Graphs and Combinatorics 5, 307–314 (1989)

Graphs and Combinatorics

Sub-Ramsey Numbers for Arithmetic Progressions

Noga Alon[1]*, Yair Caro[1] and Zsolt Tuza[2]**

[1] Department of Mathematics, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel Aviv University, Ramat-Aviv, Tel Aviv 69978, Israel

[2] Computer and Automation Institute, Hungarian Academy of Sciences, H-1111 Budapest, Kende u. 13–17, Hungary

Abstract. Let $m \geq 3$ and $k \geq 1$ be two given integers. A *sub-k-coloring* of $[n] = \{1, 2, \ldots, n\}$ is an assignment of colors to the numbers of $[n]$ in which each color is used at most k times. Call an $S \subseteq [n]$ a *rainbow set* if no two of its elements have the same color. The *sub-k-Ramsey number* $sr(m,k)$ is defined as the minimum n such that every sub-k-coloring of $[n]$ contains a rainbow arithmetic progression of m terms. We prove that $\Omega((k-1)m^2/\log mk) \leq sr(m,k) \leq O((k-1)m^2 \log mk)$ as $m \to \infty$, and apply the same method to improve a previously known upper bound for a problem concerning mappings from $[n]$ to $[n]$ without fixed points.

1. The Results

The aim of this note is to investigate certain Ramsey-type problems for arithmetic progressions. Throughout, m-AP abbreviates "arithmetic progression of m terms." One of the best known results of this kind is Van der Waerden's theorem [8], which states that for all natural numbers m and k there is an $n_0 = n_0(m,k)$ such that if $n \geq n_0$ and the set $[n] = \{1, 2, \ldots, n\}$ is colored by at most k colors then it contains a monochromatic m-AP. This statement was generalized for sets of positive upper density in the celebrated paper [7] of Szemerédi. A canonical theorem is also available, for details see [5, p. 59].

In this note we consider an additional Ramsey function which received a considerable amount of attention recently. We obtain various estimates for the sub-Ramsey numbers (defined below), and show how our method can be applied to improve a previous result concerning mappings from $[n]$ to $[n]$ having no fixed points.

Sub-Ramsey Numbers

A coloring of $[n]$ is called a *sub-k-coloring* if each color is assigned to at most k elements. For given m and k, define the *sub-k-Ramsey number* $sr(m,k)$ as the

* Research supported in part by Allon Fellowship and by a Bat Sheva de-Rothschild grant.

** Research supported in part by the "AKA" Research Fund of the Hungarian Academy of Sciences, grant No. 1-3-86-264.

minimum integer $n_0 = n_0(m, k)$ such that if $n \geq n_0$ then every sub-k-coloring of $[n]$ contains an m-AP with no pair of terms having the same color. Such an arithmetic progression will be called a *rainbow* m-AP.

For $sr(m, k)$ we have the following general estimates.

Theorem 1. *For every $m \geq 3$ and $k \geq 2$,*

$$\frac{1}{6}\frac{(k-1)m(m-1)}{\log(k-1)m} - k + 1 \leq sr(m,k)$$
$$\leq (1 + o(1))\frac{24}{13}(k-1)(m-1)^2\log(k-1)(m-1),$$

where the factor of $1 + o(1)$ approaches 1 as m tends to infinity.

If m is fixed and k grows, one can show that $sr(m, k)$ is linear in k.

Theorem 2. *For fixed m, as k tends to infinity,*

$$sr(m,k) \leq (1 + o(1))\frac{1}{2}m(m-1)^2(k-1).$$

The exact determination of the asymptotic behavior of $sr(m, k)$ seems to be difficult. If $m = 3$ (i.e., for 3-APs), Theorem 2 yields an upper bound of $(1 + o(1))6k$. In this particular case the following sharper estimate can be proved.

Theorem 3. *As k grows, $2k < sr(3, k) \leq (4.5 + o(1))k$.*

We note that $sr(m, k)$ and the so-called sub-Ramsey numbers of complete graphs, studied in [2], have a similar behavior. This seems to be an interesting coincidence as arithmetic progressions form a very special structure on $[n]$, whose connection with general edge-colorings of complete graphs is not clear.

Mappings

Motivated by [6] (and, later, by [1]), the second author proved in [3] the following result. For every natural number m there is an integer $n = n(m)$ such that for every $\varphi: [n] \to [n]$ without a fixed point there is an m-AP S satisfying $\varphi(i) \notin S$ for $i \in S$. Moreover, denoting by $n_0(m)$ the minimum value of n with the above property, $c_1 m^2/\log m \leq n_0(m) \leq m^2(\log m)^{(c_2 \log m)/\log\log m}$, for some absolute positive constants c_1 and c_2. Our methods here enable us to reduce the exponent of $\log m$ in the upper bound drastically, as well as to slightly improve the lower bound, as follows.

Theorem 4. *For every m,*

$$\frac{m(m-1)}{3\log m} + O(1) \leq sr(m,3) - 1 \leq n_0(m) \leq (1 + o(1))\frac{48}{13}m^2\log m.$$

We note that the same lower bounds hold for the numbers $n_0(m)$ even if we restrict our investigation to *one-to-one* mappings. This fact follows from the first part of the proof of Theorem 4.

It would be interesting to obtain an upper bound for $n_0(m)$ in terms of $sr(m,k)$, for some small k.

2. Proofs

In the proofs we apply the following well-known extension of the prime number theorem (see e.g. [4]).

Lemma 1. *Let d be a fixed nonnegative integer. For the sum of d-th powers of primes p smaller than x, we have the following asymptotic estimate as $x \to \infty$:*

$$\sum_{p<x} p^d = (1+o(1))\frac{x^{d+1}}{(d+1)\log x}.$$

This lemma will be used in the particular cases $d = 0$, 1 and 2. We shall also frequently use the following simple fact.

Lemma 2. *Let n, m, d be positive integers, $n > (m-1)d$. The number of arithmetic progressions with difference d and with m terms in $[n]$ is equal to $n-(m-1)d$. As a consequence, the number of m-APs in $[n]$ is*

$$\sum_{i=1}^{[n/(m-1)]}(n-i(m-1)) = \left[\frac{n}{m-1}\right]n-(m-1)\binom{\left[\frac{n+m-1}{m-1}\right]}{2}$$
$$\leq \frac{1}{2}\frac{n^2}{m-1}.$$

as $n/m \to \infty$.

Proof of Theorem 1

The Lower bound. Let n be the largest multiple of k which is at most $\frac{1}{6}m(m-1)(k-1)/\log m(k-1)$. We take a random coloring of $[n]$ with n/k colors, each of them assigned to precisely k elements. By a "random coloring" we mean that all possible $\binom{n}{k,k,\ldots,k}$ such colorings of $[n]$ are equally likely.

It is important to note that a random coloring can be produced by coloring the n elements one by one in the following manner. If the i-th color has been assigned to k_i elements, then we choose an arbitrary element from the remaining $n-\sum k_i$ ones, and assign to it the color i with probability $\frac{k-k_i}{n-\sum k_i}$. This is independent of the arrangement of previous assignments (and only depends on the multiplicities of colors having been assigned so far).

Put $q = n/k$, and let S be an arbitrary m-AP. We claim that the probability that S is a rainbow m-AP is

$$k^{m-1}\prod_{i=1}^{m-1}\frac{q-i}{n-i}.$$

Indeed, put $S = \{s_1, \ldots, s_m\}$, and form the random coloring by starting with the elements of S. If the t first elements of S have distinct colors, then the next one should be assigned to a $(t+1)$-th color, which can be chosen in $q - t$ different ways. For those colors, $k_i = k$, while for those that have already been used, $k_i = k - 1$. Since the steps are independent, our statement follows. (The original definition of a random coloring has nothing to do with the sequential choice of the elements, so that the probability of any set being rainbow is the same for every ordering.)

An elementary computation yields the following upper bound for the above probability:

$$k^{m-1} \prod_{i=1}^{m-1} \frac{q-i}{n-i} = \prod_{i=1}^{m-1} \left(1 - \frac{(k-1)i}{n-i}\right)$$
$$< \prod_{i=1}^{m-1} e^{-(k-1)i/(n-i)}$$
$$< e^{-(k-1)m(m-1)/2n} \le e^{-3\log m(k-1)} = (m(k-1))^{-3}.$$

By Lemma 2 the number of m-APs is less than $\frac{n^2}{2(m-1)}$. Since expectation is additive, the expected number of rainbow m-APs is less than $(72(k-1)\log^2 m)^{-1} < 1$, i.e. there is a sub-$k$-coloring of $[n]$ that does not contain any rainbow m-AP.

The Upper bound. Suppose that $[n]$ has a sub-k-coloring that does not contain any rainbow m-AP. Denote by $C_1, \ldots, C_q$ the color classes, where q is the number of colors, and assume that $n = sr(m,k) - 1$. By our lower bound, $n \ge \Omega(m^2/\log m)$. Hence, there is a function $f(n)$ with the following properties: $m/f(n) = o(1)$ as $m \to \infty$, and $f(n) < n^{2/3}$.

Let us call an m-AP a *sparse sequence* if its difference p is a *prime* such that

$$\frac{n}{f(n)} \le p < \frac{n}{m-1}.$$

We claim that any two elements $i, j \in [n]$ can belong to at most $2(m-1)$ sparse sequences. Indeed, for any fixed p, i and j can belong to at most $m-1$ m-APs of difference p. Moreover, the differences in sparse sequences are primes. Hence, if $p_1, p_2, \ldots$ are the differences of those sparse sequences containing i and j, then $j - i$ should be a multiple of the product of those differences. Since $|i - j| < n$ and $f(n) < n^{2/3}$, we obtain that the number of distinct differences is at most 2.

By the definition of sub-k-colorings, each color class C_i has cardinality at most k. Thus, the number of monochromatic pairs is

$$\sum_{i=1}^{q} \binom{|C_i|}{2} \le \frac{n}{k}\binom{k}{2} = \frac{1}{2}n(k-1).$$

Applying the previous observation, we obtain that the number of non-rainbow sparse sequences is at most

$$n(k-1)(m-1).$$

Observe further that in this computation some sparse sequences have been counted more than once: If $p \geq \frac{n}{t+m-1}(1 \leq t \leq m-1)$, then the number of sparse sequences with difference p that contain a given pair of elements is at most $t+m-2$, instead of $2(m-1)$. Introducing the weight function

$$w(S) = \begin{cases} 2\dfrac{m-1}{t+m-2} & \text{for} \quad \dfrac{n}{t+m-1} \leq p < \dfrac{n}{t+m-2},\ 1 \leq t \leq m-1 \\[2ex] 1 & \text{for} \quad \dfrac{n}{f(n)} \leq p < \dfrac{n}{2(m-1)} \end{cases}$$

for sparse sequences S (where p is the difference of S), we obtain that

$$n(k-1)(m-1) \geq \sum_S w(S).$$

In order to estimate the right-hand-side, we use Lemmas 1 and 2. We divide the sum into two terms Σ_1 and Σ_2. The first one contains those S with $w(S) = 1$, i.e., sparse sequences with difference less that $\frac{1}{2}n/(m-1)$. For this sum we obtain

$$\begin{aligned}
\Sigma_1 &= \sum_p (n - p(m-1)) = n\sum_p 1 - (m-1)\sum_p p \\
&= (1+o(1))\left\{ n\left[\frac{\frac{n}{2(m-1)}}{\log\frac{n}{2(m-1)}} - \frac{\frac{n}{f(n)}}{\log\frac{n}{f(n)}}\right]\right. \\
&\qquad \left. -(m-1)\left[\frac{\left(\frac{n}{2(m-1)}\right)^2}{2\log\frac{n}{2(m-1)}} - \frac{\left(\frac{n}{f(n)}\right)^2}{2\log\frac{n}{f(n)}}\right]\right\} \\
&= (1+o(1))\frac{3}{8}\frac{n^2}{(m-1)\log\frac{n}{m-1}}
\end{aligned}$$

by the assumption $m/f(n) = o(1)$.

When estimating Σ_2, we observe that each sequence S with difference p has weight $w(S) > 2p(m-1)/n$. Since there are $n - p(m-1)$ m-APs with difference p, we obtain

$$\begin{aligned}
\Sigma_2 &> \sum_p \frac{2p(m-1)}{n}(n - p(m-1)) \\
&= 2(m-1)\sum_p p - \frac{2(m-1)^2}{n}\sum_p p^2 \\
&= (1+o(1))\left\{ 2(m-1)\left[\frac{\left(\frac{n}{m-1}\right)^2}{2\log\frac{n}{m-1}} - \frac{\left(\frac{n}{2(m-1)}\right)^2}{2\log\frac{n}{2(m-1)}}\right]\right.
\end{aligned}$$

$$-\frac{2(m-1)^2}{n}\left[\frac{\left(\frac{n}{m-1}\right)^3}{3\log\frac{n}{m-1}}-\frac{\left(\frac{n}{2(m-1)}\right)^3}{3\log\frac{n}{2(m-1)}}\right]\Bigg]$$

$$=(1+o(1))\frac{1}{6}\frac{n^2}{(m-1)\log\frac{n}{m-1}}.$$

Thus,

$$n(k-1)(m-1)\geq(1+o(1))\frac{13}{24}n^2\Big/\left((m-1)\log\frac{n}{m-1}\right),$$

implying the upper bound for n. □

Proof of Theorem 2. For any $i \neq j$, fixing the i-th and j-th elements of an m-AP, all of its m elements are uniquely determined. Thus, each pair $i, j \in [n]$ is contained in at most $\binom{m}{2}$ m-APs. If there is a sub-k-coloring without any rainbow m-AP in $[n]$, then denoting by $C_1, \ldots, C_q$ the color classes, Lemma 2 implies

$$(1+o(1))\frac{n^2}{2(m-1)}\leq\binom{m}{2}\sum_{i=1}^{q}\binom{|C_i|}{2}\leq\binom{m}{2}\frac{n}{k}\binom{k}{2}$$
$$=\frac{1}{4}nm(m-1)(k-1),$$

which gives the required upper bound for n (as n gets large with k). □

Proof of Theorem 3. The lower bound is trivial, since $[2k]$ has a sub-k-coloring with just two colors which, of course, cannot contain any rainbow 3-AP.

To prove the upper bound we assume that $n = (4 + x)k$, where $\frac{1}{2} < x < 1$. Given a sub-k-coloring of $[m]$, it suffices to show that, if k is sufficiently large, there must be a rainbow 3-AP. Let us call a 3-AP *odd* if its difference is odd. In the proof we consider only odd 3-APs. One can easily check that there are $[n^2/8]$ such APs. We call a pair of distinct numbers $i, j \in [n]$ an *odd pair* if i and j have the same color and there is at least one odd 3-AP containing both. An odd pair $\{i, j\}$ is called *heavy* if there are two odd 3-APs containing i and j. Note that if i, j have the same color and $i - j \equiv 0 (\mathrm{mod}\, 4)$ then $\{i, j\}$ is not an odd pair, whereas if $i - j \equiv 2 (\mathrm{mod}\, 4)$ then $\{i, j\}$ is an odd pair which is not heavy. In case $i - j \equiv 1$ or $3 (\mathrm{mod}\, 4)$, then $\{i, j\}$ is odd and may be heavy. However, it is not heavy if $2i < j$ or if $2i - j > n$.

Suppose, now, that we have t_ℓ elements of color ℓ. Consider the graph whose vertices are these elements and whose edges are all the odd pairs of color ℓ. Clearly this graph is 4-colorable, since its vertices can be partitioned according to their residues modulo 4 to form a proper vertex coloring. Thus, this graph has at most $\frac{3}{8}t_\ell^2$ edges. Since the maximum of the quantity $\sum t_\ell^2$ subject to the constraints $0 \leq t_\ell \leq k$ and $\sum t_\ell = (4 + x)k$, where $\frac{1}{2} < x < 1$ is attained when $t_1 = t_2 = t_3 = t_4 = k$ and $t_5 = xk$, we obtain the following:

Fact 1. The total number of odd pairs does not exceed $\frac{3}{8}(4k^2 + x^2k^2)$.

We now estimate the number of heavy pairs. Clearly, for each i there are at most $[i/2]$ numbers $j > i$ such that $\{i, j\}$ is heavy, since for each such pair $2i \geq j > i$ and $j \not\equiv i (\mathrm{mod}\, 2)$. Similarly, for each i there are at most $[(n - i)/2]$ numbers $j < i$ such that $\{i, j\}$ is heavy. It follows that the total number of heavy pairs containing either an element in $(1, 2, \ldots, [k/2])$ or an element in $\left(n - [k/2] + 1, n - \left[\frac{k}{2}\right] + 2, \ldots, n\right)$ does not exceed $4\left(1 + 2 + 3 + \cdots + \left[\frac{[k/2]}{2}\right]\right) = (1 + o(1))\frac{k^2}{8}$. It remains to estimate the number of heavy pairs contained in the interval of the middle $(3 + x)k$ elements of $[n]$. The graph whose vertices are the numbers colored ℓ among these, and whose edges are the heavy pairs, is bipartite. Hence, if it has t_ℓ vertices then it has at most $t_\ell^2/4$ edges. It thus follows, as before, that the total number of heavy pairs contained in the middle interval does not exceed $\frac{1}{4}(3k^2 + x^2k^2)$. We have thus proved:

Fact 2. The total number of heavy pairs does not exceed

$$\frac{1}{4}(3k^2 + x^2k^2) + (1 + o(1))\frac{k^2}{8} = (1 + o(1))k^2\left(\frac{7}{8} + \frac{x^2}{4}\right).$$

Combining Facts 1 and 2 we conclude that the total number of odd 3-APs containing at least one odd pair does not exceed the number of odd, unheavy pairs plus twice that of the heavy pairs, which is at most

$$\frac{3}{8}(4k^2 + x^2k^2) + (1 + o(1))k^2\left(\frac{7}{8} + \frac{x^2}{4}\right) = (1 + o(1))k^2\left(\frac{19}{8} + \frac{5x^2}{8}\right).$$

For $\frac{1}{2} < x < 1$, (and for all sufficiently large k), this number is strictly smaller than the total number of odd 3-APs, which is $(1 + o(1))\frac{n^2}{8} = (1 + o(1))\frac{(4 + x)^2k^2}{8}$. Hence, there is at least one rainbow odd 3-AP, completing the proof. □

Remark. The constant 4.5 in the last theorem can be somewhat reduced by a more careful analysis. Since this is rather tedious and we suspect that the truth is, in fact, much closer to the trivial lower bound, we omit the details of this improvement.

Proof of Theorem 4. We first prove that $n_0(m) \geq sr(m, 3) - 1$. Take a sub-3-coloring of $[n]$ without any rainbow m-AP, for $n = sr(m, 3) - 1$, and suppose that the number of colors is as small as possible. The latter assumption implies that all but at most one of the colors occur at least twice, otherwise two colors occurring once might be identified, without producing any new rainbow m-AP. Moreover, if there is a one-element color class, then all the other colors occur three times, for the same reason.

If the coloring contains no one-element class, then we stay with $[n]$, otherwise delete the n-th element which either was the single class or its deletion yields a class of cardinality 2, and then we identify this color with the one which occurs once. In either case, we obtain a coloring of at least $n - 1$ elements without any rainbow m-AP, such that each color occurs twice or three times.

If $\{i,j\}$ is a two-element color class, then define $\varphi(i) = j$ and $\varphi(j) = i$. If $\{i,j,k\}$ is a three-element class, then put $\varphi(i) = j$, $\varphi(j) = k$, $\varphi(k) = i$. One can easily check that the m-APs covered by monochromatic pairs $\{i,j\}$ are precisely those m-APs which are covered by the pairs $\{i,\varphi(i)\}$. It is also clear that φ has no fixed points.

In order to obtain the upper bound, we refer to the proof of Theorem 1. In case of mappings $[n] \to [n]$, we can have at most n pairs $\{i,\varphi(i)\}$ which should cover all sparse sequences S and, as we have seen, each of those pairs can cover at most $2(m-1)$ of them. Thus, $2(m-1)n \geq \sum w(S)$, so that precisely the same upper bound can be deduced as the one we have obtained for $sr(m,3)$. □

References

1. Alon, N., Caro, Y.: Extremal problems concerning transformations of the set of edges of the complete graph. Europ. J. Comb. **7**, 93–104 (1986)
2. Alon, N., Caro, Y., Rödl, V., Tuza, Zs.: paper in preparation
3. Caro, Y.: Extremal problems concerning transformations of the edges of the complete hypergraphs. J. Graph Theory **11**, 25–37 (1987)
4. Davenport, H.: Multiplicative Number Theory. 2nd ed., revised by H.L. Montgomery. Berlin: Springer-Verlag 1980
5. Erdös, P.: My joint work with Richard Rado. In: Surveys in Combinatorics 1987 (C. Whitehead, ed.), London Math. Soc. Lecture Notes Ser. Vol. 123, pp. 53–80. Cambridge University Press 1987
6. Erdös, P., Hajnal, A.: On the structure of set mappings. Acta Math. Acad. Sci. Hung. **9**, 111–131 (1958)
7. Szemerédi, E.: On sets of integers containing no k elements in arithmetic progression. Acta Arith. **27**, 199–245 (1975)
8. van der Waerden, B.L.: Beweis einer Baudetschen Vermutung. Nieuw. Arch. Wiskd. **15**, 212–216 (1927)

Received: January 9, 1989