An elementary two-color obstruction for special generalized polynomials

Expanded exposition

We give an elementary two-color obstruction for finite families of special generalized polynomials. The proof uses the finite van der Waerden theorem, finite differences, polynomial interpolation, compactness, and multiplication of upper-unitriangular matrices. It gives simultaneous approximation at the reciprocal power of the progression length corresponding to each expression's assigned degree. The purpose of this exposition is to spell out the lifting, compactness, and rounding steps without introducing nilsequence equidistribution or a factorization theorem.

1 Statement and conventions

Write $[x]$ for a nearest integer to x, with any fixed convention at ties, and put $\|x\| = \operatorname{dist}(x, \mathbb{Z})$. Throughout, $\mathbb{N} = \{1, 2, \ldots\}$. A *special generalized polynomial* is an expression

$$P(t) = a_1 t^{w_1} \big[a_2 t^{w_2} \big[\cdots [a_\ell t^{w_\ell}] \cdots \big]\big], \qquad a_i \in \mathbb{R}, \quad w_i \in \mathbb{N}. \tag{1.1}$$

For $\ell = 1$, this means $P(t) = a_1 t^{w_1}$, with no bracket. Equivalently, starting from $H_\ell(t) = a_\ell t^{w_\ell}$, define $H_i(t) = a_i t^{w_i}[H_{i+1}(t)]$ downwards; then $P = H_1$. The *assigned degree* is $\deg P = w_1 + \cdots + w_\ell$. We retain this degree even when special coefficient choices simplify the function. For example, $\alpha t^2[\beta t]$ has assigned degree 3, including when $\beta = 0$.

A set $S \subseteq \mathbb{N}$ is $(2, N)$-*large* if every coloring $c : \mathbb{N} \to \{0, 1\}$ has a monochromatic progression $a, a + d, \ldots, a + (N - 1)d$ with $d \in S$. It is 2-*large* if this holds for every $N \geq 1$.

Theorem 1.1. *Fix finitely many lists of positive exponents. There are constants $C > 0$ and $N_0 \geq 1$, depending only on these lists, with the following property. For every corresponding family $P_1, \ldots, P_r$ of special generalized polynomials, there is a single two-coloring of $\mathbb{N}$ such that every monochromatic N-term progression, with $N \geq N_0$, has difference d satisfying*

$$\|P_i(d)\| \leq C N^{-\deg P_i} \qquad (1 \leq i \leq r). \tag{1.2}$$

The constants are independent of the real coefficients, and the coloring works simultaneously for all $N \geq N_0$. Consequently, for every $\varepsilon > 0$, the set

$$\{d \geq 1 : \max_{1 \leq i \leq r} \|P_i(d)\| \geq \varepsilon\} \tag{1.3}$$

is not 2-large. In particular, every $(2, N)$-large set, for $N \geq N_0$, contains a d satisfying (1.2).

The estimate concerns the progression length N, not the size of d. The coloring will be obtained by evaluating a matrix polynomial at n/L, where L is a fixed integer. Its role is to cancel a bounded arithmetic step and the factorial denominators produced by integer-valued polynomials.

Here is the proof strategy. An auxiliary finite coloring finds a short subprogression whose reduced matrix coordinates lie in a small box. Finite differences replace its integer lattice corrections by ordinary integer-valued polynomials. The actual two-coloring prevents this polynomial lift from escaping a bounded box over the rest of the progression. Interpolation then makes its weighted leading coefficients small. Finally, these coefficients remove the initial point of the progression, and matrix multiplication recovers the nested brackets at its difference.

2 Weighted polynomial matrices

For one exponent list $(w_1, \dots, w_\ell)$, let $U_{\ell+1}(\mathbb{R})$ be the group of upper-triangular matrices with diagonal entries one. Assign entry (i, j) the weight

$$w_{ij} = w_i + \cdots + w_{j-1} \qquad (i < j). \tag{2.1}$$

For the whole family take the direct product G of these groups, with one block for each expression, and let Γ be the subgroup with integer entries in every block. Let m be the total number of off-diagonal entries and D the largest weight. All coordinate norms and sums below are over these entries in all blocks. In particular, $\|A - B\|_\infty = \max_{i<j} |A_{ij} - B_{ij}|$.

Let $\mathcal{H}$ consist of matrix-valued polynomials $A(t)$ for which $\deg A_{ij} \le w_{ij}$. The diagonal is always one. The product formula

$$(AB)_{ij} = A_{ij} + B_{ij} + \sum_{i<k<j} A_{ik} B_{kj} \tag{2.2}$$

shows that $\mathcal{H}$ is closed under multiplication, since $w_{ik} + w_{kj} = w_{ij}$. It is also closed under inverses: if $A = I + V$, then $A^{-1} = I - V + V^2 - \cdots$, with a finite sum, and every chain of entries contributing to position (i, j) has total weight w_{ij}.

For $u > 0$ and $B \in G$, define a dilation by $(\delta_u B)_{ij} = u^{w_{ij}} B_{ij}$, with diagonal entries one. For $A \in \mathcal{H}$, define its *weighted leading-coefficient matrix* by

$$\tau(A)_{ij} = [t^{w_{ij}}] A_{ij}(t). \tag{2.3}$$

Here $[t^w]$ denotes coefficient extraction, not nearest-integer rounding. A coordinate of degree less than its weight contributes zero.

Lemma 2.1. *For $A, B \in \mathcal{H}$, $b \in \mathbb{R}$, and $q, u > 0$, one has*

$$\tau(AB) = \tau(A)\tau(B), \qquad \tau\big(A(b + qt)\big) = \delta_q \tau(A), \qquad \delta_u(BC) = \delta_u(B)\delta_u(C) \quad (B, C \in G). \tag{2.4}$$

Also, $\tau(A) = I$ whenever A is constant.

Proof. In a product $A_{ik}B_{kj}$ in (2.2), a term of degree $w_{ij} = w_{ik} + w_{kj}$ can only be obtained by multiplying the terms of degrees w_{ik} and w_{kj}. Taking this coefficient in (2.2) proves the first identity. If $f(t) = \sum_{j=0}^{w} c_j t^j$, the coefficient of t^w in $f(b + qt)$ is $q^w c_w$, independently of b. This proves the second identity. The third follows by multiplying (2.2) by $u^{w_{ij}}$ and using additivity of the weights. Finally, every weight is positive, so the weighted coefficient of every constant entry is zero. □

These identities are exact matrix identities, not coordinatewise approximations. In particular, all the noncommutative cross terms are retained by τ. This is what allows us to extract a leading matrix from a factorization later on.

3 An elementary shell coloring

Canonical representatives and a continuous size function

Fix the half-open interval $I = [-1/2, 1/2)$. Every right coset $A\Gamma$ has a unique representative, denoted $\mathrm{rep}(A)$, whose off-diagonal entries lie in I. Here is a coordinate construction. Seek $\lambda \in \Gamma$

such that $A\lambda$ has this property, and process entries in increasing order of the matrix gap $j-i$. Once the shorter-gap entries of λ are known, set

$$f_{ij} = A_{ij} + \sum_{i<k<j} A_{ik}\lambda_{kj}, \qquad \lambda_{ij} = -\lfloor f_{ij} + 1/2 \rfloor. \tag{3.1}$$

Then $(A\lambda)_{ij} = f_{ij} + \lambda_{ij} \in I$. Each choice is unique, which proves existence and uniqueness by induction. In particular, $\mathrm{rep}(A\gamma) = \mathrm{rep}(A)$ for $\gamma \in \Gamma$. This half-open convention is only for choosing representatives; the brackets in (1.1) retain their original convention.

Define

$$F(A) = \min_{\gamma\in\Gamma} \sum_{i<j} (A\gamma)_{ij}^2. \tag{3.2}$$

The canonical representative gives a candidate of value at most $m/4$. The minimum is attained: among candidates of value at most $m/4$, every entry of $A\gamma$ is bounded, and hence so is every entry of $\gamma = A^{-1}(A\gamma)$. There are only finitely many such integer matrices. Thus $0 \le F(A) \le m/4$, and reindexing the minimum gives $F(A\gamma) = F(A)$ for all $\gamma \in \Gamma$.

The same argument proves continuity. On a compact set of matrices A, the inverse matrices A^{-1} are uniformly bounded. Every minimizer in (3.2) therefore belongs to one fixed finite subset of Γ. On this compact set, F is the minimum of finitely many continuous functions. Applying this to a compact neighborhood of each point shows that F is continuous everywhere. Notice that we do not assert that rep is continuous or that it minimizes (3.2).

Lemma 3.1 (No polynomial curve stays on one shell)**.** *If $Q(t)$ is a nonconstant polynomial matrix with values in G, then $F(Q(t))$ is not constant on any nondegenerate real interval.*

Proof. Restrict to a compact nondegenerate subinterval. Only finitely many γ can minimize (3.2) there. For each of them put $f_\gamma(t) = \sum_{i<j} (Q(t)\gamma)_{ij}^2$. This is a nonconstant polynomial. Indeed, right multiplication by a fixed invertible matrix cannot turn a nonconstant curve into a constant one. If $d \ge 1$ is the maximum degree of its coordinate polynomials, the coefficient of t^{2d} in f_γ is the sum of the squares of their degree-d coefficients, and is positive.

If $F(Q(t))$ were the constant c, then at every point of the interval some one of these finitely many polynomials would satisfy $f_\gamma(t) = c$. Each equation has only finitely many solutions, so their union cannot cover the interval. □

Uniform oscillation and sampling

Set $\rho = 1/4$. Let $\mathcal{K}$ be the family of polynomial matrices $Q : [0,1] \to G$ whose coordinate degrees are at most D and which satisfy

$$|Q_{ij}(t)| \le 1 \quad (0 \le t \le 1), \qquad \|Q(1) - Q(0)\|_\infty = \rho. \tag{3.3}$$

This is a compact family in coefficient space. To see boundedness, interpolate each coordinate at the fixed nodes $0, 1/D, \dots, 1$. The coefficients are bounded linear combinations of its values at these nodes. The conditions in (3.3) are closed, giving compactness in the finite-dimensional coefficient space.

Every member is nonconstant. Moreover, the functional

$$Q \longmapsto \mathrm{osc}_{[0,1]}\, F(Q) := \max_{0\le t\le 1} F(Q(t)) - \min_{0\le t\le 1} F(Q(t))$$

is continuous: coefficient convergence gives uniform convergence of the curves, and F is uniformly continuous on the bounded coordinate box. Lemma 3.1 and compactness therefore give a constant $\eta > 0$ such that

$$\operatorname{osc}_{[0,1]} F(Q) \geq \eta \qquad (Q \in \mathcal{K}). \tag{3.4}$$

Fix $0 < h < \eta/10$ and define the right-Γ-invariant two-coloring

$$\kappa(A) = \left\lfloor \frac{F(A)}{h} \right\rfloor \pmod 2. \tag{3.5}$$

Thus consecutive shells of width h have opposite colors.

Lemma 3.2 (Uniform sampling)**.** *There is a real number $T_0 \geq 1$ 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 both colors.

Proof. The range of $F \circ Q$ is an interval of length at least $\eta > 10h$. It therefore contains the midpoints of two consecutive complete bands $[kh, (k+1)h)$, of opposite parities. Choose parameters at which these two midpoint values are attained. Each midpoint is a distance $h/2$ from its band's endpoints.

The coefficient bound for $\mathcal{K}$ also bounds the derivatives of its members on $[0, 1]$. Together with uniform continuity of F on the coordinate box, this gives $\xi > 0$, independent of Q, such that

$$|s - t| < \xi \quad \Longrightarrow \quad |F(Q(s)) - F(Q(t))| < h/4.$$

Choose T_0 so large that $1/T_0 < \xi$. Every parameter in $[0, 1]$ is within $1/T$ of some j/T with $0 \leq j \leq \lfloor T \rfloor$. Approximating the two selected parameters by such mesh points leaves their values strictly inside the two bands. Thus both colors occur. This argument also covers nonintegral T and avoids all shell boundaries. □

4 Interpolation estimates and the order of choices

We record both polynomial estimates needed below.

Lemma 4.1. *There are constants $B_D, K_D \geq 1$, depending only on D, with the following properties. If $M \geq D + 1$, $\deg f \leq D$, and $|f(n)| \leq 1/2$ for $n = 0, \ldots, M - 1$, then*

$$\sup_{0 \leq t \leq M-1} |f'(t)| \leq \frac{B_D}{M - 1}. \tag{4.1}$$

If $J > 0$, $\deg f \leq w \leq D$, and $|f(t)| \leq 3/4$ on $[0, J]$, then

$$|[t^w]f(t)| \leq K_D J^{-w}. \tag{4.2}$$

Proof. For the first assertion put $b = \lfloor (M-1)/D \rfloor$. Then $b \geq (M-1)/(2D)$ and the nodes $0, b, \ldots, Db$ are among the prescribed integer samples. The polynomial $H(u) = f(bu)$ has $|H(j)| \leq 1/2$ for $0 \leq j \leq D$. In the Lagrange formula

$$H(u) = \sum_{j=0}^{D} H(j) \prod_{\substack{0 \leq k \leq D \\ k \neq j}} \frac{u - k}{j - k},$$

the basis polynomials and their derivatives are bounded on $[0, 2D]$ by constants depending only on D. Since $(M-1)/b \le 2D$, rescaling bounds f' on $[0, M-1]$ by a constant times b^{-1}, proving (4.1).

For the second assertion interpolate $H(u) = f(Ju)$ at $0, 1/D, \dots, 1$. Every coefficient of H is bounded by a constant depending only on D. Its coefficient of u^w equals $J^w[t^w]f(t)$, which gives (4.2), also when this coefficient is zero. □

Now choose an integer $M > D+1$ so large that

$$M - 1 \ge T_0, \qquad \frac{B_D T_0}{M-1} < \rho. \tag{4.3}$$

Consequently, whenever a polynomial matrix Q of coordinate degrees at most D has $|Q_{ij}(n)| \le 1/2$ for $0 \le n < M$, the mean value estimate gives

$$\|Q(t) - Q(0)\|_\infty < \rho \qquad (0 \le t \le T_0). \tag{4.4}$$

This estimate concerns the ordinary real polynomial, not its values modulo Γ.

Choose an even integer K so large that $\delta = 1/K < 2^{-D-2}$. Partition I^m into the $R = K^m$ half-open coordinate boxes of side length δ. This gives an auxiliary R-coloring of G by the box containing $\mathrm{rep}(A)$. By the finite van der Waerden theorem, there is an integer $W = W(R, M)$ such that every R-coloring of $\{0, \dots, W-1\}$ has a monochromatic M-term arithmetic progression. Finally set

$$L = D!\,\mathrm{lcm}(1, \dots, W). \tag{4.5}$$

The order of choices is

$$(G, D) \;\longrightarrow\; (\eta, h, \kappa, T_0) \;\longrightarrow\; M \;\longrightarrow\; (\delta, R, W) \;\longrightarrow\; L.$$

All these objects depend only on the fixed exponent lists. They are chosen before any coefficients, starting point, common difference, or progression length enter the argument. The auxiliary R-coloring will only be used inside an already monochromatic binary progression.

5 Polynomial rounding inside one small box

The next lemma changes a finite list of discontinuously chosen lattice corrections into a single polynomial correction valid at every integer. No assertion is made that the polynomial remains the canonical representative outside the initial samples.

Lemma 5.1 (Polynomial rounding)**.** *Let $A \in \mathcal{H}$. Suppose that $\mathrm{rep}(A(0)), \dots, \mathrm{rep}(A(M-1))$ lie in one auxiliary grid cell. There exist $Q, C \in \mathcal{H}$ such that*

$$A(t)C(t) = Q(t) \quad (t \in \mathbb{R}), \qquad C(n) \in \Gamma \quad (n \in \mathbb{Z}), \tag{5.1}$$

and $Q(n) = \mathrm{rep}(A(n))$ for $0 \le n < M$.

Proof. For the initial integer samples write $r(n) = \mathrm{rep}(A(n)) = A(n)\lambda(n)$, where $\lambda(n) \in \Gamma$. We construct the entries of C and Q by increasing matrix gap, maintaining $C_{ij}(n) = \lambda_{ij}(n)$ and $Q_{ij}(n) = r_{ij}(n)$ on these samples. Suppose the shorter-gap entries have been constructed. For the current position (i, j) put $w = w_{ij}$ and

$$f_{ij}(t) = A_{ij}(t) + \sum_{i<k<j} A_{ik}(t)C_{kj}(t). \tag{5.2}$$

Its degree is at most w by additivity of weights. At the sample points, (2.2) and the induction hypothesis give $r_{ij}(n) - f_{ij}(n) = \lambda_{ij}(n) \in \mathbb{Z}$.

Write $\Delta u(n) = u(n+1) - u(n)$. Since $\Delta^{w+1} f_{ij} = 0$, it follows that $\Delta^{w+1} r_{ij}(n)$ is an integer whenever all its samples belong to $\{0, \dots, M-1\}$. On the other hand, the values $r_{ij}(n)$ lie in one interval of length δ. Subtracting its left endpoint, which does not change a positive-order difference, gives

$$|\Delta^{w+1} r_{ij}(n)| \le \sum_{v=0}^{w+1} \binom{w+1}{v} \delta = 2^{w+1}\delta < 1. \tag{5.3}$$

Thus all these differences vanish exactly.

It follows that the samples are values of a polynomial Q_{ij} of degree at most w. Explicitly, take the Newton interpolation polynomial

$$Q_{ij}(t) = \sum_{v=0}^{w} \Delta^v r_{ij}(0) \binom{t}{v}, \qquad \binom{t}{v} = \frac{t(t-1)\cdots(t-v+1)}{v!}, \tag{5.4}$$

with $\binom{t}{0} = 1$. This agrees with the first $w+1$ samples, and the recurrence $\Delta^{w+1} r_{ij} = 0$ forces agreement with all remaining samples. Set $C_{ij} = Q_{ij} - f_{ij}$. This has degree at most w and agrees with the integer-valued samples $\lambda_{ij}(n)$.

To see that C_{ij} is integer-valued on all of $\mathbb{Z}$, use its Newton expansion. The coefficients $\Delta^v C_{ij}(0)$, for $0 \le v \le w$, are integer combinations of its values at $0, \dots, w$, hence are integers. Moreover, $\binom{n}{v} \in \mathbb{Z}$ for every integer n, including negative n, since $\binom{n}{v} = (-1)^v \binom{v-n-1}{v}$ when $n < 0$. Thus $C_{ij}(n) \in \mathbb{Z}$ for all $n \in \mathbb{Z}$. The construction also gives $(AC)_{ij} = Q_{ij}$ identically by (5.2). Induction over all gaps and all blocks proves the lemma. □

We will need the precise denominator information in this proof:

$$w_{ij}!\, \tau(C)_{ij} = \Delta^{w_{ij}} C_{ij}(0) \in \mathbb{Z}. \tag{5.5}$$

The entries of C need not have integer coefficients. Their integer-valuedness is the correct conclusion, and the factorial in (5.5) explains the factor $D!$ in (4.5).

Remark 5.2 (An integer-valued correction with a denominator). In a single $U_2(\mathbb{R})$ block of weight 2, identify a matrix with its only off-diagonal entry. Consider

$$A(t) = \binom{t}{2} + \epsilon t^2, \qquad 0 < \epsilon < \frac{\delta}{(M-1)^2}.$$

At an integer n, the term $\binom{n}{2}$ is an integer, so the first M canonical representatives are $r(n) = \epsilon n^2 \in [0, \delta)$. This is one grid cell, since K was chosen even. The rounding lemma can take

$$Q(t) = \epsilon t^2, \qquad C(t) = -\binom{t}{2} = -\tfrac{1}{2}t^2 + \tfrac{1}{2}t.$$

Here matrix multiplication is addition of the displayed entries. The correction is integer-valued at every integer, but its leading coefficient is $-1/2$, not an integer. Multiplication by 2! clears that denominator. This simple example explains why integer-valued polynomials, rather than polynomials with integer coefficients, occur in the general argument.

6 Control of the weighted leading coefficients

We now prove the matrix statement from which the theorem follows.

Proposition 6.1. *Let $N \geq 2W+2$ and $p \in \mathcal{H}$. Suppose that $p(0), \dots, p(N-1)$ are monochromatic under κ. There is $\gamma \in \Gamma$ such that*

$$\delta_L \tau(p)\, \gamma = E, \qquad |E_{ij}| \leq K_D \left(\frac{L}{N-1-2W} \right)^{w_{ij}} \quad (i<j). \tag{6.1}$$

Proof. Apply the auxiliary grid coloring to $p(0), \dots, p(W-1)$. The choice of W gives integers $b \geq 0$ and $q \geq 1$ with $b + (M-1)q \leq W-1$, such that

$$\operatorname{rep}(p(b)), \operatorname{rep}(p(b+q)), \dots, \operatorname{rep}(p(b+(M-1)q))$$

lie in one grid cell. In particular, $b, q \leq W$. Apply Lemma 5.1 to $A(t) = p(b+qt)$. It gives

$$p(b+qt)C(t) = Q(t), \qquad C(\mathbb{Z}) \subseteq \Gamma. \tag{6.2}$$

Put $J = \lfloor (N-1-b)/q \rfloor$. For every integer $0 \leq j \leq J$, right-lattice invariance gives

$$\kappa(Q(j)) = \kappa(p(b+qj)). \tag{6.3}$$

Thus all these integer samples of Q have the original common color. Initially, $Q(0), \dots, Q(M-1)$ are canonical representatives, so their coordinates have absolute value at most $1/2$.

The lift cannot escape. We claim that

$$\|Q(t) - Q(0)\|_\infty < \rho \qquad (0 \leq t \leq J). \tag{6.4}$$

If not, continuity gives a first time $T \in (0, J]$ when the displacement is exactly ρ. The initial-motion estimate (4.4) implies $T > T_0$. Up to this time the displacement is at most ρ, and $Q(0)$ has coordinates in I. Hence all coordinates of $Q(t)$ have absolute value at most $3/4$ for $0 \leq t \leq T$.

Consequently, the rescaled curve $\widetilde{Q}(s) = Q(Ts)$ belongs to $\mathcal{K}$: its coordinate degrees are at most D, its coordinates are bounded by one, and its endpoint displacement equals ρ. Lemma 3.2, applied with the real mesh parameter T, says that

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

receive both colors. These are among the monochromatic samples in (6.3), a contradiction. This proves (6.4). The use of real T is essential: the first exit need not occur at an integer.

Extracting the small leading matrix. Every coordinate of Q is now bounded by $3/4$ throughout the real interval $[0, J]$. Applying (4.2) with its individual degree bound w_{ij} gives

$$|\tau(Q)_{ij}| \leq K_D J^{-w_{ij}}. \tag{6.5}$$

Taking weighted leading matrices in (6.2) yields the exact identity

$$\delta_q \tau(p)\, \tau(C) = \tau(Q). \tag{6.6}$$

The shift b has disappeared, while the step q is recorded by a dilation.

Clearing the denominators. Since $1 \le q \le W$, the number $s = L/q$ is an integer multiple of $D!$. Equation (5.5) shows that every off-diagonal entry of $\delta_s\tau(C)$ is an integer: if $\tau(C)_{ij} = a/w_{ij}!$, then

$$(\delta_s\tau(C))_{ij} = a\frac{s^{w_{ij}}}{w_{ij}!} \in \mathbb{Z}.$$

Here $w_{ij}!$ divides $D!$, and $w_{ij} \ge 1$. Thus $\gamma := \delta_{L/q}\tau(C) \in \Gamma$. Applying the automorphism $\delta_{L/q}$ to (6.6) gives

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

By (6.5),

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

Finally, the definition of J gives $qJ \ge N - 1 - b - q \ge N - 1 - 2W > 0$, which proves (6.1). □

Only the small initial subprogression lies in one auxiliary grid cell. The much longer bound (6.4) comes from the binary shell coloring, not from another application of van der Waerden's theorem.

Remark 6.2 (The scalar model of leading-coefficient extraction). For a single $U_2(\mathbb{R})$ block of weight w, write its polynomial entry as $p(t) = \theta t^w +$ lower-degree terms. Equation (6.6) is simply

$$q^w\theta + \tau(C) = \tau(Q), \qquad w!\,\tau(C) \in \mathbb{Z}, \qquad |\tau(Q)| \le K_D J^{-w}.$$

Multiplication by $(L/q)^w$ turns the correction into an integer, giving a small distance from $L^w\theta$ to $\mathbb{Z}$. For the polynomial entry $p(t) = \alpha((a + dt)/L)^w$, one has $\theta = \alpha(d/L)^w$, and hence $L^w\theta = \alpha d^w$. Thus the starting point a disappears and the estimate concerns the common difference d. Weighted matrix multiplication is the exact analogue of this computation for nested brackets; its cross terms are why one cannot extract each matrix coordinate independently.

7 Recovering the nested brackets

For one expression (1.1), define $g(t) \in U_{\ell+1}(\mathbb{R})$ by setting all entries above the first superdiagonal to zero and setting its first superdiagonal to

$$-a_1t^{w_1},\ -a_2t^{w_2},\ \ldots, -a_{\ell-1}t^{w_{\ell-1}},\ +a_\ell t^{w_\ell}. \tag{7.1}$$

For $\ell = 1$ the sole entry is $+a_1t^{w_1}$. For the family take the product of these band matrices, again denoted $g(t) \in G$. Each entry is homogeneous of its assigned weight, or identically zero. Thus, for $u > 0$ and $x \in \mathbb{R}$,

$$g(u) = \delta_u g(1), \qquad \tau\big(g(x + ut)\big) = g(u). \tag{7.2}$$

The first identity is not an assertion that g is a one-parameter subgroup; it is simply weighted homogeneity.

We first explain the sign choice in (7.1).

Lemma 7.1 (Decoding a small last column)**.** *Let $x_i = a_id^{w_i}$ and suppose that, in one block, $g(d)\gamma = E$ with γ an integer upper-unitriangular matrix. If $|E_{i,\ell+1}| < 1/4$ for $1 \le i \le \ell$, then*

$$P(d) + \gamma_{1,\ell+1} = E_{1,\ell+1}, \qquad \|P(d)\| \le |E_{1,\ell+1}|. \tag{7.3}$$

Proof. Set $H_\ell = x_\ell$ and $H_i = x_i[H_{i+1}]$ downwards, so $H_1 = P(d)$. The bottom nontrivial entry of the last column gives

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

Because its absolute value is less than $1/4$, the nearest integer to x_ℓ is uniquely $-\gamma_{\ell,\ell+1}$. For $i < \ell$, the band structure gives

$$E_{i,\ell+1} = \gamma_{i,\ell+1} - x_i\gamma_{i+1,\ell+1}. \tag{7.4}$$

If $\gamma_{i+1,\ell+1} = -[H_{i+1}]$, this becomes $E_{i,\ell+1} = \gamma_{i,\ell+1}+H_i$. Again the small residual gives $\gamma_{i,\ell+1} = -[H_i]$ uniquely. Downward induction therefore identifies all the nested brackets, and the top row proves (7.3). For $\ell = 1$ the bottom-row identity already is (7.3). □

For example, a single bracket $P(d) = x_1[x_2]$ is encoded by

$$g(d) = \begin{pmatrix} 1 & -x_1 & 0 \\ 0 & 1 & x_2 \\ 0 & 0 & 1 \end{pmatrix}, \qquad E_{23} = x_2 + \gamma_{23}, \quad E_{13} = \gamma_{13} - x_1\gamma_{23}.$$

The first equation gives $\gamma_{23} = -[x_2]$, and the second becomes $E_{13} = \gamma_{13} + x_1[x_2]$. The deeper case is the same substitution repeated up the last column. No symmetry convention such as $[-x] = -[x]$ is needed, because all the relevant nearest integers are unique.

Proof of Theorem 1.1. Use the fixed κ and L constructed above, and color $\mathbb{N}$ by

$$\boxed{c(n) = \kappa(g(n/L)).} \tag{7.5}$$

Suppose $a, a+d, \dots, a+(N-1)d$ is monochromatic. The matrix polynomial

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

belongs to $\mathcal{H}$, and its first N integer values are monochromatic under κ. By (7.2), $\tau(p) = g(d/L)$ and hence $\delta_L\tau(p) = g(d)$. Proposition 6.1 gives, for $N \geq 2W+2$,

$$g(d)\gamma = E, \qquad \gamma \in \Gamma, \qquad |E_{ij}| \leq K_D\left(\frac{L}{N-1-2W}\right)^{w_{ij}}. \tag{7.6}$$

Choose $N_0 \geq 4W+2$ large enough that $K_D(2L/N_0)^w < 1/4$ for every integer $1 \leq w \leq D$. For $N \geq N_0$, one has $N-1-2W \geq N/2$, so all off-diagonal entries of E have absolute value less than $1/4$. Applying Lemma 7.1 in every block, whose top-right weight is $\deg P_i$, gives

$$\|P_i(d)\| \leq K_D(2L)^{\deg P_i} N^{-\deg P_i}.$$

Taking $C = K_D \max_{1\leq w\leq D}(2L)^w$ proves (1.2). Every constant used to define C and N_0 depends only on the exponent lists. The real coefficients enter only through g in (7.5). In particular, this is one coloring for every $N \geq N_0$.

Given $\varepsilon > 0$, choose an integer $N \geq N_0$ so large that $CN^{-\deg P_i} < \varepsilon$ for every i. No monochromatic N-term progression for (7.5) can have its difference in (1.3); thus that set is not 2-large. Finally, applying $(2, N)$-largeness to this same coloring supplies a difference in the prescribed set satisfying (1.2). □

8 What makes the proof work

There are two different reductions, and they serve different purposes. The canonical representative puts the first M samples in a bounded coordinate box. It may be discontinuous, but the grid makes each of its coordinate ranges small enough for an integer finite difference to vanish. The continuous function F, by contrast, need not be computed at that canonical representative. Its role is to ensure that a normalized nonconstant polynomial curve cannot remain monochromatic. Right-lattice invariance is what transfers the original colors to the polynomial lift.

There are also two different scales. The fixed initial length M makes the lift move slowly enough to sample any first exit. The much longer interval $[0, J]$, whose length is comparable to N/q, then forces its weight-w leading coefficients to be $O(J^{-w})$. Extracting these coefficients removes the starting point. The single dilation L clears both the possible steps $q \le W$ and the factorial denominators. This accounts for the precise power $N^{-\deg P_i}$.

Positivity of the exponents is used in the algebra: every entry has positive weight, constant matrices disappear under τ, and the factorial-clearing dilation works in every coordinate. The statement is for the special expressions (1.1). It does not, without additional arguments, extend to arbitrary floor expressions or all possible generalized-polynomial syntax. Nor does the proof give a bound for the size of the allowed difference. Its uniformity is in the real coefficients, with the exponent patterns held fixed.

Provenance. The shorter source note describes this argument as an elementary reworking of the Alweiss–Farhangi argument, developed with Codex assistance and informed by the earlier manuscript. The present version expands the explanations and intermediate proofs of that note. Apart from the finite van der Waerden theorem and the elementary tools made explicit above, the argument invokes no result from the earlier manuscript and no nilrecurrence, nilsequence equidistribution, orbit-closure, or factorization theorem.