VAN DER WAERDEN LARGENESS AND NILRECURRENCE

PRODUCED BY CHATGPT

Abstract. A set of positive integers is 2-vdW-large if every two-coloring contains arbitrarily long monochromatic arithmetic progressions with common difference in that set. We prove that every such set is a set of multiple pointwise recurrence for all nilrotations. This proves a conjecture of Alweiss and answers a question of Host, Kra, and Maass with two colors. The proof is an elementary argument with special generalized polynomials. For each prescribed finite family and each number of colors, we construct one coloring such that every sufficiently long progression omitting a color has its difference in a shrinking simultaneous approximation set. The error in an expression of weight w is $O(k^{-w})$, uniformly in its coefficients. This also yields nilrecurrence from a fixed-density analogue of van der Waerden largeness and an asymptotically lossless comparison of progression extremal numbers with and without the corresponding restrictions on the difference.

1. Introduction

1.1. **Restricted differences and the large sets conjecture.** Van der Waerden's theorem [11] asserts that every finite coloring of $\mathbb{N} = \{1, 2, \dots\}$ contains arbitrarily long monochromatic arithmetic progressions. Brown, Graham, and Landman [5] studied the sets of positive integers to which the common differences can be restricted without losing this conclusion. For an integer $r \geq 2$, call $S \subseteq \mathbb{N}$ *r-vdW-large* if, for every coloring $c : \mathbb{N} \to \{0, \dots, r-1\}$ and every $k \geq 2$, there are $a, d \in \mathbb{N}$ with $d \in S$ such that $\{a + jd : 0 \leq j < k\}$ is monochromatic. Call S *vdW-large* if it is r-vdW-large for every $r \geq 2$. These are the notions traditionally called r-large and large. We include the prefix vdW to distinguish them from the density notion introduced below.

Conjecture (Brown–Graham–Landman) ([5])**.** *Every 2-vdW-large set is vdW-large.*

Thus the conjecture asks whether an obstruction involving finitely many colors can always be replaced by a two-color obstruction. Equivalently, does recurrence of arbitrarily high order for binary colorings force multiple topological recurrence? The equivalence between the finite-color formulation and multiple topological recurrence is discussed in [7, Theorem 2.5 and Section 7].

An elementary source of obstructions is Diophantine approximation. If all $d \in S$ satisfy $\|\alpha d\|_{\mathbb{R}/\mathbb{Z}} \geq \eta$ for some real α and $\eta > 0$, a coloring induced by rotation of the circle prevents long monochromatic progressions with differences in S. Host, Kra, and Maass [7, Lemma 7.4] showed that two colors suffice for this obstruction. Farhangi and Grytczuk [6, Theorem 2] obtained an explicit bound in terms of the separation from the integers. Alweiss [1] subsequently proved that every 2-vdW-large set is Bohr recurrent, treating simultaneous linear conditions in arbitrary finite dimension. We extend this conclusion from rotations of tori to translations on nilmanifolds of every finite step.

A *nilsystem* is a compact nilmanifold $X = G/\Gamma$ equipped with a translation $T(x) = gx$, where G is a nilpotent Lie group, Γ is a discrete cocompact subgroup, and $g \in G$. We do not require G to be connected or simply connected, nor T to be minimal. A compatible metric on X is denoted by d_X.

Date: 20 September 2026.

2020 *Mathematics Subject Classification.* Primary 05D10; Secondary 37B20, 11J54, 11B25.

Key words and phrases. restricted arithmetic progressions, van der Waerden largeness, Szemerédi largeness, nilrecurrence, generalized polynomials.

Theorem 1.1 (Nilrecurrence from two colors)**.** *Let $S \subseteq \mathbb{N}$ be 2-vdW-large. For every nilsystem (X, T), every $x \in X$, every $\ell \geq 1$, and every $\eta > 0$, there are arbitrarily large $d \in S$ such that*

$$\max_{1 \leq j \leq \ell} d_X(T^{jd}x, x) < \eta. \tag{1.1}$$

In particular, S is a set of multiple pointwise recurrence for nilrotations.

Theorem 1.1 proves Conjecture 3.6 of [2]. It also answers Question 7.8 of Host, Kra, and Maass [7] with two colors: if (X, T) is a minimal nilsystem, U is a nonempty open set, and $x_0 \notin \overline{U}$, then

$$N(x_0, U) = \{n \in \mathbb{N} : T^n x_0 \in U\} \tag{1.2}$$

is not 2-vdW-large. Indeed, a return of x_0 sufficiently close to itself cannot belong to this hitting-time set; we give the formal deduction in Corollary 5.5.

The pointwise assertion is important. Bohr recurrence already implies ordinary open-set recurrence for minimal nilsystems [7, Theorem 4.1]. Theorem 1.1 instead prescribes the point that must return, and permits any finite number of its multiples to return simultaneously. It rules out every finite-step nilpotent return-time obstruction to the Brown–Graham–Landman conjecture. It does not prove the conjecture itself: Alweiss [2] constructed a set meeting every finite-step nil-Bohr neighborhood but failing double topological recurrence.

1.2. **Generalized polynomials and an elementary approach.** The classical progression theorems have polynomial extensions due to Bergelson and Leibman [3]. Here, however, the allowed difference set S is arbitrary; the hypothesis is its Ramsey property, rather than a prescribed polynomial description. To exploit that hypothesis, we work with nested nearest-integer expressions.

The relation between generalized polynomials and nilrotations is part of the foundational work of Bergelson and Leibman [4], who represented bounded generalized-polynomial mappings by piecewise polynomial observations of nilrotations. Leibman [9] developed a canonical-form and distribution theory for generalized polynomials. For the present problem, the exact description of nil-Bohr neighborhoods by Huang, Shao, and Ye [8] is decisive: it allows the nilrecurrence theorem to be deduced from simultaneous approximation for a special class of nested expressions.

We prove that approximation theorem directly. The argument uses finite van der Waerden, finite differences, interpolation, compactness, and upper-unitriangular matrix multiplication. It does not use nilmanifold equidistribution, a factorization theorem, or a structure theorem for polynomial nilorbits. The nilsystem interpretation is made only after the elementary argument is complete, using the characterization in [8].

1.3. **A density analogue and the common strengthening.** For $A \subseteq \mathbb{N}$, write

$$\overline{d}(A) = \limsup_{N \to \infty} \frac{|A \cap [1, N]|}{N}.$$

For $0 < \delta < 1$, we call $S \subseteq \mathbb{N}$ *δ-Szemerédi-large* if every $A \subseteq \mathbb{N}$ with $\overline{d}(A) = \delta$ contains arbitrarily long arithmetic progressions with common difference in S. This is the natural restricted-difference analogue of Szemerédi's theorem [10]. Requiring upper density at least δ is equivalent; see Lemma 5.1. For a fixed k, we use the terms *(r, k)-vdW-large* and *(δ, k)-Szemerédi-large* when only k-term progressions are required in the corresponding definitions.

The two largeness notions lead to the same nilrecurrence conclusion, by means of the following stronger coloring theorem. To state it, let $\langle x \rangle$ be a nearest integer to x, with ties resolved toward the smaller integer, and set $\|x\|_{\mathbb{R}/\mathbb{Z}} = \operatorname{dist}(x, \mathbb{Z})$. Thus

$$\langle x \rangle = \lceil x - \tfrac{1}{2} \rceil, \qquad x - \langle x \rangle \in (-\tfrac{1}{2}, \tfrac{1}{2}].$$

Define recursively

$$\mathsf{L}(x_1) = x_1, \qquad \mathsf{L}(x_1, \ldots, x_s) = x_1 \langle \mathsf{L}(x_2, \ldots, x_s) \rangle.$$

A *special generalized polynomial* is an expression

$$P(n) = \mathsf{L}(\alpha_1 n^{w_1}, \dots, \alpha_s n^{w_s}), \qquad \alpha_i \in \mathbb{R}, \quad w_i \in \mathbb{N}. \tag{1.3}$$

Its *assigned weight* is $\mathrm{wt}(P) = w_1 + \cdots + w_s$. For $s = 1$ this is simply a monomial, with no rounding. All exponents are positive; the assigned weight is retained even if the coefficients simplify the expression. This is the special class in [8, Definition 4.1.2]. For example, αn, βn^2, and $\gamma n\langle \theta n\rangle$ have weights $1, 2, 2$, respectively.

Theorem 1.2 (An omitted color forces approximation)**.** *For every $r \geq 2$ and $m, D \geq 1$, there are constants $C = C(r, m, D) > 0$ and $k_0 = k_0(r, m, D) \geq 2$ with the following property. For every family of special generalized polynomials $P_1, \dots, P_m$ of assigned weights $1 \leq d_i \leq D$, there is a coloring $c : \mathbb{N} \to \{0, \dots, r-1\}$ such that, for all $a, d \in \mathbb{N}$ and all $k \geq k_0$,*

$$\{c(a + jd) : 0 \leq j < k\} \neq \{0, \dots, r-1\} \quad \Longrightarrow \quad \|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \leq Ck^{-d_i} \quad (1 \leq i \leq m). \tag{1.4}$$

The coloring is fixed independently of k, and the constants are independent of all the real coefficients.

The quantifier is the existence of a suitable coloring for each family, not a conclusion for every coloring. In the binary case, omitting a color means being monochromatic. For a larger palette, the progression may use all but one of the colors. Equivalently, every sufficiently long progression whose difference violates a target in (1.4) must meet *every* color class. The error is measured against the progression length k; no upper bound for the size of d is asserted.

Deleting one color from a suitable larger palette gives a test set of upper density as close to one as desired. We obtain the following consequence.

Theorem 1.3 (Nilrecurrence from a fixed density)**.** *If $S \subseteq \mathbb{N}$ is δ-Szemerédi-large for some $\delta \in (0, 1)$, then it is a set of multiple pointwise recurrence for all nilrotations. More precisely, the conclusion of Theorem 1.1, including the existence of arbitrarily large return times, holds for S.*

For $0 < \delta \leq 1/2$, δ-Szemerédi-largeness implies 2-vdW-largeness: a color class of upper density at least $1/2$ can be thinned to upper density δ. For $\delta > 1/2$ this argument no longer applies. The omitted-color formulation treats the entire range $0 < \delta < 1$ without imposing a new dynamical hypothesis. Its proof gives more than Theorem 1.3: one high-density set controls the shrinking targets at every sufficiently large progression length (Theorem 5.2).

1.4. **An extremal consequence.** For a finite set $\Omega \subseteq \mathbb{N}$ and $B \subseteq \mathbb{N}$, define

$$r_k(\Omega) = \max\{|A| : A \subseteq \Omega \text{ contains no } k\text{-term AP}\}, \tag{1.5}$$

$$r_k(\Omega; B) = \max\{|A| : A \subseteq \Omega \text{ contains no } k\text{-term AP with difference in } B\}. \tag{1.6}$$

Here and throughout, every progression has positive common difference. The strengthened coloring theorem yields an almost-lossless extremal comparison, even when the allowed differences lie in shrinking targets.

Corollary 1.4. *Fix $\varepsilon, c > 0$ and $m, D \geq 1$. There is $k_0 = k_0(\varepsilon, c, m, D)$ such that, for every family $P_1, \dots, P_m$ of special generalized polynomials of weights $1 \leq d_i \leq D$, every $k \geq k_0$, and every finite $\Omega \subseteq \mathbb{N}$, the set*

$$B_k(c) = \{d \in \mathbb{N} : \|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \leq ck^{-d_i} \text{ for } 1 \leq i \leq m\} \tag{1.7}$$

satisfies

$$r_k(\Omega) \leq r_k(\Omega; B_k(c)) \leq (1 + \varepsilon) r_k(\Omega). \tag{1.8}$$

The threshold is independent of the real coefficients and of Ω.

The comparison uses the same progression length on both sides, and the target constant c is prescribed in advance. The cost of making the multiplicative factor close to one is a larger threshold in k. We prove this corollary in Section 6 and keep it separate from the proof of the qualitative nilrecurrence statements.

1.5. **Outline of the argument.** For each expression we use an upper-unitriangular band matrix whose first superdiagonal contains its monomial factors. We work in the finite product of these matrix groups and reduce matrices by right multiplication with integer matrices. The discontinuous canonical reduction is used only for an auxiliary grid. The primary coloring comes instead from the continuous function

$$F(A) = \min_{\gamma\in\Gamma} \sum_{i<j} (A\gamma)_{ij}^2.$$

A nonconstant ordinary polynomial matrix curve cannot keep F constant on an interval: locally, F is the minimum of finitely many nonconstant sum-of-squares polynomials. Compactness makes the resulting oscillation uniform for normalized excursions. Coloring successive thin shells cyclically forces every such excursion, when sampled finely enough, to meet all colors.

An auxiliary application of finite van der Waerden puts a short subprogression in one small canonical box. Integer finite differences then replace its lattice corrections by integer-valued polynomials. If the original progression omits a color, a first-exit argument keeps the polynomial lift bounded on a much longer real interval. Its leading coefficients are therefore small. Extracting the coefficient prescribed by each matrix weight removes the starting point of the progression; a fixed dilation removes the auxiliary step and clears factorial denominators. An exact last-column calculation then recovers the nested nearest integers, without amplifying errors by the real coefficients.

Sections 2–4 prove Theorem 1.2. Section 5 derives the density and nilrecurrence assertions, invoking Huang–Shao–Ye only for the final identification of return neighborhoods. Section 6 proves the extremal corollary. We end with the question whether δ-Szemerédi-largeness for one $\delta \in (0,1)$ already implies it for all such δ.

2. Weighted polynomial matrices and integer corrections

Fix the exponent list of each of the m expressions for the moment. For the sth list $(w_{s,1},\dots,w_{s,\ell_s})$, let $G_s = \mathrm{UT}_{\ell_s+1}(\mathbb{R})$, the group of upper-triangular matrices with ones on the diagonal, and set

$$G = \prod_{s=1}^{m} \mathrm{UT}_{\ell_s+1}(\mathbb{R}), \qquad \Gamma = \prod_{s=1}^{m} \mathrm{UT}_{\ell_s+1}(\mathbb{Z}), \qquad \nu = \sum_{s=1}^{m} \binom{\ell_s+1}{2}. \tag{2.1}$$

Thus ν is the total number of off-diagonal coordinates, not the number of expressions. All coordinate formulas below are understood blockwise, and sums and maxima include all the blocks. In one block assign position (i,j), $i<j$, the weight

$$w_{ij} = w_i + \cdots + w_{j-1}. \tag{2.2}$$

Every weight is positive and at most D, and $w_{ik}+w_{kj} = w_{ij}$. Write $|A-B|_{\mathrm{coord}} = \max_{i<j} |A_{ij}-B_{ij}|$ for the coordinate distance. This is not asserted to be an invariant group metric.

2.1. **The leading-coefficient map.** Let $\mathcal{H}$ be the class of polynomial matrix maps $A:\mathbb{R}\to G$ with $\deg A_{ij} \le w_{ij}$. Define, with diagonal entries always equal to one,

$$(\delta_u B)_{ij} = u^{w_{ij}} B_{ij} \quad (u>0), \qquad \tau(A)_{ij} = [t^{w_{ij}}]A_{ij}(t). \tag{2.3}$$

The symbol $[t^w]$ denotes coefficient extraction, not rounding. An entry of degree smaller than its assigned weight contributes zero.

Lemma 2.1. *The class $\mathcal{H}$ is closed under multiplication, inversion, and affine changes of parameter. The map δ_u is an automorphism of G, and*

$$\tau(AB) = \tau(A)\tau(B), \qquad \tau\big(A(b+qt)\big) = \delta_q \tau(A) \quad (q>0). \tag{2.4}$$

Moreover, $\tau(A) = I$ if A is constant.

Proof. For upper-unitriangular matrices,

$$(AB)_{ij} = A_{ij} + B_{ij} + \sum_{i<k<j} A_{ik}B_{kj}. \tag{2.5}$$

Additivity of the weights bounds the degree of every term by w_{ij}. The inverse of $I+V$, with V strictly upper triangular, is the finite sum $I - V + V^2 - \cdots$; every path of entries contributing to position (i,j) again has total weight w_{ij}. This proves the closure assertions. The same additive identity proves that δ_u respects multiplication; its inverse is $\delta_{1/u}$.

In a product $A_{ik}B_{kj}$, a term of degree $w_{ij} = w_{ik} + w_{kj}$ can arise only from the terms of the respective maximal allowed degrees. Taking this coefficient in (2.5) proves the first identity in (2.4). The second follows from the scalar identity

$$[t^w]f(b+qt) = q^w[t^w]f(t) \qquad (\deg f \le w).$$

Finally, constant entries contribute zero because their weights are positive. □

For example, for adjacent weights $2, 1$, the top-right weight is 3. In a product, its degree-three coefficient retains the cross term $([t^2]A_{12})([t]B_{23})$. Thus τ is a matrix homomorphism, not an approximation that discards the noncommutative terms. The scalar identity

$$[t^w](\alpha(a+dt)^w) = \alpha d^w \tag{2.6}$$

explains why this operation will eliminate the starting point a.

2.2. **Canonical representatives.** Put $I = (-1/2, 1/2]$. Every right coset $A\Gamma$ has a unique representative $\mathrm{rep}(A)$ whose ν off-diagonal entries lie in I. To construct it, seek $\lambda \in \Gamma$ with $A\lambda \in I^\nu$ in these coordinates and process entries in increasing order of the gap $j-i$. Once the shorter-gap entries are known, put

$$f_{ij} = A_{ij} + \sum_{i<k<j} A_{ik}\lambda_{kj}, \qquad \lambda_{ij} = -\langle f_{ij}\rangle. \tag{2.7}$$

Then $(A\lambda)_{ij} \in I$, and this is the unique integer choice with that property. Induction proves existence and uniqueness, and also

$$\mathrm{rep}(A\gamma) = \mathrm{rep}(A) \qquad (\gamma \in \Gamma). \tag{2.8}$$

This representative need not depend continuously on A.

2.3. **Finite differences and integer-valued polynomials.** Write $\Delta f(n) = f(n+1) - f(n)$. We use

$$\Delta^j f(n) = \sum_{v=0}^{j} (-1)^{j-v} \binom{j}{v} f(n+v). \tag{2.9}$$

A polynomial of degree at most w has vanishing $(w+1)$st difference. Conversely, a finite sequence of length at least $w+1$ whose $(w+1)$st differences vanish agrees with a polynomial of degree at most w: interpolate its first $w+1$ values, and use the difference equation to determine each remaining value. Its Newton interpolation formula is

$$f(t) = \sum_{v=0}^{w} \Delta^v f(0) \binom{t}{v}. \tag{2.10}$$

Consequently, a polynomial f of degree at most w that is integer-valued at $0, \dots, w$ is integer-valued on all of $\mathbb{Z}$, and

$$w!\,[t^w]f(t) = \Delta^w f(0) \in \mathbb{Z}. \tag{2.11}$$

Here $\binom{n}{v}$ is an integer even when $n < 0$, by $\binom{-u}{v} = (-1)^v\binom{u+v-1}{v}$ for $u \ge 1$. The polynomial $\binom{t}{2} = (t^2-t)/2$ shows why integer-valuedness does not mean that the monomial coefficients are integers.

Lemma 2.2 (Polynomial rounding)**.** *Let $M \geq D+2$, and partition I^ν into boxes whose coordinate intervals have length at most ξ, where $2^{D+1}\xi < 1$. If $A \in \mathcal{H}$ and $\operatorname{rep}(A(0)), \ldots, \operatorname{rep}(A(M-1))$ lie in one box, then there are $Q, C \in \mathcal{H}$ such that*

$$A(t)C(t) = Q(t) \quad (t \in \mathbb{R}), \qquad C(n) \in \Gamma \quad (n \in \mathbb{Z}), \tag{2.12}$$

with $Q(n) = \operatorname{rep}(A(n))$ for $0 \leq n < M$. Furthermore,

$$w_{ij}!\,\tau(C)_{ij} \in \mathbb{Z} \tag{2.13}$$

for every off-diagonal position.

Proof. On the initial samples write $A(n)\lambda(n) = r(n) = \operatorname{rep}(A(n))$, with $\lambda(n) \in \Gamma$. We construct C_{ij} and Q_{ij} in increasing order of the gap, maintaining $C_{ij}(n) = \lambda(n)_{ij}$ and $Q_{ij}(n) = r_{ij}(n)$ for $0 \leq n < M$.

For the current position (i, j) put $w = w_{ij}$ and

$$f_{ij}(t) = A_{ij}(t) + \sum_{i<k<j} A_{ik}(t)C_{kj}(t).$$

The entries in the sum have already been constructed. Their weighted degree bounds give $\deg f_{ij} \leq w$, while the induction hypothesis and (2.5) give

$$r_{ij}(n) - f_{ij}(n) = \lambda(n)_{ij} \in \mathbb{Z} \qquad (0 \leq n < M).$$

Thus $\Delta^{w+1} r_{ij}(n)$ is an integer wherever it is defined. All the $r_{ij}(n)$ lie in an interval of length at most ξ. Subtracting an endpoint of that interval does not change a positive-order difference. Hence (2.9) yields

$$|\Delta^{w+1} r_{ij}(n)| \leq 2^{w+1}\xi < 1.$$

These integer differences therefore vanish exactly.

The finite interpolation observation gives a polynomial Q_{ij} of degree at most w agreeing with all the $r_{ij}(n)$. Set $C_{ij} = Q_{ij} - f_{ij}$. It has degree at most w and is integer-valued at $0, \ldots, w$. Formula (2.10) makes it integer-valued at every integer and gives (2.13). The equation defining C_{ij} is precisely the (i, j) entry of $AC = Q$. Induction through all gaps and blocks completes the construction. □

The identity in (2.12) holds on the whole real line, and $C(n)$ belongs to Γ at every integer. Only the first M samples of Q, however, are known to be canonical. The next section controls the remaining samples by a different argument.

3. Shell colorings and bounded polynomial lifts

3.1. **A continuous size function.** Define

$$F(A) = \min_{\gamma \in \Gamma} \sum_{i<j} (A\gamma)_{ij}^2, \qquad A \in G. \tag{3.1}$$

The sum ranges over all off-diagonal entries in every block.

Lemma 3.1. *The minimum in (3.1) is attained. The function F is continuous, satisfies $0 \leq F(A) \leq \nu/4$, and is invariant under right multiplication by Γ. On every compact set of matrices, only finitely many elements of Γ can occur as minimizers.*

Proof. The canonical representative gives a candidate with squared-coordinate sum at most $\nu/4$. If A ranges over a fixed compact set and $\sum (A\gamma)_{ij}^2 \leq \nu/4$, then $A\gamma$ lies in a bounded coordinate box. The matrices $\gamma = A^{-1}(A\gamma)$ are therefore bounded as well. Only finitely many integer matrices have this property. Thus the minimum is attained, with a common finite set of candidates on each compact set. Locally F is consequently the minimum of finitely many continuous functions. Right-lattice invariance follows by reindexing the minimum, and the stated bounds have already been proved. □

Lemma 3.2 (Nonconstancy on polynomial curves)**.** *If $Q : \mathbb{R} \to G$ is a nonconstant map with ordinary polynomial entries, then $F(Q(t))$ is not constant on any nondegenerate real interval.*

Proof. It is enough to consider a compact nondegenerate interval. Its image admits only finitely many minimizing matrices $\gamma_1, \dots, \gamma_s$. For each of them,

$$f_a(t) = \sum_{i<j} (Q(t)\gamma_a)_{ij}^2$$

is a nonconstant polynomial. Indeed, $Q(t)\gamma_a$ is nonconstant because multiplication by γ_a is invertible. If $b \geq 1$ is the maximal degree of its coordinates, then the coefficient of t^{2b} in f_a is the sum of the squares of their degree-b coefficients and is strictly positive. If $F(Q(t))$ were a constant c, the finitely many equations $f_a(t) = c$ would cover the interval. Each has only finitely many solutions, which is impossible. □

The restriction to ordinary polynomial curves matters: a general continuous curve can remain in a level set of F. The rounding lemma is what will provide the required polynomial coordinates.

3.2. **Uniform oscillation and all-color sampling.** Set $\rho = 1/4$. Let $\mathcal{K}$ be the family of polynomial matrix curves $Q : [0,1] \to G$ whose coordinate degrees are at most D and satisfy

$$\max_{0\leq t\leq 1} |Q(t)|_{\mathrm{coord}} \leq 1, \qquad |Q(1) - Q(0)|_{\mathrm{coord}} = \rho. \tag{3.2}$$

Here $|Q(t)|_{\mathrm{coord}}$ means the maximum absolute off-diagonal entry. For this auxiliary family we need only the common degree bound D. The family is nonempty: set one coordinate equal to ρt and all others to zero.

It is compact in coefficient space. Interpolation at $0, 1/D, \dots, 1$ bounds all coefficients, and the conditions in (3.2) are closed. The function

$$Q \longmapsto \operatorname{osc}_{[0,1]} F(Q(t))$$

is continuous on this compact family. Every member is nonconstant, so Lemma 3.2 gives a constant $\eta > 0$, depending only on G, D, such that

$$\operatorname{osc}_{[0,1]} F(Q(t)) \geq \eta \qquad (Q \in \mathcal{K}). \tag{3.3}$$

Fix the desired number $r \geq 2$ of primary colors and put

$$h = \frac{\eta}{20r}, \qquad \kappa(A) = \left\lfloor \frac{F(A)}{h} \right\rfloor \pmod r, \tag{3.4}$$

with values in $\{0, \dots, r-1\}$. Consecutive shells of width h cycle through all colors. Invariance of F gives

$$\kappa(A\gamma) = \kappa(A) \qquad (\gamma \in \Gamma). \tag{3.5}$$

Lemma 3.3 (Uniform sampling)**.** *There exists $T_0 \geq 1$, depending only on G, D, r, such that for every $Q \in \mathcal{K}$ and every real $T \geq T_0$, the samples*

$$Q(j/T), \qquad 0 \leq j \leq \lfloor T \rfloor, \tag{3.6}$$

receive all r colors under κ.

Proof. The range of $F \circ Q$ is a closed interval $[a, b]$ of length at least $\eta = 20rh$. With $j_0 = \lceil a/h \rceil$, the r midpoints

$$z_s = (j_0 + s + \tfrac{1}{2})h, \qquad 0 \leq s < r,$$

all belong to $[a, b]$. Their bands have distinct colors, and each midpoint is distance $h/2$ from the corresponding band boundaries. Choose $u_s \in [0,1]$ with $F(Q(u_s)) = z_s$.

The map $(Q, t) \mapsto F(Q(t))$ is uniformly continuous on the compact set $\mathcal{K} \times [0,1]$. There is therefore $\theta > 0$, independent of Q, such that $|t - u| < \theta$ implies $|F(Q(t)) - F(Q(u))| < h/4$. Choose $T_0 > 1/\theta$ and approximate u_s by $\lfloor Tu_s \rfloor / T$. The corresponding sampled value remains

strictly inside the band centered at z_s. This samples every color, also when T is not an integer. □

3.3. **Interpolation and the auxiliary parameters.** We need two elementary polynomial estimates.

Lemma 3.4. (i) *For fixed D, T_0 and $\rho > 0$, there is an integer $M \ge D+2$ such that every real polynomial f of degree at most D with $|f(n)| \le 1/2$ for $0 \le n < M$ satisfies*

$$|f(t) - f(0)| < \rho \qquad (0 \le t \le T_0). \tag{3.7}$$

(ii) *If $\deg f \le w$, $w \ge 1$, and $|f(t)| \le B$ on $[0, J]$, where $J > 0$, then*

$$|[t^w]f(t)| \le b_w B J^{-w}, \qquad b_w = \frac{2^w w^w}{w!}. \tag{3.8}$$

Proof. For (i), let $\lambda_0, \dots, \lambda_D$ be the Lagrange basis at $0, \dots, D$, and choose B_D larger than $\frac{1}{2}\sum_{j=0}^{D} \sup_{[0,1]} |\lambda_j'|$. Choose an integer $H > \max\{1, T_0, B_D T_0/\rho\}$ and put $M = DH + 1$. Interpolation of $f(Hu)$ at $u = 0, \dots, D$ bounds its derivative on $[0,1]$ by B_D. For $0 \le t \le T_0$ it follows that $|f(t) - f(0)| \le B_D T_0/H < \rho$.

For (ii), put $s = J/w$. The identity $\Delta_s^w f(0) = w! s^w [t^w] f(t)$ and the bound $|\Delta_s^w f(0)| \le 2^w B$ give the result. The identity is also valid when $\deg f < w$, in which case both sides are zero. □

Choose M from Lemma 3.4(i) for the already fixed D, T_0, ρ. Set $\xi = 2^{-D-3}$ and partition the canonical box I^ν into $R = 2^{(D+3)\nu}$ half-open boxes of side length ξ. The label of the box containing $\mathrm{rep}(A)$ is an *auxiliary* R-coloring; it is distinct from the primary r-coloring κ. Finite van der Waerden [11] supplies $W = W(R, M)$ such that every R-coloring of $\{0, \dots, W-1\}$ has a monochromatic M-term progression. Finally set

$$L = D!\ \mathrm{lcm}(1, 2, \dots, W). \tag{3.9}$$

The order of choices is

$$(\text{weight patterns}, r) \longrightarrow G, \Gamma, D, \eta, h, \kappa, T_0 \longrightarrow M, \xi, R, W \longrightarrow L. \tag{3.10}$$

These objects are fixed before any coefficients or any progression are given. The auxiliary coloring only finds a short subprogression; it does not impose an R-vdW-largeness hypothesis on the allowed differences.

3.4. **The first-exit argument.**

Proposition 3.5. *Let $p \in \mathcal{H}$ and $k \ge 4W + 2$. If $\{\kappa(p(j)) : 0 \le j < k\}$ omits a color, there are $\gamma \in \Gamma$ and $E \in G$ such that*

$$\delta_L \tau(p)\, \gamma = E, \qquad |E_{ij}| \le \frac{3}{4} b_{w_{ij}} (2L)^{w_{ij}} k^{-w_{ij}}. \tag{3.11}$$

Proof. Let e be a color omitted by these samples. Apply the auxiliary grid coloring to $p(0), \dots, p(W-1)$. The choice of W gives $b \ge 0$ and $q \ge 1$ with

$$b + (M-1)q \le W - 1$$

such that the canonical representatives of $p(b), p(b+q), \dots, p(b+(M-1)q)$ lie in one grid box. In particular, $b, q \le W$. The rounding lemma, applied to $p(b+qt)$, gives

$$p(b+qt) C(t) = Q(t), \qquad C(\mathbb{Z}) \subseteq \Gamma, \tag{3.12}$$

where $Q, C \in \mathcal{H}$ and $Q(0), \dots, Q(M-1)$ are canonical. Put

$$J = \left\lfloor \frac{k - 1 - b}{q} \right\rfloor \ge M - 1.$$

For all integers $0 \le j \le J$, right-lattice invariance gives

$$\kappa(Q(j)) = \kappa(p(b+qj)) \ne e. \tag{3.13}$$

We claim that

$$|Q(t) - Q(0)|_{\mathrm{coord}} < \rho \qquad (0 \le t \le J). \tag{3.14}$$

The first M samples have all coordinates bounded by $1/2$. Lemma 3.4(i) therefore gives this displacement bound up to time T_0. If (3.14) failed, continuity would give a first real time $T \le J$ when the displacement equals ρ. Necessarily $T > T_0$. Up to that time the coordinates of Q are bounded by $1/2 + \rho = 3/4$, since $Q(0)$ is canonical. The rescaled curve $\widetilde{Q}(u) = Q(Tu)$ belongs to $\mathcal{K}$. Lemma 3.3, with mesh parameter T, says that

$$\widetilde{Q}(j/T) = Q(j), \qquad 0 \le j \le \lfloor T \rfloor,$$

receive all r colors. These samples occur in (3.13), a contradiction. This proves (3.14).

Every coordinate of Q is now bounded by $3/4$ on the entire real interval $[0, J]$. Its individual degree bound and Lemma 3.4(ii) give

$$|\tau(Q)_{ij}| \le \frac{3}{4} b_{w_{ij}} J^{-w_{ij}}. \tag{3.15}$$

Applying τ to (3.12) yields the exact identity

$$(\delta_q \tau(p))\tau(C) = \tau(Q). \tag{3.16}$$

The matrix $\tau(C)$ need not belong to Γ, even though $C(n)$ does for every integer n. We clear its factorial denominators at the same time as we remove the auxiliary step q. Apply $\delta_{L/q}$ to (3.16) and define

$$\gamma = \delta_{L/q}\tau(C), \qquad E = \delta_{L/q}\tau(Q).$$

By (3.9), L/q is an integer multiple of $D!$. Writing $\tau(C)_{ij} = z_{ij}/w_{ij}!$ with $z_{ij} \in \mathbb{Z}$, as permitted by (2.13), gives

$$\gamma_{ij} = \frac{(L/q)^{w_{ij}}}{w_{ij}!}\, z_{ij} \in \mathbb{Z}. \tag{3.17}$$

Thus $\gamma \in \Gamma$ and $\delta_L \tau(p)\gamma = E$. Moreover, (3.15) yields

$$|E_{ij}| \le \frac{3}{4} b_{w_{ij}} \left(\frac{L}{qJ}\right)^{w_{ij}}.$$

Finally, $qJ \ge k - 1 - b - q \ge k - 1 - 2W \ge k/2$, which proves (3.11). □

Only the short initial subprogression lies in one auxiliary grid box. Boundedness on the long real interval is a consequence of the primary omitted-color hypothesis. The use of the actual real first-exit time is essential: an exit need not occur at an integer.

4. Recovering the special generalized polynomials

For real $x_1, \dots, x_\ell$, let $B(x_1, \dots, x_\ell)$ be the upper-unitriangular matrix whose first superdiagonal is

$$-x_1, -x_2, \dots, -x_{\ell-1}, +x_\ell, \tag{4.1}$$

and whose higher off-diagonal entries are zero. For $\ell = 1$ its only off-diagonal entry is $+x_1$.

Lemma 4.1 (Last-column decoding)**.** *Suppose $B(x_1, \dots, x_\ell)\gamma = E$, where $\gamma \in \mathrm{UT}_{\ell+1}(\mathbb{Z})$ and $|E_{i,\ell+1}| < 1/2$ for $1 \le i \le \ell$. Then*

$$\mathsf{L}(x_1, \dots, x_\ell) + \gamma_{1,\ell+1} = E_{1,\ell+1}, \qquad \|\mathsf{L}(x_1, \dots, x_\ell)\|_{\mathbb{R}/\mathbb{Z}} = |E_{1,\ell+1}|. \tag{4.2}$$

Proof. Put $H_\ell = x_\ell$ and $H_i = x_i\langle H_{i+1}\rangle$ downward, so $H_1 = \mathsf{L}(x_1, \dots, x_\ell)$. The last-column equation in row ℓ is

$$E_{\ell,\ell+1} = x_\ell + \gamma_{\ell,\ell+1}.$$

Its absolute value is less than $1/2$, so $\gamma_{\ell,\ell+1} = -\langle H_\ell \rangle$ uniquely. For $i < \ell$, the band structure gives

$$E_{i,\ell+1} = \gamma_{i,\ell+1} - x_i \gamma_{i+1,\ell+1}.$$

If $\gamma_{i+1,\ell+1} = -\langle H_{i+1}\rangle$, this becomes $E_{i,\ell+1} = \gamma_{i,\ell+1} + H_i$. Smallness again identifies the integer correction exactly: $\gamma_{i,\ell+1} = -\langle H_i\rangle$. Induction proves the top-row identity. Since its residual lies in $(-1/2, 1/2)$, distance to the integers is its absolute value. □

For example, start with $P(n) = \alpha n^u \langle \beta n^v \rangle$ and fix the integer d at which it is to be evaluated. With $x_1 = \alpha d^u$ and $x_2 = \beta d^v$, we have $P(d) = x_1\langle x_2\rangle$ and

$$g_P(d) = \begin{pmatrix} 1 & -x_1 & 0 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

The equations $E_{23} = x_2 + \gamma_{23}$ and $E_{13} = \gamma_{13} - x_1\gamma_{23}$ first identify $\gamma_{23} = -\langle x_2\rangle$ and then give $E_{13} = \gamma_{13} + P(d)$. The integer is recovered exactly; no inner rounding error is multiplied by the possibly large coefficient x_1.

Proof of Theorem 1.2. First fix all the weight patterns and carry out the choices in (3.10). For one expression $P(n) = \mathsf{L}(\alpha_1 n^{w_1}, \dots, \alpha_\ell n^{w_\ell})$, set

$$g_P(t) = B(\alpha_1 t^{w_1}, \dots, \alpha_\ell t^{w_\ell}).$$

Let $g(t) \in G$ be the tuple of these matrices for all m expressions. Every nonzero off-diagonal coordinate of g is homogeneous of its assigned weight, and $g \in \mathcal{H}$. Define

$$c(n) = \kappa(g(n/L)), \qquad n \in \mathbb{N}. \tag{4.3}$$

The coefficients enter only through g; the coloring κ and integer L depend on the patterns and r alone.

Suppose $\{a + jd : 0 \le j < k\}$ omits a color, and put

$$p(t) = g((a + dt)/L).$$

The first k integer samples of this member of $\mathcal{H}$ omit that same color. Entrywise coefficient extraction, using (2.6), gives

$$\tau(p) = g(d/L), \qquad \delta_L\tau(p) = g(d). \tag{4.4}$$

Indeed, a superdiagonal entry $\pm\alpha_i((a+dt)/L)^{w_i}$ contributes $\pm\alpha_i(d/L)^{w_i}$, and all higher entries remain zero. This is weighted homogeneity, not an assertion that g is a one-parameter subgroup.

Proposition 3.5 therefore gives

$$g(d)\gamma = E, \qquad \gamma \in \Gamma, \qquad |E_{ij}| \le C_* k^{-w_{ij}}, \tag{4.5}$$

where

$$C_* = \max\left\{1, \max_{1\le w\le D} \frac{3}{4} b_w (2L)^w\right\}.$$

For $k \ge \max\{4W + 2, \lceil 4C_*\rceil + 1\}$ all off-diagonal errors are less than $1/4$. Apply Lemma 4.1 in each block. The top-right coordinate in the block for P_i has weight d_i; hence

$$\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} = |E_{\mathrm{top},i}| \le C_* k^{-d_i}.$$

Finally, for each $d \le D$ there are only finitely many ordered lists of positive integers with sum d. There are therefore finitely many m-tuples of admissible weight patterns. Taking the maxima of the constants and thresholds over these tuples gives $C(r, m, D)$ and $k_0(r, m, D)$. For a given coefficient family, use the coloring (4.3) for its pattern tuple. It was fixed before k was supplied and works for all k above the common threshold. □

Remark 4.2. Positive exponents make the total weight bound the nesting depth, and hence the dimension of the matrix model. Our coefficient-uniform statement is made for precisely this syntax. It is not an assertion about arbitrary expressions with ordinary floor functions or constant shifts. For example, when α is irrational, $\frac{1}{4}(\lfloor \alpha n\rfloor + \lfloor -\alpha n\rfloor)$ vanishes at zero but equals $-1/4$ at every positive integer. Merely requiring an arbitrary floor expression to vanish at zero is thus insufficient.

5. Density test sets and nilrecurrence

5.1. **Exact upper density.**

Lemma 5.1 (Thinning)**.** *If $A \subseteq \mathbb{N}$ has $\overline{d}(A) = \alpha > 0$, then for every $0 < \delta \leq \alpha$ it has a subset A' with $\overline{d}(A') = \delta$.*

Proof. Enumerate $A = \{a_1 < a_2 < \cdots\}$ and set $\lambda = \delta/\alpha$. Retain a_j exactly when $\lfloor \lambda j \rfloor > \lfloor \lambda(j-1) \rfloor$. Telescoping gives

$$|A' \cap [1, N]| = \lfloor \lambda |A \cap [1, N]| \rfloor.$$

After division by N, taking limsup proves the assertion. □

This proves the equivalence of "upper density exactly δ" and "upper density at least δ" in both Szemerédi-largeness definitions. We also use a finite-palette observation. For every $c : \mathbb{N} \to \{0, \ldots, r-1\}$, some color e satisfies

$$\overline{d}(\{n : c(n) \neq e\}) \geq 1 - 1/r. \tag{5.1}$$

For each N, some color occurs at most N/r times in $[1, N]$. One fixed color has this property along an unbounded subsequence of cutoffs, proving (5.1). No additivity of upper densities is used.

Theorem 5.2 (A fixed high-density test set)**.** *For every $0 < \delta < 1$ and $m, D \geq 1$, there are constants $C = C(\delta, m, D)$ and $k_0 = k_0(\delta, m, D)$ such that, for every family $P_1, \ldots, P_m$ of special generalized polynomials with $1 \leq d_i = \mathrm{wt}(P_i) \leq D$, there is $A \subseteq \mathbb{N}$ of upper density exactly δ satisfying*

$$\{a + jd : 0 \leq j < k\} \subseteq A, \quad k \geq k_0 \quad \Longrightarrow \quad \|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \leq Ck^{-d_i} \quad (1 \leq i \leq m). \tag{5.2}$$

One set A works for all these lengths, and the constants are independent of the coefficients.

Proof. Choose $r \geq 2$ with $1 - 1/r \geq \delta$, and take the coloring in Theorem 1.2. By (5.1), a complement $A_e = \{n : c(n) \neq e\}$ has upper density at least δ. Every progression in A_e omits e, so it satisfies all the estimates in (5.2). Thin A_e to upper density exactly δ using Lemma 5.1. This does not create any new progressions. All choices are made independently of k. □

Corollary 5.3 (Finite approximation)**.** *Fix $m, D \geq 1$.*

(i) *There are C, k_0, depending only on m, D, such that whenever $k \geq k_0$ and S is $(2, k)$-vdW-large, every family $P_1, \ldots, P_m$ of special generalized polynomials with $1 \leq d_i \leq D$ admits $d \in S$ with $\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \leq Ck^{-d_i}$ for all i.*

(ii) *For every $\delta \in (0, 1)$, the same assertion holds for (δ, k)-Szemerédi-large sets, with C, k_0 depending only on δ, m, D.*

Proof. For (i), apply $(2, k)$-vdW-largeness to the binary coloring of Theorem 1.2. For (ii), apply (δ, k)-Szemerédi-largeness to the test set in Theorem 5.2. □

In particular, either a 2-vdW-large set or a δ-Szemerédi-large set, with $0 < \delta < 1$, meets every set

$$\{n \in \mathbb{N} : \|P_i(n)\|_{\mathbb{R}/\mathbb{Z}} < \eta_i \text{ for } 1 \leq i \leq m\}, \qquad \eta_i > 0. \tag{5.3}$$

Choose k large enough that $Ck^{-d_i} < \eta_i$ for all i and use the appropriate part of Corollary 5.3.

5.2. **The nil-Bohr characterization.**

A *finite-step nil-Bohr$_0$ set* is a set $B \subseteq \mathbb{N}$ containing

$$\{n \in \mathbb{N} : T^n x \in U\} \tag{5.4}$$

for a finite-step nilsystem (X, T), a point $x \in X$, and an open neighborhood U of x. Products show that finite intersections of such sets are again of this form. The only nilmanifold result needed in the deductions is the following precise consequence of the characterization in [8].

Theorem 5.4 (Huang–Shao–Ye)**.** *Every finite-step nil-Bohr$_0$ set contains a set of the form* (5.3)*, with finitely many special generalized polynomials and positive tolerances.*

Indeed, Theorem B of [8] identifies the nil-Bohr$_0$ family with the family generated by neighborhoods of zero for their normalized generalized polynomials. Theorem 4.2.11 there replaces those polynomials by the special class (1.3). Restricting the family identities from $\mathbb{Z}$ to $\mathbb{N}$ gives the statement above. The convention allowing nonminimal nilsystems is covered by passage to point orbit closures, as explained in [8, Section 3.1.1]. The containment is an actual set containment, with no density-zero exceptional set.

Proof of Theorems 1.1 and 1.3. Let S satisfy either largeness hypothesis. By Corollary 5.3 and Theorem 5.4, it meets every finite-step nil-Bohr$_0$ set.

Given $(X,T), x, \ell, \eta$, consider the nilsystem on X^ℓ with translation $T \times T^2 \times \cdots \times T^\ell$ and basepoint $(x,\dots,x)$. Its return set to the product of the open η-balls about x is

$$B = \{n \in \mathbb{N} : \max_{1\le j\le \ell} d_X(T^{jn}x, x) < \eta\}.$$

This is a finite-step nil-Bohr$_0$ set, so $S \cap B \neq \varnothing$.

To obtain arbitrarily large returns, fix $M \ge 1$ and choose an integer $q > M$. The set $q\mathbb{N}$ is the return set of zero under the circle rotation $y \mapsto y + 1/q$ to the ball of radius $1/(2q)$ about zero. Hence $B \cap q\mathbb{N}$ is again a finite-step nil-Bohr$_0$ set. An element of $S \cap B \cap q\mathbb{N}$ is at least $q > M$. This proves both assertions in full. □

Corollary 5.5 (The hitting-time question)**.** *Let (X,T) be a minimal nilsystem, let $U \subseteq X$ be nonempty and open, and suppose $x_0 \notin \overline{U}$. Then $N(x_0,U)$ in* (1.2) *is not 2-vdW-large. It is also not δ-Szemerédi-large for any $0 < \delta < 1$.*

Proof. Compactness gives $\eta = \operatorname{dist}(x_0, \overline{U}) > 0$. Either asserted largeness property would provide $n \in N(x_0,U)$ with $d_X(T^n x_0, x_0) < \eta$, contradicting $T^n x_0 \in U$. □

5.3. **Localization.** For completeness, the two-color conclusion has the following useful consequence, with color cost independent of dimension and nilpotency step.

Corollary 5.6. *Let $r \ge 2$, let S be $2r$-vdW-large, and let B be a finite-step nil-Bohr$_0$ set. Then $S \cap B$ is r-vdW-large.*

Proof. The complement $\mathbb{N} \setminus B$ is not 2-vdW-large by Theorem 1.1. Thus some binary coloring c_0 and some length K have the property that every monochromatic K-term progression has difference in B. Given an r-coloring c_1 and a desired length k, apply $2r$-vdW-largeness to the product coloring (c_0, c_1) at length $\max\{K,k\}$. Its difference belongs to $S \cap B$, and its first k terms are monochromatic for c_1. □

6. Almost-lossless extremal transfer

We first record that the multiplicative constant in the target can be specified without changing its exponent or losing the single-coloring property.

Lemma 6.1 (Prescribing the target constant)**.** *In Theorem 1.2, the constant C on the right of* (1.4) *may be replaced by any prescribed $c > 0$, at the cost of replacing the threshold by $k_0(r,m,D,c)$. The coloring is still fixed for every k above this threshold.*

Proof. Let $C_0 \ge 1, k_1$ be constants from Theorem 1.2 for r colors, $2m$ expressions, and maximum weight D. Choose the fixed integer $q = \max\{1, \lceil C_0/c \rceil\}$. Apply the theorem to $P_1,\dots,P_m, qP_1,\dots,qP_m$; each qP_i is a special generalized polynomial of the same assigned weight, since only its outermost coefficient changes.

On a progression omitting a color, for $k \geq k_1$, both $\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}}$ and $\|qP_i(d)\|_{\mathbb{R}/\mathbb{Z}}$ are at most $C_0 k^{-d_i}$. Write $P_i(d) = z_i + u_i$ with $z_i \in \mathbb{Z}$ and $|u_i| = \|P_i(d)\|_{\mathbb{R}/\mathbb{Z}}$. If $k \geq 4qC_0$, then $|qu_i| \leq 1/4$, so there is no wrap around the circle and

$$q|u_i| = \|qP_i(d)\|_{\mathbb{R}/\mathbb{Z}} \leq C_0 k^{-d_i}.$$

Thus $\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \leq (C_0/q)k^{-d_i} \leq ck^{-d_i}$. Both q and the coloring were fixed independently of k. □

Proof of Corollary 1.4. Choose $r \geq 2$ such that $r/(r-1) \leq 1+\varepsilon$, and take the coloring from Lemma 6.1 for the prescribed c. Fix k above its threshold. Let $A \subseteq \Omega$ contain no k-term progression with difference in $B_k(c)$. Delete a color class of smallest cardinality in A, leaving a set A' with

$$|A'| \geq \left(1 - \frac{1}{r}\right)|A|.$$

Every progression in A' omits the deleted color, so a k-term progression there would have difference in $B_k(c)$, a contradiction. Thus A' is k-AP-free and

$$|A| \leq \frac{r}{r-1} r_k(\Omega) \leq (1+\varepsilon) r_k(\Omega).$$

Taking the maximum over A proves the upper bound. The lower bound holds because forbidding every k-term progression is a stronger restriction than forbidding just the ones with difference in $B_k(c)$. □

For instance, for every $\varepsilon > 0$ there is $k_0(\varepsilon)$ such that for all $k \geq k_0(\varepsilon)$ and all real $\alpha, \beta, \gamma, \theta$, every $A \subseteq \Omega$ with $|A| > (1+\varepsilon) r_k(\Omega)$ contains a k-term progression whose difference d satisfies

$$\|\alpha d\|_{\mathbb{R}/\mathbb{Z}} \leq k^{-1}, \qquad \|\beta d^2\|_{\mathbb{R}/\mathbb{Z}} \leq k^{-2}, \qquad \|\gamma d\langle \theta d \rangle\|_{\mathbb{R}/\mathbb{Z}} \leq k^{-2}.$$

No density or regularity assumption on the finite ambient set Ω is needed.

For a fixed finite-step nil-Bohr$_0$ set B, choose a contained special-polynomial neighborhood by Theorem 5.4. For sufficiently large k, its corresponding $B_k(1)$ is contained in B. Corollary 1.4 and monotonicity in the allowed difference set therefore give

$$\lim_{k\to\infty} \sup_{\substack{\Omega \subseteq \mathbb{N} \text{ finite} \\ |\Omega|>0}} \frac{r_k(\Omega; B)}{r_k(\Omega)} = 1. \tag{6.1}$$

The limit is in the progression length, not the size of the ambient set. An ε-dependent threshold is necessary even for $B = 2\mathbb{N}$: $r_k([k]) = k-1$ but $r_k([k]; 2\mathbb{N}) = k$. The shrinking-target conclusion in Corollary 1.4 is stronger than (6.1), since its neighborhood changes with k and its threshold remains uniform in all real coefficients.

7. Further questions

Theorem 1.3 is independent of which fixed density $\delta \in (0,1)$ is assumed. This suggests asking whether the original combinatorial property is itself independent of that density.

Question 7.1. If $S \subseteq \mathbb{N}$ is δ-Szemerédi-large for some $\delta \in (0,1)$, must it be η-Szemerédi-large for every $\eta \in (0,1)$?

The implication for $\eta \geq \delta$ follows from Lemma 5.1; the issue is passage to smaller positive densities. An affirmative answer would imply that every set satisfying one such density hypothesis is vdW-large: in an r-coloring, some color class has upper density at least $1/r$, and one applies $(1/r)$-Szemerédi-largeness. Our theorem establishes the common nilrecurrence consequence but not this density amplification.

The endpoint $\delta = 1$ cannot be included. Every set of upper density one contains intervals of every finite length: otherwise, for some k, each disjoint block of k integers would contain

a missing point and the upper density would be at most $1 - 1/k$. Consequently $\{1\}$ is 1-Szemerédi-large, although it is not recurrent for an irrational circle rotation.

The Brown–Graham–Landman conjecture also remains beyond the present argument. Theorem 1.1 excludes all finite-step nilpotent obstructions, but the examples of [2] show why this class of obstructions is not exhaustive. A natural next question is whether 2-vdW-largeness forces multiple recurrence for a larger class of distal systems. The key restriction in our proof is that the lifted coordinates are ordinary polynomials of bounded degree; the nonconstancy and compactness arguments in Section 3 use this fact directly.

Finally, the constants in Theorem 1.2 involve the oscillation minimum over a compact family and a finite van der Waerden number. Although the error powers are explicit, we do not give numerical bounds for $C(r, m, D)$ or $k_0(r, m, D)$. Effective estimates for these constants, even for a small fixed weight, would clarify the finite strength of the result. Lemma 6.1 also illustrates the difference between controlling the target constant and controlling the threshold at which that target becomes available.

Acknowledgements. ChatGPT and Codex were used in the development of intermediate arguments and in the preparation of the manuscript, including assistance with exposition and literature checks.

References

1. Ryan Alweiss, 2-*large sets are sets of Bohr recurrence*, Preprint, arXiv:2512.01997v3, 2025, Version of 24 December 2025.
2. _____, *New obstacles to multiple recurrence*, Preprint, arXiv:2511.21680v3, 2025, Version of 24 December 2025.
3. Vitaly Bergelson and Alexander Leibman, *Polynomial extensions of van der Waerden's and Szemerédi's theorems*, J. Amer. Math. Soc. **9** (1996), no. 3, 725–753.
4. _____, *Distribution of values of bounded generalized polynomials*, Acta Math. **198** (2007), no. 2, 155–230.
5. Tom C. Brown, Ronald L. Graham, and Bruce M. Landman, *On the set of common differences in van der Waerden's theorem on arithmetic progressions*, Canad. Math. Bull. **42** (1999), no. 1, 25–36.
6. Sohail Farhangi and Jarosław Grytczuk, *Distance graphs and arithmetic progressions*, Integers **21A** (2021), Paper No. A11, 6 pp.
7. Bernard Host, Bryna Kra, and Alejandro Maass, *Variations on topological recurrence*, Monatsh. Math. **179** (2016), no. 1, 57–89.
8. Wen Huang, Song Shao, and Xiangdong Ye, *Nil Bohr-sets and almost automorphy of higher order*, Mem. Amer. Math. Soc. **241** (2016), no. 1143, v+83.
9. Alexander Leibman, *A canonical form and the distribution of values of generalized polynomials*, Israel J. Math. **188** (2012), 131–176.
10. Endre Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 199–245.
11. Bartel L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wiskd. (2) **15** (1927), 212–216.