Intersection properties of subsets of integers

Tibor Szabó

The Ohio State University, Columbus, Ohio,
231 W 18th Ave, Columbus, Ohio, 43210, USA and
Eötvös Loránd University, Budapest, Hungary

Abstract

Let N_k be the maximal integer such that there exist subsets $A_1, \dots, A_{N_k} \subseteq \{1, 2, \dots, n\}$ for which $A_i \cap A_j$ is an arithmetic progression of length at least k for every $1 \le i < j \le N_k$. R. L. Graham, M. Simonovits and V. T. Sós gave the exact value of N_0. For $k \ge 2$, Simonovits and T. Sós determined the asymptotic behavior of N_k.

In this paper we prove a conjecture of Simonovits and T. Sós concerning the asymptotic value of N_1. We show that

$$N_1 = \frac{n^2}{2} + O(n^{\frac{5}{3}} \log^3 n).$$

Moreover, we slightly improve the best known construction, thus disproving their conjecture on the exact extremal system.

1 Introduction

Intersection properties of sets have always been in the main focus of the theory of extremal set systems. The general question can be put the following way. Let I be a set of n elements and $\mathbf{P}$ be a property. What is the maximal number $N = N(\mathbf{P}, n)$ such that there exist N subsets $A_1, \dots, A_N$ of I for which each of the pairwise intersections $A_i \cap A_j, i \ne j$, have property $\mathbf{P}$? In classical results (e.g., [1, 2]) I is just a set having no additional structure, while property $\mathbf{P}$ is a restriction on the cardinality of the pairwise intersections.

In the late 70's Simonovits and T. Sós started to study problems where the baseset I had some structure (say it is the complete graph on n labelled vertices [8, 9] or the first n positive integers [6, 12]) and the intersection property $\mathbf{P}$ is formulated in terms of this structure. There are also nice results on some intersection properties of t-valued functions [4, 7]. A good survey of intersection theorems of structural type, along with open problems and conjectures can be found in [3, 11].

In the present paper we are considering an intersection problem on the positive integers. Let $I = [1, n]$ denote the first n positive integers. One of the basic questions is the following: what is the largest number N_0 such that we can choose N_0 subsets $A_1, \ldots, A_{N_0} \subseteq [1, n]$ and $A_i \cap A_j$ is an arithmetic progression for every $1 \leq i < j \leq N_0$.

In many of the classical problems of extremal set theory the maximal system is based on a simple constructions, where the intersection property is satisfied trivially. When trying to find extremal systems, two "straightforward" constructions seem to be natural:

1. All the subsets have ≤ 3 elements
2. All the subsets are arithmetic progressions.

Graham, Simonovits and T. Sós [6] settled the question about N_0, proving that Construction 1 gives the only maximal system. As a consequence,

$$N_0 = \binom{n}{3} + \binom{n}{2} + \binom{n}{1} + 1.$$

Empty intersections can be considered as a kind of "degeneracy". We can avoid them by requiring the intersections to be "real" arithmetic progressions, having at least $k > 0$ elements and ask for the maximal number N_k of such subsets. This question was addressed in [12] by Simonovits and T. Sós. They obtained for $k \geq 2$ that an appropriate construction of type 2 provides an asymptotically best family of sets. That is, for $k \geq 2$ we have

$$N_k = \left(\frac{\pi^2}{24} + o_k(1)\right) n^2.$$

A somewhat surprising feature of the bounds is that the asymptotic values are independent of k.

Our objective is to determine the asymptotic value of N_1, i. e. when only the empty arithmetic progression is excluded as an intersection. The definition below will be useful.

Definition 1 *We call a family $\mathcal{F}$ of subsets of* $[1, n]$ well-intersecting*, if for $A, B \in \mathcal{F}$, $A \neq B$, the subset $A \cap B$ is a non-empty arithmetic progression.*

For $k = 1$, we still have a version of Construction 1 which provides a larger family than any kind of system obtained by Construction 2. Let $c \in [1, n]$ be a fixed integer and consider the system of all subsets of $[1, n]$ with at most three elements which contain c. This is a well-intersecting family of cardinality $\binom{n}{2} + 1$. Simonovits and T. Sós conjectured [12, 3, 11] that this is an extremal system. As for an upper bound, in [12] they proved that

$$N_1 \leq \left(\frac{\pi^2}{24} + \frac{1}{2} + o(1) \right) n^2.$$

In Section 5 we give constructions having slightly more than $\binom{n}{2} + 1$ elements, disproving their conjecture on the extremal system. This also suggests that the extremal well-intersecting systems may have more subtle structure than the "straightforward" constructions. In the other direction we prove (Theorem 2.1) that

$$N_1 = \frac{n^2}{2} + O(n^{\frac{5}{3}} \log^3 n).$$

This shows that the conjecture is true in an asymptotic sense.

Our strategy is the following: in Section 2 we separate two cases and take care of the case of the big non-arithmetic progressions and of the small arithmetic progressions when their intersection is empty. This is a quite immediate consequence of results of [12]. In Section 3 we gather Lemmas needed for the second case. Finally Section 4 contains the argument for this, much longer case.

2 The Theorem

We introduce some notation: let $I = I(n) = [1, n] = \{1, 2, \ldots, n\}$. When considering intervals, we refer just to the integers contained in the interval. By the length of an interval $X = [x, y]$ we understand $|X| = y - x + 1$, that is the number of integers in X. $C_1, C_2, \ldots$ denote absolute constants. For $x_1, \ldots, x_m \in I$ we denote by $P(x_1, \ldots, x_m)$ the shortest arithmetic progression containing the numbers $x_1, \ldots, x_m$. We borrow the notion of δ- and

ν-triplets from [12]. If $\mathcal{F}$ is a family of subsets of I and $F \in \mathcal{F}$, then we say that a triplet $\{x, y, z\}$, contained in F, is a *determining*- or *δ-triplet* of F for $\mathcal{F}$, if there is no other $F' \in \mathcal{F}$ such that $\{x, y, z\} \subseteq F'$. We often drop references to F and $\mathcal{F}$, when this causes no ambiguity. If $\{x, y, z\} \subseteq F$ is not a δ-triplet of F, then we call it a *non-determining* or *ν-triplet*. Let us note here a key property of ν-triplets: if the family $\mathcal{F}$ is well-intersecting and $\{x, y, z\}$ is a ν-triplet of some $F \in \mathcal{F}$, then $P(x, y, z) \subseteq F$, since for some $F' \neq F \in \mathcal{F}$ we have $\{x, y, z\} \subseteq F \cap F'$.

The main contribution of the paper is the following.

Theorem 2.1 *If $\mathcal{F}$ is a well-intersecting family of subsets of $I(n)$, then*

$$|\mathcal{F}| < \frac{n^2}{2} + O(n^{\frac{5}{3}} \log^3 n).$$

In their paper [12], Simonovits and T. Sós obtained an upper bound by establishing that a well-intersecting system containing only non-arithmetic progressions (i.e. subsets, which are not arithmetic progressions) can have at most $(1/2 + o(1))n^2$ elements, while a well-intersecting system containing only arithmetic progressions has at most $(\pi^2/24 + o(1))n^2$ elements. To get rid of the extra term $(\pi^2/24)n^2$, we shall examine the restrictions originated from the well-intersection of arithmetic progressions and non-arithmetic progressions. These estimates, which are the crucial points of the proof, are contained in Lemma 3.2, Corollary 3.3 and in part (ii) of Theorem 4.1.

Proof of Theorem 2.1: We put $\mathcal{F} = \mathcal{F}_1 \cup \mathcal{F}_2$, where

$$\mathcal{F}_1 = \{F \in \mathcal{F} : F \text{ is an arithmetic progression}\} \text{ and}$$

$$\mathcal{F}_2 = \{F \in \mathcal{F} : F \text{ is not an arithmetic progression}\}.$$

First we are dealing with $\mathcal{F}_2$. We split $\mathcal{F}_2$ into $\mathcal{S}$mall and $\mathcal{B}$ig sets. We write $\mathcal{F}_2 = \mathcal{B} \cup \mathcal{S}$, where

$$\mathcal{B} = \{F \in \mathcal{F}_2 : |F| > n^{\frac{2}{3}}\},$$

$$\mathcal{S} = \{F \in \mathcal{F}_2 : |F| \leq n^{\frac{2}{3}}\}.$$

Big sets are handled by a Lemma of [12].

Lemma A ([12, Lemma 1]) Let $1 > c > 0$ be fixed and $\mathcal{B} = \{B_1, \ldots, B_M\} \subseteq 2^I$ be a well-intersecting family containing no arithmetic progressions. If $|B_1| = h > n^c$, then for every $x \in B_1$ and $t \leq h/20$ (for $n > n_0(c)$) either

(*i*) B_1 contains an arithmetic progression of at least $h - t$ elements, or

(*ii*) B_1 contains at least $th/50 \log h$ δ-triplets of form $\{x, y, z\}$.

Lemma 2.2 *Let $\mathcal{B} \subseteq 2^I$ be a well-intersecting family of subsets of integers such that no $B \in \mathcal{B}$ is an arithmetic progression. If $|B| > n^{\frac{2}{3}}$ for every $B \in \mathcal{B}$, then*

$$|\mathcal{B}| = O(n^{\frac{5}{3}} \log n).$$

Proof. We divide $\mathcal{B}$ into two parts:

$$\begin{aligned} \mathcal{B}' &= \{B \in \mathcal{B} : \exists\, x \in B \text{ such that } B \setminus \{x\} \text{ is an arithmetic progression}\}, \\ \mathcal{B}'' &= \mathcal{B} \setminus \mathcal{B}'. \end{aligned}$$

Using Lemma A with parameters $t = 1$ and $c = \frac{2}{3}$, we can estimate $|\mathcal{B}''|$. For every $B \in \mathcal{B}''$ and every $x \in B$, there are at least $n^{\frac{2}{3}}/50 \log n$ δ-triplets of the form $\{x, y, z\}$ contained in B. Hence every $B \in \mathcal{B}''$ contains at least $n^{\frac{4}{3}}/150 \log n$ δ-triplets. On the other hand, the number of possible δ-triplets is $\binom{n}{3}$, which implies that $\mathcal{B}''$ has at most $O(n^{\frac{5}{3}} \log n)$ elements.

We now turn to the estimation of $|\mathcal{B}'|$. One can assume that $n \geq 8$. Then for every $B \in \mathcal{B}'$, there exists a unique arithmetic progression $P_B \subset B$, such that $|P_B| + 1 = |B|$. Let d_B be the difference of P_B and let $x_B \in B \setminus P_B$ be the remaining element in B.

First fix an x and a d, and let $A, B \in \mathcal{B}'$ be such that $x = x_A = x_B$ and $d = d_A = d_B$. If $|P_A \cap P_B| \geq 3$, then there exists a $y \in A \cap B$ such that $\{x, y, y+d, y+2d\} \subset A \cap B$. This would imply that the arithmetic progression $P(x, y, y+d, y+2d)$ of difference less than or equal to d is contained in A. If this difference is d, then A itself is an arithmetic progression of difference d, which contradicts the assumption, that A is not an arithmetic progression. If the difference of $P(x, y, y+d, y+2d)$ is less than d, then there is an element of A, either between y and $y+d$, or between $y+d$ and $y+2d$, which is different from x. But this is impossible, since $|P_A| + 1 = |A|$. Thus, for fixed x and d we obtained that $|P_A \cap P_B| \leq 2$. Since $|P_A| > n^{\frac{2}{3}} - 1$, we infer that for d, x

fixed there are at most $n/(n^{\frac{2}{3}}-3) \leq 4n^{\frac{1}{3}}$ elements B of $\mathcal{B}'$ with $x_B = x$ and $d_B = d$.

We can select x n different ways and since $|P_B| \geq n^{\frac{2}{3}}$, we can choose at most $n^{\frac{1}{3}}$ different d's. We conclude that $|\mathcal{B}'| = O(n^{\frac{5}{3}})$. The lemma is proved.
□

Having the $\mathcal{B}$ig sets taken care of, we now consider $\mathcal{S}$. We distinguish two cases.

Case 1: $\bigcap \mathcal{S} = \emptyset$ or

Case 2: $\bigcap \mathcal{S} \neq \emptyset$.

The argument for *Case 1* is implicitly contained in [12].

Theorem B ([12, Theorem 4]) Let $\mathcal{S}$ be a well-intersecting family, such that $|S| \leq s$ for every $S \in \mathcal{S}$, and assume also that no $S \in \mathcal{S}$ is an arithmetic progression. In addition suppose that $\bigcap \mathcal{S} = \emptyset$. Then

$$|\mathcal{S}| \leq sn - \binom{s}{2} + O(n^{\frac{5}{3}} \log^3 n).$$

We can apply Theorem B for $\mathcal{S}$ with $s = n^{\frac{2}{3}}$ and obtain that $|\mathcal{S}| = O(n^{\frac{5}{3}} \log^3 n)$. Combining this with the result of Lemma 2.2, it gives us that $|\mathcal{F}| = |\mathcal{F}_1| + O(n^{\frac{5}{3}} \log^3 n)$. In view of the following easy Corollary 2.4 $|\mathcal{F}_1| \leq \frac{\pi^2}{24}n^2 + O(n \log n)$, hence we are done with Case 1.

We formulate the next statement in a slightly more general form than is needed here. This version will be useful later on.

For a family $\mathcal{A} \subseteq 2^I$ of arithmetic progressions we denote by $\mathcal{A}_d$ the subfamily of $\mathcal{A}$, which contains the arithmetic progressions of difference d. If $X, Y \subseteq I$, then

$$\mathcal{A}_X^Y = \{A \in \mathcal{A} : \text{ one endpoint of } A \text{ is in } X, \text{ the other is in } Y\}.$$

Lemma 2.3 *Let $\mathcal{A} \subseteq 2^I$ be a well-intersecting family of arithmetic progressions and let d be a positive integer, $X, Y \subseteq I$ intervals. Then*

$$|\mathcal{A}_{d,X}^Y| \leq \frac{|X||Y|}{d^2} + \frac{3n}{d}.$$

Proof. Since $A \cap B \neq \emptyset$ for any two $A, B \in \mathcal{A}_{d,X}^{Y}$, the elements of A and B are in the same residue class D modulo d. The two endpoints determine an arithmetic progression uniquely. The endpoint from X can be selected at most $|D \cap X| < |X|/d + 1$ different ways. Similarly we can choose at most $|D \cap Y| < |Y|/d + 1$ different endpoints from Y. Hence the lemma follows.
□

Corollary 2.4 *Let $\mathcal{A} \subseteq 2^I$ be a well-intersecting family of arithmetic progressions and let d be a positive integer. Then*

$$|\mathcal{A}_d| \leq \frac{n^2}{4d^2} + \frac{4n}{d}.$$

Proof. Since every two elements of $\mathcal{A}$ have nonempty intersections, they must be in the same residue class D modulo d. Within D, each arithmetic progression determines an interval, such that their pairwise intersection is nonempty. Thus there exists an element $a \in \bigcap \mathcal{A}_d$. We can use the preceding lemma with $X = [1, a]$ and $Y = [a, n]$.
□

3 Preparations

Before attacking *Case 2*, we prove some preparatory statements. They will be used extensively throughout the whole proof. The first one is a variant of Lemma 2 of [12], with a slightly simpler proof.

Lemma 3.1 *Let $\mathcal{S} \subseteq 2^I$ be a well-intersecting family containing no arithmetic progressions. Assume that there is an integer $c \in I$ contained by each element of $\mathcal{S}$ and that there exists a set $P \subseteq I \setminus \{c\}$, such that $S \cap P \neq \emptyset$ for every $S \in \mathcal{S}$. Then there exists a subfamily $\mathcal{H} = \mathcal{H}(P) \subseteq \mathcal{S}$, such that every $H \in \mathcal{H}$ contains a δ-triplet $\{c, x, p\}$ for $\mathcal{S}$, where $p \in P$, $x \in I \setminus \{c, p\}$ and $|\mathcal{S} \setminus \mathcal{H}| = O(n^{1+\epsilon})$ for any $\epsilon > 0$.*

Proof. For every $S \in \mathcal{S}$ we fix an element p_S of $P \cap S$. We define $\mathcal{H}$ in the following way:

$$\mathcal{H} = \{H \in \mathcal{S} : \exists\ \delta\text{-triplet } \{c, p_H, x_H\} \subseteq H, x_H \in I\}.$$

We partition the rest of $\mathcal{S}$ into two subfamilies:

$$\begin{aligned} \mathcal{S}' &= \{S \in \mathcal{S} \setminus \mathcal{H} : S \cap [c, p_S] \text{ is an arithmetic progression}\} \\ \mathcal{S}'' &= \{S \in \mathcal{S} \setminus \mathcal{H} : S \cap [c, p_S] \text{ is not an arithmetic progression}\} \end{aligned}$$

We intend to prove that $|\mathcal{S}' \cup \mathcal{S}''| = O(n^{1+\epsilon})$.

If $S \in \mathcal{S}' \cup \mathcal{S}''$, then $S \notin \mathcal{H}$. In particular, for every $z \in S \setminus \{c, p_S\}$ the subset $\{c, z, p_S\}$ is a ν-triplet, which implies that $P(c, z, p_S) \subset S$.

First let $S \in \mathcal{S}'$. Let d_S be the difference of $S \cap [c, p_S]$. Let A_S be the arithmetic progression of difference d_S, which is the extension of $S \cap [c, p_S]$ to the whole I. The difference d_z of $P(c, z, p_S)$ must be a multiple of d_S for every $z \in I \setminus \{c, p_S\}$, because $S \cap [c, p_S]$ is an arithmetic progression of difference d_S. Thus $S \subseteq A_S$. Let $y \in S$ be an element, for which $|y - c| = d_S$. As S is not an arithmetic progression, there is at least one $z = z_S \in S$ such that $P(c, y, z) \not\subseteq S$. In this case $\{c, y, z\}$ is clearly a δ-triplet of S.

Thus the pair (z_S, d_S) almost uniquely determines S (up to a factor of two). Observe that d_S must divide $|c - z|$. We infer that for every d and z, such that $d||c - z|$, there can be at most two sets $S \in \mathcal{S}'$, with $d_S = d$ and $z_S = z$ corresponding to the two possible selections of y. By [5, Theorem 320] we have

$$|\mathcal{S}'| \leq 2 \sum_{z=1}^{n} d(|z - c|) = O(n \log n).$$

Now let $S \in \mathcal{S}''$. Since $P(c, z, p_S) \subseteq S$ for every $z \in S$, $[c, p_S] \cap S$ is the union of at least two arithmetic progressions. Let d_1 and d_2 the difference of two, such that they are not contained in a third one (or in each other). Due to the well-intersecting property of $\mathcal{S}''$, S is uniquely determined , once d_1, d_2 and p_S are fixed. Also $d_1, d_2||c - p_S|$, so by [5, Theorem 315]

$$|\mathcal{S}''| \leq \sum_{p \in P} d(|c - p|)^2 \leq |P| n^{2\epsilon}.$$

□

The following lemma will often be convenient. Let $\mathcal{H} \subseteq 2^I$ be a family of subsets of integers, and let $c \in \bigcap \mathcal{H}$ be a fixed integer. Then let

$$\mathcal{H}^* = \mathcal{H}^*(c) = \{H \in \mathcal{H} : \exists\ \delta\text{-triplet } \{c, x_H, y_H\} \subseteq H \text{ for } \mathcal{H},$$
$$\text{such that } P(c, x_H, y_H) \subset H\}.$$

Lemma 3.2 *Let $\mathcal{H} \subseteq 2^I$ be a family of subsets of integers. Assume, that there is an integer $c \in I$ such that $c \in \cap\mathcal{H}$. Then*

$$|\mathcal{H}^*(c)| = O(n \log n).$$

Proof. For $H \in \mathcal{H}^*$, let d_H be the difference of $P(c, x_H, y_H)$. If $H \neq G \in \mathcal{H}^*$ such that $d_H = d_G = d$, then $P(c, x_H, y_H) \not\subseteq P(c, x_G, y_G) \subseteq G$, since $\{c, x_H, y_H\}$ is a δ-triplet for $\mathcal{H}$. In particular, $P(c, x_G, y_G)$ and $P(c, x_H, y_H)$ can not share an endpoint. This in turn implies that the number of elements of $\mathcal{H}^*$ with $d_H = d$ is at most $\lceil \frac{n}{d} \rceil / 2$. Hence $|\mathcal{H}^*| \leq \sum_{d=1}^{n} \frac{1}{2} \lceil \frac{n}{d} \rceil < n \log n$.
□

We define now the notation $\mathcal{H}_{d,X}^Y$ for a family $\mathcal{H}$ which do not contain arithmetic progressions. (For families containing only arithmetic progressions this notion was already introduced before Lemma 2.3.) Let $\mathcal{H} \subset 2^I$ be a family of non-arithmetic progressions, $c \in \bigcap \mathcal{H}$ and suppose that we have fixed a δ-triplet $\{c, p_H, x_H\}$ of every $H \in \mathcal{H}$. We call p_H and x_H the *essential elements* of H. For $X, Y \subseteq I$ we put

$$\mathcal{H}_X^Y = \{H \in \mathcal{H} : \text{ one essential element of } H \text{ is in } X, \text{ the other is in } Y\}.$$

Later we shall somehow fix a δ-triplet of every $H \in \mathcal{H}$ and the essential elements will be defined using those very triplets. If we say that the *essential elements* of an arithmetic progression are its endpoints, then the previous definition of $\mathcal{H}_X^Y$ will also be valid for families consisting of arithmetic progressions.

If d is a positive integer and $\mathcal{H} \subseteq 2^I$ is a family containing no arithmetic progressions, then we write

$$\mathcal{H}_d = \{H \in \mathcal{H} : \gcd(|p_H - c|, |x_H - c|) = d\}.$$

The next easy Corollary turns out to be very useful for estimating families which contain both arithmetic progressions and non-arithmetic progressions. In fact, this will be one of our main tools to exploit the well-intersection of progressions with non-progressions.

Corollary 3.3 *Let c be a positive integer. Let $\mathcal{H} \cup \mathcal{E} \subseteq 2^I$ be a well-intersecting family, such that each element of the family contains c. $\mathcal{E}$ consists of arithmetic progressions, while $\mathcal{H}$ does not contain any.*

Assume that we have fixed a δ-triplet $\{c, p_H, x_H\}$ of every set $H \in \mathcal{H}$ for $\mathcal{H} \cup \mathcal{E}$. If $X, Y \subseteq I$ are two intervals, then

$$|\left(\bigcup_{i=1}^{[n/d]} \mathcal{H}^Y_{id,X}\right) \cup \mathcal{E}^Y_{d,X}| \leq \frac{|X||Y|}{d^2} + \frac{3n}{d}.$$

If we require the chosen δ-triplets to be a δ-triplet only for $\mathcal{H}$ (as opposed to $\mathcal{H} \cup \mathcal{E}$) an easy upper bound, worse by a factor of 2, would follow just from the definitions. With the extra condition, the extra factor of 2 disappears. Lemma 3.2 will enable us to ensure this extra condition (and thus apply the Corollary) by paying practically no price in the main term.

Proof. Let D stand for the residue class of c modulo d.

Then for every pair $p \in Y \cap D$, $x \in X \cap D$, at most one of the two possibilities can hold:

—there is an arithmetic progression E in $\mathcal{E}^Y_{d,X}$ with endpoints p and x.

—there is an element H of $\bigcup_{i=1}^{n/d} \mathcal{H}^Y_{id,X}$, such that $p_H = p$ and $x_H = x$.

Indeed, the occurrence of both events would imply that $\{c, p_H, x_H\}$ is a ν-triplet for $\mathcal{E} \cup \mathcal{H}$.

Moreover, E or H is uniquely determined by the pair (p, x). The number of such pairs is at most $(|X|/d + 1)(|Y|/d + 1)$ and the statement follows.

□

We conclude this section by proving a number theoretical lemma. It is a slight generalization of a familiar fact on the average behavior of Euler's φ function: $\sum_{i=1}^{j} \varphi(i) = 3j^2/\pi^2 + O(j \log j)$ [5, Theorem 330].

Lemma 3.4 *Let $K, J \subseteq [-n, n]$ be two intervals. Then there exist absolute constants C_1, C_2 such that*

$$N = N(K, J, n) = \#\{(a, b) : a \in K, b \in J, \gcd\{a, b\} = 1\} \leq$$

$$\leq \frac{6}{\pi^2}|K||J| + C_1 n \log n + C_2.$$

Proof. It suffices to assume that $K, J \subseteq I$. The statement follows easily from this case.

Let $f : I^2 \to \mathbf{Z}$ be a function. We recall a version of the Möbius Inversion Formula:

$$\sum_{\substack{(x,y) \in I^2 \\ \gcd\{x,y\}=1}} f(x, y) = \sum_{d=1}^{n} \mu(d) S_d,$$

where μ is the Möbius function and

$$S_d = \sum_{\substack{(x,y)\in I^2 \\ d|x, d|y}} f(x,y).$$

We apply this to the following function f:

$$f(x,y) = \begin{cases} 1 & \text{if } x \in K, y \in J \\ 0 & \text{otherwise} \end{cases}$$

It is immediate that

$$S_d = \left(\frac{|K|}{d} + \epsilon_{K,d} \right) \left(\frac{|J|}{d} + \epsilon_{J,d} \right),$$

where $0 \leq |\epsilon_{K,d}|, |\epsilon_{J,d}| \leq 1$. We have

$$N(K,J,n) = \sum_{\substack{(x,y)\in I^2 \\ \gcd\{x,y\}=1}} f(x,y) = \sum_{d=1}^{n} \mu(d) \left(\frac{|K|}{d} + \epsilon_{K,d} \right) \left(\frac{|J|}{d} + \epsilon_{J,d} \right) =$$

$$= \sum_{d=1}^{n} \frac{\mu(d)|K||J|}{d^2} + \sum_{d=1}^{n} \frac{\mu(d)}{d} \left(|K|\epsilon_{J,d} + |J|\epsilon_{K,d} \right) + \sum_{d=1}^{n} \mu(d)\epsilon_{K,d}\epsilon_{J,d} \leq$$

$$\leq \sum_{d=1}^{n} \frac{\mu(d)|K||J|}{d^2} + 2n(\log n + 1) + n =$$

$$= \sum_{d=1}^{\infty} \frac{\mu(d)|K||J|}{d^2} - \sum_{d=n+1}^{\infty} \frac{\mu(d)|K||J|}{d^2} + 2n\log n + 3n \leq$$

$$\leq \frac{6}{\pi^2}|K||J| + 2n\log n + 4n.$$

For the evaluation of the first of the infinite sums the reader is referred to Theorem 287 of [5].

□

Let d be a positive integer. We extend the preceding lemma to count the pairs from K, J with gcd-value precisely d. The reduction to the relatively prime case is straightforward; we work with $[-[\frac{n}{d}], [\frac{n}{d}]]$ in the place of $[-n, n]$.

Corollary 3.5 *Let $K, J \subseteq [-n, n]$ be two intervals and d be positive integer. Then there exist absolute constants C_3, C_4 such that*

$$\#\{(a,b) : a \in K, b \in J, \gcd\{a,b\} = d\} \leq \frac{6}{\pi^2}\frac{|K||J|}{d^2} + \frac{C_3 n \log n}{d} + C_4.$$

$\square$

This Corollary is used to estimate $|\mathcal{H}_{d,X}^{Y}|$, where $\mathcal{H} \subseteq 2^I$ is a family containing no arithmetic progressions. (In this sense Corollary 3.5 corresponds to Lemma 2.3.)

4 Case 2: $\cap \mathcal{S} \neq \emptyset$

This section contains the remaining (and most laborious) part of the proof of Theorem 2.1. We assume throughout that $\bigcap \mathcal{S} \neq \emptyset$. We remark here that the conditions $|S| \leq n^{\frac{2}{3}}$ for $S \in \mathcal{S}$ will no longer have significance.

We fix a $c \in \bigcap \mathcal{S}$ and partition $\mathcal{F}_1$ (the subfamily of $\mathcal{F}$ containing the arithmetic progressions) into

$$\begin{aligned} \mathcal{P} &= \{P \in \mathcal{F}_1 : c \notin P\} \text{ and} \\ \mathcal{E} &= \{E \in \mathcal{F}_1 : c \in E\}. \end{aligned}$$

We set out to prove that $|\mathcal{P} \cup \mathcal{E} \cup \mathcal{S}| \leq \frac{n^2}{2} + O(n^{1+\epsilon})$. This, together with Lemma 2.2, implies Theorem 2.1.

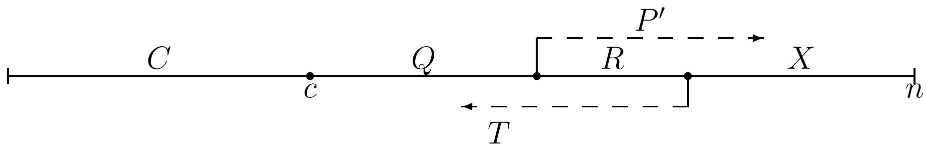
As for the strategy of our proof, we shall divide I into subintervals given by a few (3 or 4) division points. We are going to classify the elements of $\mathcal{P} \cup \mathcal{E} \cup \mathcal{S}$ by the subintervals containing their essential elements. We estimate then the cardinalities of these subsystems separately. In doing this, we ignore sets which have one (or both) of their essential elements at the division points, because there are just $O(n \log n)$ of those.

The following theorem deals with the hardest case of the argument.

Theorem 4.1 *Let $c \in [1, n]$ be an integer. Let $\mathcal{P}$ be a family of arithmetic progressions such that none of the elements of $\mathcal{P}$ contains c. Let $\mathcal{E}$ be a family of arithmetic progressions such that every element of $\mathcal{E}$ contains c. Let $\mathcal{S}$ be a family of non-arithmetic progressions, such that each of the elements contains* *c. Suppose that* $\mathcal{P}\cup\mathcal{E}\cup\mathcal{S}$ *is well-intersecting. In addition assume that there is a* $P'\in\mathcal{P}$ *such that* $\min P' > c$ *or* $\max P' < c$*. Then*

$$|\mathcal{P}\cup\mathcal{E}\cup\mathcal{S}| \le \frac{n^2}{2} + O(n^{1+\epsilon})$$

Proof. Without loss of generality we can assume that there is a $P'\in\mathcal{P}$ such that $\min P' > c$ (If $\max P' < c$, then we can reflect the set system about $(n+1)/2$). Let P' be one with the greatest left-endpoint $c+q$. Let $T\in\mathcal{P}$ be an arithmetic progression with the smallest right-endpoint $c+q+r$. Obviously, since $T\cap P'$ is nonempty, we have $0\le r$. Let $x=n-r-q-c$, $C=[1,c]$, $Q=(c,c+q]$, $R=(c+q,c+q+r]$ and $X=(c+q+r,n]$. The following figure shows the setup.



After designating the division points $c, c+q$ and $c+q+r$, we can assume without loss of generality that no element of $\mathcal{P}\cup\mathcal{E}$ has endpoints in one or both of these points. Leaving out these arithmetic progressions decreases $|\mathcal{P}\cup\mathcal{E}|$ only by $O(n\log n)$. Let us note here that although P' and T are among the arithmetic progressions "left out", we do not forget that every element of $\mathcal{P}\cup\mathcal{E}\cup\mathcal{S}$ still has to intersect P' and T nontrivially.

We can use Lemma 3.1 with $P=[c+q,n]$ to obtain a subfamily $\mathcal{K}$ of $\mathcal{S}$, such that for every $H\in\mathcal{K}$ there exist a δ-triplet $\{c,p_H,x_H\}$ of H for $\mathcal{K}$, where $p_H\in P\cap H$ and $|\mathcal{S}\setminus\mathcal{K}| = O(n^{1+\epsilon})$.

Quite possibly we have a choice in selecting the δ-triplet of an $H\in\mathcal{K}$. We fix one by the following list of preferences: first we choose δ-triplets to make $|\mathcal{K}_R^I|$ maximal. After this, we choose δ-triplets to make $|\mathcal{K}_X^{C\cup Q}|$ maximal. Finally we favor choices where $|x_H-p_H|$ is minimal.

Furthermore, we can leave out the elements of $\mathcal{K}^*(c)$ from $\mathcal{K}$ and obtain a family $\mathcal{H}$ such that $P(c,x_H,p_H)\not\subseteq H$ for any $H\in\mathcal{H}$. In particular $\{c,x_H,p_H\}$ is a δ-triplet of H for $\mathcal{H}\cup\mathcal{E}$, that is the "extra condition" of Lemma 3.3 holds. By Lemma 3.2 $|\mathcal{K}\setminus\mathcal{H}| = O(n\log n)$, hence $|\mathcal{S}\setminus\mathcal{H}| = O(n^{1+\epsilon})$.

We can also assume without loss of generality that $\mathcal{H}$ does not contain a set H, which contains a δ-triplet of the form $\{c,d_1,d_2\}$, where either d_1 or d_2 is a division point. We can simply leave out those sets from $\mathcal{H}$ which contain a δ-triplet of this kind, and repeat this procedure until we got rid of all of them. All together we left out only $O(n)$ sets. (Less than n sets for each division point.)

We estimate the sizes of the essential parts of $\mathcal{P} \cup \mathcal{E} \cup \mathcal{H}$ in six steps.

(i) $\mathcal{H}_X^X = \emptyset$.

Assume there exists $H \in \mathcal{H}_X^X$. Since $H \cap T \neq \emptyset$, there must exist a $z \in H \cap (I \setminus X)$, $z \neq c$. Let $z_H \neq c$ be an element of $H \cap (I \setminus X)$, such that $|z_H - c|$ is minimal. Here z_H is different from x_H and p_H, because both x_H and p_H are in X.

The maximality of $\mathcal{K}_R^I$ and $\mathcal{K}_X^{Q \cup C}$ implies that $\{c, z_H, x_H\}$ and $\{c, z_H, p_H\}$ were ν-triplets for $\mathcal{K}$. Hence $P(c, z_H, x_H), P(c, z_H, p_H) \subset H$, and both arithmetic progressions have difference $|z_H - c|$, since $|z_H - c|$ is minimal. Thus $P(c, x_H, p_H) \subset H$, which is a contradiction, so $\mathcal{H}_X^X = \emptyset$.

(ii) $|\mathcal{P}_1 \cup \mathcal{H}_X^Q| \leq xq + O(n)$.

Obviously we have $\mathcal{P}_1 = \mathcal{P}_{1,X}^Q$, so an upper bound of $2xq$ is trivial. In order to estimate the cardinality more carefully, we shall partition $\mathcal{H}_X^Q$ in the following way:

$$\begin{aligned}
\hat{\mathcal{H}} &= \{H \in \mathcal{H}_X^Q : [p_H, x_H] \cap H = \{p_H, x_H\}\}, \\
\bar{\mathcal{H}} &= \{H \in \mathcal{H}_X^Q : [p_H, x_H] \cap H \text{ is not an arithmetic progression}\}, \\
\dot{\mathcal{H}} &= \{H \in \mathcal{H}_X^Q : [p_H, x_H] \cap H \text{ is an arithmetic progression of length} \geq 3\}.
\end{aligned}$$

First we show that $\dot{\mathcal{H}} = \emptyset$. Assume there exists $H \in \dot{\mathcal{H}}$. Let $a = \min H \cap X$ and $b = \max H \cap (Q \cup R)$. Obviously $x_H \leq b < a \leq p_H$ and since $|H \cap [x_H, p_H]| \geq 3$, $\{x_H, p_H\} \neq \{a, b\}$.

If $b \in R$, then $\{a, b, c\}$ is a ν-triplet for $\mathcal{K}$, since first $\mathcal{K}_R^I$ was chosen to be maximal.

If $b \in Q$, then $\{a, b, c\}$ also must be a ν-triplet for $\mathcal{K}$, since $|x_H - p_H|$ was chosen to be minimal and $|x_H - p_H| > |a - b|$.

Thus in each case $P(a, b, c) \subset H$. Because of the definition of a and b, there is no element of H between a and b, which fact implies that

$$\text{difference of } (H \cap [x_H, p_H]) = |a - b| = \text{difference of } P(a, b, c).$$

Hence the arithmetic progression $P(c, x_H, p_H)$ is contained in H, which is a contradiction, so $\dot{\mathcal{H}} = \emptyset$.

Now we turn to $\hat{\mathcal{H}} \cup \bar{\mathcal{H}} \cup \mathcal{P}_1$ by defining further subdivisions. For $i \in Q$ let us define:

$$\begin{aligned}\hat{\mathcal{H}}^i &= \{H \in \hat{\mathcal{H}} : x_H = i\},\\ \bar{\mathcal{H}}^i &= \{H \in \bar{\mathcal{H}} : x_H = i\},\\ \mathcal{P}_1^i &= \{[i,l] \in \mathcal{P}_1\}.\end{aligned}$$

Let $i_1 < i_2 < \ldots < i_k \in Q$ be the indices i_j for which $\mathcal{P}_1^{i_j}$ is nonempty. We fix a $j \geq 2$. Let $[i_j, l]$ be the element of $\mathcal{P}_1^{i_j}$ with the smallest right-endpoint l. Let $p_K = \min\{p_H : H \in \bar{\mathcal{H}}^{i_j}\}$ with a corresponding set $K \in \bar{\mathcal{H}}^{i_j}$. If $\bar{\mathcal{H}}^{i_j} = \emptyset$, then we define $p_K := n+1$. Obviously $l < p_K$, otherwise $[i_j, l] \cap K$ is not an arithmetic progression.

—If $H \in \hat{\mathcal{H}}^{i_{j-1}}$, then $p_H \leq l$, otherwise $H \cap [i_j, l]$ would be empty. Thus $|\hat{\mathcal{H}}^{i_{j-1}}| \leq l - (n-x)$. (This is actually true for any $\hat{\mathcal{H}}^m$, where $m < i_j$.)

—If $[i_j, f] \in \mathcal{P}_1^{i_j}$, then $l \leq f$ by the definition of l and also $f < p_K$, since $[i_j, f] \cap K$ must be an arithmetic progression. Thus $|\mathcal{P}_1^{i_j}| \leq p_K - l$.

—If $H \in \bar{\mathcal{H}}^{i_j}$, then $p_K \leq p_H \leq n$ by the definition of K, which gives us $|\bar{\mathcal{H}}^{i_j}| \leq n - p_K + 1$.

Combining the previous three inequalities we obtain that $|\hat{\mathcal{H}}^{i_{j-1}} \cup \mathcal{P}_1^{i_j} \cup \bar{\mathcal{H}}^{i_j}| \leq x+1$ for every $2 \leq j \leq k$.

If $\mathcal{P}_1^i$ is empty for some $i \in Q$, then $|\hat{\mathcal{H}}^i \cup \bar{\mathcal{H}}^i| \leq x$, because for every element $p \in X$, $\{c, i, p\}$ can be a δ-triplet of at most one set in $\hat{\mathcal{H}}^i \cup \bar{\mathcal{H}}^i$. Summing up we have

$$|\hat{\mathcal{H}} \cup \mathcal{P}_1 \cup \bar{\mathcal{H}}| = |\bigcup_{i=c+1}^{c+q-1} (\hat{\mathcal{H}}^i \cup \mathcal{P}_1^i \cup \bar{\mathcal{H}}^i)| =$$

$$= |\bigcup_{j=2}^{k} (\hat{\mathcal{H}}^{i_{j-1}} \cup \mathcal{P}_1^{i_j} \cup \bar{\mathcal{H}}^{i_j})| + |\hat{\mathcal{H}}^{i_k} \cup \mathcal{P}_1^{i_1} \cup \bar{\mathcal{H}}^{i_1}| + |\bigcup_{i:\mathcal{P}_1^i = \emptyset} (\hat{\mathcal{H}}^i \cup \bar{\mathcal{H}}^i)| \leq$$

$$\leq (k-1)(x+1) + 3n + (q-1-k)x = qx + O(n).$$

(iii) $|\bigcup_{i=2}^{n} (\mathcal{P}_i \cup \mathcal{E}_{i,X}^C)| \leq (\frac{\pi^2}{6} - 1)x(q+c) + O(n \log n)$

By the definition of x and q obviously $\mathcal{P}_i = \mathcal{P}_{i,X}^{Q \cup C}$. We can apply Lemma 2.3 for the well-intersecting family $\mathcal{P} \cup \mathcal{E}$ of arithmetic progressions and sum up the results.

(iv) Trivially $|\mathcal{H}_X^R| \le rx$.

(v) $|\mathcal{E}_{1,X}^C \cup \mathcal{H}_X^C| \le xc + O(n)$

Since the δ-triplets $\{c, x_H, p_H\}$ of the elements of $\mathcal{H}$ are δ-triplets for $\mathcal{H} \cup \mathcal{E}$, we can apply Corollary 3.3 with $d = 1$, $Y = C$.

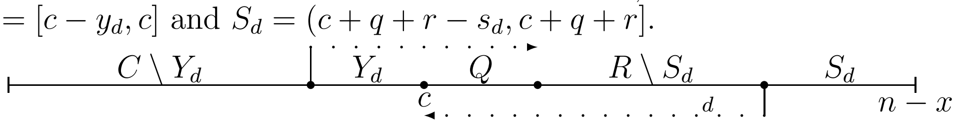
(vi)$_d$ There exist absolute constants C_5, C_6 such that for any positive integer d

$$|\mathcal{E}_{d,R}^C \cup \mathcal{H}_{d,R}^{Q\cup C\cup R}| \le \frac{1}{d^2}\frac{6}{\pi^2} f(r,q,c,s_d) + C_5 \frac{n \log n}{d} + C_6, \text{ where}$$

$c + q + r - s_d$ is the greatest right-endpoint among the elements of $\mathcal{E}_{d,R}^C$ (if $\mathcal{E}_{d,R}^C = \emptyset$, then we can set $s_d := r$), $0 \le s_d \le r$, and

$$f(r,q,c,s_d) = \frac{\pi^2}{6} c(r - s_d) + s_d(q + c + r) - \frac{s_d^2}{2}.$$

Proof. Assume first that $\mathcal{E}_{d,R}^C$ is not empty. Let $c - y_d$ be the smallest left-endpoint occurring among the elements of $\mathcal{E}_{d,R}^C$, $0 \le y_d \le c$. We define $Y_d = [c - y_d, c]$ and $S_d = (c + q + r - s_d, c + q + r]$.



To make notations easier to read, we write S instead of S_d and Y instead of Y_d. By Corollary 3.3, $|\mathcal{E}_{d,R\setminus S}^Y \cup \mathcal{H}_{d,R\setminus S}^Y| \le (r - s_d)(y_d + 1)/d^2 + 3n/d$.

By the definition of s_d for any $H \in \mathcal{H}_{d,R\setminus S}^{Q\cup(R\setminus S)}$ there exists an $E \in \mathcal{E}_{d,C}^R$ containing the δ-triplet $\{c, p_H, x_H\}$. This would mean $\{c, p_H, x_H\}$ is a ν-triplet for $\mathcal{H} \cup \mathcal{E}$, which is not possible. Thus we have $\mathcal{H}_{d,R\setminus S}^{Q\cup(R\setminus S)} = \emptyset$.

With the aid of Corollary 3.5, we estimate the following decomposition term-by-term:

$$\mathcal{E}_{d,R}^C \cup \mathcal{H}_{d,R}^{Q\cup C\cup R} = \mathcal{E}_{d,R\setminus S}^Y \cup \mathcal{H}_{d,R\setminus S}^Y \cup \mathcal{H}_{d,R\setminus S}^{C\setminus Y} \cup \mathcal{H}_{d,S}^{Q\cup C\cup(R\setminus S)} \cup \mathcal{H}_{d,S}^S.$$

We obtain that

$$|\mathcal{E}_{d,R}^C \cup \mathcal{H}_{d,R}^{Q\cup C\cup R}| \le \frac{(r - s_d)y_d}{d^2} + \frac{6}{\pi^2}\frac{(r - s_d)(c - y_d)}{d^2} +$$

$$+ \frac{6}{\pi^2}\frac{s_d(q + c + r - s_d)}{d^2} + \frac{3}{\pi^2}\frac{s_d^2}{d^2} + C_5 \frac{n \log n}{d} + C_6$$

This function is increasing in y_d. By putting in $y_d = c$, we obtain the bound of $(\mathrm{vi})_d$. In the case $\mathcal{E}_{d,R}^C = \emptyset$ the claim follows immediately from the previous estimates.

By adding up the estimates (i)-(v) and $(\mathrm{vi})_d$ for every $d \le n$ we obtain

$$|\mathcal{E} \cup \mathcal{H} \cup \mathcal{P}| \le \frac{\pi^2}{6} x(q+c) + rx + O(n \log n) +$$

$$+ \sum_{d=1}^{n} \left(\frac{1}{d^2} \frac{6}{\pi^2} f(r,q,c,s_d) + C_5 \frac{n \log n}{d} + C_6 \right).$$

To conclude the proof of our Theorem, it suffices to show that

$$\frac{\pi^2}{6} x(q+c) + rx + \sum_{d=1}^{\infty} \frac{1}{d^2} \frac{6}{\pi^2} f(r,q,c,s_d) \le \frac{n^2}{2},$$

when $q + c + x + r = n$, and $0 \le s_d \le r$.

Or, equivalently, since

$$\frac{\pi^2}{6} x(q+c) + rx = \sum_{d=1}^{\infty} \frac{1}{d^2} \frac{6}{\pi^2} \left(\frac{\pi^2}{6} x(q+c) + rx \right),$$

it is enough to show, that if $n = q + c + x + r$, $0 \le s \le r$ then

$$\frac{\pi^2}{6} x(q+c) + rx + f(r,q,c,s) \le \frac{n^2}{2}.$$

We rewrite our function using the new variable $z = r - s$. In this setting we seek the maximum of

$$\frac{\pi^2}{6} x(q+c) + (s+z)x + \frac{\pi^2}{6} cz + s(q+c+s+z) - \frac{s^2}{2}, \qquad (1)$$

subject to constraints $n = c + q + x + s + z$ and $c, q, x, s, z \ge 0$. We do not decrease the maximum of (1) replacing c by $q + c$ and q by 0:

$$\frac{\pi^2}{6} xc + (s+z)x + \frac{\pi^2}{6} cz + s(c+s+z) - \frac{s^2}{2} \le$$

$$\frac{\pi^2}{6} c(x+z) + \left(\frac{x+z}{2} \right)^2 + s(c+z+x) + \frac{s^2}{2}.$$

Here we introduce new variables again replacing x by $x + z$ and z by 0:

$$\frac{\pi^2}{6}cx + \frac{x^2}{4} + sx + sc + \frac{s^2}{2} = \frac{(s+x+c)^2}{2} - \frac{x^2}{4} - \frac{c^2}{2} + \left(\frac{\pi^2}{6} - 1\right)cx =$$

$$= \frac{n^2}{2} - \left(\frac{x}{2} - \left(\frac{\pi^2}{6} - 1\right)c\right)^2 - \left(\frac{1}{2} - \left(\frac{\pi^2}{6} - 1\right)^2\right)c^2 \leq \frac{n^2}{2}.$$

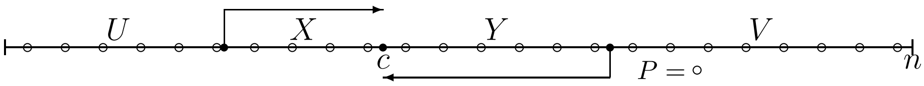
We have finished the proof of Theorem 4.1.
□

To complete our reasoning, we just have to treat the families where every element of $\mathcal{P}$ "jumps over" c, that is for each $P \in \mathcal{P}$ $\min_{x\in P} x < c < \max_{x\in P} x$. In particular $\mathcal{P}_1 = \emptyset$.

Theorem 4.2 *Assume that there exists a $P' \in \mathcal{P}_k$ for some $k \geq 2$. Then*

$$|\mathcal{E}_1 \cup \mathcal{S}| \leq \max\left\{\frac{n^2}{4}, \left(\frac{1}{k} - \frac{1}{2k^2}\right)n^2\right\} + O(n^{1+\epsilon}).$$

Proof. Let $c - x$ the smallest left-endpoint, $c + y$ the greatest right-endpoint occurring among the elements of $\mathcal{E}_1$. If $\mathcal{E}_1$ is empty, then we set $x = y = 0$. Let P be the arithmetic progression of difference k, which is the extension of P' to I, $\left[\frac{n}{k}\right] \leq |P| \leq \left[\frac{n}{k}\right] + 1$. Let $U = [1, c-x)$, $X = [c-x, c]$, $Y = [c, c+y]$ and $V = (c+y, n]$, $u = c - x, v = n - c - y$.



We can apply Lemma 3.1 with this P to obtain a subfamily $\mathcal{H}$, such that every $H \in \mathcal{H}$ contains a δ-triplet $\{c, p_H, x_H\}$ for $\mathcal{S}$ with $p_H \in P$, $x_H \in I \setminus \{c, p_H\}$ and $|\mathcal{S} \setminus \mathcal{H}| \leq O(n^{1+\epsilon})$. Let us fix such a δ-triplet of every $H \in \mathcal{H}$. With the aid of Lemma 3.2 we can assume that they are δ-triplets for $\mathcal{H} \cup \mathcal{E}$. (We can simply leave out $O(n \log n)$ "bad" sets, like we did in the proof of Theorem 4.1.)

By Corollary 3.3

$$|\mathcal{E}_{1,X}^{Y} \cup \mathcal{H}_X^Y| \leq (x+1)(y+1) + O(n) \leq \left(\frac{x+y}{2}\right)^2 + O(n). \qquad (2)$$

By the definition of X and Y, for any set $H \in \mathcal{H}_X^X \cup \mathcal{H}_Y^Y$ there is a set $E \in \mathcal{E}_1$ which contains H's δ-triplet $\{c, p_H, x_H\}$. This is a contradiction, since $\{c, p_H, x_H\}$ should be a δ-triplet for $\mathcal{H} \cup \mathcal{E}$. Hence

$$\mathcal{H}_X^X \cup \mathcal{H}_Y^Y = \emptyset. \tag{3}$$

Next we count the remaining part of $\mathcal{H}$ (see the figure):

$$|\mathcal{H}_{U\cup V}^I| \leq \frac{x+y}{k}(u+v) + \frac{u+v}{k}n - \frac{x+y}{k}\frac{u+v}{k} - \frac{\left(\frac{u+v}{k}\right)^2}{2} + O(n). \tag{4}$$

The definition of X and Y implies that $\mathcal{E}_1 = \mathcal{E}_{1,X}^Y$, hence by (3)

$$|\mathcal{E}_1 \cup \mathcal{H}| \leq (\text{r.h.s. of (2)}) + (\text{r.h.s. of (4)}). \tag{5}$$

We do not decrease the maximum of the right hand side of (5), if we substitute $y = 0, v = 0$ and replace x and u by $x + y$ and $u + v$, respectively. Up to an "error term" of $O(n)$ we obtain the expression

$$\frac{x^2}{4} + \frac{ux}{k} + \frac{u}{k}n - \frac{ux}{k^2} - \frac{u^2}{2k^2}.$$

It is a routine task to determine the maximum of this on the domain $x, u \geq 0$, with the constraint $x + u = n$. For $k \geq 4$ the maximum is attained at $u = 0, x = n$ and the value is $n^2/4$, while for $k = 2$ or 3 the maximum is at $u = n, x = 0$ and the maximal value is $(1/k - 1/2k^2)n^2$.

□

By the previous theorem and by Corollary 2.4, if $\bigcup_{i=3}^n \mathcal{P}_i \neq \emptyset$, then

$$|\mathcal{E} \cup \mathcal{P} \cup \mathcal{S}| = |\mathcal{E}_1 \cup \mathcal{S}| + |\bigcup_{i=2}^n (\mathcal{E}_i \cup \mathcal{P}_i)| \leq$$

$$\leq \left(\frac{1}{3} - \frac{1}{18} + \frac{\pi^2}{24} - \frac{1}{4}\right) n^2 + O(n^{1+\epsilon}) < \frac{n^2}{2} + O(n^{1+\epsilon}).$$

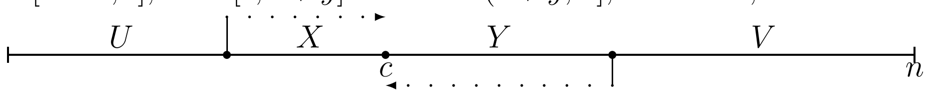
We are left with the case when $\bigcup_{i=3}^n \mathcal{P}_i \cup \mathcal{P}_1 = \emptyset$.

If there exists an element $P \in \mathcal{P}_2$, then we choose our δ-triplets by employing Lemma 3.1 with this P. If $\mathcal{P}_2 = \emptyset$, then we just use $P = I \setminus \{c\}$. We obtain a subfamily $\mathcal{H} \subseteq \mathcal{S}, |\mathcal{S} \setminus \mathcal{H}| = O(n^{1+\epsilon})$ such that every $H \in \mathcal{H}$ contains a δ-triplet $\{c, p_H, x_H\}$ for $\mathcal{S}$ with $p_H \in P$ and $x_H \in I \setminus \{c, p_H\}$. By Lemma 3.2 we can assume again that $\{c, p_H, x_H\}$ is a δ-triplet for $\mathcal{H} \cup \mathcal{E}$.

Lemma 4.3 *Let d be a positive integer. There exists absolute constants C_7, C_8 such that*

$$|\mathcal{E}_d \cup \mathcal{H}_d| \le \frac{3}{\pi^2}\frac{n^2}{d^2} + C_7\frac{n\log n}{d} + C_8.$$

Proof. If $\mathcal{E}_d = \emptyset$ then the statement follows immediately from Corollary 3.5. Otherwise let $c - x$ be the smallest left-endpoint, $c + y$ the greatest right-endpoint occurring among the elements of $\mathcal{E}_d$. Let $U = [1, c - x)$, $X = [c - x, c]$, $Y = [c, c + y]$ and $V = (c + y, n]$, $u = c - x, v = n - c - y$.



By Corollary 3.3

$$|\mathcal{E}^Y_{d,X} \cup \mathcal{H}^Y_{d,X}| \le \frac{(x+1)(y+1)}{d^2} + \frac{3n}{d} \le \left(\frac{x+y}{2d}\right)^2 + \frac{6n}{d} \le \frac{3}{\pi^2}\left(\frac{x+y}{d}\right)^2 + \frac{6n}{d}.$$

By the definition of X, Y, for any $H \in \mathcal{H}^X_{d,X} \cup \mathcal{H}^Y_{d,Y}$ there exists an $E \in \mathcal{E}_d$ which contains the δ-triplet $\{c, p_H, x_H\}$. This fact would imply that $\{c, p_H, x_H\}$ is a ν-triplet for $\mathcal{H} \cup \mathcal{E}$, so we have

$$\mathcal{H}^X_{d,X} \cup \mathcal{H}^Y_{d,Y} = \emptyset.$$

By Corollary 3.5

$$|\mathcal{E}_d \cup \mathcal{H}_d| = |\mathcal{E}^Y_{d,X} \cup \mathcal{H}^{X\cup Y}_{d,X\cup Y}| + |\mathcal{H}^{U\cup V}_{d,X\cup Y}| + |\mathcal{H}^{U\cup V}_{d,U\cup V}| \le$$

$$\le \frac{3}{\pi^2}\left(\frac{x+y}{d}\right)^2 + \frac{6}{\pi^2}\frac{x+y}{d}\frac{u+v}{d} + \frac{3}{\pi^2}\left(\frac{u+v}{d}\right)^2 + C_7\frac{n\log n}{d} + C_8 =$$

$$= \frac{3}{\pi^2}\left(\frac{n}{d}\right)^2 + C_7\frac{n\log n}{d} + C_8.$$

□

If $\mathcal{P}_2 = \emptyset$, then $\mathcal{P} = \emptyset$ and we obtain $|\mathcal{E} \cup \mathcal{H}| \le n^2/2 + O(n\log^2 n)$ by summing up the bound of the preceding lemma over $1 \le d \le n$.

If there exists an element $P \in \mathcal{P}_2$, then it must "jump over" c. This fact ensures that $\mathcal{H}_2$ is empty, since $p_H \in P$ implies that $|p_H - c|$ is not divisible by 2. Observe also, that either $\mathcal{P}_2 = \emptyset$ or $\mathcal{E}_2 = \emptyset$. We infer that $|\mathcal{H}_2 \cup \mathcal{E}_2 \cup \mathcal{P}_2| \le n^2/16 < \frac{3}{\pi^2}\frac{n^2}{2^2}$. Combining this with the bound of the previous lemma in cases $d \ne 2$, we obtain that $|\mathcal{P} \cup \mathcal{E} \cup \mathcal{H}| \le n^2/2 + O(n\log^2 n)$.

This finishes the proof of Theorem 2.1. □

5 Constructions

Here we present constructions giving a lower bound on N_1. The systems to be described constitute a modest improvement over the best known families. We start out with

$$\mathcal{C}_1 = \{S \subset [1, n];\ |S| \leq 3 \text{ and } c \in S\},$$

where $c = \lceil n/2 \rceil$. Obviously $\mathcal{C}_1$ is a well-intersecting system and $|\mathcal{C}_1| = \binom{n}{2} + 1$. It was conjectured that $\mathcal{C}_1$ is an extremal family for N_1. We show, however, that simple alterations of $\mathcal{C}_1$ lead to larger families.

To this end, we add to $\mathcal{C}_1$ the sets of the form $\{c - 2x, c - x, c, c + x, c + 2x\}, \{c - 2x, c - x, c, c + x\}$, and $\{c - x, c, c + x, c + 2x\}$ for $1 \leq x \leq [\frac{n-1}{4}]$. We have to take out the triplets that cause trouble: $\{c - 2x, c, c + x\}$ and $\{c - x, c, c + 2x\}$. We denote the new system by $\mathcal{C}$. This way for every x we put in one more set into $\mathcal{C}_1$ than we took out. We have

$$|\mathcal{C}| = |\mathcal{C}_1| + \left[\frac{n-1}{4}\right] = \binom{n}{2} + \left[\frac{n-1}{4}\right] + 1.$$

Several other equally good constructions can be obtained along similar lines. We just mention one of them, which contains arithmetic progression even of length nine. For $X \subseteq [1, [\frac{n-1}{8}]]$ we construct a well-intersecting system $\mathcal{C}_X$ in the following way: for every $x \in X$ we add the arithmetic progressions $\{c - 4x, c - 3x, c - 2x, c - x, c, c + x, c + 2x, c + 3x, c + 4x\}$ to $\mathcal{C}$, together with all its sub-arithmetic progressions containing c. This means 16 new sets for every x. (Some of the subprogressions were already contained in $\mathcal{C}$; this ensures that different values of x produce disjoint collections of new sets.) Also, we have to leave out all the 16 triplets which would violate the well-intersecting property. There is no triplet in $\mathcal{C}$ which gives a bad intersection with progressions belonging to two different values of x, hence we do not leave out the same triplet for two different elements of X. Thus, we have $|\mathcal{C}_X| = |\mathcal{C}|$.

6 Some open problems

It would be interesting to know the exact value of N_1, ideally with a description of the extremal systems. Our families $\mathcal{C}$ and $\mathcal{C}_X$ were obtained by taking in (and leaving out) $O(n)$ elements to (from) $\mathcal{C}_1$. With this in sight, it seems likely that even if $\mathcal{C}$ is not an extremal family, the extremal systems differ only slightly from a family of type $\mathcal{C}_1$. Here are two precise questions pointing to this direction: can one prove that every element of an extremal system contains a fixed integer c? Is it true that $N_1 \leq n^2/2 + O(n)$?

Finally we mention that to our knowledge, the following attractive question from [12] is still open: is it true that the extremal systems for $N_k, k \geq 2$ contain arithmetic progressions only.

ACKNOWLEDGMENT

I would like to thank my advisor, Ákos Seress for the many fruitful discussions and Lajos Rónyai for valuable comments on the manuscript and for a simpler proof of Lemma 3.4.

References

[1] N. G. de Bruijn and P. Erdős, On a combinatorial problem, *Nederl. Akad. Wetensch. Proc.* **51** (1948), 1277-1279.

[2] P. Erdős, C. Ko, R. Rado, Intersection theorems for systems of finite sets, *Quart. J. Math. Oxford Ser. (2)* **12** (1961), 313-320.

[3] P. Erdős, V. T. Sós, Problems and results on intersections of set systems of structural type, *Utilitas Math.* **29** (1986), 61–70.

[4] R. J. Faudree, R. H. Schelp, V. T. Sós, Some intersection theorems on two-valued functions, *Combinatorica* **6** (1986), 327–333.

[5] G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers,* Clarendon Press, Oxford, 1968.

[6] R. L. Graham, M. Simonovits and V. T. Sós, A note on the intersection properties of subsets of integers, *J. Combinatorial Theory Ser. A* **28** (1980), 106-110.

[7] R. H. Schelp, M. Simonovits, V. T. Sós, Intersection theorems for t-valued functions, *Europ. J. Comb.* **9** (1989), 531–536.

[8] M. Simonovits and V. T. Sós, Graph intersection theorems, *Proc. Colloq. Combinatorics and Graph Theory, Orsay, Paris,* 1976, 389-391.

[9] M. Simonovits and V. T. Sós, Intersection theorems for graphs II, *Colloq. Math. Soc. J. Bolyai, Combinatorics, Keszthely, Hungary* **18** (1976), 1017-1030.

[10] M. Simonovits and V. T. Sós, Intersection theorems for subsets of integers, *Notices American Math. Soc.* **25** (1978), A-33.

[11] M. Simonovits and V. T. Sós, Intersection theorems on structures, *Annals of Discrete Mathematics* **6** (1980), 301-313.

[12] M. Simonovits and V. T. Sós, Intersection properties of subsets of integers, *Europ. J. Combinatorics* **2** (1981), 363-372.