Sets of integers and permutations avoiding solutions to linear equations

Jonas Knape and Urban Larsson

March 2004

Department of Mathematics

Göteborgs University, Göteborg Sweden

Abstract

A considerable part of this thesis dwells on a famous result in combinatorial number theory called Szemerédi's theorem. Another recurrent theme is various applications of greedy algorithms in constructing sets of integers, or permutations of sets of integers, avoiding certain arithmetical patterns.

In Chapter 2 we investigate how dense a subset of the natural numbers can be, given that it is void of arithmetic progressions. Then we discuss how sparse such sets can be, which are saturated in the sense that they are not properly contained in any other set of positive integers avoiding arithmetic progressions.

The same questions are dealt with in another related and well-known situation, namely that of Sidon sets.

A natural generalisation is to consider sets avoiding solutions to any homogenous linear equation with integer coefficients. The question of maximizing the density of such sets gives rise to a partition of linear equations into two separate classes.

In Chapter 3, we define the concept of a permutation of (a subset of) the natural numbers avoiding solutions to a linear equation, and prove results for each of the above classes of equations.

In the last chapter we discuss a permutation π of the natural numbers, generated by a greedy algorithm refusing repetition of differences $\pi(x) - x$. We show that π has three other interpretations, namely by Fibonacci sequences, Beatty sequences and by a greedy arithmetic integer mean algorithm.

Acknowledgement We would like to thank our supervisor Peter Hegarty for the many inspiring discussions we had and for an overall understanding and generous attitude.

Also we would like to thank Fredrik Lindblad for help with code-writing and our friends and families for their fine support.

Contents

1 Introduction

We say that a triple of natural numbers (x, y, z) forms an arithmetic progression if it satisfies the equation $x - y = y - z$ with $x \neq y$. How dense can a subset of an interval of natural numbers be if it contains no arithmetical progressions? Clearly $[1, 2]$ contains no arithmetic progression, so that in this case our density is one. The largest subset of $[1, 5]$ that is void of arithmetic progressions is $\{1, 2, 4, 5\}$. Here we get density $4/5$. After these trivial examples the complexity of the problem increases. However one might be able to convince oneself that the maximal density for a subset of $[1, 9]$ is $5/9$. The following four sets are the only subsets of $[1, 9]$ with this density, void of arithmetic progressions: $\{1, 2, 6, 8, 9\}, \{1, 3, 4, 8, 9\}, \{1, 2, 4, 8, 9\}$ and $\{1, 2, 6, 7, 9\}$. There is no room in any of these sets to introduce another number from the interval $[1, 9]$. By trying to rearrange any of these sets in order to fit in an extra number, one runs into severe difficulties, it is impossible.

What is the upper asymptotic density of such sets as $n \to \infty$? This question has been answered, it is zero, as Roth showed in 1953. Even if one only wishes to avoid arithmetic progressions of fixed length greater than three, the asymptotic density is still zero. This result is due to Szemerédi (1975).

One purpose of this thesis is to attempt to give a survey of the most important results from this branch of combinatorial number theory, as well as results on similar topics, for example Sidon sets.

The second half of the thesis revolves around a topic that has not been widely researched, namely permutations avoiding arithmetical patterns.

One can define avoidance of arithmetic progressions for permutations of the natural numbers as well. Let π be a permutation that for each $i \in \mathbb{N}$ inductively takes the least value it can, refusing arithmetic progressions in triples $(\pi(i), \pi(i + j), \pi(i + 2j))$ for each $i, j \in \mathbb{N}$. We get the sequence $(\pi(i))_{i \in \mathbb{N}} = 1, 2, 4, 3, 5, 6, 8, 7, \ldots$, so that the ratio $\pi(i)/i$ takes the values $1, 1, 4/3, 3/4, 1, 1, 8/7, 7/8, \ldots$. Hegarty has recently proved that $3/8 \leq \pi(i)/i \leq 3/2$ for all i and we wonder if $\lim \pi(i)/i$ exists and equals 1 ?

The current literature on AP-avoiding permutations is thin and we have only found one different construction due to Sidorenko [27]. Is there another AP-avoiding permutation τ of $\mathbf{N}$ such that $\limsup \tau(i)/i < \limsup \pi(i)/i$? We believe that these are open questions. We have included a related result that says that, for any constant $C > 0$, there is no AP-avoiding permutation of the natural numbers ϕ such that $|\phi(i) - i| \leq C$ for every $i \in \mathbb{N}$. This is an easy consequence of Szemerédi's theorem.

2 Subsets of $\mathbb{N}$ avoiding an arithmetic condition

2.1 What is size?

If $A \subset \mathbb{N}$ is a finite set with n numbers, then the size of A is n. When $\mathcal{A} \subset \mathbb{N}$ is an infinite set, there is no obvious way of measuring "the size" of $\mathcal{A}$. However, we can measure how "dense" $\mathcal{A}$ is in $\mathbb{N}$.

If $\mathcal{A} \subset \mathbb{N}$, let $\mathcal{A}(n) = |\mathcal{A} \cap [1, n]|$. We define $\overline{d}(\mathcal{A})$ and $\underline{d}(\mathcal{A})$, the upper and lower asymptotic density of $\mathcal{A}$ as

$$\overline{d}(\mathcal{A}) = \limsup_{n\to\infty} \frac{\mathcal{A}(n)}{n} \tag{1}$$

and

$$\underline{d}(\mathcal{A}) = \liminf_{n\to\infty} \frac{\mathcal{A}(n)}{n}. \tag{2}$$

If $\overline{d}(\mathcal{A}) = \underline{d}(\mathcal{A})$ we define $d(\mathcal{A})$, the asymptotic density of $\mathcal{A}$ as

$$d(\mathcal{A}) = \lim_{n\to\infty} \frac{\mathcal{A}(n)}{n}.$$

There are other ways of measuring the density of an infinite subset of the natural numbers. If $\overline{d}(\mathcal{A}) = 0$ or $\underline{d}(\mathcal{A}) = 0$ we might want to compare the upper and lower asymptotic behaviour of $\mathcal{A}(n)$ to some function $f : \mathbb{N} \to \mathbb{R}$ that increases "more slowly" than n. We do this by looking at the values of

$$\limsup_{n\to\infty} \frac{\mathcal{A}(n)}{f(n)}$$

or

$$\liminf_{n\to\infty} \frac{\mathcal{A}(n)}{f(n)}.$$

We will also look at some cases where $B_n \subset [1, n]$ is a sequence of sets such that $B_i \not\subset B_{i+1}$ for some $i \in \mathbb{N}$. The lower and upper asymptotic behaviour of $|B_n|$ can be studied in a similar way to the equations (1) and (2).

2.2 Subsets of $\mathbb{N}$ avoiding arithmetic progressions

In combinatorial number theory one usually says that three (non-negative) integers a, b and c form an *arithmetic progression*, an AP, if $a \neq b$ and $a-b = b-c$. One defines a set of integers to be *AP-avoiding* if no three integers in the set form an AP. More generally we have the following definition:

Definition 2.1 *Let $k \in \mathbb{N}$ with $k \geq 3$. The k integers $x_1, x_2, \ldots, x_k$ form a k-term arithmetic progression, a k-AP, if $x_{i+1} - x_i = x_i - x_{i-1} \neq 0$ for $1 < i \leq k-1$. A subset X of the integers avoids k-term arithmetic progressions, or is simply k-avoiding, if no k elements of X form a k-AP.*

It is clear that a k-avoiding subset of the integers is also $(k+1)$-avoiding.

Let us fix an interval of the natural numbers. How large can a k-avoiding subset of this interval be? There is presently no complete answer to this question even for $k = 3$. It is therefore natural to look for good lower and upper bounds on the maximal size of such a set.

In the coming presentation we will use the following notation.

Definition 2.2 *Let $\nu_k(n)$ denote the maximal number of elements in a k-avoiding subset of $[1, n]$. When $k = 3$ we drop the index and write $\nu(n)$.*

2.2.1 k-term Greedy

If one wants to avoid AP:s everywhere in a subset of the natural numbers, one might want to choose the elements of the set systematically in order to construct a "fairly large/dense set". One obvious way is to apply a "greedy procedure" and recursively include "the next" natural number into the AP-avoiding set if one can. More precisely we have the following definition:

Definition 2.3 *Put $G_1^k = \{1\}$. For $n > 1$, let $G_n^k = G_{n-1}^k \cup \{n\}$ if $G_{n-1}^k \cup \{n\}$ does not contain any k-term arithmetic progression and let $G_n^k = G_{n-1}^k$ otherwise. Let $G^k(n)$ denote the number of elements in G_n^k and let $G^k = \cup_{i=1}^{\infty} G_i^k$. If $k = 3$ we put $G(n) = G^3(n)$ and $G_n = G_n^3$.*

For the rest of this section, let p be a prime. The most straighforward way to describe p-term greedy is by looking at base$-p$ expansion. Let $x_k x_{k-1} \ldots x_0$, where $0 \leq x_j \leq p-1$, mean the number $x_k p^k + x_{k-1} p^{k-1} + \ldots + x_0$.

Theorem 2.1 *Let $k, a \in \mathbb{N}_0 = \mathbb{N} \cup \{0\}$ and let p be a prime number. In base-p expansion, let $a = a_k a_{k-1} \ldots a_0$. Then*

$$a \in G^p \Leftrightarrow \; a_j \neq p-1 \quad \textit{for } 1 \leq j \leq k \textit{ and } a_0 \neq 0. \tag{3}$$

Proof. In order to simplify the proof we use the notation

$$g^k = \{x - 1 : x \in G^k\}$$

so that $g^k \subseteq \mathbb{N}_0$. We prove

$$a \in g^p \Leftrightarrow a_j \neq p-1 \text{ for } 0 \leq j \leq k. \tag{4}$$

which is equivalent to (3).

If $X = x_k x_{k-1} \ldots x_0 \geq p-1$ is not of the form (4), then there is a p-term arithmetic progression, $X > X_1 > \ldots > X_{p-1} \geq 0$ such that all the X_i are of the form (4). For example, let $X_i = X - iy$ where $y = y_k y_{k-1} \ldots y_0$ is defined by $y_j = 1$ if $x_j = p-1$, else $y_j = 0$. It is clear that in this case $g^p_X = g^p_{X-1}$, where g^p_X is the analog of G^p_X in Definition 2.3.

If on the other hand $X > p-1$ is of the form (4) and $X > X_1 > \ldots > X_{p-1} \geq 0$ is a p-term arithmetic progression, there has to be a least j such that the digit at position j in X differs from the digit at the same position in X_1 with a constant say c, where $0 < |c| < p-1$. The digit at position j in X_i is then the unique integer $t_i \in [\,0,\ p-1\,]$ such that $t_i \equiv x_j - ic \pmod{p}$. Since there are p terms in the progression and since c is relatively prime to p, we have that for some i with $1 \leq i \leq p-1$ the digit t_i is equal to $p-1$ and then $X_i \notin g^p$. In this case it follows that $g^p_X = g^p_{X-1} \cup \{X\}$ and since $x < p-1$ implies $x \in g^p$ we are done. □

The greedy algorithm gives us a construction that is best possible for certain values of the natural numbers. In Table 1 on page 64 and 65 we present computations of $\nu(n)$ for $1 \leq n \leq 128$. For large n there is another construction, see Theorem 2.6, that gives larger AP-avoiding subsets of $[1, n]$. In fact $n = 128$ is the largest natural number where we know that $G(n) = \nu(n)$.

We are now going to evaluate the number $G^p(x)$ for all natural numbers x.

Corollary 2.2 *Let p be a prime. In base-p expansion, let $x_k x_{k-1} \ldots x_0 = x \in \mathbb{N}$ where $k \in \mathbb{N}_0$ and let j be the largest index such that $x_j = p-1$ or let $j = 0$ if no such index exists. The number of elements in G^p_x is then*

$$G^p(x) = x_k(p-1)^k + x_{k-1}(p-1)^{k-1} + \ldots + x_j(p-1)^j. \tag{5}$$

Proof. By Theorem 2.1 we need to count all numbers $y \leq x$ such that $y = \sum_{i=0}^{l} y_i p^i$ where $y_0 \neq 0$ and $y_i \neq p-1$ for $i > 0$. We do this by induction over k.

If $k = 0$ we have $1 \leq x \leq p-1$ and $G^p(x) = x = x_0$.

Suppose $k \geq 1$.

If $x_k = p-1$, then for all numbers $y \in G^p$ with $y < p^{k+1}$ we have $y < x$. There are $(p-1)^{k+1}$ such numbers and so $G^p(x) = (p-1)^{k+1} = x_k(p-1)^k$. In this case equation (5) holds.

Suppose $x_k \neq p-1$ and let $y_k \ldots y_0 = y \in G^p$ with $y \leq x$. If we impose the restriction $0 \leq y_k < x_k$, then we count $x_k(p-1)^k$ such numbers y. If on the other hand $y_k = x_k$ we find $G^p(\sum_{i=0}^{k-1} x_i p^i)$ such y:s. Altogether we have

$$G^p(x) = x_k(p-1)^k + G^p\Big(\sum_{i=0}^{k-1} x_i p^i\Big) = \sum_{i=j}^{k} x_i(p-1)^i,$$

where $j < k$ is the largest index such that $x_j = p-1$ or $j = 0$ if there is no such index.

$\square$

2.2.2 The recurrence formula for $G^p(n)$

Corollary 2.3 *Let p be a prime and let k be any natural number. Then*

$$\begin{cases} G^p(pk) & = & (p-1)G^p(k) & & \\ G^p(pk+1) & = & (p-2)G^p(k) & + & G^p(k+1) \\ G^p(pk+2) & = & (p-3)G^p(k) & + & 2G^p(k+1) \\ & . & & & \\ & . & & & \\ & . & & & \\ G^p(pk+p-2) & = & G^p(k) & + & (p-2)G^p(k+1) \\ G^p(pk+p-1) & = & 0 & + & (p-1)G^p(k+1). \end{cases}$$

Proof. Let $x \geq 1$. Notice that
(i) $G^p(x) = G^p(x+1)$ if and only if x is not of the form (4).
(ii) Else $G^p(x) + 1 = G^p(x+1)$.

Let $0 \leq \alpha \leq p-1$ and take an x such as in (i). Then, by inspection $px + \alpha$ is not of the form (4) which in turn implies $G^p(px+\alpha) = G^p(px)$. Also, by

(5), $G^p(px) = (p-1)G^p(x)$. Since x was choosen as in (i) it follows in this case that $G^p(px+\alpha) = (p-1-\alpha)G^p(x) + \alpha G^p(x+1)$.

Now, let x be such as in (ii). If $0 \le \alpha < p-1$, then $px+\alpha$ is of the form (4) and $G^p(px+\alpha)+1 = G^p(px+\alpha+1)$. It follows that $G^p(px)+\beta = G^p(px+\beta)$ for $0 \le \beta \le p-1$ and therefore, by (5), $G^p(px+\beta)-\beta = (p-1)G^p(x)$. Since in this case $G^p(px+\beta) = (p-1)G^p(x)+\beta = (p-1-\beta)G^p(x)+(\beta)G^p(x+1)$ we are done. $\square$

In base-p expansion, again let $a = a_k a_{k-1} \ldots a_0$, where $a \in \mathbb{N}$. If we like, the notation of Corollary 2.3 can now be shortened. It says:

$$G^p(a) = (p - a_0 - 1)G^p(a_k a_{k-1} \ldots a_1) + a_0 G^p(a_k a_{k-1} \ldots a_1 + 1).$$

Corollary 2.4 *Let p be a prime and n a non negative integer. Then*

$$G^p\left(\frac{(p-2)p^n+1}{p-1}\right) = (p-1)^n.$$

Proof. We use induction over n. Clearly $G^p(1) = 1$. Put $x := \frac{(p-2)p^n+1}{p-1}$. Notice that

$$p\left(\frac{(p-2)p^{n-1}+1}{p-1}\right) = x+1.$$

By Corollary 2.3, for a natural number k, we have $G^p(pk) = (p-1)G^p(k)$, which in turn gives

$$G^p(x+1) = (p-1)G^p\left(\frac{(p-2)p^{n-1}+1}{p-1}\right) = (p-1)^n \tag{6}$$

where the last equality follows from the induction hypothesis assuming the result for $n-1$. Since

$$\begin{aligned} x &= (p-2)\frac{p^n-1}{p-1} + 1 = (p-2)\sum_{i=0}^{n-1} p^i + 1 \\ &= \sum_{i=1}^{n-1}(p-2)p^i + p - 1, \end{aligned}$$

the base-p expansion of $x+1$ has the digit $p-1$ at position two from the right, which by Theorem 2.1 implies $x+1 \notin G^p$ and so $G^p(x) = G^p(x+1)$, so that (6) gives the result. $\square$

We finish off this chapter with a remark. If $k \geq 4$ is not a prime, then the numbers in G^k will not satisfy the analogue of (3), the reason being the following:

If a is a prime dividing k, then the arithmetic progression $1, 1+a, 1+2a, \ldots, 1+(k-1)a$ is wholly contained in the set S consisting of elements of (the analogue of) the form (3). Therefore $G^k \neq S$ since S is not even AP-avoiding.

2.2.3 A three term AP-avoiding "Cantor-like" set

We will now describe an inductive algorithm that initially removes roughly the middle third of the interval $[1, n]$ and then recursively removes the middle third of each remaining interval until the cardinality of each interval is strictly less than three. The remaining set is subsequently called a *Cantor-like* set and it is not surprising that this set avoids AP:s which we will show by an inductive argument. At first we need a strict definition.

Definition 2.4 *Let $C_1 = \{1\}$ and $C_2 = \{1, 2\}$. For $n \geq 3$ we have three cases,*

(i) if $n = 3t$, let $C_n = C_t \cup \{x + 2t \mid x \in C_t\}$.

(ii) if $n = 3t+1$, let $C_n = C_{t+1} \cup \{x + 2t + 1 \mid x \in C_t\}$.

(iii) if $n = 3t+2$, let $C_n = C_{t+1} \cup \{x + 2t + 1 \mid x \in C_{t+1}\}$.

The sets C_1 and C_2 are AP-avoiding. Suppose $n \geq 3$ and C_i is AP-avoiding for $i < n$. An AP in C_n must contain an element a from the "lower third" of $[1, n]$ and an element b from the "upper third" of $[1, n]$. Then $2b - a > n$, $2a - b < 1$ and $(a+b)/2$ is in the excluded "middle third" of $[1, n]$. Thus, there are no AP's in C_n.

Let $C(n)$ be the number of elements of C_n. The definition of C_n gives a recursive formula for $C(n)$, namely

$$\begin{aligned} C(3t) &= 2C(t), \\ C(3t+1) &= C(t) + C(t+1), \\ C(3t+2) &= 2C(t+1), \end{aligned}$$

where $C(1) = 1$ and $C(2) = 2$ with $t > 0$.

We will now show that for certain values of n the Cantor-like set C_n coincides with the set G_n produced by our greedy algorithm described in the previous section. By induction over t we prove that $G_{\frac{3^t+1}{2}} = C_{\frac{3^t+1}{2}}$. The base condition is satisfied for $t = 0$. Suppose $G_{\frac{3^{i-1}+1}{2}} = C_{\frac{3^{i-1}+1}{2}}$. By Theorem 2.1 we have that the interval $[\frac{3^{i-1}+3}{2}, 3^{i-1}]$ is excluded from $G_{\frac{3^i+1}{2}}$ and so it is clear that

$$\begin{aligned} G_{\frac{3^i+1}{2}} &= G_{\frac{3^{i-1}+1}{2}} \cup \left\{ x + 3^{i-1} \mid x \in G_{\frac{3^{i-1}+1}{2}} \right\} \\ &= C_{\frac{3^{i-1}+1}{2}} \cup \left\{ x + 3^{i-1} \mid x \in C_{\frac{3^{i-1}+1}{2}} \right\} = C_{\frac{3^i+1}{2}}. \end{aligned}$$

There are similar Cantor-like constructions of k-term AP-avoiding sets, at least for $k > 3$ a prime. ¿From the presentation below one can check that these constructions generate sets that are less dense than G_n^k for most natural numbers $n > k$. Let k be a prime. We have the following definition

Definition 2.5 *Let $C_1^k = \{1\}$, $C_2^k = \{1,\ 2\}$, $\ldots$, $C_{k-1}^k = \{1,\ 2,\ \ldots,\ k-1\}$. For $n \geq k$ and for $0 < s < k/2$ we have these cases*

(i) If $n = kt$, let $C_n^k = C_{\frac{k-1}{2}t}^k \cup \left\{ x + \frac{k+1}{2}t \mid x \in C_{\frac{k-1}{2}t}^k \right\}$,

(ii) If $n = kt+2s-1$, let $C_n^k = C_{\frac{k-1}{2}t+s}^k \cup \left\{ x + \frac{k+1}{2}t + s \mid x \in C_{\frac{k-1}{2}t+s-1}^k \right\}$,

(iii) If $n = kt+2s$, let $C_n^k = C_{\frac{k-1}{2}t+s}^k \cup \left\{ x + \frac{k+1}{2}t + s \mid x \in C_{\frac{k-1}{2}t+s}^k \right\}$.

The recursion formula for $C^k(n) := |C_n^k|$ is

$$\begin{aligned} &C^k(kt) = 2C^k(\tfrac{k-1}{2}t), \\ &C^k(kt + 2s - 1) = C^k(\tfrac{k-1}{2}t + s) + C^k(\tfrac{k-1}{2}t + s - 1), \\ &C^k(kt + 2s) = 2C^k(\tfrac{k-1}{2}t + s) \end{aligned}$$

and where $C(1) = 1$, $\ldots$, $C(k-1) = k-1$ with $t > 0$.

The proof of k-term AP avoidance for the generalised Cantor-like sets is a direct analog to the 3-term case.

2.2.4 Infinte vs. finite constructions

We have just seen that G is a construction of an infinite AP-avoiding set. The Cantor-like sets C_i are constructions that works for all $i \in \mathbb{N}$, but we have problems with infinity here. In fact $\cup_{i=1}^{\infty} C_i = \mathbb{N}$. If we have a sequence of AP avoiding sets, $\{B_n\}_{n=1}^{\infty}$, such that $B_n \subset [1, n]$, we get a lower bound for $\nu(n)$, namely $\nu(n) \geq |B_n|$. But from this we can not necessarily tell if there exists an AP-avoiding set $A \subset \mathbb{N}$ with $A(n) \geq |B_n|$ for all n.

2.3 Subsets of $[1, n]$ avoiding AP:s modulo n

Let $k, n \in \mathbb{N}$. How large can a subset of $[1, n]$ be, provided it avoids k-term arithmetic progressions modulo n? Here we naturally need to define what we mean by a k-term arithmetic progression modulo a positive integer n.

Definition 2.6 *Let $k, n \in \mathbb{N}$. The non-negative integers $x_1, x_2, \ldots, x_k$, form a k-AP modulo n if*

$$x_{i+1} - x_i \equiv x_i - x_{i-1} \not\equiv 0 \pmod{n},$$

for $1 < i \leq k - 1$.

Note that if $x_1, x_2, \ldots, x_k$ form a $k-$term arithmetic progression modulo n, we do not know if $x_1, x_2, \ldots, x_k$ is a k-AP on the natural numbers. On the other hand, if $x_1, x_2, \ldots, x_k$ is a $k-$term arithmetic progression on the natural numbers, it is clear that $x_1, x_2, \ldots, x_k$ is a $k-$term arithmetic progression modulo n only if $x_1, x_2, \ldots, x_k$ are not all congruent modulo n.

Denote by $\eta_k(n)$ the maximal size of a subset of $[1, n]$ that is k-avoiding modulo n. From the above remarks, it is clear that $\eta_k(n) \leq \nu_k(n)$. Also, if $S \subset [1, n]$ is k-avoiding modulo n, then the set $S \cup \{x + n \mid x \in S\}$ is k-avoiding (in fact even the set $S \cup_{i=0}^{k-2} \{x + in \mid x \in S\}$ is k-avoiding). Thus $2\eta_k(n) \leq \nu_k(2n)$.

If we go in the other direction and start with a set $T \subset [1, n]$ that is k-avoiding, then one can see that T is k-avoiding modulo $2n$, so $\nu_k(n) \leq \eta_k(2n)$. Similarily, the sets $T \cap [1, \lfloor \frac{n}{2} \rfloor]$ and $T \cap [\lceil \frac{n}{2} \rceil, n]$ must be k-avoiding modulo n and then $\eta_k(n) \geq \lfloor \frac{\nu_k(n)}{2} \rfloor$.

A famous theorem (Theorem 2.9), which we will discuss on page 15 and 17, says that $\nu_k(n)/n \to 0$ as $n \to \infty$. For infinitely many n we then must have $\nu_k(2n)/2n < \nu_k(n)/n$ and for those n,

$$\frac{\eta_k(n)}{\nu_k(n)} \leq \frac{\nu_k(2n)}{2\nu_k(n)} < 1.$$

Define $c_1 := \limsup \frac{\eta(n)}{\nu(n)}$ and $c_2 := \liminf \frac{\eta(n)}{\nu(n)}$. By the above reasoning, we know that

$$1/2 \leq c_2 \leq c_1 \leq 1.$$

This leads us to asking whether or not these inequalities are strict:

Question 2.1 *(I) Is $c_1 < 1$? (II) Is $c_2 > 1/2$? (III) Is $c_1 = c_2$?*

Table 2 on page 66 shows what is going on modulo n for $n \leq 64$ in the three term case.

The following lemma gives us a means of constructing AP-avoiding sets on the natural numbers from AP-avoiding sets modulo n.

Lemma 2.5 *For $k, n \in \mathbb{N}$ and $X \subset [0,\ n-1]$, let*

$$A = \Big\{ x \mid x = \sum_{i=0}^{m} x_i n^i,\ x_i \in X,\ m \in \mathbb{N} \Big\}.$$

If X is a k-avoiding set modulo n, then A is a k-avoiding set in $\mathbb{N}_0$.

Proof. If $x^{(1)} < x^{(2)} < \ldots < x^{(k)}$ form a k-term arithmetic progression in A we can find $\delta \in \mathbb{N}$ such that $x^{(i)} = x^{(1)} + i\delta$ for $0 < i \leq k$. Then there also exists a non-negative integer α such that $n^\alpha \mid \delta$ but $n^{\alpha+1} \nmid \delta$. Observe that the integers

$$\frac{x^{(1)} - x^{(1)}}{n^\alpha}, \frac{x^{(2)} - x^{(1)}}{n^\alpha}, \ldots, \frac{x^{(k)} - x^{(1)}}{n^\alpha}$$

form a k-AP modulo n, but these numbers modulo n are just a linear translation of the digits in position α of $x^{(1)}$, $x^{(2)}$, $\ldots$, $x^{(k-1)}$ and $x^{(k)}$ respectively. Hence the digits at position α of $x^{(1)}$, $x^{(2)}$, $\ldots$, $x^{(k)}$ form a k-AP modulo n and since these digits are all members of X, we are done. $\square$

The converse of Lemma 2.5 is not true in general, which we prove by presenting a counter example, namely: The numbers 0 and 2 form a k-term AP modulo 4 for any natural number $k \geq 3$. Suppose that the set

$$A = \{\, x \mid x_i = 0 \text{ or } 2, \text{ where } i \text{ is a position in the base-4 expansion of } x \,\}$$

contains a 3-term AP, say (a, b, c), with $a < b < c$. then there is a highest position i in the base-4 expansion of a such that $b_i \neq a_i$. Notice that $c = b + (b - a)$, where $b - a > 0$, so we have that $2b_i - a_i = 10$ (in base-4 expansion) which implies either

I) $c_i = 3$, in case $\sum_{l=0}^{l=i-1}(2b_l - a_l)4^l < 0$, or

II) $c_{i+1} \equiv 1 + 2b_{i+1} - a_{i+1} \equiv 1 \pmod 2$, since $2b_{i+2} - a_{i+2} \geq 0$, so that in both cases we contradict the definition of A.

In Section 2.5.2 on page 19 we will examine a special case where the converse of Lemma 2.5 is true.

2.4 History

2.4.1 The search for a maximal AP-avoiding set

In [9] P. Erdős and P. Turán quoted a conjecture by G. Szerkeres that $\nu(\frac{3^t+1}{2}) = G(\frac{3^t+1}{2}) = 2^t$ holds for $t \in \mathbb{N}$. However, in [26], R. Salem and D. C. Spencer disproved $\nu(N) = O(N^\alpha)$ for every $\alpha < 1$ by giving an explicit lower bound on $\nu(N)$, namely

$$\nu(N) > N^{(1 - \frac{\log 2 + \epsilon}{\log \log N})}$$

for any $\epsilon > 0$ and all sufficiently large N. This bound was improved by F. A. Behrend in the article [3] by a modification of Salem and Spencer's method, and more than half a century later, Behrend's lower bound has not been improved. We include here a proof that partly follows that given by T. Tao in [31], and partly Behrend's original article.

Theorem 2.6 *For $\epsilon > 0$ and sufficiently large $N \in \mathbb{N}$,*

$$\nu(N) > \frac{N}{e^{(C+\epsilon)\sqrt{\log N}}}, \tag{7}$$

where $C = 2\sqrt{2\log 2}$ and $\nu(N)$ is as in Definition 2.2.

Proof. For any natural numbers d, n and R, let

$$S_{d,n,R} = \{x \in \{0,1,\ldots,n-1\}^d : \|x\|^2 = R\},$$

where $\|\cdot\|$ is the ordinary norm in $\mathbb{R}^d$. Thus $S_{d,n,R}$ is a set of integer points on a sphere of radius $\sqrt{R}$. If $x, y, z \in S_{d,n,R}$ and $x + y = 2z$, where addition is defined componentwise, then by the triangle inequality

$$2\sqrt{R} = \|2z\| = \|x+y\| \leq \|x\| + \|y\| = 2\sqrt{R}$$

and so $\|x+y\| = \|x\| + \|y\|$. Since $\|x\| = \|y\|$ and the coordinates of x and y are non-negative, this gives $x = y$. Thus, no three distinct points on a sphere in d dimensions can form an arithmetic progression.

We have that

$$\bigcup_{R=0}^{d(n-1)^2} S_{d,n,R} \supset \{0,1,\ldots,n-1\}^d$$

so there are $d(n-1)^2 + 1$ sets $S_{d,n,R}$ that cover the cube $\{0,1,\ldots,n-1\}^d$ which has n^d points. By the pigeonhole principle there must then exist an R' with $0 \leq R' \leq d(n-1)^2$ and such that

$$(d(n-1)^2+1)\,\left|\, S_{d,n,R'} \,\right| \geq n^d. \tag{8}$$

We will now define a map $\phi : \{0,1,\ldots,n-1\}^d \to \mathbb{N}_0$ such that if there is an AP in $\phi(M)$ for some set M, then there is also an AP in M. Let

$$\phi(x_1, x_2, \ldots, x_d) = x_1 + x_2(2n-1) + \cdots + x_d(2n-1)^{d-1}.$$

Since $0 \leq x_i \leq n-1$ for all i it is clear that ϕ is injective. Let $x = (x_1, x_2, \ldots, x_d)$, $y = (y_1, y_2, \ldots, y_d)$ and $z = (z_1, z_2, \ldots, z_d)$ and suppose

$$\phi(x) + \phi(y) = 2\phi(z)$$

then

$$\sum_{i=1}^{d} (x_i + y_i)(2n-1)^{i-1} = \sum_{i=1}^{d} 2z_i(2n-1)^{i-1}.$$

Since $0 \leq x_i + y_i < 2n-1$ and $0 \leq 2z_i < 2n-1$ we have $x_i + y_i = 2z_i$ for all i. From this it follows that x, y and z form an AP in $\{0,1,\ldots,n-1\}^d$. By these properties of ϕ, the set $\phi(S_{d,n,R'})$ is AP-avoiding and $|\phi(S_{d,n,R'})| = \left|S_{d,n,R'}\right|$. For $x \in S_{d,n,R'}$ we have

$$\phi(x) \leq \sum_{i=0}^{d-1} (n-1)(2n-1)^i = \frac{(2n-1)^d - 1}{2} < (2n-1)^d,$$

and so by (8)

$$\nu((2n-1)^d) \geq \left|\phi(S_{d,n,R'})\right| \geq \frac{n^d}{d(n-1)^2+1} \geq \frac{n^{d-2}}{d}.$$

Let N be a given large number and set $d = \left\lfloor \sqrt{\frac{2\log N}{\log 2}} \right\rfloor$ and $n = \left\lfloor \frac{N^{1/d}+1}{2} \right\rfloor$. Then $(2n-1)^d \leq N < (2n+1)^d$ and

$$\begin{aligned}\nu(N) &\geq \nu((2n-1)^d) > \frac{n^{d-2}}{d} > \frac{(N^{1/d}-1)^{d-2}}{d2^{d-2}} \\ &= \frac{N^{1-2/d}}{d2^{d-2}}(1-N^{-1/d})^{d-2}.\end{aligned}$$

We want to get rid of the factor $(1-N^{-1/d})^{d-2}$ and we do this by looking at its' logarithm

$$\begin{aligned}\log(1-N^{-1/d})^d &\simeq c\sqrt{\log N}\log(1-e^{-\sqrt{\log N}/c}) \\ &= c\sqrt{\log N}e^{-\sqrt{\log N}/c}(1-e^{-\sqrt{\log N}/c})^{e^{\sqrt{\log N}/c}} \\ &\to c\cdot 0\cdot e^{-1} = 0 \text{ as } N\to\infty.\end{aligned}$$

Here $c = \sqrt{2/\log 2}$. By this we see that for sufficiently large N we have $\log(1-N^{-1/d})^d > \log(1/2)$ and then

$$(1-N^{-1/d})^{d-2} > (1-N^{-1/d})^d > \frac{1}{2}.$$

Hence for sufficiently large N

$$\nu(N) > \frac{N^{1-2/d}}{d2^{d-1}} = N^{1-\frac{2}{d}-\frac{\log d}{\log N}-\frac{(d-1)\log 2}{\log N}}.$$

For $\epsilon > 0$ and N large enough we estimate

$$\begin{aligned}&\frac{2}{d}+\frac{\log d}{\log N}+\frac{(d-1)\log 2}{\log N} \\ &\leq \frac{\sqrt{2\log 2}}{\sqrt{\log N}-1}+\frac{1}{2}\frac{\log\frac{2\log N}{\log 2}}{\log N}+\frac{\sqrt{2\log 2}}{\sqrt{\log N}}-\frac{\log 2}{\log N} \\ &\leq \frac{2\sqrt{2\log 2}}{\sqrt{\log N}}+\frac{\epsilon}{\sqrt{\log N}}.\end{aligned}$$

Finally this implies $\nu(N) > N^{1-\frac{2\sqrt{2\log 2}+\epsilon}{\sqrt{\log N}}}$. □

The first nontrivial upper bound for $\nu(N)$ was given by K. F. Roth in [21], where he proved the following theorem.

Theorem 2.7 *There is a constant $C > 0$ such that $\nu(N) < C\frac{N}{\log\log N}$ for all natural numbers N.*

If we use the notation from the end of Section 2.1 we get that $\limsup_{n\to\infty}\frac{\nu(n)}{f(n)}$, where $f(n) := \frac{n}{\log\log n}$, is bounded.

Theorem 2.7 has been improved and the up-to-date best upper bound on $\nu(N)$ is due to Bourgain [5],

$$\nu(N) < C\frac{N\sqrt{\log\log N}}{\sqrt{\log N}}, \tag{9}$$

so that $\limsup_{n\to\infty}\frac{\nu(n)}{f(n)}$ is bounded even with $f(n) := \frac{n\sqrt{\log\log n}}{\sqrt{\log n}}$.

For a long time it was conjectured that $\nu_k(n)/n \to 0$ as $n \to \infty$ for any k and Roth's proof was the first step in this direction. In order to prove this in general, it is enough to prove that the corresponding result for AP-avoiding infinite subsets of $\mathbb{N}$ holds, as is shown by the following proposition.

Proposition 2.8 *For $k \in \mathbb{N}$, the statements i) and ii) below are equivalent.*

i) If $\mathcal{A} \subset \mathbb{N}$ and $\overline{d}(\mathcal{A}) > 0$ then $\mathcal{A}$ contains k-term AP:s.

ii)

$$\lim_{n\to\infty}\frac{\nu_k(n)}{n} = 0.$$

Proof. It is immediate that *ii)* implies *i)*. To show that *i)* implies *ii)*, suppose

$$\frac{\nu_k(n)}{n} \not\to 0 \quad \text{as} \quad n \to \infty.$$

Then there is an $\epsilon > 0$, a sequence $\{N_i\}_{i=1}^{\infty}$ and sets $B_{N_i} \subset [1, N_i]$ with $|B_{N_i}| \geq \epsilon N_i$ and B_{N_i} does not contain any k-term AP. Let $n_j = N_{i_j}$ define an infinite subsequence of $\{N_i\}_{i=1}^{\infty}$ such that $n_{j+1} \geq 3n_j$. Now define $C_j := \{x + 2n_j \mid x \in B_{n_j}\}$ and let

$$\mathcal{C} = \bigcup_{j=1}^{\infty} C_j. \tag{10}$$

$\mathcal{C}$ has upper asymptotic density at least $\epsilon/3$, since $C_j \subset [2n_j, 3n_j]$ and $|C_j| > \epsilon n_j$. Also $\mathcal{C}$ is k-term AP avoiding, since

(i) a k-term AP cannot be wholly contained in one C_j, and

(ii) a three term AP has to be contained in a single C_j. For suppose (x, y, z) is an AP in $\mathcal{C}$ with $x < y < z$. Then if $y, z \in C_l$ for some l, we have

$$z - y \leq \max C_l - \min C_l = 3n_l - (2n_l + 1) = n_l - 1.$$

But if $x \notin C_l$, then

$$\begin{aligned} y - x &\geq \min C_l - \max C_{l-1} = 2n_l + 1 - 3n_{l-1} \\ &\geq 2n_l + 1 - n_l = n_l + 1 > z - y, \end{aligned}$$

which contradicts that (x, y, z) is an AP. This implies that if $y, z \in C_l$, then $x \in C_l$. If $z \in C_l$ and $y \in C_j$ where $j < l$, then

$$\begin{aligned} z - y &\geq \min C_l - \max C_{l-1} = 2n_l + 1 - 3n_{l-1} \\ &\geq 2n_l + 1 - n_l = n_l + 1, \end{aligned}$$

and

$$y - x \leq \max C_{l-1} - 1 = 3n_{l-1} - 1 \leq n_l - 1 < z - y.$$

Again, this contradictis that (x, y, z) is an AP, so we must have $x, y, z \in C_l$ for some l.

The Proposition now follows, since *i)* denies the existence of any such set $\mathcal{C}$. $\square$

By adopting the idea in the above proof, Behrend's construction can be used to obtain a "dense" infinite AP-avoiding subset X of $\mathbb{N}$. As $N \to \infty$, the set $[1, N] \cap X$ contains at least $N^{1-\frac{d}{\sqrt{\log N}}}$ elements (d is a constant).

In [30] Szemerédi proved that statement *i)* in Proposition 2.8 holds for any k. This much celebrated result, which is known as Szemerédi's theorem, was preceded by Van der Waerden's theorem in 1927, which states that if we partition the natural numbers into finitely many classes, then at least one of these classes has arithmetic progressions of arbitrary length (for a proof see for example [20], or for an interesting outlining of the way Van der Waerden discovered the proof together with Artin and Schreier, see [32]). It is easily seen that Van der Waerden's theorem follows from Szemerédi's theorem, which we now formally state in its own right and then go ahead to indicate the existent proofs (which are all a bit to long and elaborate to fit into this thesis).

Theorem 2.9 *Let $\mathcal{A} \subset \mathbb{N}$. If $\mathcal{A}$ has positive upper asymptotic density, then $\mathcal{A}$ contains arithmetic progressions of arbitrary length.*

i) The proof by Szemerédi, which is purely combinatorial and uses, among other things, the so-called "Regularity Lemma" (see for example [4]), recursion of a "tower" function (for a brief description see the first pages of [14]) and Van der Waerden's theorem.

ii) H. Furstenberg's proof published in [12], which uses ergodic theory and recurrence in dynamical systems.

iii) Gowers' proof [14] for k-term AP avoiding sets (four term case [15]), extends the ideas that Roth used in his proof of an upper bound for $\nu(n)$. He uses Fourier analysis on cyclic groups, concepts of randomness and uniform distribution of elements in AP-avoiding subsets of a cyclic group and thereby attains an upper bound for the size of a k-term AP avoiding set, namely:

$$\nu_k(n) < n(\log \log n)^{-c(k)}, \text{ where } c(k) = \frac{1}{2^{2^{k+9}}}. \tag{11}$$

2.5 Saturated sets

Definition 2.7 *Let $S \subset \mathbb{Z}$. A k-term AP-avoiding set $A \subset S$ is said to be k-saturated in S (or just saturated if it is obvious to which k, and to which set S we are referring) if $A \cup \{x\}$ contains a k-term AP for all $x \in S \setminus A$.*

2.5.1 A lower bound on the size of a saturated set in $[1, n]$

Let M_n be a 3-saturated set in $[1, n]$.

Proposition 2.10 $|M_n| > \sqrt{\frac{2n}{3}}$.

Proof. Let $x = |M_n|$. There are $\frac{x(x-1)}{2}$ distinct couples in M_n and every couple can form an AP with at most three elements in $[1, n] \setminus M_n$. Therefore the maximal number of elements in $[1, n] \setminus M_n$ that form an AP with elements of M_n is $3\frac{x(x-1)}{2}$. Since M_n is saturated it follows that

$$n - x \leq 3\frac{x(x-1)}{2}$$

and so

$$|M_n| \geq \frac{\sqrt{24n-1}+1}{6} > \sqrt{\frac{2n}{3}}.$$

□

2.5.2 Some special saturated subsets of $[1, n]$

With notation as in Lemma 2.5, if the set A is saturated in $\mathbb{N}_0$, then it is clear that X is saturated modulo n (if X is AP-avoiding), but the converse is not true in general. For example $\{0,1\}$ is 3-saturated modulo 5 in $[0,4]$, but $A = \{0,1,5,6,25,26,\ldots\}$ is not saturated in $\mathbb{N}_0$. (The number 8 is the least number that does not form a 3-AP with elements in A). However, in Theorem 2.11 we describe a special case when the converse is true.

In the coming, if $k, t \in \mathbb{N}_0$ with $k \geq 2$, we will use standard notation for base$-k$ expansion, with the addition that $xx \overset{t}{\ldots} x$ denotes the number $xk^{t-1} + xk^{t-2} + \ldots + xk^0$.

Definition 2.8 *Let $x, t \in \mathbb{N}_0$ and let $k \in \mathbb{N}$ with $k \geq 3$. If the base-$(k+1)$ expansion of x is $x = x_{t-1} \ldots x_0$ so that $0 \leq x < (k+1)^t$, then R_t^k is defined by*

$$x \in R_t^k \Leftrightarrow x_i \leq k-2, \text{ for } 0 \leq i \leq t-1. \tag{12}$$

We define R^k by

$$x \in R^k \Leftrightarrow x \in R_t^k \text{ for some } t \in \mathbb{N}. \tag{13}$$

Notice that the maximal element of R_t^k is $(k-2)(k-2) \overset{t}{\ldots} (k-2)$ and that $|R_t^k| = (k-1)^t$. Let $I_t^k := [-\frac{(k+1)^t-1}{k}, (k-1)\frac{(k+1)^t-1}{k}]$. In base-$(k+1)$ expansion this interval is simply $I_t^k = [-11 \overset{t}{\ldots} 1,\ (k-1)(k-1) \overset{t}{\ldots} (k-1)]$.

Theorem 2.11 *Let $t \in \mathbb{N}_0$ and let $q = 4$ or q a prime greater than four. Then R_t^{q-1} is $(q-1)$-saturated in I_t^{q-1}.*

Proof. We start by showing that R_t^{q-1} is $(q-1)$-avoiding. The set $\{0,1\}$ avoids 3-term AP:s modulo 4, so we move on to show that $K := \{0,\ 1,\ \ldots,\ q-3\}$ avoids $(q-1)$-AP:s modulo q when q is a prime greater than four. Then Lemma 2.5 will give that R_t^{q-1} is $(q-1)$-avoiding. Suppose $(x_1, x_2, \ldots, x_{q-1})$ forms an $(q-1)$-AP modulo q with $x_i \in K$. Define $\Delta := x_2 - x_1$. Evidently,

$0 < |\Delta| \leq q-3$ and $gcd(\Delta, q) = 1$. For at least one couple (i, j), with $i > j$ we have that $x_i = x_j$. This follows because the AP contains $(q-1)$ terms and there are only $q-2$ elements in K, so $0 = x_i - x_j \equiv d\Delta \pmod{q}$, where $0 < d \leq q-2$, and so $\Delta \equiv 0 \pmod{q}$ which contradicts the definition of Δ.

Define a translation τ by

$$\tau(x) = x + \frac{q^t - 1}{q - 1} = x + 11 \overset{t}{\ldots} 1.$$

Then R_t^{q-1} is $(q-1)$-saturated in I_t^{q-1} if and only if $\tau(R_t^{q-1})$ is $(q-1)$-saturated in $\tau(I_t^{q-1}) = [\,0,\ q^t - 1\,]$, and we prove the latter statement.

The set $\tau(R_t^{q-1})$ contains exactly the numbers y whose base-q expansion is $y = y_{t-1} \ldots y_0$ with $1 \leq y_i \leq q-2$ for $0 \leq i \leq t-1$.

Suppose

$$y \in \tau(I_t^{q-1}) \setminus \tau(R_t^{q-1})$$

and that the base-q expansion of y is $y_{m-1}y_{m-2}\ldots y_0$. If $m < t$, put $y_i = 0$ for $m-1 < i \leq t-1$ so that $y = y_{t-1}y_{t-2}\ldots y_0$. We will now construct a $(q-1)$-AP in $\tau(R_t^{q-1}) \cup \{y\}$. To this purpose, for $1 \leq i \leq q-2$ and $0 \leq j \leq t-1$, let the integer $z^{(i)}$ be defined by its' digits

$$z_j^{(i)} := \begin{cases} y_j & \text{if } 1 \leq y_j \leq q-2, \\ i & \text{if } y_j = 0, \\ q-i-1 & \text{if } y_j = q-1. \end{cases}$$

Then $1 \leq z_j^{(i)} \leq q-2$ so that $z^{(i)} = z_{t-1}^{(i)} z_{t-1}^{(i)} \ldots z_0^{(i)} \in \tau(R_t^{q-1})$. Also, the tuple $(y_j,\ z_j^{(1)},\ z_j^{(2)}, \ldots, z_j^{(q-2)})$ is either a $(q-1)$-AP or all the numbers in the tuple are equal. Thus $(y,\ z^{(1)},\ z^{(2)}, \ldots, z^{(q-2)})$ is a $(q-1)$-AP and $\tau(R_t^{q-1})$ is saturated in $\tau(I_t^{q-1})$. □

Suppose $2 < k \neq q-1$ with q as in Theorem 2.11. If δ is the least divisor of $k+1$ strictly greater than 2 (the existence is clear) then R_t^k is not k-saturated in I_t^k if $t \geq 2$ because in this case R_t^k is not even $k-$avoiding. We prove this as follows: Let δ be as described and define the arithmetic progression $\alpha = (\alpha_{-1},\ \alpha_0,\ \ldots,\ \alpha_{k-2})$ where $\alpha_i = k - 2 + i\delta$. Since $\delta \not\equiv 0 \pmod{k+1}$ we conclude that α also forms a k-term arithmetic progression modulo $(k+1)$. Suppose that for some $i \in [1,\ k]$ either

$$\alpha_i \equiv k \pmod{k+1}$$

or

$$\alpha_i \equiv k-1 \pmod{k+1}.$$

Then, either $i\delta \equiv 2$ or $i\delta \equiv 1 \pmod{k+1}$, and this contradicts $\delta \mid k+1$ and $\delta > 2$. ¿From $0 \leq \alpha_i < (k-1)(k+1)$, we get that no digit of α_i is equal to k or $k-1$. Thus the arithmetic progression α is wholly contained in R_t^k for $t \geq 2$.

Definition 2.9 *Let $n \in \mathbb{N}$. We denote by $\mu_k(n)$ the minimal number of elements of a set that is k-saturated in $[1, n]$. If $k = 3$, we simply write $\mu(n)$.*

Notice that Theorem 2.11 gives the following result.

Corollary 2.12 *Let $t \in \mathbb{N}$ and let $x = (k+1)^t$, where $k = 3$ or $k = p-1$ with $p > 4$ a prime. Then*

$$\mu_k(x) \leq x^{\frac{\log(k-1)}{\log(k+1)}}.$$

By computer calculations, presented in Table 1 on page 64, we know the minimal size of a 3-saturated subset of $[1, n]$ for $n \leq 64$. In fact our construction in Theorem 2.11 is best possible for $n = 16$ and $n = 64$. By some manipulation of this construction we can find an upper bound, denoted by $\gamma(n)$, on $\mu(n)$ for arbitrary n and Table 1 presents these values for $n \leq 128$. Roughly speaking, we get the number $\gamma(n)$ by translating the set $R_{\lceil \log_4 n \rceil}$ with an appropriate constant for each n and "glueing" iteratively "as well as possible" with R_m , where $m < \lceil \log_4 n \rceil$, into the non-saturated part of $[1, n]$.

From Proposition 2.10, Theorem 2.11 and in view of our computer simulations we can make the following conjecture:

Conjecture 2.1 *Let $t \in \mathbb{N}$. Then $\mu_3(4^t) = |R_t^3| = 2^t$.*

2.5.3 Minimal saturation in $\mathbb{N}$

We can find arbitrarily sparse sets that are 3-saturated in the set of natural numbers. More precisely: If g is an increasing bijection on $\mathbb{R}^+$ then there exists a set A that is 3-saturated in $\mathbb{N}$ and such that $A(n) < g(n)$ for all n.

Proof. Let $j \geq 1$ and $a_0 = 0$. Define a_{2j-1} and a_{2j} recursively so that

$$a_{2j} > a_{2j-1} > \max\{g^{-1}(2j),\ 2a_{2j-2}\} \tag{14}$$

and so that a_{2j-1} and a_{2j} form an AP with the least element of

$$\mathbb{N} \setminus \{a_i \mid 1 \leq i < 2j-1\},$$

which does not already form an AP with some pair of elements in

$$\{a_i \mid 1 \leq i < 2j-1\}.$$

Put $A = \{a_j \mid j \in \mathbb{N}\}$.

If $a,\ b,\ c \in A$ with $a < b < c$ then either $2b - c \in \mathbb{N} \setminus A$ or, by (14), $2b < c$. In both cases $(a,\ b,\ c)$ is not an AP, so A is AP-avoiding. Hence, by construction, A is saturated.

Let $n \in \mathbb{N}$ and let k be the largest index such that $a_k \leq n$. Then, by (14), $a_k > g^{-1}(k)$ and so

$$g(n) \geq g(a_k) > k = A(n).$$

□

2.5.4 Cantor-like and saturation

With notation as in Section 2.2, we now turn our attention to the question whether Cantor-like C_n is saturated in $[1, n]$?

We know that G_n is saturated for all $n \in \mathbb{N}$, since the greedy algorithm always gives saturated AP avoiding subsets of the natural numbers. Since $C_n = G_n$ for $n = (3^t + 1)/2$, where $t \in \mathbb{N}$, C_n is certainly saturated here. We can also show that C_n is saturated for $n = 3^t$ (this is the 3-term Cantor-like set with least density among Cantor-like sets C_m with $m \leq \frac{3^{t+1}+1}{2}$). Namely we have that $x \in C_{3^t}$ if and only if the base three expansion of $x - 1$ consists of only $0 : s$ and $2 : s$. If we take any number $y \in [1, 3^t] \backslash C_{3^t}$, we can find y_1 respectively y_2 in C_3^t such that $y_1 < y < y_2$ is an arithmetic progression, namely if i is a position in y where $(y-1)_i = 1$, let $(y_1 - 1)_i = 0$ and $(y_2 - 1)_i = 2$, otherwise let $y_i = (y_1)_i = (y_2)_i$.

So far we have proved that C_{3^t} and $C_{(3^t+1)/2}$ are saturated, but one can find infinitely many "Cantor-like" sets that are not saturated. By inspection we have that $C_{25} = \{1, 3, 7, 9, 18, 20, 23, 25\}$ is the smallest Cantor-like set that is not saturated, precisely, no two numbers in C_{25} are in an arithmetic progression with the number 10. Therefore it is clear that for example $(C_{25\times 3^i})_{i=0}^{\infty}$ is an infinite sequence of Cantor-like sets that are not saturated.

2.6 Sidon sets

In the previous chapters we have been looking at sets that avoid arithmetic progressions and foremost been interested in bounds for the maximal size of an AP-avoiding set. One can ask the same question for sets that avoid other kinds of arithmetical patterns and our next example is the Sidon condition.

Definition 2.10 *Let $S \subset \mathbb{N}$ and let $x,\ y,\ z,\ w \in S$. Then S is a Sidon set if $z - w = y - x$ implies $z = y$ and $w = x$.*

Note that since the elements of an AP are a solution to $x - y = z - w$, every Sidon set is AP-avoiding.

Definition 2.11 *Let $\Phi(n)$ denote the maximal number of elements in a Sidon set which is also a subset of $[1, n]$.*

By analogy with the notion of a saturated AP-avoiding set, we say that a subset A of B is *Sidon saturated* in B if A is a Sidon set and $A \cup \{b\}$ is not a Sidon set for any $b \in B \setminus A$.

Definition 2.12 *Denote by $\xi(n)$ the minimal size of a set that is Sidon saturated in $[1, n]$.*

Define $GS_n \subset [1, n]$ to be the Sidon set generated by the greedy algorithm, which is analogous to the algorithm used to construct G_n. The following proposition gives a lower bound on $|GS_n|$.

Proposition 2.13 *If $A \subset [1, n]$ is Sidon saturated in $[1, n]$, then $|A| \geq \sqrt[3]{n}$. That is, $\xi(n) \geq \sqrt[3]{n}$.*

Proof. Let $t = |A|$ and let $b \in [1, n] \backslash A$. Then either there are $a_1,\ a_2,\ a_3 \in A$ where $a_1 \neq a_2$ and $b + a_1 = a_2 + a_3$, or there are $a_4,\ a_5 \in A$ with $a_4 \neq a_5$

and $2b = a_4 + a_5$. The number of such triples (a_1, a_2, a_3) in A is $t^2(t-1)$ and the number of such tuples (a_4, a_5) in A is $t(t-1)$. We get

$$|[1,n] \setminus A| = n - t \leq t^2(t-1) + t(t-1) = t^3 - t$$

and then

$$|A| = t \geq \sqrt[3]{n}.$$

□

2.6.1 History

The best known upper bound for $\Phi(n)$ was found by Erdős and Turán [10], and is given below, together with its' original proof.

Theorem 2.14 *There is a constant c, such that*

$$\Phi(n) < \sqrt{n} + c\sqrt[4]{n}$$

for all n.

Proof. Let $S \subset [1,n]$ be a Sidon set and let $m \in \mathbb{N}$ with $m < n$. For $1 \leq u \leq n+m$, let A_u denote the number of elements of S that also lie in the interval $I_{m,u} = [-m+u, u)$. If $x \in S$, then x lies in precisely those intervals $I_{m,u}$, where $x < u \leq x + m$. This gives

$$\sum_{u=1}^{m+n} A_u = m\,|S|.$$

The number of pairs $x,\ y \in S$ with $x < y$ which lie in the interval $I_{m,u}$ is $\frac{1}{2}A_u(A_u - 1)$. We can estimate the total number of pairs in all the intervals $I_{m,u}$, counting repetitions, by using the Cauchy-Schwarz inequality,

$$\begin{aligned}
\frac{1}{2}\sum_{u=1}^{m+n} A_u(A_u-1) &= \frac{1}{2}\left(\sum_{u=1}^{m+n} A_u^2 - \sum_{u=1}^{m+n} A_u\right) \\
&\geq \frac{1}{2}\left(\sum_{u=1}^{m+n}\left(\frac{m\,|S|}{m+n}\right)^2 - \sum_{u=1}^{m+n} A_u\right) \qquad (15) \\
&= \frac{1}{2}\left(\frac{(m\,|S|)^2}{m+n} - m\,|S|\right).
\end{aligned}$$

For any pair $x,\ y \in S$ in the same interval $I_{m,u}$ and with $x < y$, we have $y - x = r$ and $1 \leq r \leq m - 1$. Since S is a Sidon set, each such value r corresponds to at most one pair in S and each pair that corresponds to the value r appears in exactly $m - r$ of the intervals $I_{m,u}$. This gives

$$\sum_{r=1}^{m-1}(m-r) = \frac{m(m-1)}{2} \geq \sum_{i=1}^{m+n}\frac{1}{2}A_u(A_u - 1)$$

and with (15) we get

$$\frac{1}{2}\left(\frac{(m\,|S|)^2}{m+n} - m\,|S|\right) \leq \frac{m(m-1)}{2}.$$

This implies

$$|S|(m\,|S| - m - n) \leq (m+n)(m-1)$$

and since $m < n$,

$$|S|(m\,|S| - 2n) \leq |S|(m\,|S| - m - n) \leq (m+n)(m-1) < m(m+n).$$

By solving this quadratic inequality in $|S|$, we have

$$|S| < \frac{n}{m} + \sqrt{m + n + \frac{n^2}{m^2}}.$$

If we choose $m = \lfloor n^{\frac{3}{4}} \rfloor$, then

$$|S| < O(\sqrt[4]{n}) + \sqrt{n^{\frac{3}{4}} + n + n^{\frac{1}{2}}} = \sqrt{n^{\frac{3}{4}} + n} + O(\sqrt[4]{n}),$$

and since

$$\sqrt{n^{\frac{3}{4}} + n} = \sqrt{n}\sqrt{n^{-\frac{1}{4}} + 1} \leq \sqrt{n}\left(1 + n^{-\frac{1}{4}}\right) = \sqrt{n} + \sqrt[4]{n},$$

we get

$$|S| < \sqrt{n} + O(\sqrt[4]{n}).$$

□

A few years later Erdős [11] and Chowla [8] independently found a lower bound for $\Phi(n)$. In Section 2.6.2 we will prove a generalised version of the theorem below.

Theorem 2.15 *There is a constant c such that*

$$\Phi(n) > \sqrt{n} - cn^{\frac{5}{16}}$$

for all n.

Theorems 2.14 and 2.15 give the nice result that

$$\lim_{n\to\infty} \frac{\Phi(n)}{\sqrt{n}} = 1,$$

which can be compared to the corresponding results for AP-avoidance, the big gap between Behrend's and Bourgain's bounds.

We now turn our attention to infinite Sidon sets. The set $GS = \bigcup_n GS_n$ is a Sidon set, and by Proposition 2.13 we have $GS(n) = |GS_n| \geq \sqrt[3]{n}$. A more dense infinite Sidon set was found by Ajtai, Komlós and Szemerédi [1].

Theorem 2.16 *There is a Sidon set $\mathcal{A} \subset \mathbb{N}$ such that*

$$\mathcal{A}(n) > 10^{-3}(n \log n)^{\frac{1}{3}}$$

for sufficiently large n.

The bound in Theorem 2.16 has since been improved according to Theorem 2.17. For a proof we refer to Ruzsa's article in [25]. The main observation in Ruzsa's article is that the numbers $\log p$, for p prime, form a Sidon set of reals. He then proves how one can take the binary developement of these numbers and rearrange the digits to form a Sidon set of integers. He obtains the following bound.

Theorem 2.17 *There exists an infinite Sidon set $\mathcal{A} \subset \mathbb{N}$ satisfying*

$$\mathcal{A}(n) = n^{\sqrt{2}-1+o(1)}.$$

The bound in Theorem 2.17 can not be substituted by any bound greater than $\sqrt{n/\log n}$, as Erdős showed in [29].

Theorem 2.18 *There is a constant C such that if $\mathcal{A} \subset \mathbb{N}$ is a Sidon set, then*

$$\liminf_{n\to\infty} \frac{\mathcal{A}(n)}{\sqrt{\frac{n}{\log n}}} \leq C.$$

On the other hand, there are Sidon sets that at infinitely many n have density greater than the one given by the theorem above. This is due to Krückeberg [19].

Theorem 2.19 *There is a Sidon set $\mathcal{A} \subset \mathbb{N}$ such that*

$$\limsup_{n\to\infty} \frac{\mathcal{A}(n)}{\sqrt{n}} \geq \frac{1}{\sqrt{2}}.$$

We finish off this chapter by stating a theorem which gives an upper bound for the minimal size of a set that is Sidon saturated in $[1, n]$. The proof is by Ruzsa and is given in [24].

Proposition 2.20 *There is a constant C such that $\xi(n) < C\sqrt[3]{n \log n}$.*

Sketch of Proof. Select a prime p and put $q = 1 + p + p^2$. In [28], Singer proved that there is a set of size $p + 1$, say

$$B = \{b_0, b_1, \ldots, b_p\}, \tag{16}$$

such that the sums $b_i + b_j$ all have distinct residues modulo q. Let B be such a set and let $M = \lfloor \frac{N}{q} \rfloor$, where $N \in \mathbb{N}$. If $d_0, d_1, \ldots, d_p$ are arbitrary integers in the interval $[1, M]$, then

$$A_0 = \{a_i \mid a_i = b_i + d_i q\} \subset [1, N] \tag{17}$$

forms a Sidon set. Ruzsa shows, using probabilistic methods and classical results on the distribution of primes, that one can find a p such that $|A_0| < C\sqrt[3]{N \log N}$. He moves on to show that there is a set $A_1 \subset [1, N]$ such that $A_1 \cup A_0$ is saturated and $|A_1| < CN^{1/3}$. We omit further details.

2.6.2 A Galois field construction of a generalised Sidon set

A solution to the equation $x_1 + \ldots + x_r = y_1 + \ldots + y_r$ is *trivial* if there is a permutation π on $[1, r]$ such that $x_j = y_{\pi(j)}$ for all j.

Definition 2.13 *Let $X \subset \mathbb{N}$, then X is an r-Sidon set if every solution in X to the equation $x_1 + \ldots + x_r = y_1 + \ldots + y_r$ is trivial.*

Note that an $(r + 1)$-Sidon set is an r-Sidon set.

Definition 2.14 *Let $\Phi_r(n)$ denote the maximal number of elements in an r-Sidon set consisting of elements only in the interval $[1, n]$.*

We have the following theorem due to Bose and Chowla, [7].

Theorem 2.21

$$\liminf_{n\to\infty} \frac{\Phi_r(n)}{\sqrt[r]{n}} \geq 1.$$

The proof of this result depends on Lemma 2.22, which gives the construction of certain "large" Sidon sets, and Theorem 2.24 below.

Lemma 2.22 *Let $m = p^n$ where p is a prime and $r \geq 1$ an integer. Then there are m different integers d_i, $1 \leq i \leq m$, such that $1 \leq d_i < m^r$ and all the sums*

$$d_{i_1} + d_{i_2} + \cdots + d_{i_r}$$

where $i_1 \leq i_2 \leq \ldots \leq i_r$ are different $\bmod (m^r - 1)$.

Proof. Let α_i with $1 \leq i \leq m$ be the elements of the Galois field of order p^n, which we denote $GF(p^n)$, and let x be a generator for the extended field $GF(p^{nr})$. Then x is not a root of any polynomial of degree less than r and with coefficients in $GF(p^n)$ (if it was, then the degree of the extension generated by x would be less than r, and x could not generate $GF(p^{nr})$). Define the integers d_i by letting

$$x^{d_i} = x + \alpha_i$$

with $1 \leq d_i < m^r$. This is possible, since $x + \alpha_i \neq 0$ for all i. With this definition, all the integers d_i are different.

Suppose that

$$d_{i_1} + d_{i_2} + \ldots + d_{i_r} \equiv d_{j_1} + d_{j_2} + \ldots + d_{j_r} \pmod{m^r - 1}$$

where $i_1 \leq i_2 \leq \ldots \leq i_r$ and $j_1 \leq j_2 \leq \ldots \leq j_r$. Since $x^{m^r-1} = 1$, we then have

$$x^{d_{i_1}} x^{d_{i_2}} \cdots x^{d_{i_r}} = x^{d_{j_1}} x^{d_{j_2}} \cdots x^{d_{j_r}}$$

and so

$$(x + \alpha_{i_1})(x + \alpha_{i_2}) \cdots (x + \alpha_{i_r}) = (x + \alpha_{j_1})(x + \alpha_{j_2}) \cdots (x + \alpha_{j_r}). \qquad (18)$$

In this equation, the terms x^r cancel out and what is left is a polynomial equation of degree less than r with coefficients in $GF(p^n)$. But x can not solve such an equation, unless both sides of equation (18) are identical when seen as polynomials. But since factorization of polynomials over a field is unique, this means that the α_i's are a permutation of the α_j's, and then $i_k = j_k$ for $1 \leq k \leq m$. □

Since the integers d_i form an r-Sidon set $\mod (m^r - 1)$, they also form an r-Sidon set considered as a subset of $\mathbb{N}$.

Corollary 2.23 *If $m = p^n$, where p is a prime and $r \geq 2$ is an integer, then $\Phi(m^r - 1) \geq m$.*

Before we move on to prove Theorem 2.21. we will state a result on the difference between consecutive primes which was first proved by Ingham [18] and which we will need in the proof.

Theorem 2.24 *There is a constant C, such that if p and p' are consecutive primes with $p < p'$, then*

$$p' - p < Cp^{\frac{5}{8}}.$$

Proof of Theorem 2.21. Let $y \geq 2^r$ be an integer and let p and p' be consecutive primes such that $p^r \leq y \leq (p')^r$. By Corollary 2.23,

$$\Phi(y) \geq \Phi(p^r) \geq \Phi(p^r - 1) \geq p.$$

Using Theorem 2.24 we have

$$\Phi(y) \geq p = p' - (p' - p) \geq p' - Cp^{\frac{5}{8}} \geq \sqrt[r]{y} - Cy^{\frac{5}{8r}}$$

and then

$$\frac{\Phi(y)}{\sqrt[r]{y}} \geq 1 - Cy^{\frac{5}{8r} - \frac{1}{r}} = 1 - Cy^{-\frac{3}{8r}}$$

which proves the theorem. □

2.7 Avoiding other arithmetic conditions

Definition 2.15 *Let $k, l \in \mathbb{N}$. A linear arithmetic condition (a LAC) is a linear equation*

$$a_1x_1 + \ldots + a_ix_i = b_1y_1 + \ldots + b_jy_j$$

where $a_1, \ldots, a_k, b_1, \ldots, b_l \in \mathbb{N}$ have no common divisor > 1. We denote this condition $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$ and we say that $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$ is defined on the (i, j) indeterminates $x_1, \ldots, x_i, y_1, \ldots, y_j$. The set of all linear arithmetic conditions will be denoted by $\mathcal{L}$.

Let $A := (a_1, \ldots, a_i \mid b_1, \ldots, b_j)$ be a linear arithmetic condition. We define $d(A) := \max\{a_1 + \ldots + a_i, b_1 + \ldots + b_j\}$ and we say that $d(A)$ is the *degree* of the LAC A. We also define the *discriminant* of A as $\Delta(A) := |a_1 + \ldots + a_i - (b_1 + \ldots + b_j)|$. If $\Delta(A) = 0$ then A is a translation invariant arithmetic condition, a TAC. The set of all translation invariant arithmetic conditions will be denoted by $\mathcal{T}$. It is obvious that if $(x_1, \ldots, x_i, y_1, \ldots, y_j)$ is one solution to a TAC then $(\gamma + x_1, \ldots, \gamma + x_i, \gamma + y_1, \ldots, \gamma + y_j)$ is also a solution if γ is any integer. On the other hand, if an LAC is not a TAC then this property does not hold if $\gamma \neq 0$.

Definition 2.16 *Let $i, j \in \mathbb{N}$ with $i \leq j$. A subset of the natural numbers X avoids a linear arithmetic condition $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$ if $a_1x_1 + \ldots + a_ix_i = b_1y_1 + \ldots + b_jy_j$ with $x_1, \ldots, x_i, y_1, \ldots, y_j \in X$ implies*

$$\sum_{l \,:\, x_l = x_k} a_l = \sum_{l \,:\, y_l = x_k} b_l \tag{19}$$

for all k with $1 \leq k \leq i$.

We say that any solution to the LAC A that satisfies (19) is *trivial.* Notice that if S is a nonempty subset of the natural numbers and if A is a translation invariant arithmetic condition, then there is a trivial solution to A in S. On the other hand if A is not translation invariant, then A can not be trivially satisfied by elements from S. We have the following lemma.

Lemma 2.25 *If $A \in \mathcal{L} \backslash \mathcal{T}$ then there is an A-avoiding set with positive asymptotic density.*

Proof. Let $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$ be a LAC with $\Delta \neq 0$ and let α be a natural number relatively prime to Δ. The set

$$X = \{ x \in \mathbb{N} \mid x \equiv 1 \pmod{\alpha} \}$$

has asymptotic density $\frac{1}{\alpha}$. We prove that X is $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$-avoiding in the following way. Let $x_1, \ldots, x_i, y_1, \ldots, y_j \in X$. It is clear that $a_1x_1 + \cdots + a_ix_i - (b_1y_1 + \cdots + b_jy_j) \equiv a_1 + \cdots + a_i - (b_1 + \cdots + b_j) \pmod{\alpha}$. Since $a_1 + \cdots + a_i - (b_1 + \cdots + b_j) = \Delta$ is relatively prime to α it follows that $a_1x_1 + \cdots + a_ix_i \not\equiv b_1y_1 + \cdots + b_jy_j \pmod{\alpha}$ and so $a_1x_1 + \cdots + a_ix_i \neq b_1y_1 + \cdots + b_jy_j$. □

The above lemma says that we can always find an A-avoiding subset of the natural numbers with upper asymptotic density $1/p$, where p is the least prime not dividing the discriminant of A. Denote by $\nu(A, n)$ the maximal size of an $A-$avoiding subset of the interval $[1, n]$. The "simplest" examples of an LAC A with nonzero discriminant are maybe $(1, 1 \mid 1)$ or $(2 \mid 1)$. For both these conditions the above Lemma gives the A-avoiding set of all odd positive numbers less than or equal to n. This is actually best possible for a subset of $[1, n]$ avoiding $(1, 1 \mid 1)$ and we prove it below.

Proposition 2.26 *We have that* $\nu(\,(1, 1 \mid 1)\,,\; n\,) = \lceil \frac{n}{2} \rceil$.

Proof. Let $n \in \mathbb{N}$. The existence of an $(1, 1 \mid 1)$-avoiding set of size $\lceil \frac{n}{2} \rceil$ is clear, namely the set of all odd positive numbers up to n. Suppose $X \subset [1, n]$ is $(1, 1 \mid 1)$-avoiding and let m be the largest number in X. Then at least one element of each one of the $\lfloor \frac{m}{2} \rfloor$ couples $(1, m-1), (2, m-2), \ldots, (\lfloor \frac{m}{2} \rfloor, \lceil \frac{m}{2} \rceil)$ has to be excluded from X, so $|X| \leq m - \lfloor \frac{m}{2} \rfloor = \lceil \frac{m}{2} \rceil \leq \lceil \frac{n}{2} \rceil$. □

For the LAC $(2 \mid 1)$ we can improve on this and the result generalises.

2.7.1 The LAC $(k \mid l)$

Let $k, l, n \in \mathbb{N}$ where $k > l \geq 1$ and $\gcd(k, l) = 1$. We are going to present a situation when a greedy procedure avoiding the LAC $kx = ly$ gives a largest $(k \mid l)$-avoiding subset of each interval $[1, n]$ and hence of $\mathbb{N}$. What is more, a sort of "reversed greedy" algorithm, where you recursively include a natural number to your $(k \mid l)-$avoiding sequence "only if you have to", gives the least dense $(k \mid l)$-saturated subset of the natural numbers.

Definition 2.17 *For $k, l \in \mathbb{N}$, $k > l \geq 1$ and with $gcd(k,l) = 1$, let $K_1^{k,l} = \{1\}$ and for $n > 1$ let $K_n^{k,l} = K_{n-1}^{k,l} \cup \{n\}$ if $K_{n-1}^{k,l} \cup \{n\}$ is $(k \mid l)$-avoiding. Otherwise, let $K_n^{k,l} = K_{n-1}^{k,l}$. Finally, put $K^{k,l} = \bigcup_{n \in \mathbb{N}} K_n^{k,l}$.*

Sometimes, when it is clear from the context which k and l we are referring to, we simply write K instead of $K^{k,l}$. The following proposition shows that $K^{k,l}$ is in fact independent of l.

Proposition 2.27 *For $k, l \in \mathbb{N}$, $k > l \geq 1$ with $gcd(k,l) = 1$,*

$$K^{k,l} = \{ ak^{2i} \mid k \nmid a \in \mathbb{N}, i \in \mathbb{N}_0 \}. \tag{20}$$

Proof. For simplicity, denote the set on the right hand side of (20) by P. The proof goes by induction over n. We have $K_1 = P \cap [1,1]$. For $n > 1$, suppose $K_{n-1} = P \cap [1, n-1]$. Now $n = ak^j$ for some $j \geq 0$ and some $a \geq 1$ such that $k \nmid a$. If j is even, then by the induction hypothesis, $ak^{j-1}l = ln/k \notin K_{n-1}$ so that $n \in K_n$ and $K_n = P \cap [1,n]$. If j is odd, then $ak^{j-1}l \in K_{n-1}$ and $n \notin K_n$. So that also in this case we get $K_n = P \cap [1,n]$. $\square$

The greedy algorithm gives the largest $(k \mid l)$-avoiding subset of $[1,n]$.

Proposition 2.28 *For each $n \in \mathbb{N}$ and for each $k, l \in \mathbb{N}$ with $k > l \geq 1$ and $gcd(k,l) = 1$,*

$$\nu((k \mid l), n) = |K_n|.$$

Proof. Let B be a $(k \mid l)$-avoiding subset of $[1,n]$. We prove the proposition by giving an injection $f : B \setminus K_n \to K_n \setminus B$. If $m \in B \setminus K_n$, then $m = ak^{2i+1}$ for some $i \geq 0$ and $a \geq 1$ where $k \nmid a$. Let $f(m) = lak^{2i} = lm/k$. Then $f(m) \in K_n \setminus B$, since $kf(m) = lm$ and $gcd(k,l) = 1$. Also, f is injective. $\square$

We now turn our attention to a sparse $(k \mid l)$-saturated set. Notice that in Definition 2.18, if $l = 1$, then $l \mid (n+1)$ is always true, so here we can leave out case *I)*.

Definition 2.18 *Let $k > l \geq 1$ and let $L_1 = \{1\}$. For $n \geq 1$ put $L_{n+1} = L_n \cup \{n+1\}$ if*

I) $l \nmid (n+1)$ and there is no $i \in L_n$ such that $ki = l(n+1)$, or if

II) $l \mid (n+1)$ and $k \mid (n+1)$ and there is no $i \in L_n$ such that $ki = l(n+1)$ or $k^2 i = l^2(n+1)$,

otherwise put $L_{n+1} = L_n$. Finally let $L^{k,l} = \bigcup_{n \in \mathbb{N}} L_n$.

It is clear by the definition that $L^{k,l}$ is $(k \mid l)-$avoiding. We will now outline why $L^{k,l}$ is saturated in the natural numbers. Clearly we only need to check the case $L_{n+1} = L_n$. If $l \nmid (n+1)$ then we are done, so suppose $l \mid (n+1)$. In this case, if $k \mid (n+1)$ then we are done, so suppose $l \mid (n+1)$ but $k \nmid (n+1)$. In this case, by applying the definition on $\frac{k}{l}(n+1)$ (rather than on $(n+1)$), this forces $L_{\frac{k}{l}(n+1)} = L_{\frac{k}{l}(n+1)-1} \cup \{\frac{k}{l}(n+1)\}$ (because $(n+1) \notin L_{\frac{k}{l}(n+1)-1}$ and $\frac{l}{k}(n+1) \notin \mathbb{N}$).

The following proposition can be proved analogously to Proposition 2.27.

Proposition 2.29 *If $k > l \geq 1$, then*

$$L^{k,l} = \{\, ak^{3i+1} \mid k \nmid a \in \mathbb{N}, i \in \mathbb{N}_0 \} \cup \{\, bk^{3i} \mid k,\, l \nmid b,\ i \in \mathbb{N}_0 \}.$$

Further if $n \in \mathbb{N}$, we define

$$M_n^{k,l} := \big(L^{k,l} \cap [1,n] \big) \cup \{\, ak^{3i}l \mid k \nmid a,\ ak^{3i}l \leq n < ak^{3i+1} \}.$$

As before, we usually omit the index k and write L and M_n instead of $L^{k,l}$ and $M_n^{k,l}$. It is not hard to see that M_n is saturated in $[1,n]$. Part *ii)* of the following proposition says that the "reversed greedy" algorithm described in Definition 2.18 gives the sparsest $(k,l)-$saturated subset of $\mathbb{N}$.

Proposition 2.30 *Let $k > l \geq 1$ and $n \in \mathbb{N}$. Then*

i) $\mu((k \mid l), n) = |M_n^{k,l}|$.

ii) If $A \subset \mathbb{N}$ is $(k \mid l)$-saturated in $\mathbb{N}$, then $|A \cap [1,n]| \geq |L^{k,l} \cap [1,n]|$.

Proof. We prove *i)* in the case $l = 1$. The proof of the other cases in *i)* and *ii)* is similar but somewhat more technical if $l > 1$.

Let B be a $(k \mid 1)$-saturated subset of $[1, n]$. It is then enough to give a surjection $f : B \to M_n$. Suppose $k \nmid a$. If $ak^{3i+1} \leq n$, then if m is one of the numbers ak^{3i-1}, ak^{3i} or ak^{3i+1} and $m \in B$, put $f(m) = ak^{3i+1}$. If $ak^{3i} \leq n < ak^{3i+1}$, $m \in B$ and m is either ak^{3i-1} or ak^{3i}, put $f(m) = ak^{3i}$. If $m = ak^{3i-1} \leq n < ak^{3i}$ and $m \in B$, put $f(ak^{3i-1}) = f(ak^{3i-2})$.

By this definition of f, $f(m)$ is in M_n. We will show that f is surjective. Let $r \in M_n$. If $r = ak^{3i+1}$ for some i and some a such that $k \nmid a$, then one of ak^{3i-1}, ak^{3i} and ak^{3i+1} must be in B since otherwise B is not saturated. Hence there is an m such that $f(m) = r$. If $r = ak^{3i}$ and $kr > n$, then ak^{3i} or ak^{3i-1} is in B since B is saturated. So in this case too, there is an m such that $f(m) = r$. This shows that f is surjective. □

Proposition 2.31 *The asymptotic density of K, L and M is*

$$\lim_{n\to\infty} \frac{|K_n|}{n} = \frac{k}{k+1},$$
$$\lim_{n\to\infty} \frac{|L \cap [1,n]|}{n} = \frac{k^2 l - k^2 + kl}{k^2 + k + 1},$$
$$\lim_{n\to\infty} \frac{|M_n|}{n} = \frac{k^2 l}{k^2 + k + 1}.$$

Proof. The number of elements in K_n can be evaluated as follows:

$$|K_n| = |\{xk^{2i} \mid k \nmid x\} \cap [1,n]| = \sum_{i=0}^{\lfloor \frac{\log n}{2\log k} \rfloor} |\{x \leq \frac{n}{k^{2i}} \mid k \nmid x\}|$$

$$= \sum_{i=0}^{\lfloor \frac{\log n}{2\log k} \rfloor} \lfloor \frac{n}{k^{2i}} \rfloor - \lfloor \frac{n}{k^{2i+1}} \rfloor.$$

We get the following estimate

$$-1 - \frac{\log n}{2\log k} + \sum_{i=0}^{\lfloor \frac{\log n}{2\log k} \rfloor} (\frac{n}{k^{2i}} - \frac{n}{k^{2i+1}}) \leq |K_n| \leq 1 + \frac{\log n}{2\log k} + \sum_{i=0}^{\lfloor \frac{\log n}{2\log k} \rfloor} (\frac{n}{k^{2i}} - \frac{n}{k^{2i+1}}),$$

so that

$$\sum_{i=0}^{\infty}(\frac{1}{k^{2i}}-\frac{1}{k^{2i+1}}) \leq \lim_{n\to\infty}\frac{|K_n|}{n} \leq \sum_{i=0}^{\infty}(\frac{1}{k^{2i}}-\frac{1}{k^{2i+1}})$$

and the first result follows. The asymptotic density for L and $\lim \frac{|M_n|}{n}$ can be evaluated similarily. $\square$

By the poset construction and by Proposition 2.31, we get the following result.

Corollary 2.32 *Let A be an LAC with degree d and discriminant $\Delta \neq 0$. Let $n \in \mathbb{N}$ and let X_n be an A-avoiding subset of $[1, n]$. Then if $c = \gcd(\Delta, d)$ we get $\lim_{n\to\infty} \frac{|X_n|}{n} \leq \frac{d}{d+c}$.*

Let A be an LAC of degree d. If the discriminant $\Delta(A) := \Delta$ is (strictly) greater than $d/2$, then it is easy to construct a denser A-avoiding set than in Lemma 2.25. For example the set $[\lceil\frac{(d-\Delta)n}{d}\rceil, n]$ avoids A. We believe that this lower bound on the size of a maximal LAC-avoiding set can be improved for many different types of LAC:s [1].

2.7.2 LAC posets

Let $\binom{a_1,\ldots,a_i}{b_1,\ldots,b_j}$ denote the family of all subsets of the natural numbers that avoid the LAC $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$.

Fix two indices $1 \leq \alpha < \beta \leq i$ in an LAC $(a_1, \ldots, a_i \mid b_1, \ldots, b_j)$ of degree d. By setting $x_\alpha = x_\beta$ we get a new LAC of either the same degree d, or if the new coefficients have greatest common divisor $\gamma > 1$, then the new degree is $\frac{d}{\gamma}$. On the other hand, if we fix $1 \leq \alpha \leq i$ and $1 \leq \beta \leq j$ and set $x_\alpha = y_\beta$ then we get a LAC with degree $\frac{d-\min\{a_\alpha, b_\beta\}}{\gamma}$, where γ is the gcd of the coefficients of the new LAC.

If $a_1x_1 + \ldots + a_ix_i = b_1y_1 + \ldots + b_jy_j$, where WLOG $a_1 \geq b_1$, we can construct a new LAC by setting $x_1 = x_2$, or by setting $y_1 = x_1$. Our new LAC is either $\left(\frac{a_1+a_2}{\gamma}, \frac{a_3}{\gamma}, \ldots, \frac{a_i}{\gamma} \mid \frac{b_1}{\gamma}, \ldots, \frac{b_j}{\gamma}\right)$ if $\gamma = \gcd(a_1 + a_2, a_3, \ldots, a_i, b_1, \ldots, b_j)$, or $\left(\frac{a_1-b_1}{\omega}, \frac{a_2}{\omega}, \ldots, \frac{a_i}{\omega} \mid \frac{b_2}{\omega}, \ldots, \frac{b_j}{\omega}\right)$ if $\omega = \gcd(a_1 - b_1, a_2, \ldots, a_i, b_2, \ldots, b_j)$.

[1] For example for LAC:s of the form $(k \mid 1, 1)$, where $k \geq 3$, we have succeded in proving this in a recent joint paper with Hegarty [16].

Also it is clear from Definition 2.16 that $A = \left(\frac{a_1,\dots,a_i}{b_1,\dots,b_j}\right)$ is a subset of both

$$B = \left(\frac{\frac{a_1+a_2}{\gamma}, \frac{a_3}{\gamma}, \dots, \frac{a_i}{\gamma}}{\frac{b_1}{\gamma}, \dots, \frac{b_j}{\gamma}}\right) \tag{21}$$

and

$$C = \left(\frac{\frac{a_1-b_1}{\omega}, \frac{a_2}{\omega}, \dots, \frac{a_i}{d}}{\frac{b_2}{\omega}, \dots, \frac{b_j}{\omega}}\right). \tag{22}$$

Hence we can define an order of a poset $(P, \leq)$ whose elements, indexed by all possible LAC:s L, are sets consisting of all $L-$avoiding subsets of natural numbers for the different L. It is clear that $(0) := \left(\frac{1}{1}\right) = \left(\frac{1}{0}\right) = 2^{\mathbb{N}}$ is the greatest element of this poset. In fact, since the above gives

$$\Delta(A) = 0 \Leftrightarrow \Delta(B) = 0 \Leftrightarrow \Delta(C) = 0,$$

it is clear that the arithmetic conditions in $\mathcal{T}$ give rise to one poset that we denote $\mathcal{T}^\star$ and the family $\mathcal{L}\backslash\mathcal{T}$ gives rise to another poset that we denote $\mathcal{G}^\star$. We have indicated a proof of the following proposition.

Proposition 2.33 *With notation as above and with $\mathcal{L}$ analogously giving rise to the poset $\mathcal{L}^\star$, we have that*

$$\mathcal{T}^\star \cup \mathcal{G}^\star = \mathcal{L}^\star$$

and

$$\mathcal{T}^\star \cap \mathcal{G}^\star = \{\,(0)\,\}.$$

It is natural to ask for which LAC:s L does an analog to Szemeredi's theorem hold, in other words, for which LAC:s does a subset $\mathcal{A}$ of the natural numbers having positive upper asymptotic density imply existence of elements in $\mathcal{A}$ satisfying the LAC non-trivially? It is clear from Lemma 2.25 that L is necessarily contained in $\mathcal{T}$. Let A be a LAC and let $P(A)$ be the statement that a subset $\mathcal{X}$ of the natural numbers contains a non-trivial solution of A whenever $\mathcal{X}$ has positive upper asymptotic density. We have the following proposition.

Proposition 2.34 *The statement $P(A)$ is true if and only if $A \in \mathcal{T}$.*

Proof. It remains to show that $A \in \mathcal{T}$ implies $P(A)$. It is clear from the identification process above that a set $\mathcal{B}$ which avoids an arithmetic condition A with more than three terms and with $d := d(A) \geq 2$ also avoids

$$f(x, y, z) = ax - by - cz = 0 \tag{23}$$

for some a, b and c relatively prime natural numbers and $d = a = b + c$. So if $P(A)$ holds for an arithmetic condition of the form (23), then $P(A)$ holds for any TAC with $d \geq 2$. Let $x_0 \in \mathbb{N}$ and $k \in \mathbb{Z}$. It is clear that $x_0,\ x_0 + (a-b)k,\ x_0 + ak$ is a subsequence of the arithmetic progression $x_0, x_0 + k, \ldots, x_0 + ak$ and also that

$$f(x_0,\, x_0 + (a-b)k,\, x_0 + ak) = a(x_0 + (a-b)k) - bx_0 - (a-b)(x_0 + ak) = 0.$$

We may conclude that if $\mathcal{B}$ avoids an arithmetic condition of the form (23), then in particular $\mathcal{B}$ avoids $(d+1)-$term arithmetic progressions. Therefore Szemeredi's Theorem 2.9 implies that $P(A)$ holds for an arithmetic condition of the form (23) and we are done. □

Let $T^d \subseteq \mathbb{N}$ be a set avoiding some TAC of degree d and let $\tau^d(N)$ denote the maximal size of $T^d_N = T^d \cap [1, N]$. Then Proposition 2.33 and Theorem 2.21, combined with Gowers' result (11) on page 17 give

$$N^{1/d} \leq \tau^d(N) \leq \frac{N}{(\log\log N)^{\frac{1}{2^{2^{d+10}}}}}.$$

2.7.3 AC-saturated sets

We have included two inclusion diagrams, on page 67 and 68, of families of LAC-avoiding sets for all LAC:s of degree less than 4, the first one consists of families of sets avoiding TAC:s and the second one of families of sets avoiding LAC:s that are not translation invariant.

We believe that most questions related to these diagrams are still open questions. For example, we have not been able to find much information on the relative sizes/densities of "maximal" or "minimal" LAC-avoiding sets, depending on degree, discriminant, or/and number of terms of respective LAC. We may refer to the articles [22] and [23] by I. Ruzsa for more information on subjects related to LAC:s. However, what aroused our interest in these diagrams in the first place was the following question:

If A is a TAC with, say, n terms and K is an infinite set, which is $A-$saturated in the natural numbers, can one find another TAC, say B, also with n terms such that K is also $B-$saturated in the natural numbers. The answer is yes for the following example:

Proposition 2.35 *The set R^3 in Proposition 2.10 is $(\,2 \mid 1,1\,)$-saturated and $(\,3 \mid 2,1\,)-$saturated in the natural numbers.*

Proof. The first part is done. We need to show that R^3 is $(\,3 \mid 2,1\,)-$saturated. We show that R^3 is $(\,3 \mid 2,1\,)-$avoiding from looking at the base-4 expansion of three elements $x, y, z \in R^3$. Since every digit of an element in R^3 is either 0 or 1, we have that $3x = 2y+z$ implies $3x_i = 2y_i+z_i$ for every position i in the 4-expansion of x, y and z, and so $x = y = z$. Now we show "the saturated part". Let $z \in \mathbb{N} \setminus R^3$ and let i be any position in the base-4-expansion of x, y respectively z. The triples $(x_i, y_i, z_i) = (0,0,0), (1,1,1), (0,1,2)$ and $(1,0,3)$ satisfy $3x_i - 2y_i \equiv z_i \pmod 4$ and the triples $(x_i, y_i, z_i) = (0,0,3), (1,1,0), (0,1,1)$ and $(1,0,2)$ satisfy $3x_i - 2y_i - 1 \equiv z_i \pmod 4$. This proves that there exist y and $z \in R^3$ such that $z = 3x - 2y$. □

3 Permutations avoiding an arithmetic condition

We start this chapter by defining the concept of AP-avoiding functions, but our main concern will be AP-avoiding permutations. Later on we will discuss permutations avoiding solutions to general homogenous linear equations, in the sense that whenever we have a solution in the range of a permutation, then we cannot have a non-trivial solution in the corresponding values of the domain.

Definition 3.1 *Let $f : S \to S$ be a function where $S = \mathbb{N}$, $S = [1, N]$ or $S = \mathbb{Z}/n\mathbb{Z}$. For $k \geq 3$, we say that f is k-term AP-avoiding if, whenever $x_1 < x_2 < \ldots < x_k$ and $f(x_{i+1}) - f(x_i) = \delta$ is constant for $i = 1, \ldots, k-1$, then $x_1, \ldots, x_k$ does not form an arithmetic progession.*

Sidon functions can be defined in a similar way.

Definition 3.2 *Let $S = [1, N]$, $S = \mathbb{N}$ or $S = \mathbb{Z}/n\mathbb{Z}$. A function $f : S \to S$ is a Sidon function, if $f(x) + f(y) \neq f(z) + f(w)$ when $x + y = z + w$, $(x, y) \neq (z, w)$ and $(x, y) \neq (w, z)$. If f is also a permutation, we say that f is a Sidon permutation.*

3.1 The number of values of AP-avoiding functions

By using the upper bounds for the size of an AP-avoiding set, we can get lower bounds on the number of values an AP-avoiding function takes. Suppose $f : [1, N] \to \mathbb{N}$ is an AP-avoiding function. For every $i \in \mathbb{N}$, the set $f^{-1}(\{i\})$ must be AP-avoiding (e.g. in the three term case, if $x + z = 2y$ with $x, y, z \in f^{-1}(\{i\})$, then $f(x) + f(z) \neq 2f(y) = 2i$).

Put
$$V = \{i \mid f^{-1}(\{i\}) \neq \emptyset\},$$
then $v = |V|$ is the number of values that f takes. Since
$$\bigcup_{i \in V} f^{-1}(\{i\}) = [1, N],$$
the pigeonhole principle gives that there is an r such that
$$|f^{-1}(\{r\})| \geq \frac{N}{v}. \tag{24}$$

If f is three-term AP-avoiding and N is large enough, then Roth's theorem (Theorem 2.7) gives

$$|f^{-1}(\{r\})| < C\frac{N}{\log\log N}$$

and we get a bound on v,

$$v > C^{-1}\log\log N.$$

This bound can of course be improved by using Bourgain's result (9) instead of Roth's, and for k-term AP-avoiding functions we can use Gowers' upper bound (11) for k-term AP-avoiding sets. When f is an AP-avoiding function on $\mathbb{N}$ we can apply the above reasoning to $f|_{[1,N]}$. If we let $v(N)$ denote the number of values that $f|_{[1,N]}$ takes, then in the three-term case, $v(N) > C^{-1}\log\log N$. We can also use Szemerédi's theorem directly to show that f takes infinately many values. For if f only took finitely many values, there would only be finitely many sets $f^{-1}(\{i\})$ which were non-empty, and since these sets form a partition of $\mathbb{N}$, at least one of them would have positive upper asymptotic density. This contradicts Szemerédi's theorem.

Let f be a Sidon function, then $f^{-1}(\{i\})$ is a Sidon set for all $i \in \mathbb{N}$. If the domain of f is $[1, N]$, we can define v as for AP-avoiding functions, namely as the number of values that f takes. The inequality (24) still holds for some r and according to Theorem 2.14,

$$|f^{-1}(\{r\})| < \sqrt{N} + c\sqrt[4]{N}.$$

Then

$$v > \frac{N}{\sqrt{N} + c\sqrt[4]{N}} > \sqrt{N} - c\sqrt[4]{N}$$

for large N.

3.2 AP-avoiding permutations of $\mathbb{N}$

3.2.1 Sidorenko's AP-avoiding permutation

Let φ_4 be the following permutation of the set $\{0, 1, 2, 3\}$

$$\varphi_4(0) := 0, \varphi_4(1) := 2, \varphi_4(2) := 1, \varphi_4(3) := 3.$$

It is clear by inspection that φ_4 is AP-avoiding in the sense of Definition 3.2, if we take the set of definition and the set of values modulo 4. Let $s : \mathbb{N} \to \mathbb{N}$

be the permutation

$$s(n) := \sum_{i=0}^{\lfloor \log_4 n \rfloor} \varphi_4(x_i) 4^i, \text{ where } n = \sum_{i=0}^{\lfloor \log_4 n \rfloor} x_i 4^i.$$

The permutation s is known as Sidorenko's permutation from the article [27]. Let $a < b < c$ be natural numbers with $2b = a + c$. It is clear that s avoids AP:s since there has to be a least $i \in \mathbb{N}_0$ such that $\varphi_4(\lfloor \frac{a}{4^i} \rfloor) + \varphi_4(\lfloor \frac{c}{4^i} \rfloor) \not\equiv 2\varphi_4(\lfloor \frac{b}{4^i} \rfloor) \pmod 4$.

We will now prove two facts about Sidorenko's AP-avoiding permutation. Then we will ask a couple of questions about AP-avoiding permutations in general relating to what we know (or conjecture) about a greedy AP-avoiding permutation, defined on page 42, and about Sidorenko's permutation, including a recognition of an infinite number of "cousins" of the latter.

Let us introduce the following notations:

Definition 3.3 *If $\tau : \mathbb{N} \to \mathbb{N}$ is a permutation, then*

$$A(\tau) = \{ k \in \mathbb{N} \mid [1, k] = \tau([1, k]) \},$$
$$A(\tau, n) = |A(\tau) \cap [1, n]|,$$
$$C(\tau, n) = A(\tau, n)/\sqrt{n}.$$

Theorem 3.1 *With the above notation, $k \in A(s)$ if and only if*

$$k = 3(4^{a_1} + 4^{a_2} + \ldots + 4^{a_m}) + c, \tag{25}$$

where $c = 0$ or -1 and $0 \leq a_1 < a_2 < \ldots < a_m$, which can be written in base-4 expansion as

$$k = k_n k_{n-1} \ldots k_1 + c, \tag{26}$$

where each $k_i = 0$ or 3, $c = 0$ or -1, and at least one $k_i \neq 0$, $i \in [1, n]$.

Proof. $\Rightarrow$) If $k \in A(s)$ then either $k = s(k)$ or $k > s(k)$. If $k = s(k)$, then k obviously is of the form (26) with $c = 0$. If $k > s(k)$ then we have to show that k is of the form (26) and $c = -1$. Now let $k = \alpha_n \alpha_{n-1} \ldots \alpha_1$ in base four notation. Look at the leftmost (the highest) position where $\alpha_a \neq 0$ or 3, (there has to be one, otherwise $k = s(k)$). We know that $\alpha_a \neq 1$, since $\alpha_a = 1$ would imply $k < s(k)$. Thus $\alpha_a = 2$. Now look at the positions to the right of a. Assume that one of these $a-1$ positions has value < 3. Then we can find $x \in [1,k]$ such that $s(x) > k$, namely $x = \alpha_n \alpha_{n-1} \dots \alpha_{a+1} 133 \dots 3$, which is a contradiction, so k is of the form (26) with $c = -1$ in this case.

$\Leftarrow$) If k has the form (25) and $c = -1$, then there is an integer $a \geq 1$, such that

$$k = k_n k_{n-1} \dots k_{a+1} 233 \dots 3, \tag{27}$$

where $k_i = 0$ or 3 for $a+1 \leq i \leq n$ and there are $a-1$ 3:s at the lowest positions. If $x = k$, then clearly $s(x) < k$. Thus let $1 \leq x = x_n x_{n-1} \dots x_1 < k$. Then there is a highest position $j \in [1,n]$ such that $x_j < k_j$. If $a+1 \leq j \leq n$ then $s(x) < k$ (Since $k_j = 3 \Rightarrow x_j = 0, 1$ or $2 \Rightarrow s(x_j) = 0, 1$ or 2.) If $j = a$ then 0 or $1 = x_j < k_a = 2$, so $s(x) \leq k$ (with equality if $x_j = 1$ and $x_l = k_l$, $j \neq l$.) Finally, if $1 < j < a$ then $s(x) < k$ (since $x_j < 3 \Rightarrow s(x_j) < 3$).

Therefore $s(x) \leq k$ if $x \in [1,k]$ which implies $k \in A(s)$ in the case $c = -1$, since $s : \mathbb{N} \to \mathbb{N}$ is injective. If $c = 0$ apply the same reasoning as above, just skip the last a positions in (27).

Corollary 3.2 *With notation as in Definition 3.3,*

$$A(s, 4^t) = 2^{t+1} - 2.$$

Proof. We count all elements of the set of k:s of the form (25). There are 2^t subsets consisting of 3:s in the positions $\log_4 4^t = t, t-1, \dots, 1$. Subtract the empty set (which we should not count) and multiply by 2 for $c = 0$ and $c = -1$ respectively. Hence we get $2(2^t - 1)$ such k:s up to $n = 4^t$. $\square$

Here we present an immediate consequence of Corollary 3.2, namely

$$C(s, 4^t) = A(s, 4^t)/\sqrt{4^t} \to 2, \text{ as } t \to \infty.$$

Question 3.1 *Is there any AP-avoiding permutation $\tau : \mathbb{N} \to \mathbb{N}$ such that*

$$\lim_{t \to \infty} C(\tau, 4^t) > 2 \text{ ?}$$

3.2.2 Sidorenko's "cousins"

By [17] Lemma 3.5, we have the following result: There are permutations

$$\varphi_k : [1, k] \to [1, k]$$

that avoid AP:s modulo k for infinitely many $k \in \mathbb{N}$.

Let $s_k : \mathbb{N} \to \mathbb{N}$ be a permutation defined by analogy with Sidorenko's permutation as

$$s_k(n) := \sum_{i=0}^{\lfloor \log_k n \rfloor} \varphi_k(x_i) k^i, \quad \text{where} \quad n = \sum_{i=0}^{\lfloor \log_k n \rfloor} x_i k^i.$$

Then each such s_k avoids AP:s by the same arguments as given before for $s = s_4$.

3.2.3 The greedy AP-avoiding permutation

We will now define the greedy AP-avoiding permutation of $\mathbb{N}$.

Definition 3.4 *Let $\pi_g(1) = 1$, and for $n > 1$, let $\pi_g(n)$ be the least natural number r such that $r \neq \pi(i)$ for all $i < n$ and $r \neq 2\pi(n-s) - \pi(n-2s)$ for all $1 \leq s \leq \lfloor \frac{n-1}{2} \rfloor$.*

To prove that π_g is a permutation, we need to show that it is surjective. The proof follows that of Hegarty in [17].

Theorem 3.3 *The function π_g is surjective.*

Proof. We start by proving that

$$\pi_g(n) < \frac{3n}{2}. \tag{28}$$

If (x, y, n) is an AP with $x < y < n$, then $x = 2y - n > 0$ and $y > \frac{n}{2}$. So there are at most $\frac{n-1}{2}$ such AP:s (x, y, n) which means that at most $\frac{n-1}{2}$ numbers can be rejected by π_g at position n. There are $n-1$ numbers $\pi_g(i)$ with $i < n$ so there must be at least

$$\frac{3n}{2} - 1 - (n-1) = \frac{n}{2}$$

numbers less than $\frac{3n}{2}$ that are different from $\pi_g(i)$ for all $i < n$. So there is a number less than $\frac{3n}{2}$ that will not be rejected for $\pi_g(n)$ and $\pi_g(n) < \frac{3n}{2}$.

We will now use (28) to show that $n \in X$, where

$$X = \{\pi_g(i) \mid 1 \leq i \leq 4n\}.$$

Suppose $n \notin X$ and let X_0 be the set of elements in X that exceeds n. Then $|X_0| \geq 4n - (n-1) = 3n + 1$. Let $x = \pi_g(i_x) \in X_0$. Then since $\pi_g(i_x) \neq n$ there must be an s_x with $1 \leq s_x \leq \lfloor \frac{i_x - 1}{2} \rfloor$ such that

$$n = 2\pi_g(i_x - s_x) - \pi_g(i_x - 2s_x).$$

But if $x, y \in X_0$ and $x \neq y$, we have $\pi_g(i_x - 2s_x) \neq \pi_g(i_y - 2s_y)$. This is because $\pi_g(i_x - 2s_x) = \pi_g(i_y - 2s_y))$ forces $\pi_g(i_x - s_x) = \pi_g(i_y - s_y)$ and then $i_x = i_y$ and $x = y$. Since $|X_0| \geq 3n + 1$, there are at least $3n + 1$ numbers $\pi_g(i_x - 2s_x) \in X$ with $x \in X_0$, and all of these are congruent to $n \mod 2$. This implies that there is a $\pi_g(m) \in X$ with $\pi_g(m) > 6n$. But $m < 4n$, and by (28), $\pi_g(m) < \frac{3m}{2} < 6n$. We have a contradiction, and $n \in X$.

By this we also get

$$\frac{1}{4} \leq \frac{\pi_g(n)}{n} < \frac{3}{2}. \tag{29}$$

$\square$

In [17], Hegarty shows that the lower bound in (29) can be improved to $\frac{3}{8}$.

The conjecture below is supported by computations up to $n = 560000$.

Conjecture 3.1 *The set $A(\pi_g)$ (with notation as in Definition 3.3) is finite with largest element $n = 136$.*

The computations tell us that

$$A(\pi_g) \cap [1, 560000] = \{1, 2, 4, 5, 6, 8, 10, 14, 26, 28, 29, 135, 136\}.$$

Let us introduce the notation

$$r(n) = |\{\, \pi_g(i) \mid \; 1 \leq i \leq n \quad \textit{and} \quad \pi_g(i) > n \,\}|.$$

In view of two diagrams presented on page 69 we "sharpen" the conjecture with the following question.

Question 3.2 *Is $r(n) > cn + d$ for all n and some $c, d \in \mathbb{R}$, $c > 0$?*

On page 70 we present the plot of $\pi_g(n)/n$ for $n \leq 560000$, and wonder:

Question 3.3 *Is $\limsup \pi_g(n)/n = \liminf \pi_g(n)/n = 1$?*

3.2.4 A consequence of Szemerédi's Theorem

Let $\mathcal{A} \subset \mathbb{N}$ and let $\mathcal{A}(n)$ denote the cardinality of $\mathcal{A} \cap [1, n]$. Let $\mathcal{S}$ be the set of all permutations of the natural numbers and let $\mathcal{P}_k \subset \mathcal{S}$ be the set of all permutations of the natural numbers that avoid k-term arithmetic progressions.

Corollary 3.4 *Let $\tau \in \mathcal{S}$ and let $\alpha \in \mathbb{N}$. If $|\tau(x) - x| \leq \alpha$ for all $x \in \mathbb{N}$, then $\tau \notin \mathcal{P}_k$.*

Proof. The finite version of Theorem 2.9, that is part (ii) of Proposition 2.8, says that there is an $N_0 \in \mathbb{N}$ such that $|A| < \frac{N_0}{2\alpha+1}$ for any $A \subset [1, N_0]$ that does not contain a k-term AP. Suppose $\phi : \mathbb{N} \to \mathbb{N}$ is an AP-avoiding permutation such that $|\phi(x) - x| \leq \alpha$. Let $A_i = \{x : \ \phi(x) = x + i, \ x \leq N_0\}$. Then

$$\{1, \ldots, N_0\} = \bigcup_{i=-\alpha}^{\alpha} A_i$$

and $A_{-\alpha}, \ldots, A_\alpha$ are all disjoint. If A_i contains an arithmetic progression, say $x, x + a, \ldots, x + ka$ with $x, a \in \mathbb{N}$, then $\phi(x), \phi(x + a), \ldots, \phi(x + ka)$ is also an arithmetic progression. So A_i does not contain any AP:s which, as we saw earlier, means $|A_i| < \frac{N_0}{2\alpha+1}$. Therefore

$$N_0 = \sum_{i=-\alpha}^{\alpha} |A_i| < \sum_{i=-\alpha}^{\alpha} \frac{N_0}{2\alpha + 1} = N_0.$$

We have a contradiction and the corollary is proved. □

For $k = 3$ we can improve Corollary 3.4 slightly by using Roth's theorem 2.7. Let ϕ be an AP avoiding permutation of $\mathbb{N}$ and let A_i be as in the proof of Corollary 3.4. Then there is a C such that for sufficiently large N_0 we have that $|A_i| \leq \frac{CN_0}{\log \log N_0}$ for all i. Define $S = \{i : A_i \neq \emptyset\}$. We estimate $|S|$ by

$$N_0 = \Big| \bigcup_{i \in S} A_i \Big| = \sum_{i \in S} |A_i| \leq |S| \frac{C N_0}{\log \log N_0} \tag{30}$$

so that

$$| \, S \, | \;\; \geq \;\; \frac{\log \log N_0}{C}. \tag{31}$$

There are $|S|$ different values of $\phi(x) - x$ with $x \leq N_0$ and hence at least $|S|/2$ different values of $|\phi(x) - x|$. Therefore there exists an $x_0 \leq N_0$ such that

$$|\phi(x_0) - x_0| \geq |S|/2 > \frac{\log \log N_0}{2C} = K \log \log N_0 \tag{32}$$

where $K > 0$. For further improvements in the three term case and in the general case we can use the equations (9) and (11) respectively.

3.3 AP-avoiding permutations of $[1, n]$

Proposition 3.5 *Let $n \in \mathbb{N}$, then there is an AP-avoiding permutation of $[1, n]$.*

Proof. Let $P(n)$ be the proposition that there exists an AP-avoiding permutation of $[1, n]$. $P(1)$ is trivially true. We will show that

(i) $P(n)$ implies $P(2n)$,

(ii) $P(n)$ and $P(n + 1)$ implies $P(2n + 1)$

which suffices, by induction over n, to prove the proposition.

For (i) we use the following construction: let π_n be an AP-avoiding permutation of $[1, n]$. Then the permutation $\pi : [1, 2n] \to [1, 2n]$ given by

$$\pi(i) := \begin{cases} 2\pi_n(i), & \text{if } i \leq n, \\ 2\pi_n(i - n) - 1, & \text{if } i > n, \end{cases}$$

avoids AP:s since if there was an AP, say $(\pi(x), \pi(x + a), \pi(x + 2a))$, then it had to take values out of both intervals $[1, n]$ and $[n + 1, 2n]$, which means that $\pi(x) + \pi(x + 2a)$ is odd which is impossible.

For (ii), let π_n and π_{n+1} be AP-avoiding perumtations of $[1, n]$ and $[1, n+1]$ respectively. Then the permutation $\tau : [1, 2n+1] \to [1, 2n+1]$ given by

$$\tau(i) := \begin{cases} 2\pi_{n+1}(i), & \text{if } i \leq n+1, \\ 2\pi_n(i-n-1)-1, & \text{if } i > n+1, \end{cases},$$

is seen to avoid AP:s in a similar way. □

3.3.1 Unimodal permutations of $[1, n]$

Definition 3.5 *If π is a permuation of $[1, n]$ then π is k-modal if the graph of π has at most k turning points in the interior of $[1, n]$.*

In case $k = 1, k = 2$ or $k = 3$ in the above definition, we say that π is *unimodal*, *bimodal* respectively *trimodal*.

The proof of the following theorem is a computer computation and the code, a C++ program, can be found on

http://www.mdstud.chalmers.se/˜ md0larur/unimodal.cpp .

Theorem 3.6 *There are no unimodal AP-avoiding permutations of $[1, n]$ for $n > 14$.*

"Proof." We assume WLOG that 1 is the turning point of a unimodal AP-avoiding permutation of $[1, n]$. Then clearly $\pi(n) = n$ or $\pi(1) = n$. In the first case, $\pi|_{[1,n-1]}$ is a unimodal AP-avoiding permutation of $[1, n-1]$. In the latter case, the function ϕ defined by $\phi(i) = \pi(i+1)$ for $i \in [1, n-1]$, is a unimodal AP-avoiding permutation of $[1, n-1]$.

So every unimodal AP-avoiding permutations of $[1, n]$ can be constructed from a unimodal AP-avoiding permutation of $[1, n-1]$, by "putting n at either end". The C++ program computed all unimodal AP-avoiding permutations of $[1, n]$ for $1 \leq n \leq 15$. But there were no such permutations of $[1, 15]$, and hence there are no unimodal AP-avoiding permutations of $[1, n]$ for any $n \geq 15$. □

Below is the list of all unimodal AP-avoiding permutations of $[1, n]$ for any $n \in \mathbb{N}$, that cannot be extended by putting $n+1$ at either the right- or the left-hand side:

$(\pi(1), \pi(2), \ldots, \pi(n)) =$
$(12, 11, 9, 8, 5, 4, 2, 1, 3, 6, 7, 10)$,
$(11, 9, 8, 5, 4, 2, 1, 3, 6, 7, 10, 12)$,
$(13, 11, 10, 8, 5, 4, 2, 1, 3, 6, 7, 9, 12, 14)$,
$(9, 8, 6, 5, 2, 1, 3, 4, 7)$,
$(12, 11, 8, 6, 5, 2, 1, 3, 4, 7, 9, 10)$,
$(13, 11, 8, 6, 5, 2, 1, 3, 4, 7, 9, 10, 12)$,
$(6, 5, 2, 1, 3, 4, 7, 8)$,
$(10, 8, 7, 5, 2, 1, 3, 4, 6, 9, 11, 12)$,
$(12, 11, 8, 7, 5, 2, 1, 3, 4, 6, 9, 10, 13)$,
$(11, 8, 7, 5, 2, 1, 3, 4, 6, 9, 10, 12, 13)$,
$(13, 12, 10, 9, 6, 4, 3, 1, 2, 5, 7, 8, 11)$,
$(13, 10, 9, 6, 4, 3, 1, 2, 5, 7, 8, 11, 12)$,
$(12, 11, 9, 6, 4, 3, 1, 2, 5, 7, 8, 10)$,
$(8, 7, 4, 3, 1, 2, 5, 6)$,
$(12, 10, 9, 7, 4, 3, 1, 2, 5, 6, 8, 11, 13)$,
$(10, 9, 7, 4, 3, 1, 2, 5, 6, 8, 11, 12)$,
$(7, 4, 3, 1, 2, 5, 6, 8, 9)$,
$(14, 12, 9, 7, 6, 3, 1, 2, 4, 5, 8, 10, 11, 13)$,
$(12, 10, 7, 6, 3, 1, 2, 4, 5, 8, 9, 11)$,
or
$(10, 7, 6, 3, 1, 2, 4, 5, 8, 9, 11, 12)$.

3.4 Sidon permutations

Proposition 3.7 *There is no Sidon permutation of $\mathbb{Z}/n\mathbb{Z}$ for any $n \in \mathbb{N}$.*

Proof. Suppose there is a Sidon permutation of $\mathbb{Z}/n\mathbb{Z}$. Then for $1 \leq k \leq n$, the elements $\pi(k) - \pi(k-1)$ are all incongruent modulo n, and since there are altogether n of them, we have that $\pi(k) - \pi(k-1) \equiv 0 \pmod{n}$ for some k, so that π is not even a permutation, a contradiction. $\square$

3.4.1 There is a Sidon permutation of the natural numbers

Proposition 3.8 *There exists a Sidon permutation of* $\mathbb{N}_0$.

Proof. Let π be any permutation of $\mathbb{N}_0$ and define $\xi : \mathbb{N}_0 \to \mathbb{Z}$ by

$$\xi(n) = \begin{cases} 0, & n = 0; \\ \pi(n) - \pi(n-1), & \text{otherwise;} \end{cases}$$

Then π is a Sidon permutation if and only if the following two conditions are satisfied:

(i) the map

$$n \mapsto \pi(n) - \pi(0) = \sum_{i=0}^{n} \xi(i)$$

from $\mathbb{N}_0$ to $\mathbb{Z}$ is injective with image $\mathbb{N}_0$.

(ii) for all s, t, l, whenever

$$\sum_{i=s}^{s+l} \xi(i) = \sum_{j=t}^{t+l} \xi(j),$$

then $s = t$. We now describe a function ξ for which these requirements will be fulfilled. First we put $\xi(0) := 0$. We define a sequence of integers $(p_k)_0^\infty$ as follows. Set $p_0 = 0$. For $k \in \mathbb{N}$ set

$$p_k := \frac{3}{2}(3^k - 1) = \sum_{i=1}^{k} 3^i.$$

Then, if $p_{k-1} < n < p_k$ we put $\xi(n) := 2^{n-1}$. Finally, if $n = p_k$ for some k, we assign $\xi(n)$ so that

$$\pi(n) := \sum_{i=0}^{n} \xi(i)$$

is the least element of $\mathbb{N}_0$ not already appearing among $\pi(0), \pi(1), \ldots, \pi(n-1)$. □

3.4.2 Costas arrays

Definition 3.6 *Let $n \in \mathbb{N}$. A Costas array is simply a Sidon permutation of $[1, n]$.*

Proposition 3.9 *Let p be a prime and i a natural number. Then there are Costas arrays for $n = p^i - 2$ and for $n = p - 1$.*

The above proposition is settled by the following constructions, the proof of which can be found in [13]. Let $\pi : [1, n] \to [1, n]$.

Welch construction. Let p be prime and α a primitive element in the field $\mathbb{F}_p$. Let $n = p - 1$. Then a Costas array of order n is obtained by putting $\pi(i) := j$ iff $i = \alpha^j$ for $1 \leq i, j \leq n$.

Golomb construction. Let α and β be primitive elements, not necessarily distinct, in the field $\mathbb{F}_q$ for q a prime power. Let $n = q - 2$. Put $\pi(i) := j$ iff $\alpha^i + \beta^j = 1$, where $1 \leq i, j \leq n$.

Costas arrays of order n are known to exist for all $1 \leq n < 359$ except for $n \in \{32, 33, 43, 48, 49, 54, 63, 73, 74, 83, 84, 85, 89-93, 97, 103, 109, 113-117, 120, 121, 131-133, 139-143, 151-153, 157-159, 163, 168, 169, 173, 174, 181-186, 193, 199-207, 211-219, 223, 229, 233, 234, 242-246, 251, 257-259, 263, 271-273, 277, 283-285, 288, 289, 293-303, 313, 317-327, 331-333, 337-339, 342, 343, 349, 353, 354\}$. This result is presented in section IV.7.13 of [6], where one also says that almost all known Costas arrays of order $n \geq 30$ are derived from algebraic constructions like Welch or Golumb constructions. The few exceptions to this are when the order is $n = 31$ or 53 and when one has an explicit algebraic construction where $\pi(n) = n$ (or $\pi(1) = 1$), which clearly gives a Costas array of order $n - 1$ as well.

3.5 LAC-avoiding permutations

Definition 3.7 *Let $A = (a_1, \ldots a_l \mid b_1, \ldots, b_m)$ be a linear arithmetic condition. A permutation π of the natural numbers avoids A if whenever $(x_1, \ldots x_l, y_1, \ldots, y_m)$ is a nontrivial solution to A, then*

$$(\pi(x_1), \ldots \pi(x_l), \pi(y_1), \ldots, \pi(y_m))$$

is not a solution to A.

A permutation π of $\mathbb{Z}/n\mathbb{Z}$ avoids A if whenever $(x_1, \ldots x_l, y_1, \ldots, y_m)$ is a nontrivial solution to A modulo n, then

$$(\pi(x_1), \ldots \pi(x_l), \pi(y_1), \ldots, \pi(y_m))$$

is not a solution to A modulo n.

We call a TAC, say $A = (a_1, \ldots a_u \mid b_1, \ldots, b_v)$, a *Sidon-type* arithmetic condition if the following criterion is fulfilled:

There is a partitioning (I, J, K, L) and a partitioning (M, N, P, Q) respectively of two sets $U \subset [1, u]$ and $V \subset [1, v]$ of the LAC A such that

$$0 \neq \sum_{i \in I} a_i - \sum_{m \in M} b_m = \sum_{j \in J} a_j - \sum_{n \in N} b_n = \sum_{p \in P} b_p - \sum_{k \in K} a_k = \sum_{q \in Q} b_q - \sum_{l \in L} a_l.$$

Corollary 3.10 *There is no permutation of $\mathbb{Z}/n\mathbb{Z}$ that avoids a Sidon-type TAC.*

Proof. By identifying the indeterminates in $I \cup M$, $J \cup N$, $K \cup P$, $Q \cup L$ with the single variables x, y, z, w respectively, the Sidon type TAC reduces to the ordinary Sidon condition $(1, 1|1, 1)$. Hence the result follows immediately from Proposition 3.7. □

Proposition 3.11 *Let $\sum_{i=1}^{n} a_i x_i = \sum_{i=1}^{m} b_i y_i$ be a LAC, with $\Delta > 0$ and let p be a prime such that $\Delta \not\equiv 0 \mod p$. If π is a permutation of $\mathbb{N}$ such that $\pi(x) \equiv x + 1 \pmod p$, then π is $(a_1, \ldots, a_n \mid b_1, \ldots, b_m)$-avoiding.*

Proof. Suppose π is not $(a_1, \ldots, a_n \mid b_1, \ldots, b_m)$-avoiding. Then there is a tuple

$$(x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_m)$$

such that $\sum_{i=1}^n a_i x_i = \sum_{i=1}^m b_i y_i$ and $\sum_{i=1}^n a_i \pi(x_i) = \sum_{i=1}^m b_i \pi(y_i)$. Thus

$$\sum_{i=1}^{n} a_i(x_i+1) \equiv \sum_{i=1}^{m} b_i(y_i+1) \pmod p$$

and since

$$\sum_{i=1}^{n} a_i x_i \equiv \sum_{i=1}^{m} b_i y_i \pmod p$$

we get

$$\Delta = \sum_{i=1}^{n} a_i - \sum_{i=1}^{m} b_i \equiv 0 \pmod p,$$

which contradicts the definition of p. □

Definition 3.8 *Let p be a prime and $n \in \mathbb{N}$. Let $n \equiv r \pmod p$ where $1 \le r \le p$. If $r < p$, let $\pi_p(n) = n+1$ and if $r = p$, let $\pi_p(n) = n - p + 1$.*

Proposition 3.12 *Let A be a linear arithmetic condition. There is an $A-$avoiding permutation π of the natural numbers such that $|\pi(n) - n| < C$ for all natural numbers n and some constant C if and only if $\Delta(A) > 0$.*

Proof. $\Leftarrow$*)* The permutation π_p in Definition 3.8 satisfies $\pi_p(x) \equiv x + 1 \pmod p$, so by Proposition 3.11, π_p avoids any LAC with $\Delta > 0$ and $\Delta \not\equiv 0 \mod p$. Further, $|\pi_p(x) - x| < p$ for all x.

$\Rightarrow$*)* If A is a TAC then, as we showed in Proposition 2.34 any subset of $\mathbb{N}$ with positive upper asymptotic density contains a non-trivial solution of A. But if τ is a permutation of $\mathbb{N}$ such that $|\,\tau(x) - x\,| < c$ for some constant c and all x, then, as in the proof of Corollary 3.4 it follows that for some i, $\{x : \tau(x) = x + i\}$ has positive density. Thus this set contains a non-trivial solution to A and applying τ yields another solution since A is translation invariant. Hence τ does not avoid A as required. □

4 Another greedy permutation

In this chapter we will study one and the same permutation from three different perspectives. First we make a formal definition.

Definition 4.1 *Let $\pi : \mathbb{N} \to \mathbb{N}$ and let $x, y \in \mathbb{N}$. We define $\pi(x)$ recursively as the least natural number such that*

$$\pi(x) \notin \{ \pi(y) \mid 1 \leq y < x \}$$

and such that $\pi(x) - x \neq \pi(y) - y$ for $1 \leq y < x$.

4.1 Fibonacci sequences

Notice that π is immediately from the definition an injection into the natural numbers. We are going to show that π is surjective. It is clear that $\pi(1) = 1$. Let $x \in \mathbb{N}$ with $x \geq 2$. Let $\alpha \leq x$ be the least natural number that is not contained in $\{ \pi(i) \mid i \in [1, x] \}$ if it exists. It is clear that

$$\Delta = \min_{y \in [1,x]} \{\pi(y) - y\} \leq 0.$$

We deduce from Definition 4.1, that our greedy procedure will not accept the value of α any earlier then at position $\alpha + (1 - \Delta)$, but indeed here it is fine. Hence, we get $\pi^{-1}(\alpha) = \alpha + 1 - \Delta$ $(> x)$ and so π is a permutation of the natural numbers. We will proceed to show that π has a special property that in a sense defines it unikely.

Lemma 4.1 *Let $f : \mathbb{N} \to \mathbb{N}$ be a permutation and let $g : \mathbb{N} \to \mathbb{Z}$ be defined by $g(x) = f(x) - x$, for $x \in \mathbb{N}$. Let $x_0 < x_1$ be natural numbers. If g is a bijection such that*

$$g(x_0) > 0,\ g(x_1) > 0 \quad \Rightarrow \quad g(x_0) < g(x_1) \tag{33}$$

and

$$g(x_0) < 0,\ g(x_1) < 0 \ \Rightarrow \ g(x_0) > g(x_1), \tag{34}$$

then $f \equiv \pi$.

Proof. At first we show that π in Definition 4.1 satisfies the conditions (33) and (34). If the conditions (33) and (34) has been satisfied for all natural numbers $< a$, then by Definition 4.1 (and by the proof for the surjectivity of π) $\pi(a) - a$ will either take the value of

$$\min_{x \in [1,a]} \{\pi(x) - x\} - 1 \quad \text{or} \quad \max_{x \in [1,a]} \{\pi(x) - x\} + 1.$$

There is clearly no other choice. From this it is not hard to see that $\lambda(x) = \pi(x) - x$ is a bijection $\lambda : \mathbb{N} \to \mathbb{Z}$, so that $\pi(1) - 1 = 0$ and the base cases $\pi(2) - 2 = 1,\ \pi(4) - 4 = 2$ and $\pi(3) - 3 = -1, \pi(6) - 6 = -2$ give the result.

Suppose that $\tau \neq \pi$ is another permutation of the natural numbers that satisfies (33) and (34). Then there exists a least natural number x_0 such that $\pi(x) = \tau(x)$ if $x < x_0$ but $\pi(x_0) \neq \tau(x_0)$.

Let $\varphi : \mathbb{N} \to \mathbb{Z}$ be the bijection $\varphi(x) = \tau(x) - x$.

Case 1: $\pi(x_0) - x_0 < 0$. Then $\tau(x_0) - x_0 > 0$, because otherwise a bijective φ would have contradicted the condition (34). Since τ is surjective we can put $y_0 = \tau^{-1}(\pi(x_0))$. We have that $x_0 + \alpha = y_0$ for some $\alpha \geq 1$ (suppose on the contrary that $x_0 > y_0$, then $\tau(y_0) = \pi(y_0)$ gives $y_0 = x_0$, which is a contradiction) and so we get $\tau(y_0) - y_0 = \pi(x_0) - x_0 - \alpha$. Since φ is a bijection it follows from (33) that the $\alpha + 1$ strictly decreasing negative numbers $\pi(x_0) - x_0,\ \pi(x_0) - x_0 - 1,\ \ldots,\ \pi(x_0) - x_0 - \alpha$ has to take values only from the numbers in the sequence $\varphi(x_0 + 1),\ \varphi(x_0 + 2),\ \ldots,\ \varphi(y_0)$. This is impossible.

Case 2: $\pi(x_0) - x_0 > 0$. Then $\tau(x_0) - x_0 < 0$, because otherwise a bijective φ would have contradicted (33). But since the greedy algorithm chooses $\pi(x_0)$, which is strictly greater than $\tau(x_0)$, it means that there is a natural number $z_0 < x_0$ such that $\pi(z_0) - z_0 = \tau(x_0) - x_0$. Then $\varphi(x_0) = \varphi(z_0)$ which contradicts that φ is a bijection. □

Proposition 4.2 *Let π be as in Definition 4.1, then the natural numbers can be partitioned into two subsets I and J, such that*

$$\lim_{i \to \infty} \frac{\pi(i)}{i} = \frac{\sqrt{5} + 1}{2} \ \textit{and} \ \lim_{j \to \infty} \frac{\pi(j)}{j} = \frac{2}{\sqrt{5} + 1},$$

whenever $i \in I$ and $j \in J$.

Proof. *Step 1, "Fibonacci sequences":* Let $F_1 = F : \mathbb{N} \to \mathbb{N}$ be the "Fibonacci sequence", defined by $F(1) = F(2) = 1$ and $F(n + 2) = F(n + 1) +$

$F(n)$ for $n \geq 1$. Further, for $\alpha \geq 2$, let $F_\alpha : \mathbb{N} \to \mathbb{N}$ be defined by

$$F_\alpha(n+2) = F_\alpha(n+1) + F_\alpha(n), \tag{35}$$

for $n \geq 1$ and where $F_\alpha(1)$ is the least natural number excluded from the sequences F_β with $1 \leq \beta < \alpha$. Define $F_\alpha(2)$ by

$$F_\alpha(2) := F_\gamma(k+1) + 1, \tag{36}$$

where $1 \leq \gamma < \alpha$ is the greatest natural number such that $F_\alpha(1) = F_\gamma(k)+1$, with $k \in \mathbb{N}$. The existance of γ and k is clear from the definition of $F_\alpha(1)$.

Step 2, "relation into involution": We introduce a function $r : \mathbb{N} \to \mathbb{N}$, defined by

$$r(F_\alpha(2i-1)) := F_\alpha(2i) \tag{37}$$

and

$$r(F_\alpha(2i)) := F_\alpha(2i-1) \tag{38}$$

for $i, \alpha \in \mathbb{N}$. In order to show that r is welldefined (which we will do in *Step 3*), it suffices to prove that

$$F_\alpha(x) \neq F_\beta(y), \tag{39}$$

for each $x, y, \alpha, \beta \in \mathbb{N}$ with $\alpha \neq \beta$. Observe that r is defined on each natural number (by the greedy choice of $F_\alpha(1)$) and notice that if r is well defined as a function, then r is automatically an involution of the natural numbers.

Step 3, "increasing in couples": Let us define a property $\mathcal{V}(n)$ as follows: For each couple of natural numbers $\beta < \alpha \leq n \in \mathbb{N}$, there is exactly one natural number $j > 1$ such that

$$F_\alpha(1) < F_\beta(j) < F_\alpha(2) \tag{40}$$

and for this j we have that $F_\beta(j-1) < F_\alpha(1)$.

Let $\mathcal{V}$ denote the statement that $\mathcal{V}(n)$ holds for each $n \in \mathbb{N}$. With the notation as in (36), property $\mathcal{V}$ implies that the γ in *Step 1* "is unique", (i.e. if $F_\beta(x) + 1 = F_\gamma(y) + 1 = F_\alpha(1)$ with $\beta \leq \gamma$, then $\mathcal{V}$ gives $\beta = \gamma$ and $x = y$). Also, $\mathcal{V}$ evidently implies (39). To prove $\mathcal{V}$ we use induction over $n \in \mathbb{N}$.

Put $\alpha = 2, \beta = 1, j = 5$ in (40). By *Step 1*, $\mathcal{V}(2)$ holds. Suppose that $\mathcal{V}(n)$ holds for some $n > 2$. Define η such that $F_\eta(l) + 1 = F_{n+1}(1)$ for some

$l \in \mathbb{N}$. By (35) and by $\mathcal{V}(n)$ we get that for each $\alpha \leq n$ there is a $k \geq 2$ such that

$$F_\eta(l) < F_\alpha(k) < F_\eta(l+1).$$

By the induction hypothesis, the value $F_{n+1}(2) = F_\eta(l+1)+1$ has not been choosen by any Fibonacci sequence F_m with $\eta \neq m \leq n$, since for $a, b \in \mathbb{N}$, if $F_\eta(l-1) < a < F_\eta(l) < b < F_\eta(l+1)$ we get $a+b > F_\eta(l+1)+1$.

It follows that $F_\eta(l) < F_{n+1}(1) < F_\alpha(k) < F_\eta(l+1) < F_{n+1}(2) < F_\alpha(k+1)$. But this gives $F_\alpha(k-1) < F_{n+1}(1)$. Therefore $\mathcal{V}(n+1)$ holds and so does $\mathcal{V}$, by induction.

Step 4, "hooked in a bijection": Define a function $f : \mathbb{N} \to \mathbb{Z}$ by

$$f(x) := r(x) - x. \tag{41}$$

We want to show that f is bijective. For every $\alpha, i \in \mathbb{N}$ with the exception of $\alpha = i = 1$, we have

$$f(F_\alpha(2i-1)) = r(F_\alpha(2i-1)) - F_\alpha(2i-1) = F_\alpha(2i) - F_\alpha(2i-1) > 0$$

and

$$f(F_\alpha(2i)) = r(F_\alpha(2i)) - F_\alpha(2i) = F_\alpha(2i-1) - F_\alpha(2i) < 0.$$

For $i \geq 2$ and any $\alpha \in \mathbb{N}$ we have

$$f(F_\alpha(2i-1)) = F_\alpha(2i) - F_\alpha(2i-1) = F_\alpha(2i-2) = -f(F_\alpha(2i),$$

so that f takes both positive and negative values from each even position of each F_α. Similarily, we claim the following:

For all natural numbers β and j, there is an $\alpha > \beta$ such that

$$f(F_\alpha(1)) = F_\alpha(2) - F_\alpha(1) =$$

$$F_\beta(2j+1) - F_\beta(2j) = F_\beta(2j-1) = -f(F_\alpha(2)). \tag{42}$$

Since $f(1) = F_1(2) - F_1(1) = F_1(1) - F_1(2) = 0$, the above statements settle the bijectiveness of f and so we will use the remainder of *Step 4* to prove that (42) holds. It suffices to show that the following proposition holds: Let $k \geq 2$. If $\beta \in \mathbb{N}$ and $j \geq 2$ and if $F_\beta(2j) < F_k(1)$, then there is an $\alpha \in [1, k]$ such that

$$F_\alpha(1) = F_\beta(2j) + 1.$$

This proposition holds for $k = 2$. Suppose it holds for a fixed $k > 2$. We need to show that $F_{k+1}(1) = F_\eta(2j) + 1$ for some $j \in \mathbb{N}$ and where $\eta \leq k$. Suppose that β and j are such that $F_\beta(2j+1) < F_k(2)$. Then the induction hypothesis gives $F_{k+1}(1) \neq F_\beta(2j+1) + 1$. Suppose on the other hand that $F_{k+1}(1) > F_k(2)$. Then, by $\mathcal{V}$, this forces for each $\alpha \in [2,\, k]$, that $F_\alpha(2) - F_\alpha(1) = \alpha$. This is not true if $k > 2$. So $F_{k+1}(1) \neq F_\beta(2j+1) + 1$ for any $j \in \mathbb{N}$.

Hence there are only two possibilities left for $F_{k+1}(1)$, namely:

Alt 1. $F_{k+1}(1) = F_k(1) + 1$

Alt 2. $F_{k+1}(1) = F_\gamma(2j) + 1$, for some $\gamma, j \in \mathbb{N}$ where $\gamma \leq k$.

We are going to rule out *Alt 1.* We do this by showing that $F_k(1)+1 = F_x(y)$, where $x < k$ and y is such that $F_x(y) - F_x(y-1) = k$. Then property $\mathcal{V}$ gives that $F_{k+1}(1) \neq F_x(y)$.

Suppose $F_k(1) + 1 > F_x(y)$. Then $\mathcal{V}$ implies that exactly one number in the interval $[F_x(y-1)+1, F_x(y)]$ is not choosen by any of the Fibonacci sequences F_i, for $i \in [1,k]$. This means that r is not surjective, which is nonsense.

Suppose $F_k(1) + 1 < F_x(y)$. Then by $\mathcal{V}$, there would need to be a couple (z, w) with $x \neq z \in [1,k]$ and $w > 2$ such that $F_z(w) = F_k(1) + 1$. But then $F_z(w) - F_z(w-1) < k$ which contradicts $\mathcal{V}$ (exactly one number from each F_i with $i \in [1,k]$ had to be in the interval $[F_z(w-1), F_z(w)]$). We have shown that the initial value of a Fibonacci sequence F_i has to be of the form suggested in **Alt 2.** and it follows that f is bijective.

Step 5, "identification with greedy": By *Step 1-4* each criteria in Lemma 4.1 is fulfilled by r. Therefore $r \equiv \pi$. It remains to check the convergence of the limit values in the proposition.

Step 6, "convergence properties": Let $k \in \mathbb{N}$. It is a wellknown result that

$$\lim_{k\to\infty} \frac{F_1(k)}{F_1(k-1)} = \frac{\sqrt{5}+1}{2}$$

and this holds also for any $\alpha \geq 2$ so that

$$\lim_{k\to\infty} \frac{F_\alpha(k)}{F_\alpha(k-1)} = \frac{\sqrt{5}+1}{2}.$$

Our task is to show that $\frac{r(F_\alpha(k))}{F_\alpha(k)}$ converges as requested. Therefore it only remains to verify that

$$\lim_{\alpha\to\infty}\frac{F_\alpha(i+1)}{F_\alpha(i)}=\frac{\sqrt{5}+1}{2},$$

for a fixed $i\in\mathbb{N}$.

By property $\mathcal{V}$, there is an increasing sequence $a_x\to\infty$ as $x\to\infty$ such that for each $x\in\mathbb{N}$ we have that $F_x(i)\in[F_1(a_x-1),F_1(a_x)]$. Hence, it follows from $\mathcal{V}$ that

$$\frac{F_1(a_x)}{F_1(a_x-1)}<\frac{F_x(i+1)}{F_x(i)}<\frac{F_1(a_x+1)}{F_1(a_x)}$$

and since

$$\lim_{x\to\infty}\frac{F_1(a_x)}{F_1(a_x-1)}=\lim_{x\to\infty}\frac{F_1(a_x+1)}{F_1(a_x)}=\frac{\sqrt{5}+1}{2},$$

we are done.

□

4.2 Beatty sequences

In this section we will derive a more explicit representation of π, and for this we need a theorem known as Beatty's theorem. It appeared as a problem in *American mathematical monthly* in 1926 [2] and a solution was given in the same publication just a few months later.

Theorem 4.3 *Let r and s be a irrational numbers such that*

$$\frac{1}{r}+\frac{1}{s}=1.$$

Further, let $A=\{\lfloor nr\rfloor \mid n\in\mathbb{N}\}$ and $B=\{\lfloor ns\rfloor \mid n\in\mathbb{N}\}$. Then $A\cap B=\emptyset$ and $A\cup B=\mathbb{N}$.

Proof. The idea is to count the number of elements of A and B that is contained in the interval $[1,N]$ for some $N\in\mathbb{N}$. We have

$$\big|A\cap[1,N]\big|=\Big\lfloor\frac{N+1}{r}\Big\rfloor$$

and

$$\big|B \cap [1, N]\big| = \Big\lfloor \frac{N+1}{s} \Big\rfloor.$$

Since r and s are irrational we get he following inequalities,

$$\Big\lfloor \frac{N+1}{r} \Big\rfloor - 1 < \Big\lfloor \frac{N+1}{r} \Big\rfloor < \frac{N+1}{r}$$

and

$$\Big\lfloor \frac{N+1}{s} \Big\rfloor - 1 < \Big\lfloor \frac{N+1}{s} \Big\rfloor < \frac{N+1}{s}.$$

Adding these inequalities, we get

$$\begin{aligned} N - 1 = \frac{N+1}{r} + \frac{N+1}{s} - 2 &< \Big\lfloor \frac{N+1}{r} \Big\rfloor + \Big\lfloor \frac{N+1}{s} \Big\rfloor \\ &< \frac{N+1}{r} + \frac{N+1}{s} = N + 1 \end{aligned}$$

which implies

$$|A \cap [1, N]| + |B \cap [1, N]| = N.$$

By putting $N = 1$, we see that $1 \in A$ or $1 \in B$, but $1 \notin A \cap B$. In general, going from N to $N + 1$, exactly one element is added to either A or B but not to both. □

Theorem 4.4 *Let ϕ be the golden ratio, i.e. ϕ is the positive solution to the equation $x^2 = x+1$. Then for $n \in \mathbb{N}_0$, $\pi(\lceil n\phi \rceil) = \lceil n\phi^2 \rceil$ and $\pi(\lceil n\phi^2 \rceil) = \lceil n\phi \rceil$.*

Proof. Let us first note that since $\frac{1}{\phi} + \frac{1}{\phi^2} = 1$, Theorem 4.3 tells us that

$$\{\, \lceil n\phi \rceil \mid n \in \mathbb{N} \} \cup \{\, \lceil n\phi^2 \rceil \mid n \in \mathbb{N} \} = \mathbb{N} \setminus \{1\}$$

and that

$$\{\, \lceil n\phi \rceil \mid n \in \mathbb{N} \} \cap \{\, \lceil n\phi^2 \rceil \mid n \in \mathbb{N} \} = \emptyset.$$

We will prove the theorem by induction. So let $k \geq 2$ be a natural number and suppose the theorem is true for all positions of π strictly less than k. We have two possibilities for the position k:

Case 1: There is an $n \in \mathbb{N}$ such that $k = \lceil n\phi \rceil$. By the induction hypothesis we have

$$\pi(\lceil i\phi \rceil) - \lceil i\phi \rceil = \lceil i\phi^2 \rceil - \lceil i\phi \rceil = \lceil i\phi + i \rceil - \lceil i\phi \rceil = i + \lceil i\phi \rceil - \lceil i\phi \rceil = i \tag{43}$$

for $i \in \mathbb{N}$ with $i < n$. Correspondingly, if $i \in \mathbb{N}$ and $i < \frac{n}{\phi}$, then by using Theorem 4.3 we get $\lceil i\phi^2 \rceil < \lceil n\phi \rceil$ and the induction hypothesis gives

$$\pi(\lceil i\phi^2 \rceil) - \lceil i\phi^2 \rceil = -i. \tag{44}$$

Combining equations (43) and (44) yields

$$\{\pi(j) - j \mid 1 \leq j < k\} = [-\lfloor n/\phi \rfloor,\ n-1].$$

To prove that $\pi(\lceil n\phi \rceil) = \lceil n\phi^2 \rceil$, it suffices, by the definition of π, to prove that

$$\pi(\lceil n\phi \rceil) - \lceil n\phi \rceil \geq \lceil n\phi^2 \rceil - \lceil n\phi \rceil = n.$$

To arrive at a contradiction, suppose that

$$\pi(\lceil n\phi \rceil) - \lceil n\phi \rceil < n. \tag{45}$$

Then we must have

$$\pi(\lceil n\phi \rceil) - \lceil n\phi \rceil < -\lfloor n/\phi \rfloor$$

and so

$$\pi(\lceil n\phi \rceil) < \lceil n\phi \rceil - \lfloor \frac{n}{\phi} \rfloor = \lceil n\phi \rceil - \lfloor n(\phi - 1) \rfloor = n - 1,$$

say $\pi(\lceil n\phi \rceil) = c < n - 1$. By Theorem 4.3, there is an $l \in \mathbb{N}_0$ such that $c = \lceil l\phi \rceil$ or $c = \lceil l\phi^2 \rceil$. If $c = \lceil l\phi^2 \rceil$, then $\lceil n\phi \rceil > n - 1 > \lceil l\phi \rceil$ and the induction hypothesis gives

$$\pi(\lceil l\phi \rceil) = \lceil l\phi^2 \rceil = c = \pi(\lceil n\phi \rceil).$$

This means that π is not injective, but this contradicts the definition of π. Therefore, $c = \lceil l\phi \rceil$ and $\lceil l\phi \rceil < n - 1$ which implies $\lceil l\phi^2 \rceil < \lceil n\phi \rceil$. Then

$$\pi(\lceil l\phi^2 \rceil) = \lceil l\phi \rceil = c = \pi(\lceil n\phi \rceil)$$

which again contradicts the injectivity of π. The assumption in equation (45) must be false and $\pi(k) = \pi(\lceil n\phi \rceil) = \lceil n\phi^2 \rceil$.

Case 2: There is an $n \in \mathbb{N}$ such that $k = \lceil n\phi^2 \rceil$. Then $\lceil n\phi \rceil$ is the least number which is not equal to $\pi(i)$ for any $i < \lceil n\phi^2 \rceil$. We also have

$$\pi(i) - i > -n = \lceil n\phi \rceil - \lceil n\phi^2 \rceil$$

for $i < \lceil n\phi^2 \rceil$ and so by the definition of π, $\pi(k) = \pi(\lceil n\phi^2 \rceil) = \lceil n\phi \rceil$. □

4.3 An integer-average greedy permutation

The permutation π is identical to another permutation that is defined in a greedy way as taking the least natural number such that the average of the "initial segment" is an integer. More precisely:

Definition 4.2 *For $n \in \mathbb{N}$, let $\sigma(n)$ be the least natural number such that $\sigma(n) \neq \sigma(i)$ for all $i < n$ and such that $(\sum_{i=1}^{n} \sigma(i))/n$ is an integer.*

Lemma 4.5 *For each $n \in \mathbb{N}$ with $n > 1$, let $m_n := (\sum_{i=1}^{n} \sigma(i))/n$. Then*

i) $\sigma(n) = n(m_n - m_{n-1}) + m_{n-1}$,

ii) $\sigma^{-1}(i) \leq n$ *if* $i \leq m_{n-1}$,

iii) *either* $m_n = m_{n-1}$ *or* $m_n = m_{n-1} + 1$.

Proof. To show *i)*,

$$\sigma(n) = \sum_{i=1}^{n} \sigma(i) - \sum_{i=1}^{n-1} \sigma(i) = nm_n - (n-1)m_{n-1} = n(m_n - m_{n-1}) + m_{n-1}.$$

For *ii)* and *iii)* we use induction. The lemma is true for $n = 2$, because $\sigma(1) = m_1 = 1$ and $m_2 = 2$. Now suppose it is true for all $k < n$. Then if $i \leq m_{n-2}$ we have $\sigma^{-1}(i) \leq n - 1$ and so $\sigma(n) \geq m_{n-2} + 1$. But by the induction hypothesis, $m_{n-1} \leq m_{n-2} + 1$ and

$$\sigma(n) = n(m_n - m_{n-1}) + m_{n-1} \geq m_{n-2} + 1 \geq m_{n-1},$$

which implies $m_n \geq m_{n-1}$. If m_{n-1} is not already chosen by σ, then $\sigma(n) = m_{n-1}$ and $m_n = m_{n-1}$. If on the other hand $\sigma^{-1}(m_{n-1}) < n$, then if $k < n$ the induction hypothesis gives $m_k - m_{k-1} \leq 1$ and

$$n + m_{n-1} > k + m_{k-1} \geq k(m_k - m_{k-1}) + m_{k-1} = \sigma(k).$$

This implies $\sigma(n) = n + m_{n-1}$ and $m_n = m_{n-1} + 1$. Both cases above also give $\sigma^{-1}(m_{n-1}) \leq n$. $\square$

Note that the lemma shows that σ is a permutation.

Proposition 4.6 *The permutations σ and π are equal.*

Proof. The proof is by induction. For the base case, we see that $\pi(1) = \sigma(1) = 1$, $\pi(2) = \sigma(2) = 3$ and $\pi(3) = \sigma(3) = 2$.

Suppose $n > 3$ and that $\sigma(i) = \pi(i)$ for $i < n$.

Case 1, $\sigma(n) = m_{n-1}$: Since $\sigma^{-1}(i) = \pi^{-1}(i) < n$ if $i < m_{n-1}$, we only need to show that $\sigma(n) - n = m_{n-1} - n \neq \sigma(i) - i$ for any $i < n$. Now let $j < n$ be the position with the lowest value of $\sigma(j) - j$. The difference $\sigma(i) - i$ is negative if and only if $i > 1$ and $\sigma(i) = m_{i-1}$. E.g. $\sigma(3) = m_2$, so we must have $\sigma(j) = m_{j-1}$ when $n > 3$ and then $m_j = m_{j-1}$. By part *iii)* of Lemma 4.5,

$$m_y \leq m_x + (y - x) \tag{46}$$

for $y \geq x$, and using this we have

$$m_{n-1} - n \leq m_j + n - 1 - j - n < m_j - j = \sigma(j) - j.$$

This shows that $\sigma(n) = \pi(n)$ if $\sigma(n) = m_{n-1}$.

Case 2, $\sigma(n) = n + m_{n-1}$: We first show that $\sigma(n) - n \neq \sigma(i) - i$ for any $i < n$. Suppose $\sigma(i) - i = m_{n-1}$, then $\sigma(i) \neq m_{i-1}$, so $\sigma(i) = i + m_{i-1}$ and $m_i = m_{i-1} + 1$. But in this case $m_{n-1} = m_{i-1} = m_i - 1$ which implies $m_i > m_{n-1}$ and this is a contradicition.

We claim that $\pi(n) - n > 0$. Since $m_1 = 1$ and $m_i = m_{i-1}$ or $m_i = m_{i-1} + 1$ we have

$$\begin{aligned} n - m_{n-1} &= \left|\left\{2 \leq i < n \mid m_i = m_{i-1}\right\}\right| = \left|\left\{2 \leq i < n \mid \sigma(i) = m_{i-1}\right\}\right| \\ &= \left|\left\{i < n \mid \sigma(i) - i < 0,\ \right\}\right| = \left|\left\{i < n \mid \pi(i) - i < 0,\right\}\right| \\ &= \left|\left\{\pi(i) - i \mid \pi(i) - i < 0,\ i < n\right\}\right| \\ &= \left|\left\{\sigma(i) - i \mid \sigma(i) - i < 0,\ i < n\right\}\right|. \end{aligned}$$

If $i < n$ and $\sigma(i) - i < 0$ then by using inequality (46), we have

$$\sigma(i) - i = m_{i-1} - i = m_i - i > m_i - (i+1) \geq m_{n-1} - n.$$

Therefore

$$\left\{\sigma(i) - i \mid \sigma(i) - i < 0,\ i < n\right\} = \left(m_{n-1} - n, -1\right]. \tag{47}$$

By part *ii)* of Lemma 4.5, $\pi(n) \geq m_{n-1}$ and since $\sigma(n) = n + m_{n-1}$ we must have $\sigma^{-1}(m_{n-1}) < n$ which implies $\pi(n) > m_{n-1}$. So we have $\pi(n) - n > m_{n-1} - n$ which by (47) shows $\pi(n) - n \geq 0$. Since $\pi(1) - 1 = 0$ we must have $\pi(n) - n > 0$ and the claim is proved.

Suppose $n < x < n + m_{n-1}$. What remains in order to complete the proof, is to show that in this case $\pi(n) \neq x$. Since $0 < x - n < m_{n-1}$, there is an $i < n$ such that $x - n = m_{i-1}$. But then either $\sigma(i) = i + m_{i-1}$, which would imply $\pi(n) \neq x$, or $\sigma(i) = m_{i-1}$. In the latter case, $m_i = m_{i-1}$ and $\sigma(i+1) = i + 1 + m_i = i + 1 + m_{i-1}$, which also gives $\pi(n) \neq x$ unless $i + 1 = n$. But if $m_i = m_{i-1}$ and $x - n = m_{i-1} < m_{n-1}$, then $m_i < m_{n-1}$. So $i \neq n - 1$ and $i + 1 \neq n$.

We conclude that $\pi(n) \geq n + m_{n-1}$ and that $\pi(i) - i \neq m_{n-1}$ for any $i < n$. Hence $\pi(n) = n + m_{n-1}$. □

4.4 Another question

Proposition 4.7 *If f is a permutation of the natural numbers such that $f(x) - x \neq f(y) - y$ whenever $x \neq y$, then* $\limsup_{n\to\infty} \frac{f(n)}{n} \neq \liminf_{n\to\infty} \frac{f(n)}{n}$.

Proof. Let $n \in \mathbb{N}$ and define

$$A_n := \{\, i \mid f(i) - i \geq 0,\ i \in [1, n] \,\}$$

and

$$B_n := \{\, i \mid f(i) - i \leq 0,\ i \in [1, n] \,\}.$$

Then one can find a strictly increasing sequence of natural numbers (n_j), such that for each $j \in \mathbb{N}$ either $|A_{n_j}| \geq \frac{n_j}{2}$ or $|B_{n_j}| \geq \frac{n_j}{2}$.

Let 's say that the first case holds, i.e. for each $j \in \mathbb{N}$, $|A_{n_j}| \geq n_j/2$. Then for each $j \in \mathbb{N}$, we can find a least $i := i_j > j$ such that there is an $x_i \in [1, n_i] \setminus [1, n_j]$ with $f(x_i) - x_i \geq n_i/2$ (if there are several x_i with this property for a given i then take, say the least one).

There has to be such an i for each $j \in \mathbb{N}$, because there is only a finite number of elements in the interval $[1, n_j]$ (and therefore only a finite number of positive differences $f(y) - y > 0$ with $y \in [1, n_j]$).

Since $x_{i_j} \leq n_{i_j}$, for each i_j defined as above, we get

$$\limsup_{i\to\infty} \frac{f(x_i) - x_i}{x_i} \geq \frac{1}{2}.$$

The second case (for each $j \in \mathbb{N}$, $|B_{n_j}| \geq \frac{n_j}{2}$) is dealt with analogously. Hence we get either

$$\limsup_{i \to \infty} \frac{f(i)}{i} \geq \frac{3}{2}$$

or

$$\liminf_{i \to \infty} \frac{f(i)}{i} \leq \frac{1}{2}.$$

Since $\limsup_{i \to \infty} \frac{p(i)}{i} \geq 1$ and $\liminf_{i \to \infty} \frac{p(i)}{i} \leq 1$ for any permutation p of the natural numbers, the proposition follows. $\square$

Exercise 1: Is it true, that if $f : \mathbb{N} \to \mathbb{N}$ is as in Proposition 4.7 then, for any $n \in \mathbb{N}$, the size of the set B_n, (as defined in the above proof) is less than or equal to $n/2 + 1$?

Question 4.1 *Is there any permutation g on the natural numbers such that $g(x) - x \neq g(y) - y$ whenever $x \neq y$ and such that*

$$\limsup_{i \to \infty} \frac{g(i)}{i} - \liminf_{i \to \infty} \frac{g(i)}{i} < 1\ ?$$

5 Diagrams

n	$\mu(n)$	$\gamma(n)$	$G(n)$	$\nu(n)$	n	$\mu(n)$	$\gamma(n)$	$G(n)$	$\nu(n)$
1	1	1	1	1	33	8	8	12	13
2	2	2	2	2	34	8	8	12	13
3	2	2	2	2	35	8	8	12	13
4	2	2	3	3	36	8	8	12	14
5	3	3	4	4	37	8	8	13	14
6	4	4	4	4	38	8	8	14	14
7	4	4	4	4	39	8	8	14	14
8	4	4	4	4	40	8	8	15	15
9	4	4	4	5	41	8	8	16	16
10	4	4	5	5	42	8	8	16	16
11	4	4	6	6	43	8	8	16	16
12	4	4	6	6	44	8	8	16	16
13	4	4	7	7	45	8	8	16	16
14	4	4	8	8	46	8	8	16	16
15	4	4	8	8	47	8	8	16	16
16	4	4	8	8	48	8	8	16	16
17	5	5	8	8	49	8	8	16	16
18	6	6	8	8	50	8	8	16	16
19	6	6	8	8	51	8	8	16	17
20	6	6	8	9	52	8	8	16	17
21	6	7	8	9	53	8	8	16	17
22	6	8	8	9	54	8	8	16	18
23	6	8	8	9	55	8	8	16	18
24	6	8	8	10	56	8	8	16	18
25	6	8	8	10	57	8	8	16	18
26	6	8	8	11	58	8	8	16	19
27	7	8	8	11	59	8	8	16	19
28	7	8	9	11	60	8	8	16	19
29	7	8	10	11	61	8	8	16	19
30	7	8	10	12	62	8	8	16	19
31	7	8	11	12	63	8	8	16	20
32	7	8	12	13	64	8	8	16	20

n	$\mu(n)$	$\gamma(n)$	$G(n)$	$\nu(n)$	n	$\mu(n)$	$\gamma(n)$	$G(n)$	$\nu(n)$
65		9	16	20	97		16	24	26
66		10	16	20	98		16	24	26
67		10	16	20	99		16	24	26
68		10	16	20	100		16	24	27
69		11	16	20	101		16	24	27
70		12	16	20	102		16	24	27
71		12	16	21	103		16	24	27
72		12	16	21	104		16	24	28
73		12	16	21	105		16	24	28
74		12	16	22	106		16	24	28
75		12	16	22	107		16	24	28
76		12	16	22	108		16	24	28
77		12	16	22	109		16	25	28
78		12	16	22	110		16	26	28
79		12	16	22	111		16	26	29
80		12	16	22	112		16	27	29
81		13	16	22	113		16	28	29
82		14	17	23	114		16	28	30
83		14	18	23	115		16	28	30
84		14	18	24	116		16	28	30
85		15	19	24	117		16	28	30
86		16	20	24	118		16	29	30
87		16	20	24	119		16	30	30
88		16	20	24	120		16	30	30
89		16	20	24	121		16	31	31
90		16	20	24	122		16	32	32
91		16	21	24	123		16	32	32
92		16	22	25	124		16	32	32
93		16	22	25	125		16	32	32
94		16	23	25	126		16	32	32
95		16	24	26	127		16	32	32
96		16	24	26	128		16	32	32

Table 1: This table presents the values of $\mu(n), \gamma(n), G(n)$ and $\nu(n)$ for $n \leq 128$, with notation as in Definition 2.9, 2.3, 2.2. For an explanation of $\gamma(n)$, see page 20.

n	$\eta(n)$	$\frac{\eta(n)}{\nu(n)}$	n	$\eta(n)$	$\frac{\eta(n)}{\nu(n)}$
1	1	1.000	33	8	0.615
2	1	0.500	34	8	0.615
3	2	1.000	35	9	0.615
4	2	0.667	36	8	0.571
5	2	0.500	37	10	0.714
6	2	0.500	38	8	0.571
7	3	0.750	39	10	0.714
8	3	0.750	40	9	0.600
9	4	0.800	41	9	0.563
10	4	0.800	42	9	0.563
11	4	0.667	43	9	0.563
12	4	0.667	44	9	0.563
13	4	0.571	45	10	0.600
14	4	0.500	46	10	0.600
15	4	0.500	47	10	0.600
16	4	0.500	48	10	0.600
17	5	0.625	49	10	0.600
18	5	0.625	50	10	0.600
19	6	0.750	51	11	0.647
20	5	0.556	52	12	0.706
21	6	0.667	53	11	0.647
22	6	0.667	54	11	0.611
23	6	0.667	55	11	0.611
24	6	0.600	56	11	0.611
25	7	0.700	57	12	0.667
26	7	0.636	58	11	0.579
27	8	0.727	59	12	0.632
28	8	0.727	60	12	0.632
29	8	0.727	61	13	0.684
30	8	0.750	62	12	0.632
31	8	0.750	63	13	0.650
32	8	0.615	64	13	0.650

Table 2: This table presents the maximal size, $\eta(n)$, of a modulo n AP-avoiding subset of $[1, n]$, for $n \leq 64$. With the notation as in Definition 2.2, $\nu(n)$ denotes the maximal size of an AP-avoiding set.

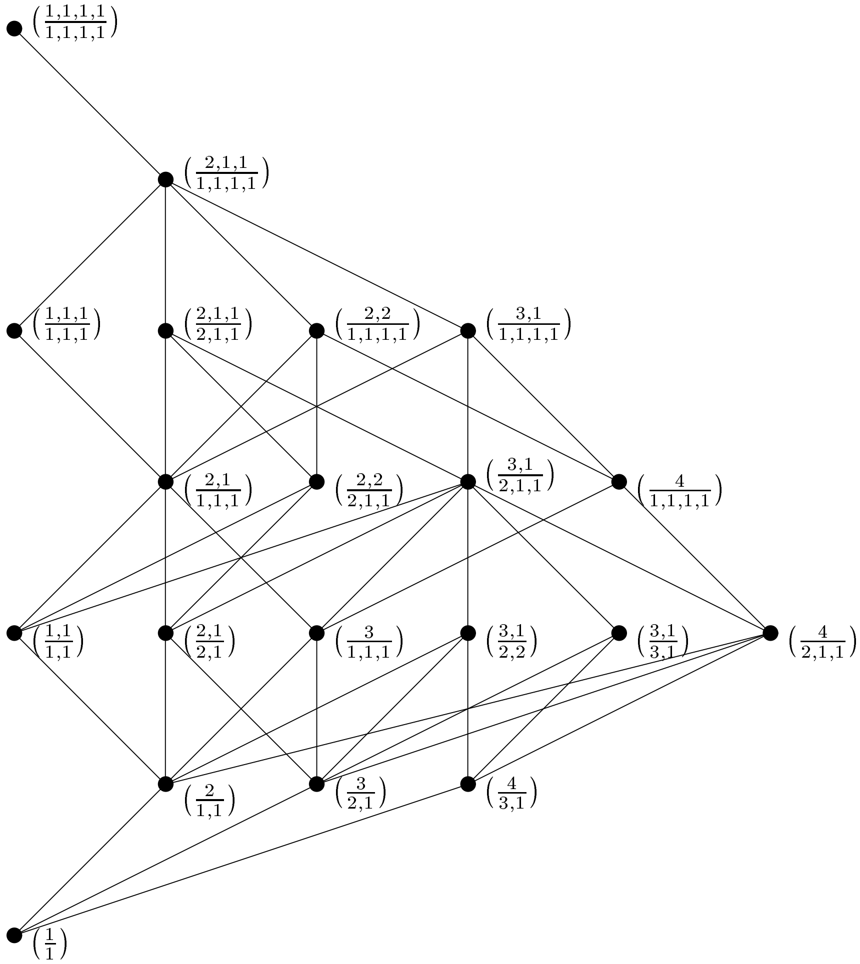

Figure 1: This diagram shows the poset of each family of sets avoiding a TAC of degree less than or equal to 4.

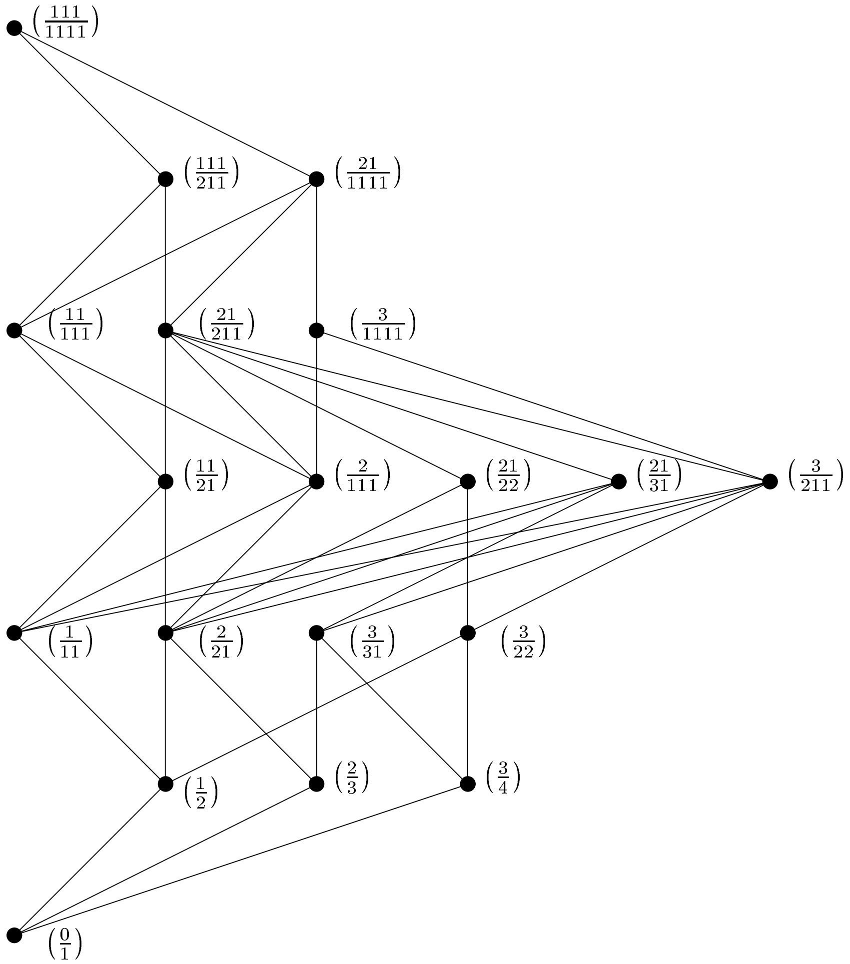

Figure 2: This diagram shows the poset of each family of sets avoiding a non-invariant LAC of degree less than or equal to 4.

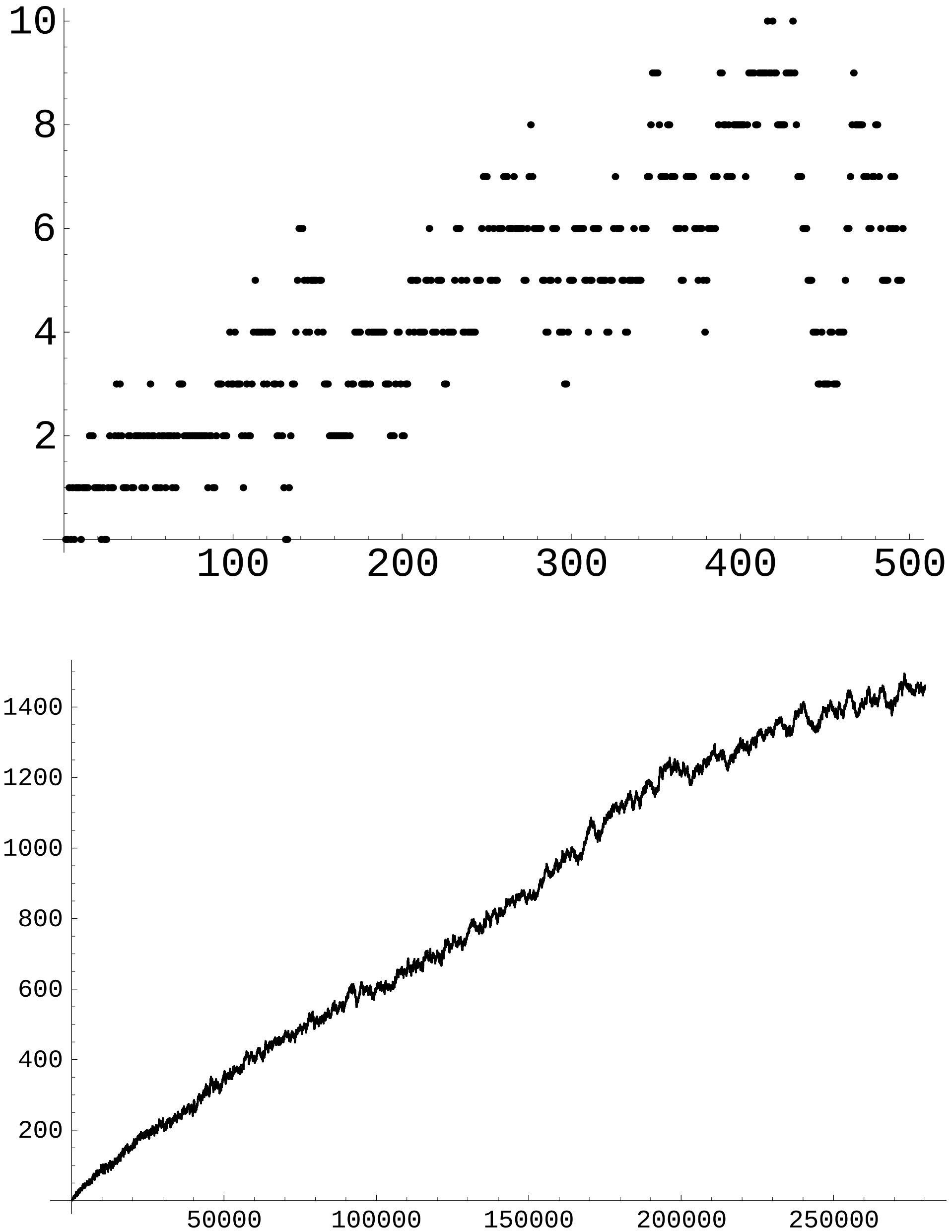

Figure 3: Let $r(n) = |\{\,\pi_g(i)\,|\,1 \leq i \leq n \quad and \quad \pi_g(i) > n\,\}|$. These two diagrams represent the values of $r(n)$ for $n \leq 500$ and 280000 respectively.

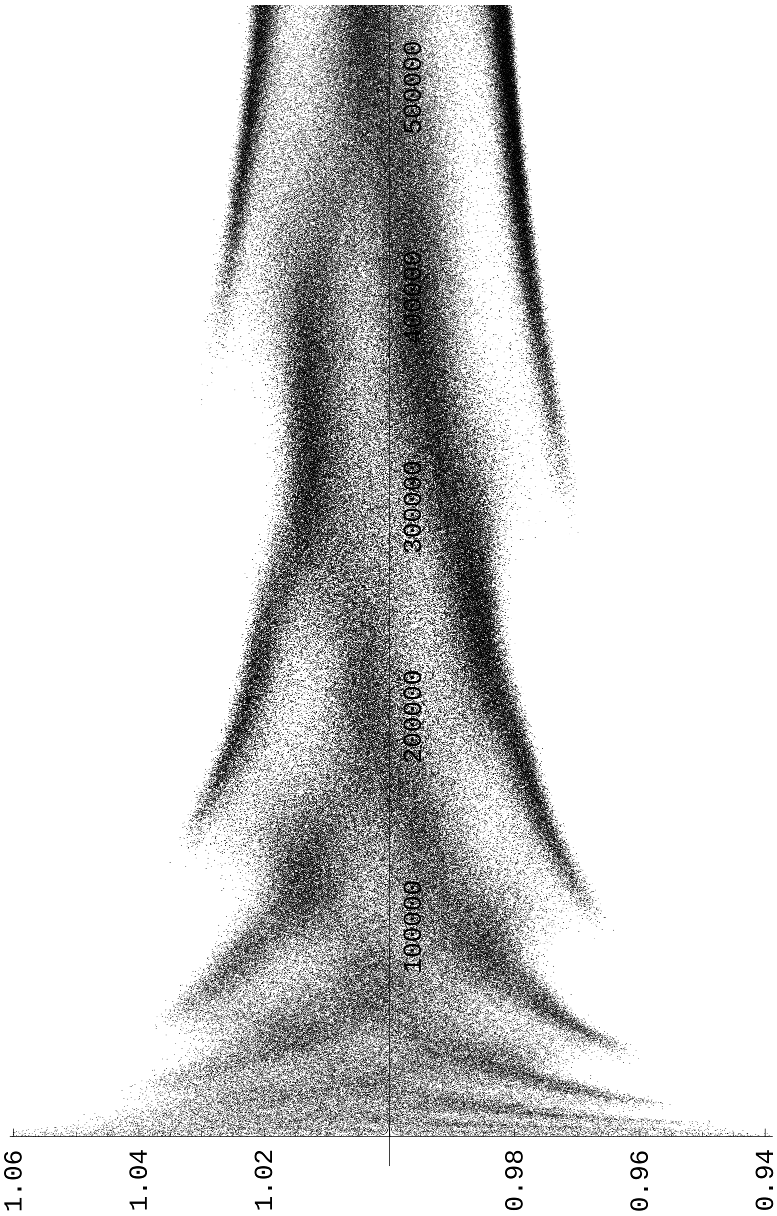

Figure 4: This plot represents the values of $\pi_g(n)/n$, for $1 \leq n \leq 560000$.

References

[1] M. Ajtai, J. Komlós and E. Szemerédi, *A dense infinite Sidon sequence*, European J. Combin. **2** (1981), no. 1, 1-11.

[2] S. Beatty, *Problem 3173*, Amer. Math. Monthly, **33** (1926).

[3] F. A. Behrend, *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. **23** (1946), 331-332.

[4] B. Bollobás, *Modern Graph Theory*, Springer Verlag, New York (1998)

[5] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), 968-984.

[6] C. J. Colbourn and J. H. Diniz *The CRC handbook of combinatorial designs*, CRC Press (1996).

[7] R. C. Bose and S. Chowla, *Theorems in the additive theory of numbers*, Comment. Math. Helv., **37** (1962-63), 141-147.

[8] S. Chowla, *Solution of a problem of Erdős and Turán in additive number theory*, Proc. Nat. Acad. Sci. India Sect. A **14** (1944), 1-2.

[9] P. Erdős and P. Turán, *On some sequences of integers*, J. London Math. Soc. **11** (1936), 261-264.

[10] P. Erdős and P. Turán, *On a problem of Sidon in additive number theory, and on some related problems*, J. London Math. Soc. **16** (1941), 212-215.

[11] P. Erdős, *On a problem of Sidon in additive number theory, and on some related problems. Addendum*, J. London Math. Soc. **19** (1944), 208.

[12] H. Furstenberg, *Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Analyse Math. **31** (1977), 204-256.

[13] S. W. Golomb, *Algebraic constructions for Costas arrays*, J. Combin. Theory A 37 (1984), 13-21.

[14] W. T. Gowers, *A new proof of Szemerédi's theorem*, Geom. Funct. Anal. **11** (2001), 465-588.

[15] W. T. Gowers, *A new proof of Szemerédi's theorem for arithmetic progressions of length 4*, Geom. Funct. Anal. **8** (1998), 529-551.

[16] P. Hegarty, J. Knape and U. Larsson, *On the equation $kx=y+z$,* Preprint, Göteborg, 2004. Postscript file available at http://www.math.chalmers.se/Math/Research/Preprints/

[17] P. Hegarty, *Permutations avoiding arithmetic progressions*, Department of Mathematics Göteborg University, June, 2003. Postscript file available at http://www.math.chalmers.se/Math/Research/Preprints/

[18] A. E. Ingham, *On the difference between consecutive primes*, Quarterly J. Math., Oxford series, **8** (1937), 255-256.

[19] F. Krückeberg, *B_2-Folgen und verwandte Zahlenfolgen*, J. Reine Angew. Math. **206** (1961), 53-60.

[20] C. Pomerance and A. Sarkőzy, *Chapter 20, Combinatorial Number Theory in Handbook of Combinatorics*, Elsevier Science B. V. (1995)

[21] K. F. Roth, *On certain sets of integers*, J. London Math. Soc. **28** (1953), 245-252.

[22] I. Z. Ruzsa *Solving a linear equation in a set of integers I*, Acta Arith. **65** (1993) 259-282.

[23] I. Z. Ruzsa *Solving a linear equation in a set of integers II*, Acta Arith. **72** (1995) 385-397

[24] I. Z. Ruzsa *A small maximal Sidon set*, Ramanujan J., **2** (1998), 55-58.

[25] I. Z. Ruzsa *An infinite Sidon sequence*, J. Number Theory **68** (1998), 63-71.

[26] R. Salem and D. C. Spencer, *On sets of integers which contain no three terms in arithmetic progression*, Proc. Nat. Acad. Sci. **28** (1942), 561-563.

[27] A. F. Sidorenko, *An infinite permutation without arithmetic progressions*, Discrete Math, **69** (1988), 211.

[28] J. Singer, *A theorem in finite projective geometry and some applications to number theory*, Trans. Amer. Math. Soc. **43** (1938), 377-385.

[29] A. Stöhr, *Gelöste und ungelöste Fragen über Basen der natürlichen Zahlenreihe, II*, J. Reine Angew. Math. **194** (1955), 111-140.

[30] E. Szemerédi, *On sets of integers containing no k elements in arithmetic progression*, Acta Arith. **27** (1975), 299-345.

[31] T. Tao, *Lecture notes 5 for 254A*, www.math.ucla.edu/~tao/254a.1.03w/notes5.dvi.

[32] How the proof of Baudet's conjecture was found, *Studies in Pure Mathematics presented to Richard Rado*, L.Mirsky, ed. Academic Press, London, 251-60.