Europ. J. Combinatorics (1981) **2**, 363–372

Intersection Properties of Subsets of Integers

MIKLÓS SIMONOVITS AND VERA T. SÓS

Let $\{A_1, \ldots, A_N\}$ be a family of subsets of $\{1, 2, \ldots, n\}$. For a fixed integer k we assume that $A_i \cap A_j$ is an arithmetic progression of $\geq k$ elements for every $1 \leq i < j \leq N$. We would like to determine the maximum of N. For $k = 0$, R. L. Graham and the authors have proved that

$$N \leq \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + 1.$$

For $k \geq 2$, the extremal and asymptotically extremal systems have

$$\left(\frac{\pi^2}{24} + o(1)\right) n^2 \text{ sets.}$$

For $k = 1$, the maximum is between

$$\binom{n}{2} + 1 \quad \text{and} \quad \left(\frac{\pi^2}{24} + \frac{1}{2} + o(1)\right) n^2.$$

We conjecture that the lower bound is sharp.

INTRODUCTION

Intersection properties of sets have been widely investigated by many authors. One type of theorems proved for them has the following form [9]. Let S be an n-element set and $A_1, \ldots, A_N \subseteq S$, $I \subseteq [1, n]$. Assume that $|A_i \cap A_j| \in I$ for $1 \leq i < j \leq N$. How large can N be under this condition, depending on n and I? Thus, e.g., the de Bruijn–Erdös theorem [1] asserts that if $|A_i \cap A_j| = 1$ for all $i \neq j$, then $N \leq n$. There are different extremal systems in the de Bruijn–Erdös theorem, but all they are known.

Another typical example of intersection theorems is the well known result of Erdös, Ko and Rado [2]: If $|A_i| = k$ and

$$A_i \cap A_j \neq \emptyset, \quad \text{then } N \leq \binom{n-1}{k-1},$$

assuming that $n \geq 2k$.

In these and many other similar examples S has no structure. We were looking for intersection theorems where S is endowed with some structure and $A_i \cap A_j$ has some prescribed substructure: instead of having conditions on the cardinality of $A_i \cap A_j$ we have conditions on its structure [4, 5, 6].* Thus in [5, 6] we had graphs on S and assumed that, e.g., $G_i \cap G_j$ is a path or a cycle, ... and so on. Here we assume that $S = \{1, 2, \ldots, n\}$ and that $A_i \cap A_j$ is an arithmetic progression—the simplest non-trivial structure on the integers. Let $\mathbb{P}_k$ denote the family of arithmetic progressions of at least k elements.

PROBLEM. If k is fixed and $A_1, \ldots, A_N \subseteq \{1, 2, \ldots, n\} := [1, n]$ and if $A_i \cap A_j \in \mathbb{P}_k$, how large can N be? The maximum N is denoted by $f(n, \mathbb{P}_k)$ and the systems attaining the maximum are called extremal. (In heuristic arguments, if N is near to the maximum, then $A_1, \ldots, A_N$ will be called almost extremal.)

* See also our survey paper [8].

0195-6698/81/040363+10 $01.00/0

For $k=0$, R. L. Graham and the authors [4] proved that

$$f(n, \mathbb{P}_0) = \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + 1$$

and the only extremal system consists of the subsets of $[1, n]$ of at most three elements. Here the intersections have at most two elements. If we wish to have non-degenerate arithmetic progressions as intersections, we have to assume that $k \geq 3$. Surprisingly, $f(n, \mathbb{P}_k)$ depends on k very weakly if $k \geq 2$.

THEOREM 1. *Let $k \geq 2$ be fixed and $A_1, \ldots, A_N \subseteq [1, n]$. Let $A_i \cap A_j \in \mathbb{P}_k$ for every $1 \leq i < j \leq N$. Then*

$$N \leq \left(\frac{\pi^2}{24} + o(1)\right) n^2, \tag{1}$$

and (1) *is sharp, for any $k \geq 2$.*

REMARK 1. Let $A_1, \ldots, A_N$ be arithmetic progressions of form

$$A_i = \left\{\left[\frac{n}{2}\right] + jd : j = -a, -a+1, \ldots, -1, 0, 1, 2, \ldots, b\right\}$$

for some $d \leq n^{\frac{1}{3}}$, $\sqrt{n} \leq b \leq n/2d$ and $a \leq n - 1/2d$. Obviously, $A_i \cap A_j$ is an arithmetic progression with $\geq n^{\frac{1}{6}}$ elements and

$$N = \frac{n^2}{4}\left(\sum_1^\infty \frac{1}{d^2} + o(1)\right) = \left(\frac{\pi^2}{24} + o(1)\right) n^2, \tag{2}$$

showing that (1) is sharp.

In the above example we constructed a system of subsets where the intersections were arithmetic progressions just because the sets themselves were also arithmetic progressions. What happens, if we exclude this? The next theorem shows that in any almost extremal system almost all the sets are arithmetic progressions, the number of non-arithmetic progressions is $O(n^{\frac{5}{3}} \log^3 n)$. Even more, if we have a system of subsets of $[1, n]$, say $A_1, \ldots, A_M$, no one of which is an arithmetic progression and the intersection of any two of which is an arithmetic progression, then $M = O(n^{\frac{5}{3}} \log^3 n)$, even if these sets do not belong to an almost extremal system.

THEOREM 2. *Let $k \geq 2$ and $A_1, \ldots, A_N \subseteq [1, n]$. Assume that no A_i is an arithmetic progression but $A_i \cap A_j \in \mathbb{P}_k$ for every $1 \leq i < j \leq k$. Then*

$$N = O(n^{\frac{5}{3}} \log^3 n). \tag{3}$$

Here (3) can be improved but since we do not know whether the exponent $\frac{5}{3}$ is sharp or not, we do not care about getting rid of $\log^3 n$. On the other hand we did care about getting $n^{\frac{5}{3}}$ instead of $n^{\frac{5}{3}+\varepsilon}$: some tricks in the proof were needed just for this purpose.

It would be interesting to know what happens if we allow $|A_i \cap A_j| = 1$. It is natural to conjecture that one extremal system for $\mathbb{P}_1$ is the system of sets of form $\{c, x, y\}$ where c is fixed; $c, x, y \in [1, n]$ are not necessarily different. This construction yields that

$$f(n, \mathbb{P}_1) \geq \binom{n-1}{2} + n = \binom{n}{2} + 1. \tag{4}$$

There exist some other equally good constructions, given at the end of this paper. Therefore, if we have equality in (4), then there are many extremal systems. Unfortunately, we can prove only a weaker bound on $f(n, \mathbb{P}_1)$, namely, the following theorem.

THEOREM 3. *If $A_1, \ldots, A_N \subseteq [1, n]$ and $A_i \cap A_j \in \mathbb{P}_1$ for every $1 \leq i < j \leq N$, then*

$$N \leq \binom{n-1}{2} + \frac{\pi^2}{24} \cdot n^2 + O(n^{\frac{5}{3}} \log^3 n). \tag{5}$$

REMARK 2. We can improve the upper bound in Theorem 3, but we cannot prove the conjecture stated above.*

The following theorem is slightly technical. Its main point is that if A_is are not too large and their intersection is empty, then Theorem 3 can be improved in term of the upper bound on $|A_i|$.

THEOREM 4. *Let $A_1, \ldots, A_N \subseteq [1, N]$, $|A_i| \leq a$ for $i = 1, 2, \ldots, n$, and assume also that no A_i is an arithmetic progression. If the intersection $A_i \cap A_j$ is always an arithmetic progression $(1 \leq i < j \leq N)$ and $\bigcap_{i=1}^{N} A_i = \varnothing$, then*

$$N \leq an - \binom{a}{2} + O(n^{\frac{5}{3}} \log^3 n). \tag{6}$$

Now, our plan is the following: first we prove Theorem 4, then observe that Theorem 2 follows by exactly the same argument. To prove Theorem 1 we shall separate the arithmetic progressions and estimate their number by $\pi^2/24 \,.\, n^2 + O(n \,.\, \log n)$. Finally we prove Theorem 3 using Theorem 4. In this last step we shall restrict our considerations primarily to those A_is which have at most $n^{\frac{2}{3}}$ elements. Then Theorem 4 (applied with $a = n^{\frac{2}{3}}$) yields an $O(n^{\frac{5}{3}} \log^3 n)$ term. This is why we needed it.

PROOF OF THEOREM 4

Our basic tool is the notion of the ν and δ-triplets.

DEFINITION 1. Given a system $A_1, \ldots, A_N \subseteq [1, n]$, triplet $\{x, y, z\}$ is called *determining* or δ-triplet if it belongs to exactly one A_i. Otherwise it is a non-determining, or ν-triplet.

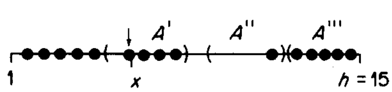
LEMMA 1. *Let $1 > c > 0$ be fixed and $A_1, \ldots, A_M \subseteq [1, n]$ be given sets. Let $A_i \cap A_j \in \mathbb{P}_1$ for $1 \leq i < j \leq M$. If $|A_1| = h > n^c$, then for every $x \in A_1$ and $t \leq h/20$ (for $n > n_0(c)$) either*

(i) *A_1 contains an arithmetic progression of at least $n - t$ elements or*

(ii) *A_1 contains at least $th/50 \log h$ δ-triplets of form $\{x, y, z\}$.*

PROOF. Let $A_1 = \{u_1, u_2, \cdots, u_h\}$ where $u_1 < u_2 < \cdots < u_h$. If $x = u_{i_0}$, we may assume that $i_0 < h/2$, that is, at least $h/2$ elements of A_1 are above x. Now these elements are split into three segments:

$$A' = \{u_i\colon i_0 < i \leq \tfrac{3}{5}h\},$$
$$A'' = \{u_i\colon \tfrac{3}{5}h < i \leq \tfrac{7}{10}h\},$$
$$A''' = \{u_i\colon i > \tfrac{7}{10}h\}.$$



(a) *If for every $y_0 \in A''$ there are at least $h/\log h$ δ-triplets of form (x, y_0, z), then we have at least $\frac{1}{20}(h^2/\log h)$ δ-triplets and we are home.*

(b) *Assume that there is a $y_0 \in A''$ for which all but $(h/\log h)$ z define a ν-triplet (x, y_0, z). Then all but $h/\log h$ integers $z \in [x, y_0]$ can be covered by at most n^{ε} arithmetic progressions*

* In [7], overlooking a term, we stated that we can.

$P_1, \ldots, P_\lambda$ for $\varepsilon = c/10$ and n sufficiently large. Indeed, since (x, y_0, z) is a ν-triplet, x, y_0 and z are contained not only by A_1 but also by some other $A_{i(z)}$, $i(z) \neq 1$. Therefore, if

$P(u, v, w)$ denotes the minimal arithmetic progression containing u, v and w,

then $P(x, y_0, z) \subseteq A_1 \cap A_{i(z)} \subseteq A_1$. The segment $P(x, y_0, z) \cap [x, y_0]$ is also an arithmetic progression with fixed endpoints x and y_0, therefore it is determined by its difference d. Since $x - y_0$ has at most n^ε divisors (if n is sufficiently large) [3], the assertion is proved.

(c) Now we shall construct many δ-triplets (x, y, z) by taking large arithmetic progressions $R_i \subseteq A_i$, containing x, a $z \notin R_i$ and $y = x + pd_1$, where d_1 is the difference of R_1 and p is coprime with $z - x$. Two cases will be distinguished: first we consider the case when A_1 contains a relatively long arithmetic progression, then the situation when it does not.

Assume that $Q_1 \in \mathbb{P}_1$, $Q_1 \subseteq A_1$, $\min Q_1 = x$, and $|Q_1| \geq h/20$. Let the difference d_1 of Q_1 be the minimum under these conditions. Let R_1 be a maximal arithmetic progression in A_1 containing Q_1. We shall prove that

if $z \in A_1 - R_1$, $p > h/40$ and p is coprime with $z - x$, then $(x, x + pd_1, z)$ is a δ-triplet, $(p \leq h/20!)$.

Assume the contrary. Then the induced arithmetic progression $Q' = P(x, x + pd_1, z)$ is in A_1. Let d' be the difference of Q'. We know that $d' \neq d_1$, since $z \in Q'$ but $z \notin R_1$. Further, $d' | d_1$, since $x, x + pd_1 \in Q'$, that is, $pd_1 = qd'$, but p is relative prime to $z - x = kd'$. Hence $d' \leq \frac{1}{2} d_1$, and therefore $Q'' = \{x + jd' : j = 0, 1, 2, \ldots, [h/20]\} \subseteq A_1$: Indeed, it is at least twice as dense as Q_1, and contains $y = x + pd_1 = x + rd'$ for $p > h/40$, that is, for $r > h/20$. This is a contradiction: Q'' should have been chosen instead of Q_1.

Let us count these δ-triplets. If $|A_1 - R_1| = t$, then for each $z \in A_1 - R_1$ we exclude $O(\log n)$ primes dividing $z - x$ and $\leq h/35 \log h$ other primes. However, we still have at least $h/50 \log h$ primes for which $(x, x + pd_1, z)$ is a δ-triplet. This proves the assertion.

(d) Assume now that each arithmetic progression containing x has at most $h/20$ elements from $[x, n]$. In (b) we have covered all but at most $h/\log h$ elements of $[x, y_0]$ by at most n^ε arithmetic progressions $P_1, \ldots, P_\lambda$. We may assume that none of these P_js is contained in some other P_i. Some of them may be extended beyond $[x, y_0]$: let R_j be the maximal arithmetic progression in $[x, n]$ containing P_j. If s is the minimum integer for which

$$|R_1 \cup R_2 \cup \cdots \cup R_s| \geq \frac{h}{20}, \tag{7}$$

then for $A^+ := A''' - \bigcup_{j \geq s} R_j$ we know that

$$|A^+| \geq \frac{h}{5}, \tag{8}$$

since $|R_1 \cup R_2 \cup \cdots \cup R_j|$ increases by at most $h/20$ when we pass from j to $j+1$, and hence $|\bigcup_{j \leq s} R_j| \leq h/10$; further, $|A'''| \geq \frac{3}{10} h$.

Take all the triplets $(x, x + pd_j, z)$ for $j = 1, \ldots, s$, $z \in A^+$, where d_j is the difference of R_j and p is a prime not dividing $y_0 - x$ and $z - x$.

We show that *these triplets are all δ-triplets, each counted only once and that their number is at least*

$$\left(\frac{1}{100} - o(1)\right) \cdot \frac{h^2}{\log h}. \tag{9}$$

(i) Indeed, a slight modification of the argument of (c) yields that they are all δ-triplets: if not, then we would get a $Q'' := \{x + ld' : l = 0, 1, 2, \ldots\}$ joining x to y_0, and containing P_j and being at least twice as dense.

(ii) In theory it could happen that

$$(x, x+pd_i, z) = (x, x+p'd_j, z) \qquad (i \neq j).$$

However, this would imply that $pd_i = p'd_j$. If $p = p'$, then $d_i = d_j$, therefore $i = j$, and we are home. If $p \neq p'$, then, by symmetry, we may assume that $p|d_j$. But we know that $d_j|y_0 - x$ and we assumed that p is relative prime to $y_0 - x$, a contradiction.

(iii) Each R_j with $k_j := |R_j|$ elements yields $(1-o(1))\, k_j/\log k_j$ primes p, by the prime number theorem $(j = 1, 2, \ldots, s)$. This is the point where we use (b): since the number of P_js covering $[x, y_0]$ is only $o(h)$, we may easily see that $|P_j| \to \infty$ for $j \leqslant s$, if, e.g., we take that very permutation of $P_1, \ldots, P_\lambda$, for which the size is decreasing. If $|P_j| \to \infty$, then the prime number theorem can be applied. Thus, we get altogether at least

$$(1-o(1))\frac{h}{5}\sum \frac{k_j}{\log k_j} \geqslant (1-o(1))\frac{h}{5}\cdot\frac{1}{\log h}\sum k_j \tag{10}$$

δ-triplets: $z \in A^+$ can be chosen in at least $h/5$ ways, and excluding the primes dividing $y_0 - x$ or $z - x$ means only excluding $2 \,.\, \log n = o(k_j/\log k_j)$ primes. (This is, where we use that $h > n^c$). Since $k_1 + \cdots + k_s \geqslant h/20$, (10) immediately implies (9). This completes the proof of the lemma.

LEMMA 2. *Let $A_1, \ldots, A_M \subseteq [1, n]$, and assume that no A_i is an arithmetic progression but $A_i \cap A_j$ is an arithmetic progression, $1 \leqslant i < j \leqslant M$. Assume that there is an integer $c \in [1, n]$ contained by each A_i and an s-element set $S \subseteq [1, n] - \{c\}$, such that $S \cap A_i \neq \emptyset$ $i = 1, 2, \ldots, M$. Then for every $\varepsilon > 0$*

$$M \leqslant sn - \binom{s}{2} + O(n^{1+\varepsilon}). \tag{11}$$

We need the following definition.

DEFINITION 2. Let $P \subseteq A$ be a maximal arithmetical progression (with respect to $\subseteq$) and $z \in A - P$. We denote the infinite extension of P by $P^\infty(P)$ and call z *external* or *internal* according to whether $z \in P^\infty(P)$ or not.

REMARK. Observe that if $z \in P^\infty$, then we call it external. The reason for this is, that in this case it makes no real trouble. We could say that it does not belong to P only because we have forgotten some elements of A. The following example illustrates the definition:

$A = \{1, 3, 4, 5, 7, 15, 20\}$ and $P = \{1, 3, 5, 7\}$. Now 15 is external, 20 and 4 are internal.

PROOF. We shall subdivide $\mathbb{F} = \{A_1, \ldots, A_M\}$ into four subfamilies $\mathbb{F}_1$, $\mathbb{F}_2$, $\mathbb{F}_3$ and $\mathbb{F}_4$ as follows:

$\mathbb{F}_1$ is the family of A_is containing a maximal arithmetic progression $P_i \ni c$ and an *external* z_i.

For all the other A_is we fix a $y_i \in A_i \cap S$.

$\mathbb{F}_2 = \{A_i \in \mathbb{F} - \mathbb{F}_1$: there exists a $z_i \in A_i - \{y_i, c\}$ such that (c, y_i, z_i) form a δ-triplet for $\mathbb{F}\}$.

$\mathbb{F}_3 = \{A_i \in \mathbb{F} - \mathbb{F}_1 - \mathbb{F}_2$: $A_i \cap [c, y_i]$ is an arithmetic progression and if P_i denotes the maximal arithmetic progression containing $A_i \cap [c, y_i]$, then there is an internal z_i for this A_i and $P_i\}$.

If $A_i \in \mathbb{F}_3$, let us fix a corresponding P_i and a z_i. Finally,

$\mathbb{F}_4 = \mathbb{F} - \mathbb{F}_1 - \mathbb{F}_2 - \mathbb{F}_3$.

(a) We give an upper bound on $|\mathbb{F}_1|$. Given z_i, the difference d_i of P_i divides $z_i - c$. We put $d_i < 0$ if $c = \max P_i$ and $d_i > 0$ otherwise. Clearly, c, d_i and z_i determine A_i uniquely by $A_i \cap A_j \in \mathbb{P}_1$. Hence

$$|\mathbb{F}_1| \leq 2 \sum_{k \leq n} d(k) \leq 2n^{1+\varepsilon},$$

where $d(k)$ is the number of divisors of k. As a matter of fact, $\sum_{k \leq n} d(k) = O(n \log n)$, [3].

(b) Let $D(A_i)$ be a triple defined for each $A_i \in \mathbb{F}_2 \cup \mathbb{F}_3$ as follows:

if $A_i \in \mathbb{F}_2$, we choose a determining triplet (c, y_i, z_i) for $\mathbb{F}$. This is $D(A_i)$.

If $A_i \in \mathbb{F}_3$, we choose a P_i and z_i according to the definition above. Again, $D(A_i) = (c, y_i, z_i)$.

We show that for $\mathbb{F}_2 \cup \mathbb{F}_3$ $D(A_i)$ is a determining triplet (though it may happen that for $A_i \in \mathbb{F}_3$ $D(A_i)$ is not a δ-triplet for the whole $\mathbb{F}$). The only case to be considered is when $A_i, A_j \in \mathbb{F}_3$. Clearly, if $y_i = y_j$, then $z_i \neq z_j$.

Indeed, if (as above), $P^\infty(P)$ denotes the doubly infinite extension of the arithmetic progression P, by definition

$$z_i \notin P^\infty(P_i) \supseteq P^\infty(A_i \cap A_j \cap [c, y_i]) \supseteq A_i \cap A_j.$$

The other case, $y_i = z_j$, $y_j = z_i$ can be eliminated as well: let d^* be the difference of $P^* = A_i \cap A_j$. Then $d_j | d^*$, since, by definition, A_j is an arithmetic progression between c and y_j, namely, $P_j \cap [c, y_j] \supseteq P^* \cap [c, y_j]$. But $d_j | d^*$ implies that $z_j = y_i \in P^*$ is external, a contradiction. Hence $D(A_i) = (c, y_i, z_i)$ is different for different is $(A_i \in \mathbb{F}_2 \cup \mathbb{F}_3)$. Therefore

$$|\mathbb{F}_2 \cup \mathbb{F}_3| \leq s(n-1) - \binom{s}{2}.$$

(c) The basic idea we use to estimate $|\mathbb{F}_4|$ is that an $A_i \in \mathbb{F}_4$ is the union of arithmetic progressions and the set of their differences characterize A_i.

Let $P_i \subseteq A_i \in \mathbb{F}_4$ be a maximal arithmetic progression joining c and y_i and having the smallest difference. Since $A_i \notin \mathbb{F}_1$, if $A_i \supseteq \tilde{P}_i \supseteq P_i$ is a maximal arithmetical progression, all $z \in A_i - \tilde{P}_i$ are internal. Since $A_i \notin \mathbb{F}_2$, all $z \in A_i - \tilde{P}_i$ yield ν-triplets with c and y_i. Since $A_i \notin \mathbb{F}_3$, $A_i \cap [c, y_i] \neq P_i$. Choose a $z_i \in A_i \cap [c, y_i] - P_i$. Now, (c, y_i, z_i) is a ν-triplet: there exists a maximal $P_i^* \in \mathbb{P}_1$, $P_i^* \subseteq A_i$, joining c to y_i through z_i. It is trivial that $(P_i \cup P_i^*) \cap [c, y_i]$ cannot be contained by any arithmetic progression $\subseteq A_i$ which means that for given y_i $P_i \cap [c, y_i]$ and $P_i^* \cap [c, y_i]$ uniquely determine $A_i \in \mathbb{F}_4$. Since they are determined by the corresponding differences d_i and d_i^* which divide $y_i - c$, therefore

$$|\mathbb{F}_4| = O(sn^{2\varepsilon}).$$

PROOF OF THEOREM 4. We split the A_is into the classes

$$\mathbb{F}^* = \{A_i : |A_i| \leq n^{\frac{1}{3}}\},$$
$$\mathbb{F}_\nu = \{A_i : 2^\nu \leq |A_i| < 2^{\nu+1}\} \quad \text{if } n^{\frac{1}{3}} < 2^\nu < n^{\frac{2}{3}}$$

and

$$\mathbb{F}^{***} = \{\text{the other } A_i\text{s}\} \subseteq \{A_i : |A_i| \geq n^{\frac{2}{3}}\}.$$

(a_1) If each $A_i \in \mathbb{F}^*$ intersects $A_1 \in \mathbb{F}^*$ in at least two points, then $|\mathbb{F}^*| = O(n^{\frac{5}{3}})$. Indeed, we fix a $c_j \in A_1$ and put $S_j = A_1 - \{c_j\}$. Then we apply Lemma 2 with this c_j and S_j and sum up the results. Each $A_i \in F^*$ is counted at least twice:

$$|\mathbb{F}^*| < |A_1|(|A_1|n + O(n^{1+\varepsilon})) = O(n^{\frac{5}{3}}).$$

(a_2) If on the other hand, e.g., $|A_1 \cap A_2| = 1$ ($A_1, A_2 \in \mathbb{F}^*$), then we put $\{c\} = A_1 \cap A_2$ and choose an $A_3 \not\ni c$ ... ($A_3 \in \mathbb{F}^*$ is not necessarily true). We apply Lemma 2 first to the A_is containing c, with $S = A_3$: their number is at most

$$an - \binom{a}{2} + O(n^{\frac{5}{3}}).$$

Then we apply Lemma 2 to each $c_j \in (A_1 \cup A_2) - \{c\}$ and $S_j = (A_1 \cup A_2) - \{c, c_j\}$ and get $\leqslant 2n^{\frac{4}{3}} + O(n^{1+\varepsilon})$ sets for each j. This yields altogether

$$|\mathbb{F}^*| \leqslant an - \binom{a}{2} + O(n^{\frac{5}{3}}):$$

each $A_i \in \mathbb{F}^*$ was counted above at least once and, obviously, $|A_1 \cup A_2| \leqslant 2n^{\frac{1}{3}}$. Thus the $|S_j|$s and the number of js are $\leqslant 2n^{\frac{1}{3}}$.

(b) We fix a $\nu \in [\frac{1}{3}\log_2 n, \frac{2}{3}\log_2 n)$. Let $H \in \mathbb{F}_\nu$, $h = |H| = \max_{A_i \in \mathbb{F}_\nu} |A_i|$ and $x_i \in A_i \cap H$ be fixed for each $A_i \in \mathbb{F}_\nu$. Let

$$\mathbb{F}'_\nu =: \{A_i \in \mathbb{F}_\nu : A_i \text{ contains} \geqslant \frac{h^2}{\log^2 h}\ \delta\text{-triplets } (x_i, u, v)\},$$

$$\mathbb{F}''_\nu =: \mathbb{F}_\nu - \mathbb{F}'_\nu.$$

Clearly, if $\tilde{x} \in H$ is fixed, we have at most

$$\binom{n-1}{2} \Big/ \frac{h^2}{\log^2 h} = O\left(\frac{n^2 . \log^2 n}{h^2}\right),$$

$A_i \in \mathbb{F}'_\nu$ with $x_i = \tilde{x}$. Thus

$$|\mathbb{F}'_\nu| = O\left(\frac{n^2}{h} . \log^2 n\right) = O(n^{\frac{5}{3}} \log^2 n). \tag{12}$$

To estimate $|\mathbb{F}''_\nu|$ we apply Lemma 1 to each $A_i \in \mathbb{F}''_\nu$, $x_i \in A_i \cap H$ and $t = [h/\log^3 h]$. Thus we get a maximal arithmetic progression $P_i \subseteq A_i$ with

$$|P_i| \geqslant |A_i| - \frac{h}{\log^2 h}. \tag{13}$$

By the way, P_i is independent of $x_i \in A_i \cap H$, since A_i can contain at most one such P_i. Let $d_i > 0$ denote the difference of P_i. We show that *for given $\tilde{x}$ and d there exist at most $2n$ A_is for which $x_i = x_j = \tilde{x}$ and $d_i = d_j = d$.*

Indeed, in this case we shall call a P_i "high" if at least half of its element are above $\tilde{x}$, otherwise P_i is "low". If, e.g., both A_i and A_j are "high", then

$$\Delta_{ij} = (A_i - P_i) \cap (A_j - P_j) = \varnothing, \tag{14}$$

otherwise a $z \in \Delta_{ij}$ and $P_i \cap P_j$ (both contained in $A_i \cap A_j$) would induce an arithmetic progression P^* with difference $\leqslant d/2$. It is easy to see that this is impossible. Since $A_i - P_i \neq \varnothing$, we get at most n "high" A_is for each $\tilde{x}$ and d.

Unfortunately, we thus get only that

$$|\mathbb{F}_\nu| \leqslant h\frac{2n}{h} . 2n = O(n^2) \tag{15}$$

since $\tilde{x}$ can be chosen in $\leqslant h$ ways, d in $\leqslant 2n/h$ ways. To improve (12) we shall prove that

> One can choose $x_i \in A_i \cap H$ so that ($x_i \in P_i$, (9) holds and) for each $d \leqslant 2n/h$ and for all but at most two $x \in H$ there are at most $4h$ As with $x_i = \tilde{x}$ and $d_i = d$. (*)

This will imply

$$|\mathbb{F}''_\nu| \le \frac{2n}{h} \cdot h \cdot 4h + \frac{2n}{h} \cdot 4n = O(n^{\frac{5}{3}}) \tag{16}$$

since d and $\tilde{x}$ can be chosen in $\le 2n/h \cdot h$ ways and then A_i can be chosen *mostly* in $\le 4h$ ways. This yields the first term. The second one comes from the two exceptional $\tilde{x}_d$ and $\tilde{\tilde{x}}_d$ contained in at most $2n$ A_is. Clearly, (16) will suffice.

(c) Assume, indirectly, that (*) does not hold: for some $\tilde{d}$ there are at least three exceptional xs say $\tilde{x}$, $\tilde{\tilde{x}}$, $\tilde{\tilde{\tilde{x}}}$.

(c₁) We may assume that at least $2h$ A_is in (*) are of the same type, e.g. for $\tilde{x}$ and $\tilde{\tilde{x}}$ the P_is are "high", for $\tilde{\tilde{\tilde{x}}}$ "low". Thus for the index sets

$$I = \{i : x_i = \tilde{x},\quad d_i = \tilde{d},\quad A_i \text{ is "high"}\},$$

$$J = \{i : x_i = \tilde{\tilde{x}},\quad d_i = \tilde{d},\quad A_i \text{ is "high"}\},$$

$$K = \{i : x_i = \tilde{\tilde{\tilde{x}}},\quad d_i = \tilde{d},\quad A_i \text{ is "low"}\},$$

we have $|I| > 2h$, $|J| > 2h$, $|K| > 2h$.

We prove that $\tilde{x} \equiv \tilde{\tilde{x}} \equiv \tilde{\tilde{\tilde{x}}} \pmod{\tilde{d}}$. Assume the contrary, e.g. $\tilde{x} \not\equiv \tilde{\tilde{x}} \pmod{\tilde{d}}$. Hence $P_i \cap P_j = \emptyset$ if $i \in I$, $j \in J$.

Restricting ourselves to the set $\{\tilde{x} + j\tilde{d} : j = 0, \pm 1, \pm 2, \ldots\} = \tilde{D}$, the sets P_i $(i \in I)$ are subintervals of it and $|P_i| < h$, $P_i \ni \tilde{x}$. Hence $|\bigcup_{i \in I} P_i| \le 2h - 1$. The sets $A_j - P_j$ are pairwise disjoint while $j \in J$ and $|J| > 2h$, hence we can find a $j_0 \in J$ for which $(A_{j_0} - P_{j_0}) \cap (\bigcup_{i \in I} P_i) = \emptyset$. This means that $A_{j_0} \cap (\bigcup_{i \in I} P_i) = \emptyset$ (since $P_{j_0} \cap P_i = \emptyset$ by $\tilde{x} \not\equiv \tilde{\tilde{x}} \pmod{\tilde{d}}$). Thus the $|I|$ pairwise disjoint sets $A_i - P_i$ $(i \in I)$ do intersect A_{j_0}, i.e. $|A_{j_0}| \ge |I| > 2h$. This is a contradiction, proving that $\tilde{x} \equiv \tilde{\tilde{x}} \equiv \tilde{\tilde{\tilde{x}}}$. We may assume that $\tilde{x} < \tilde{\tilde{x}} < \tilde{\tilde{\tilde{x}}}$.

(c₂) Observe that above we used only that

(i) $\{P_i : i \in I\}$ consists of P_is of the same type: either they are all "high" or they are all "low". (This ensures that $(A_i - P_i)$s are disjoint.) Of course, we used this for J and K as well.

(ii) Further, we used that if $i \in I$ and $j \in J$, then $P_i \cap P_j = \emptyset$.

Now let us replace $x_i = \tilde{x}$ and $x_i = \tilde{\tilde{\tilde{x}}}$ by $x_i = \tilde{\tilde{x}}$ whenever we can: if $\tilde{\tilde{x}} \in P_i$. (The P_is remain unchanged, since they are uniquely determined by A_is as observed earlier.) After this step for every P_i and P_j *corresponding* to $\tilde{x}$ and $\tilde{\tilde{\tilde{x}}}$ respectively $P_i \cap P_j = \emptyset$, since

$$\max P_i < \tilde{\tilde{x}} < \min P_j.$$

As we have mentioned, if even after the alteration both $\tilde{x}$ and $\tilde{\tilde{\tilde{x}}}$ remain exceptional, that is belong to $\ge 4h$ A_is, then we may take $2h$ A_is corresponding to $\tilde{x}$ and $2h$ A_js corresponding to $\tilde{\tilde{\tilde{x}}}$ of the same type ("high" or "low") and repeat the argument of (c₁). This contradiction shows that either $\tilde{x}$ or $\tilde{\tilde{\tilde{x}}}$ ceased to be exceptional after the alteration. We can iterate this step until at most two $\tilde{x}$s are exceptional. This completes the proof of (16).

(d) If $|A_i| \ge n^{\frac{2}{3}}$, then by Lemma 1 (applied with $t = 1$) A_i contains at least $\frac{1}{100} n^{\frac{4}{3}}/\log n$ δ-triplets: for each $x \in A_i$ we have at least $n^{\frac{2}{3}}/100 \log n$ δ-triplets. Since the total number of δ-triplets is at most $\binom{n}{3}$, the number of these A_is is $\le 20 n^{\frac{5}{3}} \log n$.

(e) Adding up the estimates in (a), (b) and (c) we obtain the upper bound needed.

Proofs of Theorems 1 and 2

Proof of Theorem 2. We can repeat the proof of Theorem 4 word by word with the only exception that in (a) we always have the first case: each $A_i \in \mathbb{F}^*$ intersects A_1

in at least two points. Thus

$$|\mathbb{F}^*| = O(n^{\frac{5}{3}}).$$

To prove Theorem 1 we need the following lemma.

LEMMA 3. *If $A_1, \dots, A_M \in \mathbb{P}_3$ and $A_i \cap A_j \neq \emptyset$, then*

$$M \leq \frac{\pi^2}{24} n^2 + O(n \log n).$$

Obviously, Lemma 3 and Theorem 2 imply Theorem 1.

PROOF OF LEMMA 3. For a fixed d we consider the A_is with the difference d. Clearly, they belong to the same congruence class $I_{d,a} = \{l \in [1, n]: l \equiv a \pmod d\}$, since $A_i \cap A_j \neq \emptyset$. In this $I_d = I_{d,a(d)}$ each A_i is an interval, hence $\cap A_i \neq \emptyset$ for them. Let $c_d \in A_i$, then their number is at most $\frac{1}{4}(|I_d|+1)^2$. Hence

$$M \leq \sum_1^n \frac{n^2}{d^2} + O(n \log n) \leq \frac{\pi^2}{24} n^2 + O(n \log n).$$

PROOF OF THEOREM 3

In the proof of Theorem 4 (a_2) is the only step allowing more than $O(n^{\frac{5}{3}} \log n)$ A_is and all the A_is have a common element c. Thus we may neglect all the $A_i \notin \mathbb{P}_1$, $A_i \not\ni c$ and restrict our considerations to the following two families:

$$\mathbb{F}_1 =: \{A_i \ni c: A_i \notin \mathbb{P}_1\},$$
$$\mathbb{F}_2 =: \{A_i: A_i \in \mathbb{P}_1\}.$$

By Lemma 3,

$$|\mathbb{F}_2| \leq \frac{\pi^2}{24} n^2 + O(n \log n).$$

$|\mathbb{F}_1|$ will be estimated as follows:

If $A_i \in \mathbb{F}_1$, then $c \in A_i$. Let $c_i \in A_i$ be one of the neighbours of c (i.e. $(c, c_i) \cap A_i = \emptyset$). Let $P_i \subseteq A_i$ be a maximal arithmetic progression of the form $\{c + l(c_i - c)\}$. Since $A_i \notin \mathbb{P}_1$, we can fix a $z_i \in A_i - P_i$. Trivially, $\{c, c_i, z_i\}$ is a δ-triplet. Hence

$$|\mathbb{F}_1| \leq \binom{n-1}{2}.$$

OPEN PROBLEMS

Among the many open problems connected with these theorems two seem to be the most closely related to them.

PROBLEM 1. Can one prove that $N \leq \binom{n-1}{2} + n$ in Theorem 3 if $n > n_0$?

The estimate of Theorem 3 can easily be improved but we think one should be able to prove, that the best choice of $A_1, \dots, A_N$ in Theorem 3 is if they are all the subsets of $[1, n]$ containing a fixed c and having at most three elements. If this is true, then there are other extremal systems as well, e.g. $\{c\}$ can be replaced by $[1, n]-\{c\}$ or some triplets $\{c, c+x, c+4x\}$ can be replaced by $\{c, c+x, c+2x, c+4x\}$, and some triplets $\{c-x, c, c+2x\}$ can be replaced by $\{c-x, c+x, c+2x\}$. Probably these are all the extremal systems.

PROBLEM 2. Let $A_1, \ldots, A_N$ be an extremal system for some $d \geq 2$ in Theorem 2. Is it true, that each A_i is in $\mathbb{P}_2$?

REFERENCES

1. N. G. de Bruijn and P. Erdös, On a combinatorial problem, *Nederl. Akad. Wetensch. Proc.* **51** (1948), 1277–1279.
2. P. Erdös, C. Ko and R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. (2)* **12** (1961), 313–320.
3. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, Clarendon Press, Oxford, 1971.
4. R. L. Graham, M. Simonovits and V. T. Sós, A note on the intersection properties of subsets of integers, *J. Combin. Theory Ser. A* **28** (1980), 106–110.
5. M. Simonovits and V. T. Sós, Graph intersection theorems, *Proc. Colloq. Combinatorics and Graph Theory, Orsay, Paris*, 1976, 389–391.
6. M. Simonovits and V. T. Sós, Intersection theorems for graphs II, *Colloq. Math. Soc. J. Bolyai, Combinatorics, Keszthely, Hungary* **18** (1976), 1017–1030.
7. M. Simonovits and V. T. Sós, Intersection theorems for subsets of integers, *Notices Amer. Math. Soc.* **25** (1978), A-33.
8. M. Simonovits and V. T. Sós, Intersection theorems on structures, *Combinatorial Mathematics, Optimal Design and their Applications, Ann. Discrete Math.* **6** (1980), 301–314.
9. V. T. Sós, Some remarks on the connection of graph theory, finite geometry and block designs, *Proc. Colloq. Int. Teorie Combinatorie, Roma, 1973, Accad. Naz. Lincei*, 223–233.

Received 10 April 1981

M. SIMONOVITS AND V. T. SÓS
Eötvös Loránd University, Analizis I,
1088 Budapest, Muzeum Krt 6–8, Hungary