

RAINBOW THREE-TERM PROGRESSIONS IN $\mathbb{Z}_n$ WITH FOUR OR MORE COLOURS

KOREA SUPERINTELLIGENCE LABS

ABSTRACT. For a surjective r -colouring of $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, call an ordered solution (x, y, z) of $x + y = 2z$ *rainbow* if its three coordinates carry three different colours. Elvin, Gonzales, Rodriguez and Wilbur determined the largest possible proportion of rainbow solutions for $r = 3$ — it is $2/3$, attained exactly when $3 \mid n$ — and asked what happens for $r \geq 4$, where their counting argument gives only $1 - 1/r$. We show that the whole gap at $r = 4$ collapses to a single inequality about the colour classes: some pair of classes has size product at most the total *defect* $\sum_i |A_i|^2 - M$, where M counts the monochromatic solutions. That inequality implies the proportion $2/3$ for every odd modulus, and we prove it unconditionally at every odd prime and at the composite moduli 9, 15, 21, 27, 33, 39, 51, 57 and 69. The modulus 27 is the first for which the extension $0 \rightarrow \mathbb{Z}_9 \rightarrow \mathbb{Z}_{27} \rightarrow \mathbb{Z}_3 \rightarrow 0$ does not split; we give an exact fibre decomposition of the defect twisted by the 2-cocycle of that extension and show the twist is not removable and is exactly what makes the modulus 27 work. Along the way we record an exact identity for odd n relating the rainbow count to the class sizes, two unconditional bounds strictly below $3/4$, and the exact four-, five-, six- and seven-colour maxima at small moduli, from which the maxima appear to be $1 - 1/p(r)$ with $p(r)$ the largest odd integer at most r — so that only an odd number of colours ever helps. All statements are machine-checked in Lean 4.

1. INTRODUCTION

Let $n \geq 1$, write $\mathbb{Z}_n = \mathbb{Z}/n\mathbb{Z}$, and let

$$T_n = \{(x, y, z) \in \mathbb{Z}_n^3 : x + y = 2z\}.$$

The triple $(x, y, z) \in T_n$ records the three-term progression $\{x, z, y\}$ with z in the middle. Ordered triples over the whole cube are counted, so the two orderings of a progression are two solutions and degenerate triples such as (a, a, a) are counted as well; $|T_n| = n^2$ for every n , since $z \mapsto 2z$ is either a bijection (odd n) or two-to-one onto an index-two subgroup (even n).

Following [5], an r -colouring of $\mathbb{Z}_n$ is a surjective map $c: \mathbb{Z}_n \rightarrow \{1, \dots, r\}$, and a solution $(x, y, z) \in T_n$ is *rainbow* for c if $c(x), c(y), c(z)$ are pairwise distinct. Write $R(c)$ for the number of rainbow solutions and

$$m_r(n) = \max\{R(c) : c \text{ an } r\text{-colouring of } \mathbb{Z}_n\}, \quad \rho_r(n) = m_r(n)/n^2.$$

Elvin, Gonzales, Rodriguez and Wilbur [5] proved that $\rho_3(n) \leq 2/3$ for every n , with equality exactly when $3 \mid n$, the extremal colouring being $x \mapsto x \bmod 3$. Their argument is a Cauchy–Schwarz count that generalises verbatim to give $\rho_r(n) \leq 1 - 1/r$, and they close the paper by asking:

Question 1.1 ([5]). What are some bounds on the maximum number of rainbow three-term progressions when more than three colours are used?

The mod-3 construction already shows $\rho_4(n) \geq 2/3$ whenever $3 \mid n$: recolour a single point of one class with the fourth colour, and nothing is lost, because a solution with two points in a coset of the index-3 subgroup has its third point there too. So for $r = 4$ the truth lies between $2/3$ and $3/4$, and Question 1.1 is the question of which end.

What this paper does. Every exact four-colour maximum we can compute equals $2n^2/3$ when $3 \mid n$ and is smaller otherwise, so the answer at $r = 4$ appears to be $2/3$ again: *a fourth colour buys nothing*. The bulk of the paper is a route to that statement.

Section 3 records an exact identity, valid for odd n and any number of colours,

$$R(c) + 3S(c) = n^2 + 2M(c), \quad S(c) = \sum_i |A_i|^2, \quad A_i = c^{-1}(i),$$

with $M(c)$ the number of monochromatic solutions. The identity has been published before (Section 3 gives the references), and it makes the whole problem a question about the *defect* $D(c) = S(c) - M(c)$. From it we obtain two unconditional bounds strictly below $3/4$ — one by an additive constant, one by a term linear in n — and the exact value $2/3$ for every four-colouring two of whose classes have reflection-closed union, which is the entire conjectured extremal family.

Section 5 isolates the inequality that closes the gap completely. Call a modulus n *defect-bounded* if every four-colouring of $\mathbb{Z}_n$ has two colour classes whose size product is at most $D(c)$. For odd n , defect-boundedness implies $\rho_4(n) \leq 2/3$, and when a class is empty the hypothesis is free, so the whole of the $r = 4$ problem sits inside this one inequality. We prove it at every odd prime, by a per-class bound coming from Pollard's theorem: summing a bound $\frac{1}{4}|A|^2$ over *all four* classes suffices, because $(a^2 + b^2 + c^2 + d^2)/4 \geq ab$ for $a \leq b \leq c \leq d$, with equality only when all four are equal — which four class sizes summing to an odd number never are.

Composite moduli are genuinely harder, and not for a reason of technique: no bound depending only on the four class sizes can settle $n = 15, 21$ or 27 (Proposition 5.4). Sections 6 and 7 therefore work with the finer datum of the 4×3 matrix of class sizes on the three fibres of $\mathbb{Z}_n \rightarrow \mathbb{Z}_3$. For $n = 3m$ with $\gcd(3, m) = 1$ this reduces defect-boundedness to a statement internal to $\mathbb{Z}_m$; the case $m = 9$, i.e. $n = 27$, is the first where the extension

$$0 \longrightarrow \mathbb{Z}_9 \longrightarrow \mathbb{Z}_{27} \longrightarrow \mathbb{Z}_3 \longrightarrow 0$$

does not split and the reduction is unavailable. Section 6 supplies a *twisted* decomposition instead: the fibre blocks pick up the 2-cocycle of the extension, and the resulting shifts are not removable by any change of section. They are also not a nuisance but the mechanism: a single block can be empty, whereas two blocks aiming at the same fibre cannot both be, and that one gain of 8 units is exactly what a modulus divisible by 9 was missing. Section 7 spends the symmetry group $S_3 \times S_4$ of the resulting finite check on the enumeration itself, which is what puts $n = 69$ in range.

Conventions. Everything below is about $\mathbb{Z}_n$. The behaviour over an interval $[n]$ is different — four colours on $[9]$ already reach $28/41 > 2/3$ — so nothing here answers Question 1.1 in that setting. Rainbow counts, class sizes and defects are natural numbers throughout, and inequalities between them are stated without division wherever a denominator would be the only reason to leave $\mathbb{Z}$.

2. PRELIMINARIES

Definition 2.1. For $n \geq 1$ and $r \geq 1$ a *colouring* is a map $c: \mathbb{Z}_n \rightarrow \{1, \dots, r\}$, and an r -*colouring* is a surjective one. A triple $(x, y, z) \in \mathbb{Z}_n^3$ is a *solution* if $x + y = 2z$, and it is *rainbow* for c if $c(x) \neq c(y)$, $c(x) \neq c(z)$ and $c(y) \neq c(z)$. We write $R(c) = |\{t \in T_n : t \text{ rainbow}\}|$ and $m_r(n) = \max_c R(c)$ over r -colourings.

These are the conventions of [5]: surjectivity is part of the definition of a colouring there, ordered triples are counted, and the middle coordinate is the third one.

Fix a colouring c and write $A_i = c^{-1}(i)$, $a_i = |A_i|$, and

$$S(c) = \sum_i a_i^2, \quad M(c) = |\{(x, y, z) \in T_n : c(x) = c(y) = c(z)\}|, \quad D(c) = S(c) - M(c).$$

$S(c)$ counts ordered same-coloured pairs and $M(c)$ counts monochromatic solutions; the difference $D(c) \geq 0$ is the *defect* of c . For a single set $A \subseteq \mathbb{Z}_n$ put

$$(1) \quad d(A) = |\{(x, z) \in A \times A : 2z - x \notin A\}|,$$

the *defect* of A : the number of pairs of A whose reflection escapes A . Since $(x, z) \mapsto (x, 2z - x, z)$ is a bijection $\mathbb{Z}_n^2 \rightarrow T_n$, summing (1) over the classes gives $\sum_i d(A_i) = S(c) - M(c) = D(c)$. Call A *reflection-closed* if $2z - x \in A$ for all $x, z \in A$; equivalently, for odd n , if A is a coset of a subgroup of $\mathbb{Z}_n$. Thus $d(A) = 0$ exactly for reflection-closed A .

The source's theorem, and the construction attaining it, are the starting point.

Theorem 2.2 ([5]). *For every n and every map $c: \mathbb{Z}_n \rightarrow \{1, 2, 3\}$,*

$$3R(c) \leq 2n^2.$$

Consequently $m_3(n) \leq 2\lfloor n/3 \rfloor n$ whenever $3 \mid n$.

Surjectivity is not needed for the inequality. The proof is two steps: every rainbow solution is determined by the pair (first endpoint, midpoint), and that pair is bichromatic, so $R(c) \leq n^2 - S(c)$; and Cauchy–Schwarz gives $n^2 \leq 3S(c)$ for three classes.

Proposition 2.3. *Let $3 \mid n$ and let $c(x) = x \bmod 3$. Then c is a 3-colouring and $R(c) = 2\lfloor n/3 \rfloor n = \frac{2}{3}n^2$.*

The mechanism is that 2 is invertible modulo 3, so $x + y = 2z$ forces $z \equiv (x + y)/2 \pmod{3}$ and the solution fails to be rainbow exactly when $x \equiv y \pmod{3}$ — a $1/3$ fraction. The same computation with 3 replaced by any odd $p \mid n$ gives proportion $1 - 1/p$, which is the construction behind every lower bound in Section 4. The mechanism is folklore; see [1, Prop. 3.14].

Corollary 2.4. *For every n divisible by 3, $m_3(n) = 2\lfloor n/3 \rfloor n$, i.e. $\rho_3(n) = 2/3$.*

3. THE MIDPOINT IDENTITY AND TWO UNCONDITIONAL BOUNDS

Theorem 3.1. *Let n be odd and $c: \mathbb{Z}_n \rightarrow \{1, \dots, r\}$ any colouring. Then*

$$(2) \quad R(c) + 3S(c) = n^2 + 2M(c).$$

Proof sketch. Parametrise solutions by their first two coordinates (x, z) , so that the third is $2z - x$; this is a bijection $\mathbb{Z}_n^2 \rightarrow T_n$. For each of the three pairs of coordinates let $E_1, E_2, E_3 \subseteq \mathbb{Z}_n^2$ be the sets of solutions on which that pair is monochromatic. Because two colour coincidences among three points force the third, the indicator identity $\mathbf{1}_{E_1} + \mathbf{1}_{E_2} + \mathbf{1}_{E_3} + \mathbf{1}_{\text{rainbow}} = 1 + 2 \cdot \mathbf{1}_{E_1 \cap E_2 \cap E_3}$ holds pointwise, and summing gives $|E_1| + |E_2| + |E_3| + R(c) = n^2 + 2M(c)$. Each $|E_j|$ equals $S(c)$: for two of the three pairs this is immediate from the parametrisation, and for the pair of *endpoints* it uses that n is odd, so that the midpoint $(x + y)/2$ exists and is unique and the endpoints too parametrise T_n bijectively. $\square$

Oddness is used exactly once, and the identity is false without it: at $n = 6$ the optimal four-colouring 012013 has $R(c) = 24$ and $S(c) = M(c) = 10$, while $n^2 - 3S(c) + 2M(c) = 26$. The correct general form replaces one of the three coincidence counts by a count of pairs congruent modulo 2. Every statement in this section is therefore for odd n .

Identity (2) is not new. It is the specialisation to $x + y = 2z$ of Montejano and Serra's count of patterns in coloured orthogonal arrays [10]: the solutions of $x + y = 2z$ in $\mathbb{Z}_n$ form an $OA(3, 2)$ exactly when n is odd, and [10] names this application. The half of it that splits a colouring along cosets has been published several times over [4, 2, 6, 13]. We note in passing that the bound $1 - 1/r$ of [5] is a one-line consequence of (2) together with Cauchy–Schwarz.

Write (2) as $3R(c) = 3n^2 - 9S(c) + 6M(c)$ and compare with the target $3R(c) \leq 2n^2$: what has to be shown is

$$(3) \quad n^2 \leq 3(S(c) + 2D(c)).$$

Cauchy–Schwarz on four classes gives only $n^2 \leq 4S(c)$, which is weaker. But merging two classes turns four classes into three, and Cauchy–Schwarz on those three gives

$$(4) \quad n^2 \leq 3(S(c) + 2a_i a_j) \quad \text{for every } i \neq j.$$

Comparing (4) with (3): it suffices to find one pair $i \neq j$ with $a_i a_j \leq D(c)$. That observation is the subject of Section 5.

Two unconditional bounds come out of (2) directly. Both use the same structural fact. For odd n a reflection-closed set is a coset, hence has odd size; four odd numbers cannot sum to an odd n ; so *some class of a four-colouring is not reflection-closed*, and its defect is positive. It is in fact at least 2, because the defect set of (1) carries a fixed-point-free involution. With three colours no such obstruction exists, and this asymmetry is the mechanism behind the whole phenomenon that only an odd number of colours helps.

Theorem 3.2. *Let n be odd and c a surjective four-colouring of $\mathbb{Z}_n$. Then*

$$4R(c) + 16 \leq 3n^2.$$

The counting argument of [5] gives $4R(c) \leq 3n^2$ for every n ; the additive 16 is the entire gain, and it comes from the parity obstruction above through $D(c) \geq 2$. Surjectivity is needed: with three colours the value $2/3$ is attained and all classes can be cosets.

The additive constant can be replaced by a term growing with n , by bounding the defect of a single class from below rather than by the constant 2. Call $z \in A$ a *centre* of A if $2z - A = A$. Every non-centre of A contributes a defect pair, so $|A| \leq d(A) + |\text{centres}(A)|$; for odd n a centre gives a fixed-point-free involution of $A \setminus \{z\}$, so a class of even size has no centre at all and $|A| \leq d(A)$; and $\text{centres}(A)$ is itself reflection-closed, hence a coset of a subgroup T , and A is a union of at least two T -cosets unless A is reflection-closed, giving $|A| \leq 2d(A)$ in general. Since four class sizes summing to an odd n cannot all be odd, some class has even size; combining that with the divisor constraint on the sizes of cosets gives the following.

Theorem 3.3. *Let n be odd and c a surjective four-colouring of $\mathbb{Z}_n$. Then*

$$60R(c) + 16n \leq 45n^2, \quad \text{i.e.} \quad \frac{R(c)}{n^2} \leq \frac{3}{4} - \frac{4}{15n}.$$

This is an improvement in the count, not in the proportion: the right-hand side still tends to $3/4$. The conjectured $2/3$ needs a defect quadratic in n , and no bound on the defect of one class at a time can supply one — see Proposition 5.4.

A quadratic defect is available as soon as two classes interact, and for one family of colourings that is enough. If A and B are disjoint with $A \cup B$ reflection-closed, send $(p, q) \in A \times B$ to the midpoint $z = (p + q)/2$, which lies in $A \cup B$ because a reflection-closed set is a coset and cosets are closed under halving; if $z \in A$ then (p, z) is a defect pair of A , and if $z \in B$ then (q, z) is one of B . Hence $|A||B| \leq d(A) + d(B) \leq D(c)$, and (4) applies.

Theorem 3.4. *Let n be odd and let c be a four-colouring of $\mathbb{Z}_n$ two of whose classes have reflection-closed union. Then $3R(c) \leq 2n^2$. No surjectivity hypothesis is needed.*

The conjectured extremal family for $r = 4$ and $3 \mid n$ is: two full cosets of the index-3 subgroup, and the third coset split arbitrarily into two classes. The two pieces of the split have reflection-closed union by construction, so Theorem 3.4 covers the whole family, and Proposition 2.3 shows the family attains $2/3$. Exhaustive search confirms that every colouring attaining $3R(c) = 2n^2$ satisfies the hypothesis, at $n = 9$ (all 9 of them) and at $n = 15$ (all 45).

4. EXACT MAXIMA AT SMALL MODULI

The exact values below are obtained by evaluating $R(c)$ at every one of the r^n maps $\mathbb{Z}_n \rightarrow \{1, \dots, r\}$, discarding the non-surjective ones, and taking the maximum; the largest such search is $r = 4$, $n = 8$, i.e. $4^8 = 65\,536$ colourings against $8^2 = 64$ solutions each. The lower bounds at larger n evaluate a single named colouring.

Theorem 4.1. *The four-colour maxima at the moduli $4 \leq n \leq 8$ are*

$$m_4(4) = 8, \quad m_4(5) = 14, \quad m_4(6) = 24, \quad m_4(7) = 24, \quad m_4(8) = 32.$$

In particular

- (1) $m_4(6) = m_3(6) = 24 = \frac{2}{3} \cdot 6^2$: at $n = 6$ a fourth colour gains nothing;
- (2) $\rho_4(6) = 2/3$ and $\rho_4(8) = 1/2$;
- (3) no four-colouring of $\mathbb{Z}_6$ has $R(c) \geq 27$, i.e. $\rho_4(6) < 3/4$.

Moreover $m_4(9) \geq 54$, $m_4(12) \geq 96$ and $m_4(15) \geq 150$, each equal to $\frac{2}{3}n^2$ and witnessed by the mod-3 colouring with one point recoloured.

Remark 4.2. Across the machine-checked range of Theorem 4.1 one has $m_4(n) \leq \frac{2}{3}n^2$ at every n , with equality exactly when $3 \mid n$. The same two-sided conclusion is available at the moduli where $\text{DB}(n)$ is proved (Section 7), the upper bound coming from Proposition 5.2 and, when $3 \mid n$, the matching lower bound from the construction of Proposition 2.3 with one class split in two; in particular $m_4(9) = 54$ and $m_4(15) = 150$. The statement “four colours gain nothing over three” is therefore a statement about the supremum over n and not about each n separately: when $3 \nmid n$ four colours can be strictly better than three — at $n = 5$ the four-colour maximum is 14 against a three-colour maximum of 8, a recorded computation of which only the value 14 is machine-checked — while the supremum is $2/3$ either way and is attained only when $3 \mid n$.

The picture for $r \geq 5$ is different, and it separates odd from even r .

Theorem 4.3. *The exact maxima*

$$m_5(5) = 20, \quad m_5(6) = 24, \quad m_5(7) = 30, \quad m_6(6) = 24$$

hold, so that $\rho_5(5) = 4/5$, strictly above the three-colour answer $2/3$, while $m_5(6) = m_6(6) = m_3(6) = 24$: at $n = 6$ neither five nor six colours — nor even the colouring that uses a different colour at every point — beats three. Moreover

$$m_5(10) \geq 80, \quad m_5(15) \geq 180, \quad m_6(10) \geq 80, \quad m_7(7) \geq 42, \quad m_7(14) \geq 168,$$

the lower bounds attained by $x \mapsto x \bmod p$ for $p = 5, 5, 5, 7, 7$ respectively, with a point recoloured where surjectivity demands it.

Every entry is an instance of the construction of Proposition 2.3: choose an odd $p \leq r$ with $p \mid n$ and colour by $x \mapsto x \bmod p$, recolouring a few points to make the colouring surjective, which the data says costs nothing. Reading the table of exact values,

r	3	4	5	6	7
$\max_n \rho_r(n)$ observed	$2/3$	$2/3$	$4/5$	$4/5$	$6/7$
attained at	$3 \mid n$	$3 \mid n$	$5 \mid n$	$5 \mid n$	$7 \mid n$

suggests the following answer to Question 1.1.

Conjecture 4.4. Let $p(r)$ be the largest odd integer with $p(r) \leq r$. Then $\sup_n \rho_r(n) = 1 - 1/p(r)$, attained exactly when $p(r) \mid n$.

The lower bound half of Conjecture 4.4 is Proposition 2.3. Note what the conjecture says: going from $r = 3$ to $r = 4$ does not help, and neither does going from $r = 5$ to $r = 6$; the gain comes only at odd r . For comparison, a uniformly random colouring gives $r(r-1)(r-2)/r^3$, which is 0.375 at $r = 4$ and 0.48 at $r = 5$; every maximum above beats it.

5. THE DEFECT INEQUALITY

Definition 5.1. A modulus n is *defect-bounded* if for every four-colouring c of $\mathbb{Z}_n$ there are two colours $i \neq j$ with

$$a_i a_j + M(c) \leq S(c), \quad \text{equivalently} \quad a_i a_j \leq D(c).$$

We write $\text{DB}(n)$ for this statement.

Proposition 5.2. Let n be odd. If $\text{DB}(n)$ holds then $3R(c) \leq 2n^2$ for every four-colouring c of $\mathbb{Z}_n$; and if moreover $3 \mid n$ then $m_4(n) \leq 2\lfloor n/3 \rfloor n$, so that $\rho_4(n) = 2/3$ by Theorem 4.1’s construction.

Proof. Immediate from (2) and (4): merging the classes i and j turns the four-class Cauchy–Schwarz $n^2 \leq 4S(c)$ into the three-class $n^2 \leq 3(S(c) + 2a_i a_j)$, and the hypothesis $a_i a_j \leq D(c)$ pays for the extra term, giving (3). $\square$

If some class is empty, the hypothesis of Definition 5.1 is free — take that class as one of the two — and Proposition 5.2 degenerates to Theorem 2.2. So $\text{DB}(n)$ is exactly the content of the step from three colours to four, and nothing else is missing.

Evidence. $\text{DB}(n)$ holds with no violation over all four-colourings of $\mathbb{Z}_n$ up to permutation of the colours for $n = 5, 7, 9, 11, 13, 15$ — 44.7 million colourings at $n = 15$ — and a steepest-descent search over four-colourings for $n \leq 45$ finds none either. Equality holds precisely on the extremal family of Theorem 3.4, so any proof of $\text{DB}(n)$ has to be tight; it is not a slack inequality that can be proved crudely. We know of no prior appearance of the inequality in the literature: it is an *upper* bound on the number of monochromatic three-term progressions in a four-colouring, whereas the published work on rainbow progressions is anti-Ramsey in character — when can a colouring *avoid* a rainbow progression — and states nothing of this shape. A literature pass over arXiv, MathSciNet, zbMATH and OEIS found no earlier statement of the formulation.

Theorem 5.3. $\text{DB}(p)$ holds at every odd prime p . Consequently $3R(c) \leq 2p^2$ for every four-colouring of $\mathbb{Z}_p$.

Proof sketch. Pollard’s theorem [12] in $\mathbb{Z}_p$ states that for $1 \leq t \leq \min(|X|, |Y|)$,

$$\sum_s \min(t, r_{X,Y}(s)) \geq t \cdot \min(p, |X| + |Y| - t), \quad r_{X,Y}(s) = |\{(x, y) \in X \times Y : x + y = s\}|,$$

which at $t = 1$ is Cauchy–Davenport. Read $d(A)$ as the representation mass of $A + A$ that misses $2A$; since $|2A| = |A|$, truncating Pollard’s sum at level t leaves $t(|A| - t)$ outside, and taking $t = \lceil |A|/2 \rceil$ gives

$$d(A) \geq \left\lceil \frac{|A|}{2} \right\rceil \cdot \min\left(p - |A|, \left\lfloor \frac{|A|}{2} \right\rfloor\right),$$

which below density $2/3$ is $\lfloor |A|^2/4 \rfloor$, degrades gracefully above it, and is 0 at $A = \mathbb{Z}_p$ as it must be. Summing this over all four classes suffices, because for sizes $a \leq b \leq c \leq d$ summing to p ,

$$\frac{a^2 + b^2 + c^2 + d^2}{4} \geq ab,$$

with equality exactly when all four are equal — which four sizes summing to an odd p never are. So a quarter of the sharp per-class constant is enough, and the parity of p supplies the slack. The margin is thin: at $p = 5$ the four-size inequality is an equality. One class can be too large for the quarter form, and is then handled alone against its small complement. $\square$

The step from a prime to a composite modulus is not a matter of effort.

Proposition 5.4. (1) *A reflection-closed set has defect 0; in particular $d(A) = 0$ whenever A is a coset, so no bound of the form $d(A) \geq \gamma|A|^2$ with $\gamma > 0$ holds for all A .*
 (2) *The strict per-class bound $|A|^2 \leq 2d(A)$ fails at every singleton, and $d(\{0, 1, 2\}) = 4$ in $\mathbb{Z}_{13}$, one short of $9/2$; so the best possible per-class constant in the sparse range is $\lfloor |A|^2/2 \rfloor$ and not $|A|^2/2$.*
 (3) *Pairing the two smallest classes is not enough: there is a four-colouring of $\mathbb{Z}_{13}$ whose two smallest classes have $d(A_0) + d(A_1) < |A_0||A_1|$, both meeting the sparse bound with equality.*
 (4) *Let $f: \mathbb{N} \rightarrow \mathbb{N}$ be any function with $f(|A|) \leq d(A)$ for every $A \subseteq \mathbb{Z}_{15}$. Then there is a four-colouring c of $\mathbb{Z}_{15}$, with class sizes $2, 3, 5, 5$, such that $\sum_k f(a_k) < a_i a_j$ for all $i \neq j$. Hence no size-only bound of any kind can prove $\text{DB}(15)$; the same holds at $n = 21$ and $n = 27$.*

Part (4) is not an artefact of a weak f : one may take f to be the true minimum of the defect over sets of each size, and the sizes 5, 5, 3 each admit a coset. What excludes the bad configuration is that three cosets of orders 5, 5, 3 cannot be pairwise disjoint in $\mathbb{Z}_{15}$, and that is a joint statement about the classes, invisible to any function of the sizes alone. Part (1) is also the reason the obstruction is not merely one of technique: Chase [3] bounds the number of three-term progressions inside a subset of any finite abelian group in terms of the relative density alone, and his bound is trivial at a coset, as it must be.

6. FIBRE DECOMPOSITION, AND THE TWIST AT $n = 27$

By Proposition 5.4 a proof at composite n must see more than the four class sizes. The right refinement is the 4×3 matrix of *fibre trace sizes*. Let $n = 3m$. Three-term progressions project to three-term progressions along $\pi: \mathbb{Z}_{3m} \rightarrow \mathbb{Z}_3$: if x lies over u and z over w then $2z - x$ lies over $2w - u$. Splitting the defect pairs of a class by the pair of fibres their two coordinates lie in therefore produces nine blocks, the block over (u, w) having its third point constrained to the fibre over $2w - u$. Each block is an *escape count*

$$(5) \quad e(X, Z, Y) = |\{(x, z) \in X \times Z : 2z - x \notin Y\}|,$$

so that $d(A) = e(A, A, A)$.

Lemma 6.1. *Let $\gcd(3, m) = 1$. Then reduction modulo m is injective on every fibre of π and maps it onto $\mathbb{Z}_m$, so every block transports to an escape count between subsets of $\mathbb{Z}_m$, and*

$$\sum_i d(A_i) = \sum_i \sum_{u, w \in \mathbb{Z}_3} e(B_{i,u}, B_{i,w}, B_{i,2w-u}),$$

where $B_{i,u} \subseteq \mathbb{Z}_m$ is the trace of the class A_i on the fibre over u . The class sizes add up the same way, and for each u the four traces partition $\mathbb{Z}_m$. Consequently the following statement, internal to $\mathbb{Z}_m$, implies DB(3m): for every family $(B_{i,u})$ with each fibre partitioned into four possibly empty parts and every colour used somewhere, some pair $i \neq j$ satisfies $(\sum_u |B_{i,u}|)(\sum_u |B_{j,u}|) \leq \sum_k \sum_{u,w} e(B_{k,u}, B_{k,w}, B_{k,2w-u})$.

The point of the reduction is that its hypothesis is a function of the 4×3 matrix of trace sizes rather than of the four class sizes, so it escapes Proposition 5.4(4).

The hypothesis $\gcd(3, m) = 1$ is not cosmetic. At $m = 9$ the extension $0 \rightarrow \mathbb{Z}_9 \rightarrow \mathbb{Z}_{27} \rightarrow \mathbb{Z}_3 \rightarrow 0$ does not split, reduction modulo 9 is not injective on a fibre, and the transport of Lemma 6.1 cannot be instantiated at all. The decomposition nevertheless exists, twisted by the 2-cocycle of the extension.

Choose the set-theoretic section $\psi(u, s) = u + 3s$ of $\mathbb{Z}_{27} \rightarrow \mathbb{Z}_3$, with $u \in \mathbb{Z}_3$ and $s \in \mathbb{Z}_9$. A direct computation gives the transport rule

$$(6) \quad 2\psi(w, t) - \psi(u, s) = \psi(2w - u, 2t - s + \sigma(u, w)),$$

where $\sigma: \mathbb{Z}_3 \times \mathbb{Z}_3 \rightarrow \mathbb{Z}_9$ is the 2-cocycle of the extension. It vanishes on the diagonal and takes the values

$$\sigma(0, 1) = 0, \quad \sigma(0, 2) = 1, \quad \sigma(1, 0) = 8, \quad \sigma(1, 2) = 1, \quad \sigma(2, 0) = 8, \quad \sigma(2, 1) = 0$$

off it. A block is therefore not an escape count but a *shifted* escape count

$$(7) \quad e_c(X, Z, Y) = |\{(x, z) \in X \times Z : 2z - x + c \notin Y\}|, \quad X, Z, Y \subseteq \mathbb{Z}_9, \quad c \in \mathbb{Z}_9,$$

and $e_0 = e$. Writing $\text{tr}(A, u) \subseteq \mathbb{Z}_9$ for the trace of $A \subseteq \mathbb{Z}_{27}$ on the fibre over u , read through ψ , we obtain the following exact identity.

Proposition 6.2. *For every $A \subseteq \mathbb{Z}_{27}$,*

$$d(A) = \sum_{u \in \mathbb{Z}_3} \sum_{w \in \mathbb{Z}_3} e_{\sigma(u,w)}(\text{tr}(A, u), \text{tr}(A, w), \text{tr}(A, 2w - u)).$$

No coprimality hypothesis is needed: the failure of the extension to split is absorbed into the shifts. Consequently, the analogue of the statement in Lemma 6.1 with e replaced by $e_{\sigma(u,w)}$ implies DB(27).

Proposition 6.3. *The cocycle σ is not a coboundary: there is no $f: \mathbb{Z}_3 \rightarrow \mathbb{Z}_9$ with*

$$\sigma(u, w) + 2f(w) - f(u) - f(2w - u) = 0 \quad \text{for all } u, w \in \mathbb{Z}_3.$$

Established by evaluating the condition at each of the $9^3 = 729$ maps f . Concretely, the two shifts of a pair of blocks aiming at the same target fibre always differ by a unit modulo 3: the three pairs, indexed by their target, carry shift pairs $(0, -1)$, $(1, -1)$ and $(1, 0)$.

So no change of section removes the shifts. The next proposition says they cannot be ignored either.

Proposition 6.4. (1) *A shift is a translation of the first set: $e_c(X, Z, Y) = e(X - c, Z, Y)$. Hence the minimum of e_c over sets of prescribed sizes does not depend on c , and no bound that sees one block at a time can see the shift at all.*
 (2) *With $H = \{0, 3, 6\} \subseteq \mathbb{Z}_9$ one has $e(H, H, H) = 0$, and $e(\{0, 3\}, \{0, 3\}, \{0, 6\}) = 1$ is the exact minimum at sizes $(2, 2, 2)$.*
 (3) *Consequently every blockwise minimum is attained simultaneously on the profile all three of whose columns are $(1, 2, 3, 3)$, where the resulting cap is $3 \cdot 2 + 6 \cdot 1 = 12$, against a minimum pair product of 18. No blockwise size-only bound can prove DB(27).*

Part (3) is the exact deficit of 6 that stopped this route before, and it is an artefact of the untwisted transport rather than of DB(27). What is needed is a bound on the six off-diagonal blocks of a class *jointly*, since they share the three traces $\text{tr}(A, 0), \text{tr}(A, 1), \text{tr}(A, 2)$ and the shifts obstruct their minima being attained at once. Two ingredients supply it. The first is an off-diagonal bound that survives a composite fibre modulus.

Proposition 6.5. *Let N be odd and $X, Z, Y \subseteq \mathbb{Z}_N$. Writing $\theta(X, Z, Y) = |X||Z| - e(X, Z, Y)$ for the number of pairs of $X \times Z$ whose reflection lands in Y ,*

$$\theta(X, Z, Y) \leq |Y| \cdot \min(|X|, |Z|).$$

In particular, for $X, Z, Y \subseteq \mathbb{Z}_9$ and any $c \in \mathbb{Z}_9$,

$$|X||Z| \leq e_c(X, Z, Y) + |Y| \cdot \min(|X|, |Z|).$$

The proof runs the injection $(x, z) \mapsto (x, 2z - x)$ in *both* coordinates, which is what removes the primality that the previously available off-diagonal bound needed; composed with Proposition 6.4(1) it applies verbatim to a shifted block. The second ingredient is where the twist becomes a number.

Theorem 6.6. *Let $X, Z, Y \subseteq \mathbb{Z}_9$ with $|X| = |Z| = |Y| = 3$. Then*

$$8 \leq e_{\sigma(0,1)}(X, Z, Y) + e_{\sigma(1,0)}(Z, X, Y).$$

Proof sketch. A single one of the two blocks can vanish: $e(H, H, H) = 0$ for the order-three subgroup H . But $e(X, Z, Y) = 0$ with $|X| = |Z| = |Y| = 3$ forces X, Z, Y to be cosets of H , with the coset of Y determined by those of X and Z ; and by Proposition 6.3 the two blocks of a pair name *different* cosets, because their shifts differ by a unit modulo 3. So at most one of the two can escape completely. That the exact joint minimum is 8 is established by evaluating the right-hand side at every one of the $\binom{9}{3}^3 = 84^3 = 592\,704$ triples of three-element subsets of $\mathbb{Z}_9$. $\square$

The single cell $(3, 3, 3) \mapsto 8$ of Theorem 6.6 is the entry that turns the $n = 27$ check from a deficit of 6 into a pass at 0. For the check itself the whole table is needed.

Proposition 6.7. (1) (Affine transport.) *Let $l \in (\mathbb{Z}_9)^\times$ and $s, t, c, c' \in \mathbb{Z}_9$ with $lc' + t = s + c$. Then for all $X, Z, Y \subseteq \mathbb{Z}_9$,*

$$e_c(lX + s, lZ + s, lY + t) = e_{c'}(X, Z, Y).$$

- (2) (The pair table.) *For all $X, Z, Y \subseteq \mathbb{Z}_9$ of sizes at most 3, the joint minimum of the two blocks aiming at one fibre is the value $P(|X|, |Z|, |Y|)$ of the table*

$(X , Z) \setminus Y $	0	1	2	3
(1, 1)	2	1	0	0
(1, 2)	4	2	1	0
(1, 3)	6	4	2	1
(2, 2)	8	5	4	2
(2, 3)	12	8	6	4
(3, 3)	18	13	10	8

(symmetric in $|X|, |Z|$; all entries with a size 0 are 0), and each of the three target shapes satisfies $P(|X|, |Z|, |Y|) \leq e_{\sigma(u,w)}(X, Z, Y) + e_{\sigma(w,u)}(Z, X, Y)$.

Part (2) is established by one exhaustive evaluation over all $130^3 = 2\,197\,000$ triples of subsets of $\mathbb{Z}_9$ of size at most 3, the 130 being $1 + 9 + 36 + 84$. Fifteen of the 64 cells beat what Proposition 6.5 gives; the bottom-right 8 is Theorem 6.6, the cell where a single block can be 0 and the pair cannot. One enumeration serves all three target shapes rather than three, and that is what part (1) is for: the three pairs carry shifts whose *differences* are 1, 2, 1, and the affine map $x \mapsto 5x + 5$ — with 5 the unit of $\mathbb{Z}_9$ that does it — converts the 2 into a 1.

Theorem 6.8. *DB(27) holds. Together with the two smaller odd composite moduli divisible by 3,*

$$\text{DB}(15) \wedge \text{DB}(21) \wedge \text{DB}(27).$$

In particular $3R(c) \leq 2 \cdot 27^2$ for every four-colouring of $\mathbb{Z}_{27}$, so $m_4(27) \leq 486$.

Proof sketch. By Proposition 6.2 it suffices to prove the twisted trace statement. The diagonal blocks are honest defects — the cocycle vanishes on the diagonal — and are bounded below as a function of the four trace sizes of a column by the per-class input of Section 5 together with the exact value for a two-part partition. Each of the three off-diagonal pairs is bounded below by the larger of Proposition 6.5 and the table of Proposition 6.7. What is left is one arithmetic inequality about 4×3 matrices of non-negative integers with all three column sums equal to 9: for every such matrix, the smallest product of two row sums is at most the assembled bound. There are $\binom{12}{3} = 220$ possible columns and hence $220^3 = 10\,648\,000$ matrices, and the inequality is verified at each of them. With the completeness of the 130-element list of subsets of size at most 3, that is three finite evaluations in all; every other step is a proof. $\square$

Three controls say that the ingredients of Theorem 6.8 are the right ones and that its conclusion is sharp.

- Proposition 6.9.** (1) (The pair table is needed.) *With only Proposition 6.5 behind the off-diagonal blocks, the assembled bound on the profile all of whose columns are (1, 2, 3, 3) is 6, against a minimum pair product of 18.*
- (2) (The box is exactly right.) *With the table of Proposition 6.7 truncated to sizes at most 2, the assembled bound on the profile with columns (0, 3, 3, 3), (0, 3, 3, 3), (3, 1, 2, 3) is 20, against a minimum pair product of 21; with the full table the same profile clears at 59.*
- (3) (The critical entry is not free.) *At sizes (3, 3, 3) the bound of Proposition 6.5 is 0, a single block really is 0, and the table gives 8.*
- (4) (Sharpness.) *For the colouring of $\mathbb{Z}_{27}$ consisting of two whole residue classes modulo 3 with the third split in two, every pair product is at least $S(c)$; so the inequality supplied by DB(27) is an equality there and cannot be strengthened.*

7. SYMMETRY, AND THE MODULI 39, 51, 57, 69

For $n = 3m$ with m an odd prime, Lemma 6.1 together with the per-class input of Section 5 reduces DB($3m$) to a purely arithmetic statement: for every 4×3 matrix $(n_{i,u})$ of non-negative

integers all of whose column sums equal m , writing $r_i = \sum_u n_{i,u}$ for the row sums,

$$(8) \quad \min_{i \neq j} r_i r_j \leq \Lambda(m; n_{i,u}),$$

where Λ is a sum of one term per column (the diagonal blocks) and one per row (the off-diagonal blocks). There is no additive combinatorics left in (8); it is one inequality about integer matrices, and the only obstacle is the number of matrices. The columns are the $\binom{m+3}{3}$ ways of writing m as an ordered sum of four non-negative integers, so a naive sweep is $\binom{m+3}{3}^3$, which is $1.76 \cdot 10^{10}$ already at $m = 23$.

Both symmetries of (8) can be spent on that enumeration. Λ is invariant under permuting the three fibres, because the row terms are symmetric in their arguments and the column terms are merely permuted; and under permuting the four colours, because the column terms are symmetric functions of a column's entries and the row terms are permuted, while the conclusion is a disjunction over *unordered* pairs. The fibre group S_3 divides the enumeration by walking the three columns in non-decreasing list position, a clean factor of 6. The colour group S_4 acts on all three columns at once and so cannot be pushed into the loop nesting; what it can do is canonicalise the entries of the *first* column.

Proposition 7.1. *Let $\text{cols}(m)$ be the set of four-tuples of non-negative integers summing to m , and let e_2 denote the second elementary symmetric function of a column's four entries. If (8) holds at every triple (A, B, C) of columns for which A is non-decreasing and $e_2(A) \leq e_2(B)$, $e_2(A) \leq e_2(C)$, then it holds at every triple.*

Proof sketch. For the colour half, put $\text{wgt}(A) = a_1 + 2a_2 + 3a_3$ for $A = (a_0, a_1, a_2, a_3)$. Each of the three adjacent transpositions, applied when the corresponding pair is inverted, strictly increases wgt , and $\text{wgt}(A) \leq 3m$ on $\text{cols}(m)$; induct downwards on $3m - \text{wgt}(A)$, the base case $\text{wgt}(A) = 3m$ forcing A to be sorted already. This replaces a four-element sorting network by one induction with four branches. For the fibre half, the remaining freedom is which of the three columns is placed first, and it may be spent on any function of a column that a colour permutation cannot change; e_2 is such a function, and requiring the first column to carry the least e_2 is a further genuine reduction. The two conditions do not interfere precisely because e_2 is colour-invariant: the descent that sorts the first column moves no e_2 at all. $\square$

Theorem 7.2. *DB(39), DB(51) and DB(57) hold.*

The three are the fibre moduli $m = 13, 17, 19$. With only the S_3 reduction, the enumerations are $2.94 \cdot 10^7$, $2.48 \cdot 10^8$ and $6.10 \cdot 10^8$ column triples respectively, and each is checked exhaustively. The modulus $m = 19$ needs one further ingredient: the bound Λ built from the per-class input alone *fails* there, on a profile whose binding feature is a fibre carrying exactly three colours. For three non-empty parts of a fibre, with sizes (x, y, z) and the fourth part empty, an exact identity plus a crude count of the rainbow solutions gives

$$2xz + 2(x+z)y \leq 2 \sum_k d(B_k) + 3 \min(x, y, z) \cdot m,$$

and turning that into a function of the size matrix — which is what (8) consumes — has one design constraint: the term must be invariant under permuting the four colours. Writing it as a single formula guarded by “some part is empty” rather than as a four-way case distinction achieves that, because $2(xy + yz + zx)$ is $2e_2$ of all four sizes when one of them is 0, and $\min(x, y, z)$ is the second smallest of all four. With this term included, $m = 19$ passes.

Theorem 7.3. *DB(69) holds. It is the first modulus past the previous enumeration frontier.*

Here $m = 23$, $|\text{cols}(23)| = 2600$, and the transversal of Proposition 7.1 visits 176 252 004 triples against the 2 932 714 200 of the S_3 reduction alone; the inequality (8) is verified at each of them.

Proposition 7.4. (1) *The guarded leaf — the form of (8) restricted to the case where colours 0 and 1 carry the two smallest classes — is not colour-invariant, so it is unsound over a colour transversal. At $m = 19$ the triple of columns $(0, 0, 0, 19)$, $(0, 8, 10, 1)$, $(19, 0, 0, 0)$ lies in the transversal, the guarded form holds vacuously there, and (8) is false at it.*

(2) *The transversal is not vacuous: run against the weaker bound Λ that omits the three-colour term, the same walk returns false at $m = 19$, exactly where that bound is known to fail.*

(3) *The transversal cannot be as short as $|\text{cols}(m)|^3/144$, because the action of $S_3 \times S_4$ is not free: at $m = 23$ one has $|\text{cols}(23)|^3 < 144 \cdot 176\,252\,004$. Triples with all three columns equal are fixed by the whole S_3 factor, and profiles whose colour-0 and colour-1 rows agree are fixed by a transposition.*

Status. Collecting Theorems 5.3, 6.8, 7.2 and 7.3, $\text{DB}(n)$ is now known at every odd prime and at $n = 9, 15, 21, 27, 33, 39, 51, 57, 69$.

Two entries of that list are not covered by the four theorems just cited but belong to the same formal development: $\text{DB}(9)$ is proved directly from the exact per-class defect table of $\mathbb{Z}_9$, and $\text{DB}(33)$ by the check (8) at fibre modulus $m = 11$. By Proposition 5.2, $3R(c) \leq 2n^2$ for every four-colouring of $\mathbb{Z}_n$ at each of these moduli; at those divisible by 3 the construction of Proposition 2.3, with one of its classes split in two, supplies the matching lower bound, so $\rho_4(n) = 2/3$ there. Three cases remain open, and they are of three different kinds. For $n = 3p$ with $p \geq 29$ prime, the route is the one used here and the obstacle is the size of the enumeration. For n divisible by 9 and larger than 27 — 45, 63, 81 — the twisted decomposition of Section 6 would have to be written at a general fibre modulus, and its finite inputs are modulus-specific and rapidly grow past what can be enumerated. For odd composite n with $3 \nmid n$, of which 25 is the smallest, there is no fibre structure over $\mathbb{Z}_3$ at all and nothing above applies.

8. VERIFICATION

Every statement in this paper is formalised and machine-checked in Lean 4 against Mathlib [8, 9]. Definitions 2.1 and 5.1, and the statements of all lemmas, theorems and propositions, are stated in the formal development in the form printed here; the proofs are complete, with no unproved assumptions and no appeals to unformalised lemmas. The finite evaluations named in the text — the exact maxima of Section 4, the 729 sections of Proposition 6.3, the 84^3 and 130^3 triples behind Theorem 6.6 and Proposition 6.7, the 220^3 matrices behind Theorem 6.8, and the enumerations of Theorems 7.2 and 7.3 — are carried out by evaluating the corresponding decision procedures — the smaller ones inside the kernel, the larger ones by compiled evaluation — against definitions that are proved to express the stated claims; every other step is checked by the kernel. The repository reference to be added at submission.

REFERENCES

- [1] S. Butler, C. Erickson, L. Hogben, K. Hogenson, L. Kramer, R. L. Kramer, J. C.-H. Lin, R. R. Martin, D. Stolee, N. Warnberg and M. Young, *Rainbow arithmetic progressions*, J. Combinatorics **7** (2016), no. 4, 595–626; arXiv:1404.7232.
- [2] P. J. Cameron, J. Cilleruelo and O. Serra, *On monochromatic solutions of equations in groups*, Rev. Mat. Iberoam. **23** (2007), no. 1, 385–395.
- [3] Z. Chase, *The maximum number of three-term arithmetic progressions, and triangles in Cayley graphs*, arXiv:1809.03729.
- [4] E. Croot, *The minimal number of three-term arithmetic progressions modulo a prime converges to a limit*, Canad. Math. Bull. **51** (2008), no. 1, 47–56; arXiv:math/0501004.
- [5] G. Elvin, A. Gonzales, A. Rodriguez and I. Wilbur, *Bounds on arithmetic rainbow Ramsey multiplicities*, preprint, arXiv:2601.05442 (2026).
- [6] B. Green and O. Sisask, *On the maximal number of three-term arithmetic progressions in subsets of $\mathbb{Z}/p\mathbb{Z}$* , Bull. Lond. Math. Soc. **40** (2008), 945–955; arXiv:0709.4432.
- [7] V. Jungić, J. Licht, M. Mahdian, J. Nešetřil and R. Radoičić, *Rainbow arithmetic progressions and anti-Ramsey results*, Combin. Probab. Comput. **12** (2003), 599–620.

- [8] L. de Moura and S. Ullrich, *The Lean 4 theorem prover and programming language*, in: Automated Deduction — CADE 28, LNCS **12699**, Springer, 2021, 625–635.
- [9] The mathlib Community, *The Lean mathematical library*, in: Proc. 9th ACM SIGPLAN Int. Conf. on Certified Programs and Proofs (CPP 2020), 367–381.
- [10] A. Montejano and O. Serra, *Counting patterns in colored orthogonal arrays*, Discrete Math. **317** (2014), 44–52; arXiv:1104.0190.
- [11] O. Parczyk and C. Spiegel, *An unsure note on an un-Schur problem*, Electron. J. Combin. **33** (2026), #P45; arXiv:2410.22024.
- [12] J. M. Pollard, *A generalisation of the theorem of Cauchy and Davenport*, J. London Math. Soc. (2) **8** (1974), 460–462.
- [13] L. Versteegen, *Common and Sidorenko equations in Abelian groups*, J. Combinatorics **14** (2023), no. 1, 53–67; arXiv:2109.04445.