On rainbow arithmetic progressions

Maria Axenovich

Department of Mathematics
Iowa State University
USA
axenovic@math.iastate.edu

Dmitri Fon-Der-Flaass

Department of Mathematics
University of Illinois at Urbana-Champaign
USA
and
Institute of Mathematics, Novosibirsk
Russia
flaass@math.uiuc.edu

Submitted: June 23, 2003; Accepted: December 2, 2003; Published ??.
MR Subject Classifications: 11B25, 11B75

Abstract

Consider natural numbers $\{1, \cdots, n\}$ colored in three colors. We prove that if each color appears on at least $(n+4)/6$ numbers then there is a three-term arithmetic progression whose elements are colored in distinct colors. This variation on the theme of Van der Waerden's theorem proves the conjecture of Jungić et al.

1 Introduction

In this paper we investigate the colorings of sets of natural numbers. We say that a subset is monochromatic if all its elements have the same colors and we say that it is rainbow is all its elements have distinct colors. A famous result of van der Waerden [3] can be reformulated the following way.

Theorem 1. *For each pair of positive integers k and r there exists a positive integer M such that in any coloring of integers $1, \cdots, M$ into r colors there is a monochromatic arithmetic progression of length k.*

This theorem was generalized by the following very strong statement of Szemerédi [2].

Theorem 2. *For every natural number k and positive real number δ there exists a natural number N such that every subset of $\{1, \cdots, N\}$ of cardinality at least δN contains an arithmetic progression of length k.*

One can ask a "dual" question. Assume again that $\{1, \cdots, n\}$ is colored into r colors. Can we find an arithmetic progression of length k so that all its elements are colored in distinct colors? Next, we call such colored arithmetic progressions *rainbow* $AP(k)$.

In general, the answer to this question is "No", for $r \leq \lfloor log_3 n + 1 \rfloor$. The following coloring c of $\{1, \cdots, n\}$, given in [1], demonstrates this fact. Let $c(i) = \max\{q : i$ is divisible by $3^q\}$. This coloring is easy to show not to have any rainbow arithmetic progressions of length at least 3.

It turns out that in order to force a rainbow AP(k) we need to ensure that for each color there are "many" elements having this color. So, while Szemerédi's theorem requires only one color class to have large cardinality to ensure the existence of monochromatic AP(k), we need each color class to have large cardinality to force rainbow AP(k).

This problem was studied by Jungić et. al. [1] in the infinite case. It was shown that if the natural numbers are colored in three colors and the upper density of each color is greater than 1/6 then there is a rainbow $AP(3)$. Similar results were obtained for $\mathbb{Z}_n$. When $\{1, \cdots, n\}$ is colored in three colors, Jungić et. al. [1] conjectured that if each color class has cardinality at least $(n+4)/6$ then there is a rainbow $AP(3)$.

In this paper we investigate the conditions on a coloring of $\{1, \cdots, n\}$ forcing rainbow $AP(3)$ and prove the conjecture of Jungić et. al. Next we denote $[n] = \{1, \cdots, n\}$.

Let $c : [n] \to \{A, B, C\}$ be a coloring of $[n]$ in three colors. Let $M(c)$ be the cardinality of the smallest color class in c. We define $M(n)$ to be the largest $M(c)$ over all colorings c of $[n]$ in three colors with no rainbow $AP(3)$. The following construction, given in [1], provides a coloring with large $M(c)$ and no rainbow $AP(3)$.

$$c(i) = \begin{cases} A & \text{if} \quad i \equiv 1 \pmod 6 \\ C & \text{if} \quad i \equiv 4 \pmod 6 \\ B & \text{otherwise.} \end{cases} \tag{1}$$

This coloring has $M(c) = \lfloor (n+2)/6 \rfloor$. When $n = 6k - 4$, there exists a slightly better coloring, with $M(c) = k = \lfloor (n+4)/6 \rfloor$:

$$c(i) = \begin{cases} A & \text{for} \quad \text{odd} \quad i \in \{1, \cdots, 2k-1\} \\ C & \text{for} \quad \text{even} \quad i \in \{4k-2, \cdots, 6k-4\} \\ B & \text{otherwise.} \end{cases} \tag{2}$$

We prove that $M(c)$ can not be made larger without forcing a rainbow $AP(3)$.

Theorem 3. $M(n) \leq (n+4)/6$.

2 Proof of theorem 3

Consider a coloring c of the interval $I = \{1, \cdots, n\}$ in three colors A, B, C such that there is no rainbow arithmetic progression of length three. We say that there is a string, or a word $\mathbf{x} = (c_1, c_2, \cdots, c_m) \in \{A, B, C\}^m$ at a position i (or that $\mathbf{x}$ occupies position i) if $c(i) = c_1$, $c(i+1) = c_2, \cdots, c(i+m-1) = c_m$. We say that there is a word $\mathbf{x}$ in the coloring if there is $\mathbf{x}$ at some position i. There is no $\mathbf{x}$ in the coloring if there is no $\mathbf{x}$ at any position i. We denote by $|A|, |B|, |C|$ the number of elements colored A, B and C respectively.

Lemma 1. *In any interval colored with $\{A, B, C\}$ and no rainbow $AP(3)$'s for any $X, Y \in \{A, B, C\}$, between any two occurrences of XX and YY, $X \neq Y$, there is an occurrence of XY or YX.*

Proof. Take two closest occurrences of XX and YY and assume without loss of generality that they occur at positions 1 and m respectively, i.e., $c_1 = c_2 = X$, $c_m = c_{m+1} = Y$. Let $J \subseteq \{2, \cdots, m\}$ be the set of positions of letters X and Y; we only need to show that J contains two adjacent positions.

If m is odd then both $(m+1)/2$ and $(m+3)/2$ are in J. So, let $m = 2(k-1)$ be even. We have $k = (2+m)/2 \in J$; without loss of generality let $c(k) = Y$. If $i < k$ is in J then $c_i = X$ (otherwise both $2i-1$ and $2i-2$ are in J). Similarly, $c_i = Y$ for each $i > k$, $i \in J$.

Define the function $f(i) = \lfloor (2k - i + 2)/2 \rfloor$. If $i \in J \cap \{3, \cdots, k\}$ then $f(i) \in J \cap \{3, \cdots, k-1\}$: indeed, $c(2k - i) = Y$ because of the progression $(i, k, 2k - i)$; and then $c(f(i)) = X$ because of the progression $(2, f(i), 2k - i)$ or $(1, f(i), 2k - i)$, depending on the parity of i.

Iteratively applying f to the initial value k, we obtain a sequence of elements of J. If the first repetition in this sequence is $f^a(k) = f^b(k)$ then $f^{a-1}(k) \neq f^{b-1}(k)$, and $f(f^{a-1}(k)) = f(f^{b-1}(k))$. This implies that $f^{a-1}(k)$ and $f^{b-1}(k)$, both elements of J, differ by 1, and the lemma is proved. □

We say that a color Y is a *dominating color* if whenever $c(i) \neq c(i+1)$, $1 \leq i \leq n-1$, either $c(i) = Y$ or $c(i+1) = Y$. Next we treat two cases: when c has a dominating color and when it does not.

2.1 There is a dominating color A in c.

This means that there are no subwords BC or CB. We treat the following two cases:
Case 1. There is no subword BB and no subword CC.
Subcase 1.0. Every B or C, except possibly the last one, is followed by at least two A's. Then $|B| + |C| \leq (n+2)/3$, and therefore either $|B|$ or $|C|$ is at most $(n+2)/6$.

Since the words BAC and CAB are forbidden, we can assume without loss of generality that there is BAB at a position i. Now, all C's must occupy positions of the opposite parity. Otherwise, take C at a position $j \equiv i \pmod 2$. Now, $(i+j)/2$ and $(i+2+j)/2$ can not be colored A, and we have two of B and C next to each other which contradicts our assumption for this case.

Subcase 1.1. There is BAB as well as CAC.
Then from the above we have that all B's occupy positions of the same parity, and all C's occupy positions of the opposite parity. Consider the minimum distance d between two numbers colored with B and C. Note that d is odd. Assume that $c(x) = B$, $c(x+d) = C$. Then, if $x + 2d \in I$ we have $c(x + 2d) = B$ (because of the arithmetic progression $(x, x+d, x+2d)$), and similarly if $x-d \in I$, $c(x-d) = C$. Continuing in the same manner in both directions, we get positions $(i_0, i_0+d, \cdots, i_0+pd)$ with $i_0 \leq d$, $i_0+pd \geq n+1-d$, at which the entries are alternatively B and C, and all other entries between i_0 and i_0+pd are A's. Assume that $c(i_0) = B$ and $i_0 \leq n - pd - i_0 + 1$.

We see that the subwords BAB and CAC can occur only in the segments $\{1, \cdots, i_0\}$ and $\{i_0+pd, \cdots, n\}$ respectively. Therefore $c(i_0+pd) = C$; so p is odd. Moreover, $p = 1$; otherwise for $i < i_0$ such that $c_i = B$ we would get a rainbow AP $(i, i_0+d, i_0+2d+(i_0-i))$. So, we have

$$|B| + |C| \leq \frac{i_0+1}{2} + \frac{n-i_0-d+2}{2} = \frac{n-d+3}{2}.$$

Since $3d \geq i_0 + 2d \geq n+1$,

$$|B| + |C| \leq \frac{n+4}{3},$$

and either $|B|$ or $|C|$ is at most $(n+4)/6$.
Subcase 1.2.
There is BAB but no CAC.
All positions of C's are of the same parity. For each $c_i = B$, take the one-element set $\{i\}$ or $\{i+1\}$: whichever of them has this parity. For each $c_i = C$, take the 2-element set $\{i, i+2\}$. By the hypotheses for this case, all these sets are disjoint. Therefore, $|B| + 2|C| \leq (n+3)/2$. It follows that at least one of $|B|$, $|C|$ must not exceed $(n+3)/6$.
Case 2. There is a subword BB but no subword CC.
We know also that the distance between any B and any C is at least 3. The main observation here is that if we have BB at a position i and C at a position j, and both $2j-i$ and $2j-i-1$ belong to $[n]$, then there is BB at a position $(2j-i-1)$. Call it the reflection of BB in C.

Now, let $J_1, \cdots, J_k$ be maximal intervals in $\{1, \cdots, n\}$ not containing BB. Clearly, J_is are disjoint. We assume that J_i starts before J_j for $i < j$. Our goal is to show that each such interval does not contain "too many" C's.

First consider an inner interval $J = J_m = [i \ldots j]$, where $i \neq 1$ and $j \neq n$. By construction, we have $c_{i-1} = c_i = c_j = c_{j+1} = B$, and J has no BB. If there are no C's in I, we are done. If $c_k = C$, $i < k < j$, and $k \neq (i+j)/2$, then the reflection of the closest to k BB in $c_k = C$ is inside I, which is impossible. So, we can have at most one C, right in the middle if I. Both $j-k$ and $k-i$ are at least 3 thus $|J| \geq 7$.

Second, consider an end-interval $J = J_1 = c_1 \ldots c_j$, with BB at a position j. The above "reflection argument" tells us that C's can appear only in the left half of J, i.e., if $c_k = C$ then $k \leq (1+j)/2$. And the distance between any two C's in I is at least 3. Indeed, if $c_k = c_{k+2} = C$, then one of j, $j+1$ is of the same parity as k, say j. Then $(k+j)/2$ and $(k+2+j)/2$ are two consecutive numbers colored with B or C, a contradiction. Treating the other end-interval J_k in a similar manner, we have the total number of C's in $[n]$ being at most $l_1/7+(l_2+2)/6$ where l_1 is the total length of inner intervals and l_2 is the total length of end-intervals. Thus the number of C's at most $(n+2)/6$.

The last possibility, when there are both BB and CC, cannot occur, by Lemma 1.

2.2 There is no dominating color in c.

Let w be the shortest subinterval of I containing all three adjacencies AB, BC, CA. To simplify the notations, in this subsection we will shift the indexing in such a way that $w = \{1, \cdots, n'\}$; and the whole word is indexed from a to b, $b-a+1=n$. We shall refer to the interval $[a,0]$ as the left part and the interval $[n'+1,b]$ as the right part if such exist. For $\{X,Y,Z\} = \{A,B,C\}$, we assume that $c_1 = X$ and $c_2 = Y$. Consider the first appearance of X after position 1. If it is preceded by Y in a position j, then the word $w \setminus \{1\}$ is a shorter word containing all adjacencies. Thus X must be preceded by Z at a position j. Again, if $j \neq n'-1$ then $w \setminus \{n'\}$ contains all three adjacences and it is shorter than w.

Thus w satisfies the following hypothesis:
$(*)$ w is colored $XY \cdots ZX$ and has no X's inside.

Now, we assume that the word satisfying (*) is the one with $X = A$, $Y = B$ and $Z = C$. We shall show that the number of A's is small.

Claim 1. For $i \geq 1$, $c(1+2^i) = B$ whenever $1+2^i < n'$. Symmetrically, $c(n'-2^i) = C$ whenever $n' - 2^i > 1$. It is easy to see by induction, successively considering $AP(3)$s $(1, 2^i+1, 2^{i+1}+1)$.

Claim 2. Let k be the first occurrence of C in w. Then k is even. Symmetrically, if l is the last occurrence of B then $n'-l$ is odd. Otherwise, $AP(3)$s $(1,(1+k)/2,k)$, $(l,(n'+l)/2,n')$ are rainbow.

Claim 3. If $n' = 2m$ then $c(m+1) = B$ and $c(m) = C$. Otherwise, $AP(3)$'s $(2, m+1, 2m)$ and $(1, m, 2m-1)$ are rainbow.

Claim 4. If $n' = 2m$ then w is the whole word, that is, it cannot be extended to either side and $a = 1$, $b = n = n'$. Indeed, using Claims 2 and 3, the following $AP(3)$s: $(0,1,2)$, $(0,k/2,k)$, $(0,m,2m)$ give $c(0) \neq C$; $c(0) \neq A$, and $c(0) \neq B$ respectively, which is impossible. The symmetric argument works for $c(2m+1)$.

Claim 5. $|w| \geq 8$. Indeed, since $c_3 = B$ and $c_{n'-2} = C$, we have $n' \geq 6$. If $n' = 7$ then positions $3,5,7$ give a rainbow $AP(3)$. When $n' = 8$, we find the unique possibility, the word $ABBCBCCA$, which satisifies the conclusion of the theorem: $|A| = 2 = (8+4)/6$.

Now, by Claims 4 and 5, the theorem is proved for even n'. So we can assume that n' is odd.

Let $n' = p \cdot 2^i + 1$ for odd p. Consider the sequence $w' = (1, 1+2^i, \cdots, 1+p\cdot 2^i)$ of length $p+1$.

Claim 6. $p \neq 1$. Assume that $n' = 2^i+1$ then $c(1+2^{i-1}) = B$ and $c(1+2^i-2^{i-1}) = C$ using Claim 1 and the fact that $i \geq 2$. Since $1+2^{i-1} = 1+2^i-2^{i-1}$, this is impossible.

Next we assume that $p \geq 3$.

Claim 7. The hypothesis $(*)$ is satisfied by w'. For $p \geq 3$, $1+2^i < n'-2^i = 1+(p-1)2^i$. Thus using Claim 1 we have that w' begins with AB and ends with CA. Obviously, there are no A's inside w', and it's rainbow $AP(3)$-free.

Claim 8. The smallest index of a letter in the original word is at least $2-2^i$; symmetrically, the largest is at most $n'+2^i-1$. Moreover, Claim 5 implies that $p \geq 7$. Since w' is of even length, it cannot be extended to either side by Claim 4. This means that $c(1-2^i)$ and $c(n'+2^i)$ cannot be defined.

Claim 9. There is no subword AA. Suppose there is, say, in the left part (for the right part the argument is symmetric, as everywhere above). By Lemma 1, between this AA and CC at a position $n'-2$ there is AC or CA; there are no such inside w, therefore there is AC or CA in a position preceding w. Consider k as in Claim 2, it is easy to see using Claim 1 that $k < n'/2$. Now, the interval $[a,k]$ has all three adjacencies AB, BC, CA and length at most $2^i + n'/2 \leq n'$. Thus $[a,k]$ is shorter than w, a contradiction.

Finally, we see that the number of A's is at most $2^i/2 + 2^i/2 = 2^i$, and the length of the word is at least $7 \cdot 2^i + 1$, as required.

3 Concluding remarks

This note settles the case when we study $[n]$ colored in three colors with no rainbow $AP(3)$. When we use $k \geq 5$ colors in $[n]$, the following construction demonstrates that no matter how large the smallest color class is, there is a coloring with no rainbow $AP(k)$.

Construction Let $n = k(2m)$, $k \geq 5$. We subdivide $[n]$ into k consecutive intervals of length $2m$ each, say $A_1, \cdots, A_k$ and let $t = \lfloor k/2 \rfloor$.

$$c(i) = \begin{cases} j & \text{if} \quad i \in A_j \quad \text{and} \quad j \neq t, \quad j \neq t+2, \\ t & \text{if} \quad i \in A_t \cup A_{t+2} \quad \text{and} \quad i \quad \text{is} \quad \text{even}, \\ t+2 & \text{if} \quad i \in A_t \cup A_{t+2} \quad \text{and} \quad i \quad \text{is} \quad \text{odd}. \end{cases} \tag{3}$$

It is easy to see that the above coloring does not contain any rainbow $AP(k)$ and the size of each color class is n/k.

The case when $k = 4$ is the only unresolved problem here. Next we provide a coloring c of $[n]$, where $n = 10m + 1$ with the smallest color class of size $\frac{n-1}{5}$ and no rainbow $AP(4)$. Let $[n] = A_1 \cup \cdots \cup A_5$ where A_is are consecutive intervals of lengths $2m$, $2m$, $2m+1$, $2m$, $2m$ respectively. Then

$$c(i) = \begin{cases} A & \text{if} \quad i \in A_1 \cup A_2 \quad \text{and} \quad i \quad \text{is} \quad \text{odd}, \\ D & \text{if} \quad i \in A_4 \cup A_5 \quad \text{and} \quad i \quad \text{is} \quad \text{even}, \\ B & \text{if} \quad i \in A_1 \quad \text{and} \quad i \quad \text{is} \quad \text{even}, \\ B & \text{if} \quad i \in A_5 \quad \text{and} \quad i \quad \text{is} \quad \text{odd}, \\ C & \text{otherwise}. \end{cases} \tag{4}$$

There are colorings of $[n]$ for $n \leq 16$ such that each color class has size $n/4$ and with no rainbow $AP(4)$ [1]. Nevertheless, we do not know whether any coloring of $[n]$ in four almost equally sized colors always has a rainbow $AP(4)$.

Acknowledgments The research of the second author was partially supported by RFBR grants $02-01-00039$ and $03-01-00796$, and by a grant of the 6th Young Scientists' Projects Expertise of Russian Academy of Sciences.

References

[1] V. Jungić, J. Licht (Fox), M. Mahdian, J. Nešetril, R. Radoičić, *Rainbow Arithmetic Progressions and Anti-Ramsey Results*, Combinatorics, Probability and Computing **12** (2003), 599-620.

[2] E. Szemerédi, *Integer sets containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299–345.

[3] B. L. van der Waerden, *Beweis einer Baudetschen Vermutung*, Nieuw Arch. Wisk. **15** (1927), 212–216.