Fixed-density largeness implies nilrecurrence

An omitted-color strengthening of the elementary shell argument
Produced by ChatGPT on September 15, 2026

Abstract

Fix $0 < \delta < 1$. Suppose that $D \subseteq \mathbb{N}$ has the following property: every subset of $\mathbb{N}$ with natural upper density exactly δ contains arbitrarily long arithmetic progressions whose common difference belongs to D. We prove that D is a set of multiple pointwise recurrence for every nilsystem. The proof establishes a stronger quantitative obstruction. For a finite family of special generalized polynomials, and any prescribed number $r \geq 2$ of colors, one coloring forces the difference of every sufficiently long progression that omits even one color to approximate all the polynomials, at the reciprocal power of the progression length prescribed by each weight. The constants are uniform in the real coefficients, and one coloring works for all lengths. Removing a suitably chosen color and thinning its complement produces a test set of upper density exactly δ. The elementary matrix, rounding, interpolation, and first-exit arguments are proved in detail. The passage to nilsystems uses the explicitly stated characterization of nil-Bohr neighborhoods by Huang, Shao, and Ye.

Contents

1 Statements, conventions, and the logical structure **2**

2 Two elementary facts about upper density **4**

3 Polynomial tools, with quantitative details **5**

4 Weighted polynomial matrices and coefficient extraction **7**

5 Canonical lattice representatives and the shell function **8**

6 A shell coloring that forces every color to appear **10**

7 The fixed auxiliary parameters **11**

8 Polynomial rounding inside a single small box **12**

9 An omitted color prevents escape of the polynomial lift **13**

10 Recovering nested nearest-integer expressions **16**

11 From omitted colors to the prescribed upper density **18**

12 Nil-Bohr neighborhoods and the nilrecurrence conclusion **19**

13 Scope, the endpoint, and checks on the quantifiers **20**

Provenance and references **22**

1 Statements, conventions, and the logical structure

Throughout, $\mathbb{N} = \{1, 2, \ldots\}$. The natural upper density of $A \subseteq \mathbb{N}$ is

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

No assertion that the corresponding limit exists is implicit in this notation. All arithmetic progressions have positive integer starting point and positive integer common difference. A k-term progression is a set $\{a + jd : 0 \le j < k\}$, with $a, d \in \mathbb{N}$.

Definition 1.1 (Fixed-density largeness)**.** Let $0 < \delta < 1$ and $k \in \mathbb{N}$. A set $D \subseteq \mathbb{N}$ is (δ, k)-*density large* if every $A \subseteq \mathbb{N}$ with $\overline{d}(A) = \delta$ contains a k-term progression with common difference in D. It is δ-*density large* if it is (δ, k)-density large for every $k \in \mathbb{N}$.

The last condition is equivalent to saying that every such A contains arbitrarily long progressions with differences in D: taking the first k terms of a longer progression preserves its difference.

A *nilsystem* is a compact nilmanifold $X = G/\Gamma$, where G is a nilpotent Lie group and Γ is a discrete cocompact subgroup, equipped with a translation $T(y) = gy$ for some $g \in G$. We do not assume that G is connected or simply connected, or that the translation is minimal. A compatible metric on X is denoted by d_X.

Theorem 1.2 (Fixed-density largeness implies nilrecurrence)**.** *Suppose that $D \subseteq \mathbb{N}$ is δ-density large for some $\delta \in (0, 1)$. For every nilsystem (X, T), every $x \in X$, every $\ell \in \mathbb{N}$, and every $\varepsilon > 0$, there is $d \in D$ such that*

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

In fact, the admissible d in (1.1) are unbounded.

Thus the conclusion is *multiple pointwise* recurrence, not merely recurrence of some point in a given nonempty open set.

1.1 Special generalized polynomials

For real u, let $\langle u \rangle$ be a nearest integer, with ties resolved toward the smaller integer. Equivalently, using the ordinary ceiling function,

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

Write $\|u\|_{\mathbb{R}/\mathbb{Z}} = \operatorname{dist}(u, \mathbb{Z})$. The recurrence

$$\mathsf{L}(u) = u, \qquad \mathsf{L}(u_1, \ldots, u_s) = u_1 \, \langle \mathsf{L}(u_2, \ldots, u_s) \rangle$$

defines the nested expressions used below.

Definition 1.3 (Special expressions and their weights)**.** A *special generalized polynomial* is an expression

$$P(n) = \mathsf{L}(\alpha_1 n^{w_1}, \ldots, \alpha_s n^{w_s}), \qquad \alpha_i \in \mathbb{R}, \quad w_i \in \mathbb{N}. \tag{1.2}$$

Its *assigned weight* is $w_1 + \cdots + w_s$. All exponents are positive. For $s = 1$, the expression is simply $\alpha_1 n^{w_1}$, with no nearest-integer operation. We retain the assigned weight even when special choices of coefficients simplify the resulting function.

This is the convention of the supplied drafts [1, 2] and of SGP_{d_*} in [3, Definition 4.1.2]. The quantitative assertions below are about this syntax, not about arbitrary expressions built with floor functions and additive constant shifts.

Theorem 1.4 (A fixed high-density quantitative test set)**.** *For every $\delta \in (0,1)$ and $m, d_* \in \mathbb{N}$, there are constants $C = C(\delta, m, d_*) > 0$ and $k_0 = k_0(\delta, m, d_*) \in \mathbb{N}$ with the following property. For every family of special generalized polynomials $P_1, \dots, P_m$ of assigned weights $1 \le d_i \le d_*$, there is a set $A \subseteq \mathbb{N}$ with $\overline{d}(A) = \delta$ such that, for all $k \ge k_0$,*

$$\{a + jd : 0 \le j < k\} \subseteq A \quad \Longrightarrow \quad \|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \le Ck^{-d_i} \quad (1 \le i \le m). \tag{1.3}$$

The set A is chosen once for the entire family of progression lengths. The constants do not depend on any of the real coefficients.

Corollary 1.5 (Finite quantitative approximation)**.** *With C, k_0 as in Theorem 1.4, let $k \ge k_0$ and let D be (δ, k)-density large. Then, for every family as above, some $d \in D$ satisfies*

$$\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \le Ck^{-d_i} \qquad (1 \le i \le m). \tag{1.4}$$

The principal new combinatorial assertion is the following strengthening of the two-color shell theorem in [2].

Theorem 1.6 (Omitted-color theorem)**.** *Fix $r \ge 2$ and $m, d_* \in \mathbb{N}$. There are constants $C = C(r, m, d_*) > 0$ and $k_0 = k_0(r, m, d_*) \in \mathbb{N}$ such that, for every family $P_1, \dots, P_m$ as in Theorem 1.4, there is a coloring*

$$c : \mathbb{N} \longrightarrow \{0, 1, \dots, r-1\}$$

with the following property: if $k \ge k_0$ and

$$\{c(a + jd) : 0 \le j < k\} \ne \{0, 1, \dots, r-1\},$$

then

$$\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \le Ck^{-d_i} \qquad (1 \le i \le m). \tag{1.5}$$

The coloring does not depend on k, and the constants are independent of the real coefficients of the P_i.

For $r = 2$, omitting a color is exactly being monochromatic. For $r > 2$, a progression may use $r - 1$ colors and still fall under the theorem. Its contrapositive says that a sufficiently long progression whose difference violates one of the required bounds must contain *every* color. It is not enough, for the application to density near one, merely to have an r-coloring with no monochromatic bad progression.

1.2 Proof architecture and quoted inputs

The elementary construction follows the weighted-matrix and shell architecture of [2]; it is reproduced here rather than cited as a black box. First, a finite-dimensional compactness argument produces a shell coloring that samples all r colors on every normalized nonconstant polynomial excursion. Second, an auxiliary grid and finite differences give an ordinary polynomial lift of a short part of a matrix orbit. A first-exit argument then keeps this lift bounded whenever the orbit omits a color. Finally, weighted leading coefficients eliminate the starting point, and a fixed dilation clears the denominators of the integer-valued lattice corrections.

The density deduction uses complements of color classes and exact thinning. The last step, from special expressions to nilsystems, uses the established nil-Bohr characterization stated in Theorem 12.2. The two external inputs quoted without proof are finite van der Waerden and that characterization. All other ingredients used in the proof are established below. In particular, the proof does not assume an equidistribution or factorization theorem, nor does it invoke the original 2-largeness result to prove its density extension.

2 Two elementary facts about upper density

These facts are independent of the matrix argument. They explain why the omitted-color statement is the right strengthening.

Lemma 2.1 (Exact thinning)**.** *If $A \subseteq \mathbb{N}$ has upper density $\alpha > 0$, then for every $0 < \delta \leq \alpha$ there is $A' \subseteq A$ with $\overline{d}(A') = \delta$. More precisely, one may arrange*

$$|A' \cap \{1, \dots, N\}| = \left\lfloor \frac{\delta}{\alpha} |A \cap \{1, \dots, N\}| \right\rfloor \qquad (N \in \mathbb{N}). \tag{2.1}$$

Proof. Enumerate $A = \{a_1 < a_2 < \cdots\}$ and put $\lambda = \delta/\alpha \in (0, 1]$. Retain a_j precisely when

$$\lfloor \lambda j \rfloor - \lfloor \lambda (j-1) \rfloor = 1.$$

Each difference is either zero or one. If A has M elements up to N, the number retained up to N is the telescoping sum

$$\sum_{j=1}^{M} (\lfloor \lambda j \rfloor - \lfloor \lambda (j-1) \rfloor) = \lfloor \lambda M \rfloor.$$

This proves (2.1). Dividing by N gives an error of absolute value at most $1/N$ from $\lambda |A \cap [1, N]|/N$, so taking limsup yields $\overline{d}(A') = \lambda \alpha = \delta$. □

Consequently the definition of (δ, k)-density largeness is unchanged if one replaces "upper density exactly δ" by "upper density at least δ." The nontrivial direction is obtained by thinning, since all progressions in a subset are still progressions in the original set.

Lemma 2.2 (A large complement of one color)**.** *If $c : \mathbb{N} \to \{0, \dots, r-1\}$, then some $e \in \{0, \dots, r-1\}$ satisfies*

$$\overline{d}(\{n : c(n) \neq e\}) \geq 1 - \frac{1}{r}. \tag{2.2}$$

Proof. For every N, the counts of the r colors in $\{1, \dots, N\}$ sum to N. At least one color therefore occurs at most N/r times. Some fixed color e has this property for infinitely many, hence unbounded, values of N. Along that subsequence its complement has proportion at least $1 - 1/r$. Taking limsup proves the assertion. □

Remark 2.3 (No subtraction of upper densities)*.* It would be incorrect to justify this by asserting that the upper densities of the color classes add to one, or that the upper density of a complement is one minus the upper density of the removed class. Neither identity is available in general. The finite counting and subsequence argument above is what is used.

For a given $\delta < 1$, choose

$$r = \max\left\{2, \left\lceil \frac{1}{1-\delta} \right\rceil \right\}. \tag{2.3}$$

Then $1 - 1/r \geq \delta$. Thus an omitted-color obstruction yields a test set of upper density at least δ, and Lemma 2.1 reduces it to exactly δ without creating any new progressions. For instance, density 0.99 is handled by 100 colors and by removing one suitable color, not by seeking a single color of density 0.99.

3 Polynomial tools, with quantitative details

3.1 Finite differences and integer-valued polynomials

For a function on the integers, put $\Delta f(n) = f(n+1) - f(n)$. Induction gives

$$\Delta^j f(n) = \sum_{a=0}^{j} (-1)^{j-a} \binom{j}{a} f(n+a). \tag{3.1}$$

For an ordinary polynomial f of degree at most w, one has $\Delta^{w+1} f = 0$.

Lemma 3.1 (Finite interpolation and Newton coefficients)**.** *Let $w \geq 0$ and $M \geq w+2$.*

(i) *If real numbers $r(0), \ldots, r(M-1)$ satisfy $\Delta^{w+1} r(n) = 0$ whenever $0 \leq n \leq M-w-2$, then they agree with a unique ordinary polynomial $Q(t)$ of degree at most w.*

(ii) *If an ordinary polynomial $C(t)$ has degree at most w and is integer-valued at $0, 1, \ldots, w$, then it is integer-valued on all of $\mathbb{Z}$ and*

$$w!\,[t^w]C(t) \in \mathbb{Z}. \tag{3.2}$$

Proof. For (i), interpolate the first $w+1$ values. The relation $\Delta^{w+1} r(n) = 0$, expanded by (3.1), determines $r(n+w+1)$ uniquely from the preceding $w+1$ values. The interpolating polynomial satisfies the same relation. Induction extends agreement to all M samples.

For (ii), the Newton identity for a polynomial of degree at most w is

$$C(t) = \sum_{j=0}^{w} \Delta^j C(0) \binom{t}{j}, \qquad \binom{t}{j} = \frac{t(t-1)\cdots(t-j+1)}{j!}. \tag{3.3}$$

Each coefficient $\Delta^j C(0)$ is an integer by (3.1). The polynomial $\binom{t}{j}$ takes integer values at every integer: for nonnegative integers this is standard, and for $u \geq 1$ one has

$$\binom{-u}{j} = (-1)^j \binom{u+j-1}{j}.$$

Thus $C(\mathbb{Z}) \subseteq \mathbb{Z}$. Since only the $j = w$ summand can contribute to t^w, its coefficient is $\Delta^w C(0)/w!$, proving (3.2). □

For example, $C(t) = \binom{t}{2} = (t^2 - t)/2$ is integer-valued on $\mathbb{Z}$, but its coefficient of t^2 is $1/2$. This is why the later proof must clear factorial denominators; integer-valued does not mean that the ordinary monomial coefficients are integers.

3.2 Small initial motion from many bounded samples

Lemma 3.2 (An explicit interpolation length)**.** *Fix $d_* \geq 1$, $T_0 \geq 1$, and $\rho > 0$. There is an integer $M \geq d_* + 2$, depending only on these data, such that every ordinary polynomial f of degree at most d_* satisfying*

$$|f(n)| \leq \tfrac{1}{2} \qquad (0 \leq n < M)$$

also satisfies

$$|f(t) - f(0)| < \rho \qquad (0 \leq t \leq T_0). \tag{3.4}$$

The same choice works simultaneously for any finite list of such polynomials.

Proof. Let

$$\lambda_j(u) = \prod_{\substack{0 \leq i \leq d_* \\ i \neq j}} \frac{u - i}{j - i} \qquad (0 \leq j \leq d_*)$$

be the Lagrange basis at $0, 1, \ldots, d_*$. Set

$$B_{d_*} = 1 + \frac{1}{2} \sum_{j=0}^{d_*} \sup_{0 \leq u \leq 1} |\lambda_j'(u)|.$$

Choose an integer

$$H > \max\{1, T_0, B_{d_*} T_0/\rho\}, \qquad M = d_* H + 1.$$

For $g(u) = f(Hu)$, interpolation gives

$$g(u) = \sum_{j=0}^{d_*} f(jH) \lambda_j(u).$$

The indices jH lie in $\{0, \ldots, M-1\}$. Thus $|g'(u)| \leq B_{d_*}$ for $0 \leq u \leq 1$. If $0 \leq t \leq T_0$, then $0 \leq t/H \leq 1$, and the mean value estimate yields

$$|f(t) - f(0)| = |g(t/H) - g(0)| \leq B_{d_*} T_0/H < \rho.$$

The estimate is coordinatewise and has no dependence on how many coordinates are being considered. □

3.3 The highest allowed coefficient of a bounded polynomial

Lemma 3.3 (Coefficient bound on a long real interval)**.** *Let $w \geq 1$, $J > 0$, and $f \in \mathbb{R}[t]$ have degree at most w. If $|f(t)| \leq B$ for $0 \leq t \leq J$, then*

$$|[t^w] f(t)| \leq b_w B J^{-w}, \qquad b_w = \frac{2^w w^w}{w!}. \tag{3.5}$$

The coefficient is interpreted as zero if the actual degree is below w.

Proof. For the difference operator with step h, one has $\Delta_h^w f(0) = w! h^w [t^w] f(t)$. This follows first for monomials, since a difference lowers their degree by one, and then by linearity. Take $h = J/w$. All the samples $0, h, \ldots, wh$ lie in $[0, J]$, and (3.1) gives

$$|\Delta_h^w f(0)| \leq B \sum_{j=0}^{w} \binom{w}{j} = 2^w B.$$

Division by $w!(J/w)^w$ proves the claim. □

4 Weighted polynomial matrices and coefficient extraction

Fix, for now, the weight pattern of each of $P_1, \dots, P_m$. For a single pattern $(w_1, \dots, w_s)$, use the group $\mathrm{UT}_{s+1}(\mathbb{R})$ of upper triangular matrices with all diagonal entries equal to one. Assign to entry (i, j), for $i < j$, the positive integer weight

$$w_{ij} = w_i + w_{i+1} + \cdots + w_{j-1}. \tag{4.1}$$

In particular,

$$w_{ik} + w_{kj} = w_{ij} \qquad (i < k < j). \tag{4.2}$$

Take the direct product G of these groups, one block for each expression, and let Γ be the product of their integer subgroups. Let ν be the total number of off-diagonal coordinates. Every weight is at most d_*, the maximum allowed weight of an expression. All statements about entries, sums of entries, and coordinate bounds are understood blockwise.

For $A, B \in G$, write

$$|A|_{\mathrm{coord}} = \max_{i<j} |A_{ij}|, \qquad |A - B|_{\mathrm{coord}} = \max_{i<j} |A_{ij} - B_{ij}|.$$

These are ordinary coordinate quantities, not invariant group distances. Diagonal entries are not included.

Definition 4.1 (The polynomial class)**.** Let $\mathcal{H}$ be the set of maps $A : \mathbb{R} \to G$ such that every $A_{ij}(t)$ is an ordinary real polynomial of degree at most w_{ij}. The diagonal entries of such a map are the constant polynomial one.

Lemma 4.2 (Algebra of the polynomial class)**.** *The class $\mathcal{H}$ is closed under multiplication, inversion, and real affine changes of parameter. For $u > 0$, define*

$$(\mathrm{dil}_u A)_{ij} = u^{w_{ij}} A_{ij} \quad (i < j),$$

with diagonal entries one. Then dil_u is a group automorphism of G. If

$$\tau(A)_{ij} = [t^{w_{ij}}] A_{ij}(t), \qquad \tau(A)_{ii} = 1, \tag{4.3}$$

then

$$\tau(AB) = \tau(A)\tau(B), \tag{4.4}$$

$$\tau\big(A(b + qt)\big) = \mathrm{dil}_q \tau(A) \quad (q > 0), \tag{4.5}$$

$$\mathrm{dil}_u(BC) = (\mathrm{dil}_u B)(\mathrm{dil}_u C). \tag{4.6}$$

For a constant matrix curve $A(t) \equiv A_0$, one has $\tau(A) = \mathrm{I}$.

Proof. Matrix multiplication gives

$$(AB)_{ij} = A_{ij} + B_{ij} + \sum_{i<k<j} A_{ik} B_{kj}. \tag{4.7}$$

The degree of each product in the sum is at most $w_{ik} + w_{kj} = w_{ij}$. This proves closure under multiplication. For inversion, write each block as $\mathrm{I} + N(t)$, with $N(t)$ strictly upper triangular. Its inverse is the finite sum

$$(\mathrm{I} + N(t))^{-1} = \mathrm{I} - N(t) + N(t)^2 - \cdots + (-1)^s N(t)^s.$$

Every nonzero summand contributing to entry (i, j) corresponds to a strictly increasing path from i to j; the weights along that path sum to w_{ij}. The degree bound follows. Closure under affine changes of parameter is immediate.

To obtain (4.4), take the coefficient of $t^{w_{ij}}$ in (4.7). In a product $A_{ik}(t)B_{kj}(t)$, a term of degree $w_{ij} = w_{ik} + w_{kj}$ can arise only by multiplying the coefficient of degree w_{ik} in the first factor and the coefficient of degree w_{kj} in the second. This is exactly the product law for $\tau(A)\tau(B)$.

If $f(t) = \sum_{a=0}^{w} c_a t^a$, then

$$[t^w]f(b + qt) = q^w c_w.$$

Indeed, terms of degree below w contribute nothing, and $[t^w](b+qt)^w = q^w$. Applying this entrywise proves (4.5). Equation (4.6) follows directly from the same additive weights in (4.7); the inverse automorphism is $\mathrm{dil}_{1/u}$. Finally, all weights are positive, so a constant curve has zero in every extracted off-diagonal position. □

Example 4.3 (What coefficient extraction does and does not mean)**.** The notation $[t^w]f(t)$ means the coefficient of the specified power:

$$[t^2](7 + 3t - \tfrac{5}{2}t^2) = -\tfrac{5}{2}, \qquad [t^3](7 + 3t - \tfrac{5}{2}t^2) = 0.$$

It does not mean nearest-integer rounding. For adjacent weights $1, 2$, the three positions in a 3×3 block have weights $1, 2, 3$:

$$A(t) = \begin{pmatrix} 1 & 5 + 2t & 11 + 7t + 3t^2 + 4t^3 \\ 0 & 1 & 6 + 9t + 8t^2 \\ 0 & 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad \tau(A) = \begin{pmatrix} 1 & 2 & 4 \\ 0 & 1 & 8 \\ 0 & 0 & 1 \end{pmatrix}.$$

If the top-right cubic term is deleted, the extracted top-right entry becomes zero, not three: the assigned weight of that position is still three. In the scalar case,

$$[t^w](\alpha(a + dt)^w) = \alpha d^w,$$

which explains how the starting point a will disappear from a progression.

5 Canonical lattice representatives and the shell function

5.1 Canonical reduction by integer matrices

Put $I = (-1/2, 1/2]$, and identify the canonical coordinate box with I^ν. We use the same nearest-integer convention as before.

Lemma 5.1 (Canonical representative)**.** *For every $A \in G$, there is a unique $\gamma_A \in \Gamma$ such that*

$$R(A) := A\gamma_A$$

has all off-diagonal entries in I. In particular, $R(A)$ is uniquely determined by the right coset $A\Gamma$, and $|R(A)_{ij}| \le 1/2$.

Proof. Work separately in each block and process positions (i, j) in order of increasing gap $j - i$. Assume the entries γ_{kj} with $i < k < j$ have already been chosen. By (4.7),

$$(A\gamma)_{ij} = A_{ij} + \gamma_{ij} + \sum_{i<k<j} A_{ik}\gamma_{kj}.$$

Let

$$f_{ij} = A_{ij} + \sum_{i<k<j} A_{ik}\gamma_{kj}.$$

There is exactly one integer choice that puts this coordinate in I, namely $\gamma_{ij} = -\langle f_{ij}\rangle$. Every entry on which this choice depends has smaller gap. This constructs γ_A and proves its uniqueness by induction.

If $A' = A\lambda$ with $\lambda \in \Gamma$, then $A'\gamma_{A'} = A(\lambda\gamma_{A'})$ is another matrix in the same right coset with entries in I. Uniqueness for A therefore gives $R(A') = R(A)$. □

For clarity, $R(A)$ need not vary continuously with A; jumps can occur at rounding boundaries. We do *not* use it as a continuous function. The continuous function used for the primary coloring is different.

5.2 The shell function and its continuity

For $A \in G$, define

$$F(A) = \min_{\gamma\in\Gamma} \sum_{i<j} (A\gamma)_{ij}^2, \tag{5.1}$$

where the sum includes all blocks.

Lemma 5.2 (Existence, continuity, and invariance)**.** *The minimum in* (5.1) *exists. The function F is continuous,*

$$0 \le F(A) \le \nu/4, \qquad F(A\lambda) = F(A) \quad (A \in G,\ \lambda \in \Gamma). \tag{5.2}$$

For each compact $K \subseteq G$, there is a finite subset $\Gamma_K \subseteq \Gamma$ containing all minimizers for every $A \in K$.

Proof. The canonical representative supplies a candidate whose squared-coordinate sum is at most $\nu/4$. For a compact set K, consider any $A \in K$ and any $\gamma \in \Gamma$ for which

$$\sum_{i<j} (A\gamma)_{ij}^2 \le \nu/4.$$

Then $A\gamma$ lies in the closed coordinate box of radius $\sqrt{\nu}/2$. This is a compact subset of the finite-dimensional space G. Since inversion and multiplication are continuous, all the matrices

$$\gamma = A^{-1}(A\gamma)$$

under consideration belong to a common compact set. A bounded set contains only finitely many matrices with integer entries. Thus only finitely many candidates with value at most $\nu/4$ can occur, uniformly for $A \in K$. At least one candidate exists at each A by canonical reduction, so the minimum is attained, and the asserted finite set of minimizers exists.

On K, the function is consequently the minimum of finitely many continuous functions $A \mapsto \sum (A\gamma)_{ij}^2$. Taking K to be a compact neighborhood of an arbitrary point proves continuity on G. Nonnegativity and the upper bound have already been shown. Finally, as γ ranges through Γ, so does $\lambda\gamma$. Substitution in (5.1) gives right-lattice invariance. □

Lemma 5.3 (No nonconstant polynomial curve has constant shell value)**.** *Let $Q : \mathbb{R} \to G$ have ordinary polynomial entries. If Q is nonconstant, then $F(Q(t))$ is not constant on any nondegenerate compact real interval.*

Proof. Suppose, to the contrary, that $F(Q(t)) = c$ on $[u, v]$, with $u < v$. The image of this interval is compact. Lemma 5.2 supplies finitely many possible minimizers, say $\gamma_1, \dots, \gamma_s$. For each a, define the ordinary polynomial

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

The curve $Q(t)\gamma_a$ is nonconstant, since right multiplication by γ_a is invertible. Let $b \geq 1$ be the largest degree of its off-diagonal coordinate polynomials. The coefficient of t^{2b} in f_a is the sum of the squares of the leading coefficients of the degree-b coordinates. It is strictly positive. Thus f_a is a nonconstant polynomial.

For every $t \in [u, v]$, at least one $f_a(t)$ equals c. But each nonzero polynomial $f_a - c$ has only finitely many roots, and a finite union of finite root sets cannot contain an interval. This contradiction proves the lemma. □

Remark 5.4 (The role of ordinary polynomial coordinates). A general nonconstant continuous curve can lie inside a level set of F. The preceding assertion is not a statement about all curves on a nilmanifold. The sum-of-squares argument uses ordinary polynomial coordinates in an essential way. Those coordinates will be obtained by the rounding lemma in Section 8.

6 A shell coloring that forces every color to appear

Fix $\rho = 1/4$. Let $\mathcal{K}$ be the set of polynomial matrix curves $Q : [0, 1] \to G$ such that every off-diagonal coordinate has degree at most d_*, and

$$\sup_{0 \leq t \leq 1} |Q(t)|_{\text{coord}} \leq 1, \qquad |Q(1) - Q(0)|_{\text{coord}} = \rho. \tag{6.1}$$

For this auxiliary family we only impose the common degree bound d_*, not the more restrictive individual weighted bounds. This makes the sampling statement applicable to every rescaled curve needed later.

Lemma 6.1 (Uniform oscillation on normalized excursions)**.** *The coefficient vectors of $\mathcal{K}$ form a nonempty compact set. There is $\eta > 0$, depending only on G and d_*, such that*

$$\operatorname{osc}_{0 \leq t \leq 1} F(Q(t)) \geq \eta \qquad (Q \in \mathcal{K}). \tag{6.2}$$

Proof. The family is nonempty: let one coordinate equal ρt and all other off-diagonal coordinates equal zero. For compactness, evaluate each coordinate at the $d_* + 1$ distinct points $0, 1/d_*, \dots, 1$. Lagrange interpolation expresses its monomial coefficients as fixed linear combinations of these bounded values. Thus all coefficient vectors are bounded. The conditions in (6.1) are closed under coefficientwise convergence, which here is equivalent to uniform convergence on $[0, 1]$. The family is therefore compact.

The map $(Q, t) \mapsto F(Q(t))$ is continuous on $\mathcal{K} \times [0, 1]$. It follows that its maximum and minimum over t depend continuously on Q, and hence so does the oscillation. Every $Q \in \mathcal{K}$ is nonconstant by its endpoint displacement. Lemma 5.3 makes its oscillation strictly positive. A strictly positive continuous function on the compact set $\mathcal{K}$ has a strictly positive minimum, which we call η. □

Now fix the desired number $r \geq 2$ of primary colors. Set

$$h = \frac{\eta}{20r}, \qquad \kappa(A) = \left\lfloor \frac{F(A)}{h} \right\rfloor \pmod{r}. \tag{6.3}$$

The residue is taken in $\{0,\dots,r-1\}$. The color is constant on each shell $jh \le F(A) < (j+1)h$, with shell colors cycling modulo r. By right-lattice invariance of F,

$$\kappa(A\gamma) = \kappa(A) \qquad (\gamma \in \Gamma). \tag{6.4}$$

Lemma 6.2 (Uniform sampling of all colors)**.** *There is $T_0 \ge 1$, depending only on G, d_*, r, such that, for every $Q \in \mathcal{K}$ and every real $T \ge T_0$, the samples*

$$Q(j/T), \qquad j = 0, 1, \dots, \lfloor T \rfloor, \tag{6.5}$$

receive all r colors under κ.

Proof. Write $f_Q(t) = F(Q(t))$. Its range is a closed interval $[a,b]$, since it is continuous on a connected compact interval. By Lemma 6.1, $b - a \ge \eta = 20rh$. Let $j_0 = \lceil a/h \rceil$. The r numbers

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

all lie in $[a,b]$. Indeed, the first lies at least $h/2$ above a, and the last lies less than $(r+1/2)h$ above a. These numbers are midpoints of r consecutive shell bands, so their colors run through all residues modulo r. Each is at distance $h/2$ from the boundaries of its own band.

For each s, choose $u_s \in [0,1]$ with $f_Q(u_s) = z_s$. Uniform continuity of $(Q,t) \mapsto f_Q(t)$ on the compact set $\mathcal{K} \times [0,1]$ gives $\theta > 0$, independent of Q, such that

$$|t - u| < \theta \quad \Longrightarrow \quad |f_Q(t) - f_Q(u)| < h/4.$$

Choose T_0 with $T_0 > 1/\theta$. For $T \ge T_0$, put $j_s = \lfloor Tu_s \rfloor$. Then $0 \le j_s \le \lfloor T \rfloor$ and $|u_s - j_s/T| < \theta$. The sampled value $f_Q(j_s/T)$ lies within $h/4$ of the band midpoint z_s, and hence strictly inside the same band. It has color $j_0 + s$ modulo r. This proves the claim. □

The proof deliberately samples *interiors* of shell bands. It does not infer colors from a limiting value on a shell boundary. It also works for noninteger T: the sample $\lfloor Tu \rfloor / T$ approximates every $u \in [0,1]$, including $u = 1$.

7 The fixed auxiliary parameters

The construction must be completed before a progression length or the real coefficients are supplied. Here is the order of all choices.

Fix the weight patterns and r. The group G, lattice Γ, coordinate count ν, and degree bound d_* are now fixed. Choose $\rho = 1/4$, then obtain η from Lemma 6.1, define h and κ by (6.3), and obtain T_0 from Lemma 6.2. Use Lemma 3.2 to choose $M \ge d_* + 2$ so that coordinate bounds by $1/2$ at the first M integers force displacement less than ρ up to time T_0.

Next set

$$\xi = 2^{-d_*-3}. \tag{7.1}$$

Partition the canonical box I^ν into finitely many boxes whose coordinate intervals have length at most ξ. For instance, subdivide each copy of I into 2^{d_*+3} half-open intervals. Let R be the number of boxes. We call the resulting finite label the *auxiliary grid color*. It is different from the primary color κ, whose palette has size r.

We quote finite van der Waerden in the following form [4].

Theorem 7.1 (Finite van der Waerden)**.** *For every $R, M \in \mathbb{N}$ there is an integer $W = W(R, M)$ such that every R-coloring of $\{0, \dots, W-1\}$ contains an M-term monochromatic arithmetic progression of positive common difference.*

Fix one such W and define

$$L = d_*! \operatorname{lcm}(1, 2, \dots, W). \tag{7.2}$$

The purpose of the least common multiple is to make L/q an integer for every auxiliary progression difference $1 \le q \le W$. The additional factor $d_*!$ will clear the factorial denominators in Lemma 3.1.

For later reference, the dependence chain is

$$\begin{aligned}(\text{weight patterns}, r) \;&\longrightarrow\; (G, \Gamma, d_*, \rho) \;\longrightarrow\; \eta, h, \kappa, T_0 \\ &\longrightarrow\; M, \xi, R, W \;\longrightarrow\; L.\end{aligned} \tag{7.3}$$

No step here depends on the coefficients of the special expressions, the initial point of a progression, its common difference, or its length. In particular, the use of the potentially very large number R does not introduce an R-largeness hypothesis on the allowed differences. It will be used only to find a short subprogression inside one already given long progression.

8 Polynomial rounding inside a single small box

The lattice correction for $A(n)$ generally depends discontinuously on n. On a sufficiently small box and a sufficiently long finite sequence, however, these corrections agree exactly with integer-valued polynomials. No limiting or approximate replacement is involved.

Lemma 8.1 (Polynomial rounding)**.** *Let $A \in \mathcal{H}$. Suppose that the canonical representatives*

$$R(A(0)), R(A(1)), \dots, R(A(M-1))$$

all lie in a single auxiliary grid box. Then there exist $Q, C \in \mathcal{H}$ such that

$$A(t)C(t) = Q(t) \qquad (t \in \mathbb{R}), \qquad C(n) \in \Gamma \qquad (n \in \mathbb{Z}), \tag{8.1}$$

and

$$Q(n) = R(A(n)) \qquad (0 \le n < M). \tag{8.2}$$

Moreover, for every off-diagonal position,

$$w_{ij}!\, \tau(C)_{ij} \in \mathbb{Z}. \tag{8.3}$$

Proof. Write

$$A(n)\gamma(n) = R(A(n)), \qquad \gamma(n) \in \Gamma, \qquad 0 \le n < M.$$

These are the canonical lattice corrections from Lemma 5.1. We will construct each polynomial entry $C_{ij}(t)$ so that $C_{ij}(n) = \gamma(n)_{ij}$ on these samples. The construction is blockwise and proceeds in increasing order of the gap $j - i$.

Fix a position (i, j) and put $w = w_{ij}$. Assume that all entries $C_{kj}(t)$ with $i < k < j$ have already been constructed, are integer-valued on $\mathbb{Z}$, and agree with $\gamma(n)_{kj}$ for $0 \le n < M$. Define

$$f_{ij}(t) = A_{ij}(t) + \sum_{i<k<j} A_{ik}(t) C_{kj}(t). \tag{8.4}$$

Each summand has degree at most $w_{ik} + w_{kj} = w$, so $\deg f_{ij} \le w$. When $j - i = 1$, the sum is empty, and this starts the induction.

Let $r_{ij}(n) = R(A(n))_{ij}$ for $0 \le n < M$. The matrix product and the induction hypothesis give

$$r_{ij}(n) - f_{ij}(n) = \gamma(n)_{ij} \in \mathbb{Z}. \tag{8.5}$$

Taking $(w+1)$-st differences where all the samples are defined, and using $\Delta^{w+1} f_{ij} = 0$, shows that $\Delta^{w+1} r_{ij}(n)$ is an integer.

On the other hand, all the $r_{ij}(n)$ lie in one real interval of length at most ξ, because the representatives lie in one grid box. Choose an endpoint b of the closure of that interval. The coefficients of a positive-order finite difference sum to zero, so subtracting b does not change the difference. By (3.1),

$$\begin{aligned} |\Delta^{w+1} r_{ij}(n)| &= |\Delta^{w+1}(r_{ij} - b)(n)| \\ &\le \sum_{a=0}^{w+1} \binom{w+1}{a} \xi = 2^{w+1}\xi \le 2^{d_*+1} 2^{-d_*-3} = \tfrac{1}{4} < 1. \end{aligned} \tag{8.6}$$

An integer of absolute value less than one is zero. Hence all these $(w+1)$-st differences vanish.

Since $M \ge d_* + 2 \ge w + 2$, Lemma 3.1(i) gives a polynomial $Q_{ij}(t)$ of degree at most w agreeing with all the $r_{ij}(n)$. Put

$$C_{ij}(t) = Q_{ij}(t) - f_{ij}(t).$$

It has degree at most w, and (8.5) shows that it agrees with $\gamma(n)_{ij}$ at the first M integers. In particular, it is integer-valued at $0, \dots, w$. Lemma 3.1(ii) makes it integer-valued at every integer and gives $w![t^w]C_{ij}(t) \in \mathbb{Z}$. This completes the induction step.

Set all diagonal entries of Q and C equal to one. The identity $Q_{ij} = f_{ij} + C_{ij}$ and (8.4) say exactly that $Q(t) = A(t)C(t)$ in every entry. Every coordinate has its prescribed degree bound. The asserted sample identities and denominator bounds were established during the induction, so all conclusions follow. □

Remark 8.2 (What is exact, and what is only initially bounded). The identity $A(t)C(t) = Q(t)$ holds for every real t, and the lattice condition $C(n) \in \Gamma$ holds for every integer, not just the first M samples. Initially, however, we only know that $Q(n)$ is canonical for $0 \le n < M$. We do not assume that $Q(n)$ remains in the canonical box later. Controlling its later behavior is precisely the purpose of the first-exit argument.

9 An omitted color prevents escape of the polynomial lift

We now prove the central matrix statement, independently of the particular band matrices that encode the special expressions.

Proposition 9.1 (Weighted leading coefficients from an omitted color)**.** *With κ, W, L chosen in Section 7, let $p \in \mathcal{H}$ and $k \ge 4W + 2$. Suppose that*

$$\{\kappa(p(n)) : 0 \le n < k\} \ne \{0, \dots, r-1\}.$$

Then there exist $\gamma \in \Gamma$ and $E \in G$ such that

$$\mathrm{dil}_L \tau(p)\, \gamma = E \tag{9.1}$$

and, for every off-diagonal position of weight $w = w_{ij}$,

$$|E_{ij}| \le \frac{3}{4} b_w (2L)^w k^{-w}, \qquad b_w = \frac{2^w w^w}{w!}. \tag{9.2}$$

Proof. Choose a primary color e absent from the k samples. We divide the argument into five steps.

Step 1: find a short grid-monochromatic subprogression. Assign to $n \in \{0, \dots, W-1\}$ the auxiliary grid box containing $R(p(n))$. Finite van der Waerden gives integers $b \ge 0$ and $q \ge 1$ such that

$$b + (M-1)q \le W - 1 \tag{9.3}$$

and the representatives at $b, b+q, \dots, b+(M-1)q$ lie in one grid box. In particular, $b \le W-1$ and $q \le W$.

The curve $A(t) = p(b+qt)$ belongs to $\mathcal{H}$. Apply Lemma 8.1 to obtain

$$p(b+qt)C(t) = Q(t), \qquad C(n) \in \Gamma \quad (n \in \mathbb{Z}), \tag{9.4}$$

with $Q, C \in \mathcal{H}$ and $Q(n)$ canonical for $0 \le n < M$.

Step 2: identify the full range of integer samples available. Put

$$J = \left\lfloor \frac{k-1-b}{q} \right\rfloor. \tag{9.5}$$

By (9.3) and $k \ge W$, we have $J \ge M-1 \ge 1$. For every integer $0 \le j \le J$, the index $b+qj$ lies between 0 and $k-1$. Equations (9.4) and (6.4) therefore give

$$\kappa(Q(j)) = \kappa(p(b+qj)) \ne e \qquad (0 \le j \le J). \tag{9.6}$$

This uses the lattice condition at *all* integers. No assertion about the size or coefficients of $C(j)$ is required.

Step 3: prove that the lift cannot leave a bounded box. We claim

$$|Q(t) - Q(0)|_{\mathrm{coord}} < \rho = \tfrac{1}{4} \qquad (0 \le t \le J). \tag{9.7}$$

At each of the first M integer samples, every coordinate of $Q(n)$ is bounded by $1/2$. By the choice of M and Lemma 3.2,

$$|Q(t) - Q(0)|_{\mathrm{coord}} < \rho \qquad (0 \le t \le T_0).$$

Suppose that (9.7) fails. The coordinate displacement is continuous and zero at time zero, so there is a first time $T \in (0, J]$ at which it equals ρ. The preceding initial-motion bound implies $T > T_0$. For $0 \le t \le T$ the displacement is at most ρ. Since $Q(0)$ is canonical, every coordinate of $Q(t)$ on this interval has absolute value at most $1/2 + \rho = 3/4$.

Define the rescaled curve

$$\widetilde{Q}(u) = Q(Tu), \qquad 0 \le u \le 1.$$

Its coordinate degrees are at most d_*, all its coordinates are bounded by $3/4 < 1$, and its endpoint displacement is exactly ρ. Thus $\widetilde{Q} \in \mathcal{K}$. Lemma 6.2, with mesh parameter T, says that

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

receive all r colors. Because $\lfloor T \rfloor \le J$, this contradicts (9.6). The claim is proved.

It follows that every off-diagonal coordinate of Q has absolute value at most $3/4$ throughout the *real* interval $[0, J]$. For a position of weight w, its degree is at most w. Lemma 3.3 therefore gives

$$|\tau(Q)_{ij}| \le \frac{3}{4} b_w J^{-w}. \tag{9.8}$$

Step 4: remove the starting point and clear denominators. Apply the weighted leading-coefficient map τ to (9.4). Lemma 4.2 gives

$$(\mathrm{dil}_q \tau(p))\tau(C) = \tau(Q). \tag{9.9}$$

Apply the automorphism $\mathrm{dil}_{L/q}$. Since dilations compose by multiplying their parameters, the resulting equation is

$$\mathrm{dil}_L \tau(p)\, (\mathrm{dil}_{L/q} \tau(C)) = \mathrm{dil}_{L/q} \tau(Q).$$

Define

$$\gamma = \mathrm{dil}_{L/q} \tau(C), \qquad E = \mathrm{dil}_{L/q} \tau(Q). \tag{9.10}$$

We must check $\gamma \in \Gamma$, rather than merely rationality of its entries. Since $1 \le q \le W$, the number L/q is an integer multiple of $d_*!$. By (8.3), for each weight w there is an integer z_{ij} with

$$\tau(C)_{ij} = \frac{z_{ij}}{w!}.$$

Consequently

$$\gamma_{ij} = \frac{(L/q)^w}{w!}\, z_{ij} \in \mathbb{Z}, \tag{9.11}$$

because $w \le d_*$ and $w!$ divides $d_*!$, hence divides L/q. This proves the required lattice membership and (9.1).

Step 5: retain the original progression length in the bound. Equations (9.8) and (9.10) imply

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

The floor in (9.5) gives

$$qJ \ge k - 1 - b - q \ge k - 1 - 2W.$$

Since $k \ge 4W + 2$, the last expression is at least $k/2$. Hence

$$|E_{ij}| \le \frac{3}{4} b_w (2L)^w k^{-w},$$

which is (9.2). □

Remark 9.2 (The exact use of an omitted color). In the proof, the primary color hypothesis is used only in (9.6) and in the contradiction at the first exit. The lift need not be monochromatic. A single fixed color missing from all its integer samples is enough, because a normalized excursion forces all colors. The auxiliary grid does not alter the primary coloring.

10 Recovering nested nearest-integer expressions

10.1 An exact last-column calculation

For $x_1, \dots, x_s \in \mathbb{R}$, let $B(x_1, \dots, x_s)$ be the upper-unitriangular matrix with first superdiagonal

$$-x_1, -x_2, \dots, -x_{s-1}, +x_s, \tag{10.1}$$

and all entries above that superdiagonal equal to zero. For $s = 1$, its sole off-diagonal entry is $+x_1$.

Lemma 10.1 (Nested rounding from a small last column)**.** *Suppose*

$$B(x_1, \dots, x_s)\gamma = E, \qquad \gamma \in \mathrm{UT}_{s+1}(\mathbb{Z}), \quad E \in \mathrm{UT}_{s+1}(\mathbb{R}),$$

and $|E_{i,s+1}| < 1/2$ *for every* $1 \le i \le s$*. Then*

$$\mathsf{L}(x_1, \dots, x_s) + \gamma_{1,s+1} = E_{1,s+1}. \tag{10.2}$$

In particular, $\|\mathsf{L}(x_1, \dots, x_s)\|_{\mathbb{R}/\mathbb{Z}} \le |E_{1,s+1}|$.

Proof. Put $z_s = x_s$ and, recursively, $z_i = x_i \langle z_{i+1} \rangle$ for $i = s-1, \dots, 1$. Thus $z_i = \mathsf{L}(x_i, \dots, x_s)$.

The bottom nontrivial entry in the last column of the matrix identity is

$$x_s + \gamma_{s,s+1} = E_{s,s+1}.$$

Since $\gamma_{s,s+1}$ is an integer and the residual has absolute value less than $1/2$, it follows that $\gamma_{s,s+1} = -\langle z_s \rangle$. There is no rounding tie.

Suppose, working upward, that $\gamma_{i+1,s+1} = -\langle z_{i+1} \rangle$. The matrix B has no nonzero entries beyond the first superdiagonal. Its row-i last-column equation is therefore exactly

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

Substitution gives

$$\gamma_{i,s+1} + x_i \langle z_{i+1} \rangle = \gamma_{i,s+1} + z_i = E_{i,s+1}.$$

Again, the residual identifies the nearest integer uniquely, so $\gamma_{i,s+1} = -\langle z_i \rangle$. Induction reaches $i = 1$ and yields (10.2). When $s = 1$, the bottom-row calculation is already the desired assertion. □

Example 10.2 (One and two nested brackets)**.** For $P(n) = \alpha n^u \langle \beta n^v \rangle$, use

$$g(n) = \begin{pmatrix} 1 & -\alpha n^u & 0 \\ 0 & 1 & \beta n^v \\ 0 & 0 & 1 \end{pmatrix}.$$

If $g(d)\gamma = E$ with a sufficiently small last column, then

$$\gamma_{23} = -\langle \beta d^v \rangle, \qquad E_{13} = \gamma_{13} + \alpha d^u \langle \beta d^v \rangle = \gamma_{13} + P(d).$$

The weight of position $(1, 3)$ is $u + v$, so a bound $|E_{13}| \le Ck^{-(u+v)}$ gives precisely that rate for $\|P(d)\|_{\mathbb{R}/\mathbb{Z}}$.

For $P(n) = \alpha n^u \langle \beta n^v \langle \theta n^w \rangle \rangle$, the 4×4 band matrix has superdiagonal $-\alpha n^u, -\beta n^v, +\theta n^w$. Its last column yields

$$\begin{aligned} \gamma_{34} &= -\langle \theta d^w \rangle, \\ \gamma_{24} &= -\langle \beta d^v \langle \theta d^w \rangle \rangle, \\ E_{14} &= \gamma_{14} + \alpha d^u \langle \beta d^v \langle \theta d^w \rangle \rangle. \end{aligned}$$

The top position has weight $u + v + w$. These identities are exact; no large coefficient is multiplied by an uncontrolled rounding error.

10.2 Proof of the omitted-color theorem

Proof of Theorem 1.6. First fix the weight patterns of all the expressions, and make the choices in Section 7. For one expression

$$P(n) = \mathsf{L}(\alpha_1 n^{w_1}, \dots, \alpha_s n^{w_s}),$$

define the ordinary polynomial matrix

$$g_P(t) = B(\alpha_1 t^{w_1}, \dots, \alpha_s t^{w_s}). \tag{10.3}$$

Its off-diagonal entries are monomials of exactly the allowed weight, or are zero. In particular, $g_P \in \mathcal{H}$ for its block. Take $g(t)$ to be the product of these block matrices over $P_1, \dots, P_m$.

Define the coloring of the positive integers by

$$c(n) = \kappa(g(n/L)). \tag{10.4}$$

Only g involves the real coefficients. The shell coloring κ and the integer L were fixed from the patterns and r alone.

Let $\{a + jd : 0 \le j < k\}$ omit a color under c, with $k \ge 4W + 2$. Consider

$$p(t) = g((a + dt)/L). \tag{10.5}$$

This belongs to $\mathcal{H}$, and its first k primary colors are the colors of the given progression. Proposition 9.1 therefore applies.

For a superdiagonal coordinate $\sigma\alpha_j t^{w_j}$ of g, where $\sigma \in \{-1, +1\}$ is its prescribed sign, the corresponding entry of p has coefficient

$$[t^{w_j}] \left(\sigma\alpha_j \left(\frac{a + dt}{L} \right)^{w_j} \right) = \sigma\alpha_j (d/L)^{w_j}.$$

Every higher off-diagonal entry of g and of p is zero. Therefore, entry by entry,

$$\tau(p) = g(d/L), \qquad \operatorname{dil}_L \tau(p) = g(d). \tag{10.6}$$

This calculation uses weighted homogeneity, not a claim that g is a homomorphism from the additive real line.

We conclude that

$$g(d)\gamma = E, \qquad \gamma \in \Gamma, \qquad |E_{ij}| \le C_* k^{-w_{ij}}, \tag{10.7}$$

where we may take the pattern-dependent constant

$$C_* = \max \left\{ 1, \max_{1 \le w \le d_*} \frac{3}{4} b_w (2L)^w \right\}. \tag{10.8}$$

Choose

$$k_* = \max\{4W + 2, \lceil 4C_* \rceil + 1\}. \tag{10.9}$$

For $k \ge k_*$, every off-diagonal error in (10.7) has absolute value less than $1/4$, because $w_{ij} \ge 1$ and $C_* k^{-w_{ij}} \le C_*/k < 1/4$.

Apply Lemma 10.1 to each block. The top-right entry for P_i has weight d_i, so

$$\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \le |E_{\mathrm{top},i}| \le C_* k^{-d_i} \qquad (1 \le i \le m).$$

This proves the assertion for the fixed patterns.

It remains to remove dependence on which weight patterns were chosen. For each integer $d \le d_*$, there are finitely many ordered lists of positive integers summing to d (indeed, 2^{d-1} such lists). Thus there are finitely many possible m-tuples of patterns with individual total weights at most d_*. For each tuple carry out the construction above. Taking the maximum of its C_* and k_* over these finitely many tuples produces $C(r, m, d_*)$ and $k_0(r, m, d_*)$.

For a prescribed coefficient family, use the coloring (10.4) belonging to its pattern tuple. That coloring was defined before k was supplied and works for every $k \ge k_0$. This proves all the stated quantifiers and uniformities. □

11 From omitted colors to the prescribed upper density

Proof of Theorem 1.4. Fix δ, m, d_*, and choose r by (2.3). Apply Theorem 1.6 with this r. Its constants now depend only on δ, m, d_*. Given $P_1, \dots, P_m$, let c be its coloring.

Lemma 2.2 gives a color e such that

$$A_e = \{n \in \mathbb{N} : c(n) \ne e\}, \qquad \alpha := \overline{d}(A_e) \ge 1 - 1/r \ge \delta.$$

Every progression in A_e omits color e, no matter how many of the other colors it uses. Therefore Theorem 1.6 gives all of the estimates (1.3) for progressions in A_e.

Use Lemma 2.1 to choose $A \subseteq A_e$ with $\overline{d}(A) = \delta$. Every progression in A is a progression in A_e, so all of the same estimates remain valid. The coloring c, the omitted color e, and the thinning are chosen once, independently of the progression length k and of any proposed set D of allowed differences. This proves the asserted uniform test-set statement. □

Proof of Corollary 1.5. For the prescribed family $P_1, \dots, P_m$, take the set A supplied by Theorem 1.4. The (δ, k)-density largeness of D applied to this particular A gives

$$\{a + jd : 0 \le j < k\} \subseteq A \quad \text{for some } a \in \mathbb{N},\ d \in D.$$

Equation (1.3) supplies (1.4). □

Corollary 11.1 (Simultaneous qualitative approximation)**.** *If D is δ-density large for some $\delta \in (0, 1)$, then for every finite family of special generalized polynomials $P_1, \dots, P_m$ and every $\varepsilon_1, \dots, \varepsilon_m > 0$, there is $d \in D$ such that*

$$\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} < \varepsilon_i \qquad (1 \le i \le m). \tag{11.1}$$

Proof. Let d_i be the assigned weights, set $d_* = \max_i d_i$, and take C, k_0 from Corollary 1.5. Choose an integer $k \ge k_0$ so large that $Ck^{-d_i} < \varepsilon_i$ for every i. Since D is (δ, k)-density large, the corollary gives the desired d. □

Example 11.2 (Simultaneous weights one, two, and three)**.** Fix any real $\alpha, \beta, \gamma, \theta$ and any $0 < \delta < 1$. Theorem 1.4, for the expressions

$$P_1(n) = \alpha n, \qquad P_2(n) = \beta n^2, \qquad P_3(n) = \gamma n \left\langle \theta n^2 \right\rangle,$$

produces one set A of upper density exactly δ such that every sufficiently long k-term progression in A has difference d satisfying

$$\|\alpha d\|_{\mathbb{R}/\mathbb{Z}} \le Ck^{-1}, \qquad \left\|\beta d^2\right\|_{\mathbb{R}/\mathbb{Z}} \le Ck^{-2}, \qquad \left\|\gamma d \left\langle \theta d^2 \right\rangle\right\|_{\mathbb{R}/\mathbb{Z}} \le Ck^{-3}.$$

The constant depends only on δ and these bounded combinatorial data, not on the four real coefficients. If D is δ-density large, the same test set supplies such a difference in D at every sufficiently large progression length.

12 Nil-Bohr neighborhoods and the nilrecurrence conclusion

12.1 The precise dynamical input

Definition 12.1 (Nil-Bohr neighborhood)**.** A set $B \subseteq \mathbb{N}$ is a *finite-step nil-Bohr$_0$ set* if there are a finite-step nilsystem (X, T), a point $x \in X$, and an open neighborhood U of x such that

$$\{n \in \mathbb{N} : T^n x \in U\} \subseteq B.$$

We use the following form of the characterization in [3, Theorem B and Theorem 4.2.11].

Theorem 12.2 (Huang–Shao–Ye: special-polynomial basis)**.** *Every finite-step nil-Bohr$_0$ set $B \subseteq \mathbb{N}$ contains a set of the form*

$$\{n \in \mathbb{N} : \|P_i(n)\|_{\mathbb{R}/\mathbb{Z}} < \varepsilon_i \text{ for } 1 \le i \le m\}, \tag{12.1}$$

where $P_1, \dots, P_m$ are special generalized polynomials as in Definition 1.3 and $\varepsilon_1, \dots, \varepsilon_m > 0$.

The source states the equality of the corresponding families on $\mathbb{Z}$; restricting its basis sets to positive integers gives the displayed form. Theorem B identifies nil-Bohr and generalized-polynomial families, and Theorem 4.2.11 replaces general normalized polynomials by special ones. The treatment of nonminimal nilsystems by passage to point orbit closures is recorded in [3, Section 3.1.1]. Thus the input applies to the convention in Definition 12.1; it requires no connectedness hypothesis here. This is a containment of actual sets, not containment after deleting an exceptional set of density zero.

12.2 A high-density obstruction to each nilpotent return condition

Theorem 12.3 (High-density nil-Bohr obstruction)**.** *Let $B \subseteq \mathbb{N}$ be a finite-step nil-Bohr$_0$ set, and let $0 < \delta < 1$. There are $A \subseteq \mathbb{N}$ and $K \in \mathbb{N}$ such that $\overline{d}(A) = \delta$ and*

$$\{a + jd : 0 \le j < K\} \subseteq A \quad \Longrightarrow \quad d \in B. \tag{12.2}$$

The same implication holds for every progression length at least K.

Proof. Choose the finite family P_i and tolerances ε_i in Theorem 12.2. If no conditions are needed, one may take the single zero expression $P_1(n) = 0 \cdot n$, of assigned weight one. Let d_i be the weights and put $d_* = \max_i d_i$. Use Theorem 1.4 to obtain A, C, k_0. Choose an integer $K \ge k_0$ such that

$$CK^{-d_i} < \varepsilon_i \qquad (1 \le i \le m).$$

If a K-term progression lies in A, its difference d satisfies $\|P_i(d)\|_{\mathbb{R}/\mathbb{Z}} \le CK^{-d_i} < \varepsilon_i$ for every i. By (12.1), this forces $d \in B$. A longer progression contains a K-term initial subprogression with the same difference, so the final assertion follows as well. □

This formulation makes the qualitative result transparent: the complement of any nil-Bohr neighborhood fails the density-largeness test even at densities arbitrarily close to one.

12.3 Proof of multiple pointwise nilrecurrence

Proof of Theorem 1.2. First, D meets every finite-step nil-Bohr$_0$ set. Indeed, given such a set B, apply Theorem 12.3 with the fixed δ from the hypothesis on D, obtaining A and K. By δ-density largeness, A contains a K-term progression with some difference $d \in D$. Equation (12.2) gives $d \in B$, so $D \cap B \neq \varnothing$.

Now fix (X, T), x, ℓ, and ε as in the theorem. On the product nilmanifold X^ℓ, consider the translation

$$S = T \times T^2 \times \cdots \times T^\ell$$

and the point $\mathbf{x} = (x, \dots, x)$. If $X = G/\Gamma$ and T is translation by g, this is the nilsystem on G^ℓ/Γ^ℓ with translating element $(g, g^2, \dots, g^\ell)$. Define the open neighborhood

$$V = \{(y_1, \dots, y_\ell) : d_X(y_j, x) < \varepsilon \text{ for all } j\}$$

of $\mathbf{x}$. Its positive return-time set is

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

It is a finite-step nil-Bohr$_0$ set by definition, so the first part of the proof supplies $d \in D \cap B$. This gives (1.1).

Finally, let $M \in \mathbb{N}$ be arbitrary, and choose an integer $q > \max\{M, 1\}$. The set $q\mathbb{N}$ is a Bohr$_0$ set: it is the return time set of 0 for the circle rotation $y \mapsto y + 1/q$ to the neighborhood $\{y \in \mathbb{T} : \|y\| < 1/(2q)\}$. Indeed, a nonzero residue modulo q has distance at least $1/q$ from zero. Taking the product of this circle rotation with the preceding product nilsystem shows that $B \cap q\mathbb{N}$ is again a finite-step nil-Bohr$_0$ set.

The first part of the proof now yields $d \in D \cap B \cap q\mathbb{N}$. Since d is a positive multiple of q, we have $d \ge q > M$. As M was arbitrary, the admissible differences are unbounded. This proves the theorem in full. □

The same product argument allows finitely many different nilsystems, points, neighborhoods, and multiples to be prescribed at once. It also explains why no separate minimality assumption is imposed on a product translation, which need not be minimal even when its factors are.

13 Scope, the endpoint, and checks on the quantifiers

13.1 The range already covered by the binary theorem

For comparison, call D 2-*large* if every two-coloring of $\mathbb{N}$ has arbitrarily long monochromatic progressions with differences in D. If $0 < \delta \le 1/2$ and D is δ-density large, then D is 2-large. For any two-coloring, at each cutoff one color occupies at least half the integers, and some fixed color does so along an unbounded sequence of cutoffs. Its class has upper density at least $1/2$, hence at least δ. Thin it to upper density δ and apply the hypothesis on D. The resulting progressions are monochromatic in the original coloring.

Thus in that range the qualitative theorem can also be deduced from the original two-color nilrecurrence result. For $1/2 < \delta < 1$, this simple deduction is unavailable. The proof here instead removes one color from a larger palette. The assertion that every bad long progression contains *all* colors is what permits this change.

13.2 Why density one must be excluded

Proposition 13.1 (Failure at the endpoint)**.** *If one extends Definition 1.1 to $\delta = 1$, then $D = \{1\}$ is 1-density large but is not a set of pointwise recurrence even for circle rotations.*

Proof. Let $A \subseteq \mathbb{N}$ have $\overline{d}(A) = 1$. Fix $k \geq 2$. If A contained no interval of k consecutive integers, each of the disjoint blocks

$$\{jk+1, \ldots, (j+1)k\}, \qquad j = 0, 1, 2, \ldots,$$

would contain at most $k-1$ elements of A. Hence

$$|A \cap \{1, \ldots, N\}| \leq (k-1)\lfloor N/k \rfloor + k,$$

which would imply $\overline{d}(A) \leq 1 - 1/k < 1$. This contradiction shows that A has an interval of every length. Each such interval is a progression with difference one, proving 1-density largeness of $\{1\}$.

For nonrecurrence, choose an irrational α, let $T(y) = y + \alpha$ on $\mathbb{T}$, and take $x = 0$. For $0 < \varepsilon < \|\alpha\|_{\mathbb{R}/\mathbb{Z}}$, the only permitted time $d = 1$ does not return x within ε of itself. □

13.3 What the quantitative result measures

The bounds Ck^{-d_i} are in terms of the *progression length* k, not the size of the common difference d. The proof gives no upper bound on that common difference. The constants depend on the fixed density, number of expressions, and maximum assigned weight; their construction involves compactness and a finite van der Waerden number. No numerical rate for that dependence is asserted.

The fixed-coloring statement in Theorem 1.6 is stronger than producing a new coloring separately for each k. Likewise, Theorem 1.4 produces one set A that works at every sufficiently large length. The finite corollary requires only the single-length hypothesis on D, whereas qualitative nilrecurrence uses the ability to choose that length after the desired return neighborhood has been specified.

Only natural *upper* density is used. We do not claim that the constructed test set has a natural density or a prescribed lower density. The thinning gives exact equality of upper densities, which is exactly the hypothesis in the question.

13.4 Dependency and error checks

The most sensitive points of the proof can be located as follows. The shell function is continuous although the canonical representative is not: these are separated in Lemmas 5.1 and 5.2. Uniform positive oscillation follows from compactness and the nonconstancy lemma, rather than from a uniform claim about arbitrary continuous curves. Lemma 6.2 uses band midpoints to avoid all color-boundary issues.

In the rounding lemma, finite differences are first shown to be *integers*, and then to have absolute value below one. This yields exact vanishing and an exact polynomial identity. The Newton argument extends lattice membership to every integer, which is essential when the lift is sampled beyond the initial M points.

At the first exit, time is rescaled by the actual real exit time T; the sampling lemma is valid for this real mesh parameter. Weighted coefficient extraction is used only on entries whose degrees are at most the assigned weights. The factorial and least-common-multiple factors in L prove actual

integrality in (9.11). The nested-bracket calculation is exact and never amplifies a small inner error by an unbounded coefficient. Finally, the upper-density argument is a finite counting argument followed by thinning, not an assumption that color classes have densities.

Provenance and references

The supplied Alweiss–Farhangi manuscript [1] contains the original 2-largeness-to-nilrecurrence result and the coefficient-uniform approximation framework. The companion elementary draft [2] supplies the weighted-matrix, polynomial-rounding, and shell-coloring architecture. The omitted-color extension and its fixed-density deductions have been proved here in full within that architecture; they are not being attributed to the external nil-Bohr characterization. No access to those working drafts is needed to follow the argument in this file. All source references are included below, and no separate bibliography file is required.

References

[1] R. Alweiss and S. Farhangi, 2-*large sets, nilrecurrence, and polynomial approximation*, working manuscript supplied in the accompanying discussion, 2026.

[2] *An elementary two-color obstruction for special generalized polynomials*, companion working draft to the Alweiss–Farhangi manuscript, supplied in the accompanying discussion, 2026.

[3] W. Huang, S. Shao, and X. Ye, *Nil Bohr-sets and almost automorphy of higher order*, Memoirs of the American Mathematical Society **241** (2016), no. 1143, v+83 pp. Preprint: arXiv:1407.1179. The numbered statements used here are Theorem B, Theorem 4.2.11, and Definition 4.1.2; the numbering is also present in the 2014 arXiv version.

[4] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Archief voor Wiskunde **15** (1927), 212–216.