lambda_1}\varphi_{\lambda_2}$. This means that for a.e. $x \in X$, $\varphi_{\lambda_1\lambda_2}(x) = \varphi_{\lambda_1}(x)\varphi_{\lambda_2}(x)$ so that if we define $\pi(x)(\lambda) = \varphi_\lambda(x)$, then for a.e. $x, \pi(x) \in \hat{\Lambda} = Z$. Moreover $\pi(Tx)(\lambda) = \varphi_\lambda(Tx) = \lambda\varphi_\lambda(x) = \alpha(\lambda)\pi(x)(\lambda) = (\alpha + \pi(x))(\lambda)$; so $\pi(Tx) = \pi(x) + \alpha$. The mapping π is measurable since all φ_λ are measurable, and so the foregoing gives an explicit map of X to its Kronecker factor. This map will play a role in §7.

Note that for $\lambda \in \Lambda$, the eigenfunction φ_λ on X can be identified with $\chi \circ \pi$, where χ is the character on Z given by $\chi(z) = z(\lambda)$ where Z is identified with $\hat{\Lambda}$, since $\chi(\pi(x)) = \pi(x)(\lambda) = \varphi_\lambda(x)$ by definition of π. Since the time series representing φ_λ is $\lambda^n = \chi(n\alpha)$, we conclude:

Proposition 6.5. *Given a scheme generic for a Kronecker-complete process $(X, \mathcal{B}, \mu, T, \Phi)$, if π is the canonical map of X to its Kronecker factor (Z, α) then for any continuous function ψ on Z, $\psi \circ \pi$ can be adjoined to Φ, and it will be represented by the time series $\{\psi(n\alpha)\}$.*

Proof. ψ can be approximated uniformly by linear combinations of $\{\varphi_\lambda\}$. □

7 Weighted Ergodic Averages for Kronecker-complete Processes

Let $(X, \mathcal{B}, \mu, T, \Phi)$ be a Kronecker-complete process and $(\{J_l\}, \{\xi_r^l\})$ a generic scheme. We shall show how to evaluate L^2-limits of weighted ergodic averages

$$\frac{1}{|J_l|}\sum_{n\in J_l}\xi^l(n)T^nf$$

for $f\varepsilon L^2(X,\mathcal{B},\mu)$ and ξ^l representing a function $\varphi\in\Phi$. By our assumption $(X,\mathcal{B},\mu,T)$ is ergodic so that $\frac{1}{N}\sum_{n=0}^{N-1}T^nf\to\int fd\mu$ in L^2. Since T is a contraction we can write

$$\frac{1}{|J_l|}\sum_{n\in J_l}T^nf\longrightarrow\int fd\mu$$

for any array $\{J_l\}$. This will be generalized for processes that are Kronecker-complete, except that the limits taken are weak L^2-limits.

Recall from §6 the notion of Kronecker factor and the canonical map $\pi\colon X\to Z$ where Z is a compact abelian group and $\pi(Tx)=\pi(x)+\alpha$. All eigenfunctions on X are, up to constant multiples, of the form $\chi\circ\pi$ where χ is a character on Z. The set of all functions in $L^2(X,\mathcal{B},\mu)$ of the form $\psi\circ\pi,\psi\in L^2(Z)$ form a subspace that is spanned by eigenfunctions. If $f\in L^2(X,\mathcal{B},\mu)$ we denote by $E(f|Z)$ the unique function in $L^2(Z)$ so that $E(f|Z)\circ\pi$ denotes the orthogonal projection of f to the subspace $L^2(Z)\circ\pi$. $E(f|Z)=0\Leftrightarrow f$ is orthogonal to all eigenfunctions in $L^2(X,\mathcal{B},\mu)$. We will make use of an operation on $L^1(Z)$ related to (but not the same as) convolution:

$$f_1\Box f_2(z)=\int_Z f_1(z+u)f_2(u)du$$

Proposition 7.1. *Let $(\{J_l\},\{\xi_r^l\})$ be generic for the Kronecker-complete process $(X,\mathcal{B},\mu,T,\Phi)$, let $f\in L^2(X,\mathcal{B},\mu)$, and let $\varphi\in\Phi$ be represented by the time series ξ^l. Then*

$$\frac{1}{J_l}\sum_{n\in J_l}\xi^l(n)T^nf\xrightarrow{w}\Big[E(f|Z)\Box E(\varphi|Z)\Big]\circ\pi\tag{16}$$

where $\xrightarrow{w}$ signifies weak convergence in $L^2(X,\mathcal{B},\mu)$.

Proof. It suffices to consider two cases: (a) $E(f|Z)=0$, (b) f is an eigenfunction.

In the first case, for any g in $L^2(X,\mathcal{B},\mu)$, the sequence $\{\int T^nf\cdot gd\mu\}$ satisfies

$$\frac{1}{N}\sum_{k=n+1}^{n+N}\left|\int T^kf\cdot gd\mu\right|^2\underset{N\to\infty}{\longrightarrow}0$$

uniformly in n, so that the left hand side of (16) goes to 0 weakly, and the proposition is verified. We turn to case (b) with $f=\varphi_\lambda$. To $\lambda\in\Lambda$ we associate the character χ on $\hat{\Lambda}$ with $\chi(z)=z(\lambda)$. Then $\chi\circ\pi(x)=\pi(x)(\lambda)=\varphi_\lambda(x)=f(x)$, and $E(f|Z)=\chi$. In this case the right hand side of (16) is $[\chi \Box E(\varphi|Z)] \circ \pi = \left(\int\limits_Z E(\varphi|Z) \chi dz \right) \chi \circ \pi$. We evaluate the left hand side of (16):

$$\frac{1}{|J_l|} \sum_{n \in J_l} \xi^l(n) T^n f = \frac{1}{|J_l|} \sum_{n \in J_l} \lambda^n \xi^{(l)}(n) f$$

which by genericity converges to $\left(\int \varphi_\lambda \varphi d\mu \right) f$. Since $\varphi_\lambda \in L^2(Z) \circ \pi$, we can replace φ by its projection to this subspace which is $E(\varphi|Z) \circ \pi$. Since $\varphi_\lambda = \chi \circ \pi$ we now have

$$\int\limits_X \varphi_\lambda \cdot \varphi d\mu = \int\limits_Z \chi E(\varphi|Z) dz$$

and since $f = \chi \circ \pi$, this proves the proposition. ◻

8 A Condition for PW-Bohr$_0$

We know from Theorem 3.5 that if $d^*(S) > 0$ for a subset $S \subset \mathbb{Z}$, then $S - S$ is PW-Bohr$_0$. We can rephrase this as saying that if for each $s \in S$, $S - s \cap B = \emptyset$ for a subset $B \subset \mathbb{Z}$, then the complement of B is PW-Bohr$_0$. In this section we show that it will suffice for this conclusion that $d^*\Big((S - s) \cap B\Big) = 0$ for each $s \in S$. In §9 we'll see how this leads to a proof of Theorem II.

Proposition 8.1. *Let $A \subset \mathbb{Z}$ and $B = \mathbb{Z} \backslash A$ and let $S \subset \mathbb{Z}$ with $d^*(S) > 0$. If for every $s \in S$, $d^*\Big((S - s) \cap B\Big) = 0$, then A is a PW-Bohr$_0$ set.*

Proof. Let $\{J_l\}$ be an array with $\frac{|J_l \cap S|}{|J_l|} \to \beta > 0$. Set $\xi_1^l(n) = 1_A(n)$, $\xi_2^l = 1_B(n), \xi_3^l(n) = 1_S(n)$ and consider the scheme $(\{J_l\}\ \{\xi_1^l, \xi_2^l, \xi_3^l\})$. By Theorem 2.9 we can find a subscheme generic for an ergodic process $(X, \mathcal{B}, \mu, T, \Phi)$ where Φ includes $\varphi_1, \varphi_2, \varphi_3$ which are respectively represented by $\xi_1^l, \xi_2^l, \xi_3^l$. By the scholium to Theorem 2.9 we can assume φ_3 is not a.e. 0. Since $(\xi_i^{(l)})^2 = \xi_i^{(l)}$ we find $\varphi_i^2 = \varphi_i$ a.e. and so φ_i take values 0, 1. We write $\varphi_1 = 1_{\tilde{A}}$, $\varphi_2 = 1_{\tilde{B}}$, $\varphi_3 = 1_{\tilde{S}}$ with $\tilde{A}, \tilde{B}, \tilde{S} \subset X$, $\mu(\tilde{S}) > 0$, and $\tilde{A} \cup \tilde{B} = X$. Using Theorem 6.2 we can also assume that the process $(X, \mathcal{B}, \mu, T, \Phi)$ is Kronecker-complete and that the eigenfunctions $\{\varphi_\lambda\}$ of the process are represented by time series $\eta_\lambda(n) = \lambda^n$. We will also make use of the canonical map $\pi: X \to Z$, where (Z, α) is the Kronecker factor of $(X, \mathcal{B}, \mu, T)$.

We now apply Proposition 7.1 to this subscheme generic for the Kronecker-complete process with $\varphi_1, \varphi_2, \varphi_3 \in \Phi$, and where we again denote the array of intervals by $\{J_l\}$. We will take $f = \varphi = 1_{\tilde{S}} = \varphi_3$ which is represented by $\xi_3^l(n) = 1_S(n)$. We conclude that in the weak L^2-topology,

$$\frac{1}{|J_l|}\sum_{n\in J_l} 1_S(n)T^n 1_{\tilde{S}} \longrightarrow (f\Box f)\circ\pi \tag{17}$$

where $f = E(1_{\tilde{S}}|Z)$. The function f is bounded and non-negative with $\int f dz = \mu(\tilde{S}) > 0$ so it is non-trivial. We note that since $f \in L^\infty(Z)$, the function $F = f\Box f$ is continuous on Z.

We turn now to the hypothesis that $d^*\Big((S-s)\cap B\Big) = 0$ for $s \in S$. This implies that

$$\frac{1}{|J_l|}\sum_{n\in J_l} 1_B(n)1_S(n+s) \longrightarrow 0$$

or

$$\int 1_{\tilde{B}}T^s 1_{\tilde{S}} d\mu = 0.$$

In particular, averaging over $s \in S$:

$$\frac{1}{|J_l|}\sum_{n\in J_l} 1_S(n)\int T^n 1_{\tilde{S}} 1_{\tilde{B}}\, d\mu \longrightarrow 0. \tag{18}$$

But by (17), the limit in (18) is

$$\int F\circ\pi\cdot 1_{\tilde{B}}\, d\mu \tag{19}$$

and so the latter integral vanishes. We again apply the generic scheme where according to Corollary 6.4, $F\circ\pi$ is represented by $\{F(n\alpha)\}$, a non-negative almost periodic sequence with

$$\frac{1}{|J_l|}\sum_{n\in J_l} F(n\alpha) \longrightarrow \int F dz > 0$$

Since the integral in (19) vanishes we can write

$$\frac{1}{|J_l|}\sum_{n\in J_l} F(n\alpha)1_B(n) \longrightarrow 0$$

Let H be the Bohr_0 set for which $F(n\alpha) > \delta$ where $\delta > 0$ is chosen so that H is non-empty. Then

$$\frac{\sum_{n\in J_l} 1_H(n)1_B(n)}{\sum_{n\in J_l} 1_H(n)} \longrightarrow 0$$

whence

$$\frac{\sum_{n\in J_l} 1_{H\cap A}(n)}{\sum_{n\in J_l} 1_H(n)} \longrightarrow 1.$$

This implies that there are arbitrarily long intervals $L_l \subset J_l$ for which $H\cap L_l = H\cap A\cap L_l \subset A$. Hence $H\cap\bigcup L_l \subset A$ from which it follows that A is PW-Bohr_0. This proves Proposition 8.1. □

9 Application to Δ_r^*-sets

We shall apply the foregoing results to prove Theorem II of §1. We recall that a subset $A \subset \mathbb{Z}$ is a Δ_r^*-set, $r = 2, 3, \ldots$ if for distinct numbers $x_1, x_2, \ldots, x_r$, some difference $x_j - x_i$, $i < j$ belongs to A. More generally we will need

Definition 9.1. *If $S \subset \mathbb{Z}$ we shall write $A \in \Delta_r^*(S)$ if for $x_1, x_2, \ldots, x_r \in S$, $x_i \neq x_j$ for $i \neq j$, there exists $i < j$ with $x_j - x_i \in A$.*

In the sequel, A and B denote complementary sets in $\mathbb{Z}$, $B = \mathbb{Z} \backslash A$. If $0 \in B$ we denote by B' the set $B \backslash \{0\}$.

Lemma 9.2. *The following are equivalent for a set $S \subset \mathbb{Z}$:*

(a) $A \in \Delta_{r+1}^*(S)$

(b) $A \in \Delta_r^*\Big(B' \cap (S - s)\Big)$ *for every $s \in S$.*

Proof. (a) $\Rightarrow$ (b): Suppose $x_1, x_2, \ldots, x_r \in B' \cap (S - s)$. Form the $(r+1)$-tuple $s, s + x_1, s + x_2, \ldots, s + x_r$ and apply (a). (b) $\Rightarrow$ (a): Let $x_0, x_1, x_2, \ldots, x_r$ be distinct elements in S. If $\{x_1 - x_0,\ x_2 - x_0, \ldots, x_r - x_0\}$ doesn't meet A, then this is an r-tuple in $B' \cap (S - x_0)$ and we can apply (b). ☐

We recall Theorem II:

Theorem II. *For any $r = 2, 3, \ldots$, if A is a Δ_r^*-set then A is a PW-Bohr$_0$.*

Proof. We assume A is not PW-Bohr$_0$. By Proposition 8.1 this will imply that whenever $d^*(S) > 0$ there must be some $s \in S$ with $d^*\Big(B \cap (S - s)\Big) > 0$. This will give us an inductive procedure to obtain sets S_i with $d^*(S_i) > 0$. Start with $S_0 = \mathbb{Z}$ and we find $d^*(B) > 0$. Set $S_1 = B'$, there exists $s_1 \in S_1$ with $d^*(B \cap (S_1 - s_1)) > 0$. Set $S_2 = B' \cap (S_1 - s_1)$ and continue with $S_{k+1} = B' \cap (S_k - s_k)$, $s_k \in S_k$. Now apply the foregoing lemma. $A \in \Delta_r^* \Leftrightarrow A \in \Delta_r^*(\mathbb{Z}) \Rightarrow A \in \Delta_{r-1}^*\Big(B' \cap (\mathbb{Z} - s_0)\Big) = \Delta_{r-1}^*(S_1) \Rightarrow A \in \Delta_{r-2}^*\Big(B' \cap (S_1 - s_1)\Big) = \Delta_{r-2}^*(S_2) \Rightarrow \cdots$ We continue with $A \in \Delta_{r-k}^*(S_k)$ for $k = 0, 1, \ldots, r-2$. Finally $A \in \Delta_2^*(S_{r-2})$. At each stage we have $d^*(S_k) > 0$. But $d^*(S_{r-2}) > 0 \Rightarrow S_{r-2} - S_{r-2}$ is PW-Bohr$_0$; and $A \in \Delta_2^*(S_{r-2}) \Rightarrow A \supset S_{r-2} - S_{r-2}$. This proves the theorem. ☐

References

[Ber96] V. Bergelson, Ergodic Ramsey theory – an update in *Ergodic Theory of $\mathbb{Z}^d$-actions*, London Math. Soc. Lecture Notes Series 228 (1996), 1-61.

[BH96] V. Bergelson, I. J. Håland, Sets of recurrence and generalized polynomials, Convergence in Ergodic Theory and Probability (Columbus, OH, 1993) 91–110, *Ohio State University Math. Res. Inst. Publ.* 5, de Gruyter, Berlin, 1996.

[BFM96] V. Bergelson, H. Furstenberg, R. McCutcheon, IP-sets and polynomial recurrence, *Ergodic Theory and Dynamical Systems* **16** (1996) no. 5, 963–974.

[Bilu97] Y. Bilu, Addition of sets of integers of positive density, *Journal of Number Theory* **64** (1997) 233-275.

[B-M86] A. Bertrand-Mathis, Ensembles intersectifs et recurrence de Poincaré, *Israel J. Math.* **55** (1986), 184–198.

[Bog39] N. Bogoliouboff, Sur quelques propriétés arithmetiques des presque periodes, *Ann. Chaire Phys. Math. Kiev* **4** (1939), 195–205.

[Bro71] T. C. Brown, An interesting combinatorial method in the theory of locally finite semigroups, *Pacific Journal of Math.* **36** (1971), 285–289.

[Føl54] E. Følner, Generalization of a theorem of Bogoliouboff to topological abelian groups, *Mathematica Scandinavica* **2** (1954), 5–19.

[Fur67] H. Furstenberg, Disjointness in ergodic theory, minimal sets, and a problem in diophantine approximation, *Math. Systems Theory* **1** (1967), 1–49.

[Fur81] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton, Princeton Univ. Press, 1981.

[GRS80] R. L. Graham, B. L. Rothschild, J. H. Spencer, *Ramsey Theory*, John Wiley, New York (1980).

[Hel83] H. Helson, *Harmonic Analysis*, Addison-Wesley Reading, MA, (1983).

[Jin02] R. Jin, The sumset phenomenon, *Proc. Amer. Math. Soc.* **130** (2002), no. 3, 855–861.

[Kre85] U. Krengel, *Ergodic Theorems*, De Gruyter Studies in Mathematics, 6, Walter de Gruyter & Co., Berlin, (1985) viii+357.

[Kříž87] I. Kříž, Large independent sets in shift-invariant graphs: solution of Bergelson's problem, *Graphs Combin.* **3** (1987) no. 2, 145–158.

[Ruz82] I. Ruzsa, Uniform distribution, positive trigonometric polynomials and difference sets, *Seminar on Number Theory*, 1981/1982, Exp. No. 18, 18 pp. Univ. Bordeaux I, Talence, 1982.

[Sár78] A. Sárközy, On difference sets of sequences of integers, III, *Acta Math. Acad. Sci. Hungar.* **31** (1978) no. 3-4, 355–386.