Sidon sets

Graaf, S. de

S. de Graaf

Sidon sets

Bachelor thesis

July 21 2022

Thesis supervisor: dr. J.H. Evertse

Leiden University
Mathematical Institute

Contents

1 Introduction

Almost 100 years ago, the Hungarian mathematician Simon Sidon introduced a concept in his investigations of Fourier series. A concept that we now call Sidon sets/ sequences. We define them here as follows.

Definition 1.0.1. *A Sidon set is a subset $S \subset \mathbb{N} := \{0, 1, \dots\}$ such that the sums $a + b$ with $a, b \in S, a \leq b$ are pairwise distinct.*

In 1932 Sidon approached Erdős to ask him about the growth of these sets when they are infinite. To discuss the growth of these Sidon sets, we will first define a counting function. For a set A we denote by $|A|$ the cardinality of the set A.

Definition 1.0.2. *Let A be a sequence of positive integers, define*

$$A(x) = |\{a \in A : a \leq x\}|$$

At this time Sidon [2] had found a set satisfying definition 1.0.1 with $A(x) \gg x^{\frac{1}{4}}$ for all large x and later Erdős [4] found one with $A(x) \gg x^{\frac{1}{3}}$ for all x, and for almost 50 years this was the best known result in the study of infinite Sidon sets. Until in 1981, Atjai, Komlos and Szemeredi [1] proved the existence of an infinite Sidon set such that $A(x) > (x \log(x))^{\frac{1}{3}}$. Later Ruzsa [9] proved the existence of a Sidon set with $A(x) > x^{\sqrt{2}-1}$ as $x \to \infty$.
This has not been improved since. Rusza's proof was non-constructive. In this thesis we present two alternative, constructive methods, both due to Cilleruelo [2] giving the same lower bound as Rusza's.
Erdős (see [6], pp. 89/90 for a proof) proved that for every Sidon set A one has $A(x) \ll \sqrt{\frac{x}{\log(x)}}$ as $x \to \infty$. Note that this grows less quickly than $\sqrt{x}$.
However, for finite Sidon sets there are different results. First we will define $\Phi(n)$, the maximum cardinality of a Sidon set where the elements are bounded by n.

Definition 1.0.3. $\Phi(n) = max\{|S| : S \subset \{1, \dots, n\},\ S \text{ Sidon}\}$

We will present a proof of the fact that $\lim_{n\to\infty} \frac{\Phi(n)}{\sqrt{n}} = 1$. We will deduce this by combining a lower bound $\lim_{n\to\infty} \Phi(n) \geq (1 - o(1))\sqrt{n}$ as $n \to \infty$ using a combinatorial argument of Erdős and Turán [5] from 1941 and a upper bound $\lim_{n\to\infty} \Phi(n) \leq (1 + o(1))\sqrt{n}$ as $n \to \infty$ from a result of Singer [8] from 1938 on difference sets. A result that Erdős and Turán unfortunately appear to have missed, because they claimed the problem was still unsolved when writing the article where they gave the aforementioned upper bound needed to solve the problem in its

entirety, more than three years after Singer published his result.
In section 2 we will discuss finite Sidon sets, before moving on to the infinite cases in sections 3 and 4. In section 3 we will discuss the original construction that Cilleruelo gave in [2] and in section 4 we will work out in detail a second construction, which was only outlined by Cilleruelo in [2].

2 Finite Sidon sets

Recall the definition of a Sidon set:

Definition 2.0.1. *A Sidon set is a subset $S \subset \mathbb{N} = \{0, 1, \dots\}$ such that the sums $a + b$ with $a, b \in S$ and $a \leq b$ are pairwise distinct.*

We will prove the following theorem:

Theorem 2.1. *Let $\Phi(n)$ be the maximum cardinality of a Sidon subset of $\{1, \dots, n\}$. Then*

$$\lim_{n\to\infty} \frac{\Phi(n)}{\sqrt{n}} = 1. \tag{2.0.1}$$

This theorem is a consequence of the following statements:

Theorem 2.2. *For every $\epsilon > 0$ there is an $n_0(\epsilon)$ such that for every $n \geq n_0(\epsilon)$, there exists a Sidon subset of $\{1, ..., n\}$ with cardinality $\geq (1 - \epsilon)\sqrt{n}$.*

Theorem 2.3. *For every $\epsilon > 0$ there is an $n_0(\epsilon)$ such that for every $n \geq n_0(\epsilon)$, every Sidon subset of $\{1, ..., n\}$ has cardinality $\leq (1 + \epsilon)\sqrt{n}$.*

Theorem 2.3 is a result of Erdős and Turán [5]. We will prove these two theorems in the following sections and combine them to form the proof of Theorem 2.1.

2.1 Proof of Theorem 2.2

This proof is based on the following result of Singer [8].

Theorem 2.4. *For every prime power t there is a subset $\{d_0, ..., d_t\}$ of $t + 1$ elements in $\{0, ..., t^2 + t\}$ such that the differences $d_i - d_j$ with $0 \leq i, j \leq t, i \neq j$ are pairwise distinct.*

Then we will obtain Theorem 2.2 by combining this with the Prime Number Theorem.

2.1.1 Proof of Singer's theorem

This proof will be based on collineations on finite projective planes. We will first give some definitions and intuition regarding these objects.

Definition 2.4.1. *A finite projective plane is a finite set $\mathbb{P}$, together with a collection of subsets of $\mathbb{P}$ of cardinality at least 2, that form the lines of $\mathbb{P}$ such that they satisfy the following properties:*

1. *Every pair of points in $\mathbb{P}$ is contained in exactly one line.*

2. *Every pair of lines in $\mathbb{P}$ has exactly one common point.*

3. *There exist four points in $\mathbb{P}$ no three of which are collinear.*

Lemma 2.4.1. *Every line of a projective plane $\mathbb{P}$ contains the same number of points.*

Proof. Let l, m be distinct lines in a projective plane. There exists a point $P \notin l, m$. Suppose this point does not exist. Then all points of $\mathbb{P}$ lie on l or m. There exist $L_1, L_2 \in l$ and $M_1, M_2 \in m$. Take the intersection of the line through L_1, M_1 and the line through L_2, M_2, this point cannot lie on l or m because then these lines would have two common points. Hence P is a point that is neither on l nor m. For any point $Q \in l$, the line through P, Q exists and intersects m at some point Q'. We define a map $f : l \to m$ by mapping $Q \in l$ to the intersection point Q' of m with the line through P and Q. This is a bijection from l to m, where f^{-1} maps $Q' \in m$ to the intersection of l with the line through P and Q'. It follows that l and m have the same number of points. □

Definition 2.4.2. *Let $\mathbb{P}$ be a finite projective plane. The order of a projective plane is equal to t if the number of points on a line in that plane is equal to $t + 1$.*

Lemma 2.4.2. *If a projective plane $\mathbb{P}$ has order t, then every point of $\mathbb{P}$ lies on precisely $t + 1$ lines of $\mathbb{P}$.*

Proof. Let P be a point in $\mathbb{P}$ and let l be a line of $\mathbb{P}$ that does not contain P. For all $t + 1$ points of l there must be a line, through P and itself. Define a map from l to the collection of lines through P, by mapping $Q \in l$ to the line through P and Q. This is a bijection, since any line through P has precisely one common point with l. □

Corollary 2.4.1. *A projective plane of* $\mathbb{P}$ *order* t *has precisely* t^2+t+1 *points.*

Proof. Let P be a point in $\mathbb{P}$. There are precisely $t+1$ lines through P, which all have t points on them distinct from P. This gives a total of t^2+t+1 points. No other point can exist that does not lie on one of these lines because there always exists a line through that point and P. □

Let t be a prime power. Then a projective plane of order t can be obtained as follows. Take a three-dimensional vector space V over $\mathbb{F}_t$. Then the points of the projective plane are the one-dimensional linear subspaces of V and the lines of the plane are obtained by taking for each two-dimensional linear subspace W of V the collection of one-dimensional linear subspaces of W.
Clearly we have that every pair of lines is contained in exactly one two-dimensional subspace, every two two-dimensional subspaces intersect in one line. Let $[x]$ represent the point in a projective plane given by a one-dimensional linear subspace x of V, and let a, b, c be a base of V. Any three of the vectors $a, b, c, a+b+c$ span V. By assumption a, b, c spans V, and any combination consisting of $a+b+c$ and any two other vectors from $\{a, b, c\}$ also spans V since the third vector from $\{a, b, c\}$ is obtained by subtracting the other two from $a+b+c$. Thus, no more than two of the one-dimensional subspaces spanned by $a, b, c, a+b+c$ can lie in one two-dimensional subspace of V. Hence, no three of the points in the projective space $[a], [b], [c], [a+b+c]$ are collinear.

We now define the projective plane $\mathbb{P}^2(\mathbb{F}_t)$ by taking for V the finite field $\mathbb{F}_{t^3}$. Recall that $\mathbb{F}_{t^3}$ is indeed a three-dimensional vector space over $\mathbb{F}_t$. Further we may view $\mathbb{F}_t$ as a subfield of $\mathbb{F}_{t^3}$ via

$$\mathbb{F}_t = \{\xi \in \mathbb{F}_{t^3} : \xi^t = \xi\}$$

The points of $\mathbb{P}^2(\mathbb{F}_t)$ are the one-dimensional $\mathbb{F}_t$-linear subspaces of $\mathbb{F}_{t^3}$, i.e. $[\alpha] = \{\xi\alpha : \xi \in \mathbb{F}_t\}$ for $\alpha \in \mathbb{F}_{t^3}^*$.
Clearly $[\alpha] = [\beta] \iff \beta = \xi\alpha$ for some $\xi \in \mathbb{F}_t^*$. Thus there is a one-to-one correspondence between $\mathbb{P}^2(\mathbb{F}_t)$ and $\mathbb{F}_{t^3}^*/\mathbb{F}_t^*$.
Recall that the multiplicative group of $\mathbb{F}_{t^3}^*$ is cyclic of order t^3-1. Let λ be a generator of $\mathbb{F}_{t^3}^*$. In other words $\mathbb{F}_{t^3}^* = \langle\lambda\rangle$.
We have

$$\begin{aligned} \mathbb{F}_t^* &= \{\lambda^i \in \mathbb{F}_{t^3} : (\lambda^i)^t = \lambda^i\} \\ &= \{\lambda^i : i(t-1) \equiv 0 \pmod{t^3-1}\} = \{\lambda^i : i \equiv 0 \pmod{q}\}, \end{aligned} \tag{2.1.1}$$

where $q = \frac{t^3-1}{t-1} = t^2 + t + 1$.

Two points $[\lambda^i], [\lambda^j]$ of $\mathbb{P}^2(\mathbb{F}_t)$ are equal if and only if $\lambda^{i-j} \in \mathbb{F}_t^*$ which is equivalent to $i \equiv j \pmod{q}$.

Definition 2.4.3. *A collineation of $\mathbb{P}^2(\mathbb{F}_t)$ is a transformation C of $\mathbb{P}^2(\mathbb{F}_t)$ with the property that it maps lines of $\mathbb{P}^2(\mathbb{F}_t)$ to lines of $\mathbb{P}^2(\mathbb{F}_t)$.*

The order of a collineation is defined to be the smallest number k such that $C^k = Id_{\mathbb{P}^2(\mathbb{F}_t)}$.

Lemma 2.4.3. *$\mathbb{P}^2(\mathbb{F}_t)$ has a collineation of order $t^2 + t + 1 = q$*

Proof. Consider the map $\phi_\lambda : \mathbb{F}_{t^3} \to \mathbb{F}_{t^3}$ given by $x \mapsto \lambda x$. This is an $\mathbb{F}_t$-linear transformation, so it induces a collineation $[\phi_\lambda]$ of $\mathbb{P}^2(\mathbb{F}_t)$, where

$$[\phi_\lambda] : \mathbb{P}^2(\mathbb{F}_t) \to \mathbb{P}^2(\mathbb{F}_t) : [x] \mapsto [x\lambda]. \tag{2.1.2}$$

Note that $[\phi_\lambda]^i$ maps $[x] \in \mathbb{P}^2(\mathbb{F}_t)$ to $[\lambda^i x]$. By (2.1.1) we have that $[\phi_\lambda]^i = Id_{\mathbb{P}^2(\mathbb{F}_t)} \iff \lambda^i \in \mathbb{F}_t^* \iff i \equiv 0 \pmod{q}$. Hence $[\phi_\lambda]$ has order q. □

2.1.2 Lines in the projective plane

Let $d_0 := 0, d_1 := 1$, and let l_0 be the line through $[\lambda^{d_0}] = [1]$ and $[\lambda^{d_1}] = [\lambda]$. Suppose that

$$l_0 = \{[\lambda^{d_0}], [\lambda^{d_1}], ..., [\lambda^{d_t}]\}. \tag{2.1.3}$$

Consider sets

$$l_i = \{[\lambda^{d_0+i}], ..., [\lambda^{d_t+i}]\} \text{ with } i \in \{0, ..., q-1\}. \tag{2.1.4}$$

Remark. *The set l_i with $i \in \{0, \ldots, q-1\}$ is a line of $\mathbb{P}^2(\mathbb{F}_t)$, since*

$$l_i = [\phi_\lambda]^i(l_0) \text{ for } i \in \{0, \ldots, q-1\}.$$

Define

$$l_i^* = \{d_0 + i, ..., d_t + i\} \subset \mathbb{Z}/q\mathbb{Z} \tag{2.1.5}$$

where for convenience we have written a for the residue class $a \pmod{q}$. These sets of exponents can uniquely be identified with the lines and will be used to represent the lines from now on.

Lemma 2.4.4. *The sets $l_0^*, ..., l_{q-1}^*$ have the following properties:*

1. $l_0^* \neq l_1^*$.

2. $l_0^* = \{d_0, ..., d_t\}$ *does not contain a pair of consecutive residue classes modulo* q *other than* $\{0, 1\}$.

3. $l_i^* \neq l_j^*$ *for* $i, j \in \{0, ..., q-1\}, i \neq j$.

Proof. We prove the three claims.

1. Suppose $l_0^* = l_1^*$. Since $d_1 = 1 \in l_0^*$, we have $2 \in l_1^*$, hence $2 \in l_0^*$, hence $3 \in l_1^*$, hence $3 \in l_0^*$ etc. We find that l_0^* contains all residue classes modulo $q = t^2 + t + 1$, but this contradicts the fact that l_0^* has cardinality $t + 1$ and $t > 0$.

2. Suppose l_0^* contains apart from $0, 1$, the consecutive residue classes $d_n, d_n + 1$. Then we know that l_1^* contains 1 and $d_n + 1$ by construction. This means that l_0^* and l_1^* would share more than one common element and this is only possible if $l_0^* = l_1^*$ contradicting claim 1.

3. Suppose $l_i^* = l_j^*$. Then there are $u, v \in 0, ..., t$ such that $d_0 + i \equiv d_u + j$ and $d_1 + i \equiv d_v + j \pmod{q}$. This implies $d_u - d_v \equiv d_1 - d_0 \pmod{q}$. We find that $d_v = d_u + 1 \pmod{q}$ contradicting claim 2. We conclude that $l_0^*, \ldots, l_{q-1}^*$ must be distinct.

□

Proposition 2.4.1. *Singer(1938) If* t *is a prime power, there exist* $t + 1$ *integers* $d_0, ..., d_t \in \{0, ..., t^2 + t\}$ *such that their* $t^2 + t$ *differences* $d_i - d_j$ *with* $0 \leq i, j \leq t, i \neq j$ *are pairwise distinct.*

Proof. Suppose $d_i - d_j = d_k - d_l$, then $d_j, d_l \in l_0^* \cap l_r^*$ □

Corollary 2.4.2. *Let* t *be a prime power. Then* $\{0, ..., t^2 + t\}$ *has a Sidon subset of* $t + 1$ *elements.*

2.1.3 Prime Number Theorem

We will now derive Theorem 2.2 by applying the Prime Number Theorem [3]. The theorem was first proved independently by Hadamard and de la Vallée Poussin in 1896.

Theorem 2.5. *(Prime Number Theorem) [Hadamard, de la Vallée Poussin (1896)] For $x \in \mathbb{R}_{\geq 0}$ denote by $\pi(x)$ the number of primes $\leq x$. Then we have:*

$$\lim_{x\to\infty} \frac{\pi(x)\log(x)}{x} = 1. \tag{2.1.6}$$

We will not prove this theorem here. The known proofs of this theorem are very extensive and require a lot of prior knowledge.

Lemma 2.5.1. *For every $\epsilon > 0$ there exists a $x_0(\epsilon)$ such that for every $x \geq x_0(\epsilon)$ there is a prime number in $[x, (1+\epsilon)x]$.*

Proof. Let $\delta > 0$. Then there exists a $x_1(\delta)$ such that $1 - \delta \leq \frac{\pi(x)\log(x)}{x} \leq 1 + \delta$ for all $x \geq x_1(\delta)$. Dividing gives us $(1-\delta)\frac{\log(x)}{x} \leq \pi(x) \leq (1+\delta)\frac{\log(x)}{x}$ for all $x \geq x_1(\delta)$.
Let $x \geq x_1(\delta)$.
Then we have $\pi((1+\epsilon)x) - \pi(x) \geq \frac{(1-\delta)(1+\epsilon)x}{\log((1+\epsilon)x)} - \frac{(1+\delta)x}{\log(x)}$.
There exists $x_2(\epsilon)$ such that $x^{1+\frac{\epsilon}{2}} > (1+\epsilon)x$ for all $x \geq x_2(\epsilon)$. For $x \geq x_2(\epsilon)$ we have

$$\pi((1+\epsilon)x) - \pi(x) \geq \frac{(1-\delta)(1+\epsilon)x}{(1+\frac{\epsilon}{2})\log(x)} - \frac{(1+\delta)x}{\log(x)}. \tag{2.1.7}$$

Choose $\delta > 0$ such that $\frac{(1-\delta)(1+\epsilon)}{1+\frac{\epsilon}{2}} > (1+\delta)$. Now for all $x > \max\{x_1(\delta), x_2(\epsilon)\}$ we have

$$\pi((1+\epsilon)x) - \pi(x) > 0. \tag{2.1.8}$$

□

Let $(p_i)_{i\in\mathbb{N}}$ be the sequence of prime numbers. Suppose $n \in \mathbb{N}, n \geq 6$. Then there is an index $i(n)$ such that $p_{i(n)}^2 + p_{i(n)} \leq n < p_{i(n)+1}^2 + p_{i(n)+1}$.

Note that the subset $\{0, ..., p_{i(n)}^2 + p_{i(n)}\} \subset \mathbb{N}$ contains a Sidon subset of $p_{i(n)} + 1$ elements. Hence,

$$\Phi(n) \geq \Phi(p_{i(n)}^2 + p_{i(n)}) \geq p_{i(n)} + 1 \geq \sqrt{p_{i(n)}^2 + p_{i(n)}} \geq \frac{\sqrt{p_{i(n)}^2 + p_{i(n)}}}{\sqrt{p_{i(n)+1}^2 + p_{i(n)+1}}}\sqrt{n} \tag{2.1.9}$$

By Lemma 2.5.1 there exists a $n_0(\epsilon)$ such that $\frac{\sqrt{p_{i(n)}^2+p_{i(n)}}}{\sqrt{p_{i(n)+1}^2+p_{i(n)+1}}} > (1-\epsilon)$ for $x \geq n_0(\epsilon)$.
So

$$\Phi(n) \geq (1-\epsilon)\sqrt{n} \text{ for } n \geq n_0(\epsilon). \tag{2.1.10}$$

This proves Theorem 2.2.

2.2 Proof of Theorem 2.3

We follow the proof of Erdős and Turán [5].
Let $\Phi(n)$ be the maximum cardinality of a Sidon subset of $\{1, ..., n\}$.
We have to show that for all $\epsilon > 0$ there is an $n_0(\epsilon)$ such that

$$\Phi(n) \leq (1+\epsilon)\sqrt{n} \text{ for all } n > n_0(\epsilon). \tag{2.2.1}$$

Note that this directly proves Proposition 2.3. This proof will be significantly shorter than the proof of Proposition 2.2 and will only require a combinatorial argument, hence no new definitions are needed.

Let $S = \{a_1 < ... < a_x\} \subset \{0, ..., n\}$ be such that the sums $a_i + a_j (1 \leq i \leq j \leq x)$ are all different. Let m be a positive integer such that $m < n$, and consider the intervals

$$[-m+1, 1), [-m+2, 2), ..., [n, m+n). \tag{2.2.2}$$

Let A_u denote the number of elements $a_i \in S$ in the interval $[-m+u, u)$. Since each a_i occurs in exactly m intervals, i.e., $[-m+u, u)$ for $u = a_i + 1, \ldots, a_i + m$ we have

$$\sum_{u=1}^{m+n} A_u = mx. \tag{2.2.3}$$

The number of pairs a_i, a_j in $[-m+u, u)$ with $i < j$ is $\frac{1}{2}A_u(A_u - 1)$. Using the fact that $f(x) = x(x-1)$ is a concave function, we get that the total number of triples (a_i, a_j, u) with $a_i, a_j \in [-m+u, u) \cap S$, $1 \leq i < j \leq x$, $1 \leq u \leq m+n$ is

$$\sum_{u=1}^{m+n} \frac{1}{2}A_u(A_u - 1) \geq \frac{1}{2}(m+n)(\frac{mx}{m+n})(\frac{mx}{m+n} - 1). \tag{2.2.4}$$

A pair (a_i, a_j) with $a_i - a_j = r$ occurs in precisely $m-r$ intervals $[-m+u, u)$. Hence the total number of triples (a_i, a_j, u) with $a_i, a_j \in [-m+u, u) \cap S, 1 \leq i < j \leq x$, $1 \leq u \leq m+n$ is at most

$$\sum_{u=1}^{m-1} (m-r) = \frac{1}{2}m(m-1). \tag{2.2.5}$$

When we compare (2.2.4) with (2.2.5) we find that

$$\frac{1}{2}(m+n)(\frac{mx}{m+n})(\frac{mx}{m+n} - 1) \leq \frac{1}{2}m(m-1).$$

This can be rewritten as

$$\begin{gathered} mx(mx-m-n) \leq m(m-1)(m+n), \\ \text{or as } x(mx-2n) < m(m+n), \\ \text{and this implies} \\ x < \frac{n}{m} + \sqrt{n+m+\frac{n^2}{m^2}}. \end{gathered} \tag{2.2.6}$$

We can rewrite the square root as

$$\sqrt{n+m+\frac{n^2}{m^2}} = \sqrt{n}\sqrt{1+\frac{m}{n}+\frac{n}{m^2}} = \sqrt{n}(1+O(\frac{m}{n}+\frac{n}{m^2})). \tag{2.2.7}$$

So $x < \frac{n}{m} + \sqrt{n} + O(\frac{m}{\sqrt{n}} + \frac{n^{\frac{3}{2}}}{m})$. Choose $m = [n^p]$, with $p \in (0,1)$. Then,

$$\begin{aligned} x &< \frac{n}{m} + \sqrt{n} + O(\frac{m}{\sqrt{n}} + \frac{n^{\frac{3}{2}}}{m}) \\ &= n^{1-p} + \sqrt{n} + O(n^{p-\frac{1}{2}} + n^{\frac{3}{2}-2p}) \\ &= \sqrt{n} + O(n^{1-p} + n^{p-\frac{1}{2}} + n^{\frac{3}{2}-2p}). \end{aligned} \tag{2.2.8}$$

The error term is asymptotically minimal for $p = \frac{3}{4}$, and with this value of p we get,

$$x < \sqrt{n} + O(n^{\frac{3}{4}}) \tag{2.2.9}$$

This inequality proves Theorem 2.3. Now by Theorem 2.2 and Theorem 2.3 we have Theorem 2.1.

3 Infinite Sidon sets

In this section we will consider infinite Sidon sets. We will study two results from Cilleruelo [2], the second of which resembles a result from Ruzsa [9]. Ruzsa proved that an infinite Sidon set A exists such that $A(x) = x^{\sqrt{2}-1+o(1)}$ as $x \to \infty$. His proof, however, was not constructive. Cilleruelo's second result will give a constructive version of Rusza's result. Cilleruelo's first result will be a weaker version of the second result, but with a shorter proof.
We will construct a set $A_{\bar{q},c} = (a_p)_{p\in\mathcal{P}}$, indexed by the set of prime numbers $\mathcal{P}$, depending on a real number c and a base $\bar{q}$ (to be defined later), which is such that $A_{\bar{q},c}(x) = x^{c+o(1)}$ as $x \to \infty$. By taking the infimum of all c such that $A_{\bar{q},c}$ is a Sidon set, i.e., $c = \frac{\sqrt{5}-3}{2}$, we obtain Cilleruelo's first result.

Theorem 3.1. *There is a Sidon set $A \subset \mathbb{N}$ such that*

$$A(x) = x^{\frac{3-\sqrt{5}}{2}+o(1)} \text{ as } x \to \infty$$

.

This theorem is very easy to prove once we have the construction of the sequence $A_{\bar{q},c}$. As mentioned before, this theorem is weaker than Cilleruelo's second theorem that we state below. To derive this theorem, we take the sequence $A_{\bar{q},c}$ as above but with $c = \sqrt{2} - 1$. For this value of c, the set $A_{\bar{q},c}$ itself is not a Sidon set, but as it will turn out, we can construct a Sidon set by taking a suitable subset of $A_{\bar{q},c}$ with about the same density.

Theorem 3.2. *There is a Sidon set $A' \subset \mathbb{N}$ such that*

$$A'(x) = x^{\sqrt{2}-1+o(1)} \text{ as } x \to \infty$$

.

3.1 Construction

We first consider the following fact, which will be used throughout the construction.

Lemma 3.2.1. *Given an infinite sequence $\bar{b} = (b_j)_{j=1}^{\infty}$, with $b_i \in \mathbb{N}_{>1}$ (the base), every positive integer n can be uniquely written in the form*

$$n = x_1 + x_2 b_1 + x_3 b_1 b_2 + \cdots + x_k b_1 \dots b_{k-1} \tag{3.1.1}$$

with digits $x_j \in \{0, \dots b_j - 1\}, 1 \le j \le k$, where k is the smallest integer such that $n < b_1 \dots b_k$.

If n is given by (3.1.1), then we represent it as $n = (x_k \dots x_1)_{\bar{b}}$.

Proof. Suppose $n < b_1 \dots b_k$, then we can write $n = xb_1 \dots b_{k-1} + y$ with $x, y \in \mathbb{Z}$ with $0 \leq y < b_1 \dots b_{k-1}$ and $0 \leq x < b_k$. We proceed inductively on k, where k is the smallest integer such that $n < b_1 \dots b_k$. First let $k = 1$. Because n is smaller than b_1, we can write $n = 0 \cdot b_1 + y$, with $0 < y < b_1$.
Suppose $k \geq 2$. Let $b_1 \dots b_{k-1} \leq n < b_1 \dots b_k$. Then $n = x_{k-1}b_1 \dots b_{k-1} + y$ with $0 < x_{k-1} \leq b_k$ and $0 \leq y < b_1 \dots b_{k-1}$. By the induction hypothesis, we can write $y = x_1 + x_2b_1 + \dots + x_{k-1}b_1 \dots b_{k-2}$. Hence $n = x_1 + x_2b_1 + \dots + x_{k-1}b_1 \dots b_{k-2} + x_kb_1 \dots b_{k-1}$. □

We consider the base $\bar{q} = (4q_j)_{j=1}^{\infty}$ where $q_1, q_2, ...$ is a given infinite sequence of prime numbers such that

$$2^{2j-1} < q_j \leq 2^{2j+1} \tag{3.1.2}$$

for all $j \geq 1$. This is possible because of Bertrand's postulate [7]. Now choose for each j a primitive root g_j of $\mathbb{F}_{q_j}^*$ (mod q_j).
Let us fix c such that $0 < c < \frac{1}{2}$ and consider the partition of the set of prime numbers,

$$\mathcal{P} = \bigcup_{k \geq 3} \mathcal{P}_k, \text{ where } \mathcal{P}_k = \{p \text{ prime } : 2^{c(k-1)^2-3} < p \leq 2^{ck^2-3}\}$$

Given $p \in \mathcal{P}$, we define a_p as follows. Choose k such that $p \in \mathcal{P}_k$ and then define

$$a_p = (x_k(p) \dots x_1(p))_{\bar{q}} \tag{3.1.3}$$

where $x_j(p)$ is the unique integer solution of

$$g_j^{x_j} \equiv p \pmod{q_j},\ q_j + 1 \leq x_j \leq 2q_j - 1. \tag{3.1.4}$$

We define $x_j(p) = 0$ when $j > k$. Define $A_{\bar{q},c} := (a_p)_{p \in \mathcal{P}}$

3.2 Properties of $A_{\bar{q},c}$

Here we will prove some properties of the sequence $A_{\bar{q},c}$ that will eventually lead to a value of c for which $A_{\bar{q},c}$ is a Sidon set.

Proposition 3.2.1. *The terms a_p of $A_{\bar{q},c}$ are pairwise distinct.*

Proof. Suppose $a_p = a_{p'}$ for some $p, p' \in \mathcal{P}$ with $p \neq p'$. Then all digits of a_p and $a_{p'}$ must be equal. That is, $x_j(p) = x_j(p')$ for all $j \geq 1$ and by construction we have that $p, p' \in \mathcal{P}_k$ such that k is the largest j such that the digit $x_j \neq 0$. We also know that

$$p \equiv g_j^{x_j(p)} \equiv g_j^{x_j(p')} \equiv p' \pmod{q_j} \text{ for } j \leq k.$$

From this we can conclude that $p \equiv p' \pmod{q_1 ... q_k}$. Suppose that $p \neq p'$. Using the fact that $p, p' \in \mathcal{P}_k$ and $q_j > 2^{2j-1}$ we find that

$$2^{ck^2} \geq |p - p'| \geq q_1 \dots q_k > 2^{1+\dots+(2k-1)} = 2^{k^2} \tag{3.2.1}$$

which is not possible because we assumed $0 < c < \frac{1}{2}$. Hence, the elements a_p of $A_{\bar{q},c}$ are pairwise distinct. □

The next proposition concerns the growth of the sequence $A_{\bar{q},c}$.

Proposition 3.2.2. *We have $A_{\bar{q},c}(x) = x^{c+o(1)}$ as $x \to \infty$.*

Proof. Let $x \in \mathbb{R}$ and consider the integer k such that

$$4q_1 \dots 4q_k < x \leq 4q_1 \dots 4q_{k+1}$$

Using (3.1.2) we find that $2^{k^2+2k} < x \leq 2^{(k+1)^2+2(k+1)}$. Hence $x = 2^{k^2(1+O(\frac{1}{k}))}$ and thus $2^{k^2} = x^{1+o(1)}$.
If $p \leq 2^{ck^2-3}$ then $p \in \mathcal{P}_l$ for some $l \leq k$ so

$$a_p = x_1(p) + \sum_{j \leq l} x_j(p)(4q_1) \dots (4q_{j-1}) \leq (4q_1) \dots (4q_k) \leq x.$$

Thus, applying the Prime Number Theorem, we find the lower bound

$$A_{\bar{q},c}(x) \geq \pi(2^{ck^2-3}) = 2^{ck^2(1+o(1))} = x^{c+o(1)}.$$

For the upper bound, we notice that if $p \in \mathcal{P}_{k+2}$ then $a_p \geq (4q_1) \dots (4q_{k+1}) \geq x$. Thus,

$$A_{\bar{q},c}(x) \leq \pi(2^{c(k+1)^2-3}) = 2^{ck^2(1+o(1))} = x^{c+o(1)}.$$

□

The following proposition is about some of the properties of repeated sums of terms of the sequence $A_{\bar{q},c}$.

Proposition 3.2.3. *Suppose that there exist* $a_{p_1}, a_{p_2}, a_{p_1'}, a_{p_2'} \in A_{\bar{q},c}$ *such that* $a_{p_1} > a_{p_1'} \geq a_{p_2'} > a_{p_2}$ *and*

$$a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}.$$

Then we have:

1. *There exist* k_1, k_2 *with* $k_2 \leq k_1$ *such that* $p_1, p_1' \in \mathcal{P}_{k_1}$ *and* $p_2, p_2' \in \mathcal{P}_{k_2}$,
2. $p_1 p_2 \equiv p_1' p_2' \pmod{q_1 \dots q_{k_2}}$,
3. $p_1 \equiv p_1' \pmod{q_{k_2+1} \dots q_{k_1}}$ *if* $k_2 < k_1$,
4. $(1-c)k_1^2 < k_2^2 < \frac{c}{1-c} k_1^2$.

Proof. We prove the four claims

1. Since $0 \leq x_j(p_1) + x_j(p_2) < 4q_j$ and $0 \leq x_j(p_1') + x_j(p_2') < 4q_j$ for all j, and $a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}$, we infer that the digits of the sums are equal, that is:

$$x_j(p_1) + x_j(p_2) = x_j(p_1') + x_j(p_2') \text{ for all } j. \tag{3.2.2}$$

   By construction, $p_1 \in \mathcal{P}_{k_1}$, where k_1 is the largest j such that

$$x_j(p_1) + x_j(p_2) \geq q_j + 1.$$

   This is because taking $j > k_1$ gives $x_j(p_1) + x_j(p_2) = 0$ and $j \leq k_1$ gives $x_j(p_1) + x_j(p_2) \geq x_j(p_1) \geq q_1 + 1$.
   We also have that $p_2 \in \mathcal{P}_{k_2}$ where k_2 is the largest j such that

$$x_j(p_1) + x_j(p_2) \geq 2q_j + 2.$$

   This is possible because taking $k_2 < j \leq k_1$ gives $x_j(p_1) + x_j(p_2) = x_j(p) \leq 2q_j - 1$ and $j < k_2$ gives that both $x_j(p_1), x_j(p_2) \geq q_j + 1$ hence $x_j(p_1) + x_j(p_2) \geq 2q_j + 2$.
   This proves claim 1.

2. To prove claim 2 we observe that (3.2.2) implies that for all j the following congruence holds:

$$g_j^{x_j(p_1)+x_j(p_2)} \equiv g_j^{x_j(p_1')+x_j(p_2')} \pmod{q_j}.$$

   If $p \in \mathcal{P}_k$, then $g_j^{x_j(p)} \equiv p \pmod{q_j}$ for $j \leq k$ and $g_j^{x_j(p)} \equiv 1 \pmod{q_j}$ when $j > k$. Thus, for all $j < k_2$ we find $p_1 p_2 \equiv p_1' p_2' \pmod{q_j}$, leading to

$$p_1 p_2 \equiv p_1' p_2' \pmod{q_1 \dots q_{k_2}}$$

3. For claim 3 we note that if $k_2 < k_1$, for $k_2 + 1 \leq j \leq k_1$ we have that $p_1 \equiv p_1'$ (mod q_j). Hence,
$$p_1 \equiv p_1' \pmod{q_{k_2+1} \dots q_{k_1}}.$$

4. Claims 2 and 3 give us the last statement

$$\begin{gathered} 2^{c(k_1^2+k_2^2)} \geq |p_1p_2 - p_1'p_2'| \geq q_1 \dots q_{k_2} > 2^{1+\dots+2k_2-1} = 2^{k_2^2}, \\ \text{implying } k_2^2 < \frac{c}{1-c} k_1^2. \end{gathered} \tag{3.2.3}$$

In particular this implies that $k_2 < k_1$, so we can apply claim 3 and obtain,

$$\begin{gathered} 2^{c(k_1^2)} \geq |p_1 - p_1'| \geq q_{k_2+1} \dots q_{k_1} > 2^{(2k_2+1)+\dots+2k_1-1} = 2^{k_1^2-k_2^2}, \\ \text{which implies } k_2^2 > (1-c)k_1^2. \end{gathered} \tag{3.2.4}$$

□

Now we can quite easily prove Theorem 3.1.

Proof of Theorem 3.1. Suppose that $A_{\bar{q},c}$ is not a Sidon set, then there exist $p_1, p_2, p_1', p_2' \in \mathcal{P}$ with $(p_1, p_2) \neq (p_1', p_2')$ such that $a_{p_1} > a_{p_1'} \geq a_{p_2'} > a_{p_1}$ and $a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}$. Now Proposition 4.1.2 implies that $1 - c < \frac{c}{1-c}$, which implies that $c < \frac{3-\sqrt{5}}{2}$. Thus, $A_{\bar{q},c}$ is a Sidon set for $c = \frac{3-\sqrt{5}}{2}$. □

In the proof of Theorem 3.2 we will choose $c = \sqrt{2} - 1$. This will require us to remove some of the items of $A_{\bar{q},c}$, because $A_{\bar{q},c}$ itself is not a Sidon for this value of c. We will remove some of the terms a_p that appear in repeated sums, and find that after having done so, the resulting set will have the desired growth and be a Sidon set.

Proof of Theorem 3.2. Proposition 3.2.3 implies that all repeated sums that appear are of the form:

$$a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'} \text{ where } (p_1, p_2) \neq (p_1', p_2'), \tag{3.2.5}$$

where $p_1, p_1' \in \mathcal{P}_{k_1}, p_2, p_2' \in \mathcal{P}_{k_2}$ and $k_2^2 < \frac{c}{1-c}k_1^2$. Let $Q_1 := q_1 \dots q_{k_2+1}$ and $Q_2 := q_{k_2+1} \dots q_{k_1}$. By substituting this we get

$$p_1(p_2 - p_2') = p_1p_2 - p_1'p_2' + (p_1' - p_1)p_2' = \frac{p_1p_2 - p_1'p_2'}{Q_1}Q_1 + \frac{(p_1' - p_1)}{Q_2}Q_2p_2'.$$

Proposition 3.2.3 also implies that if there is a repeated sum, then $s_1 = \frac{p_1p_2 - p_1'p_2'}{Q_1}$, $s_2 = \frac{(p_1' - p_1)}{Q_2}$ are non-zero integers such that

$$|s_1| = \frac{|p_1p_2 - p_1'p_2'|}{Q_1} \leq \frac{2^{c(k_1^2+k_2^2)-6}}{Q_1} \text{ and } |s_2| = \frac{|p_1' - p_1|}{Q_2} \leq \frac{2^{ck_1^2-3}}{Q_2}.$$

Hence if $p_1 \in \mathcal{P}_{k_1}$ appears in a repeated sum then it divides an integer $s \neq 0$ of the following set:

$$S_{k_2,k_1} = \{s = s_1Q_1 + s_2Q_2p_2' : 1 \leq |s_1| \leq \frac{2^{c(k_1^2+k_2^2)-6}}{Q_1}, 1 \leq |s_2| \leq \frac{2^{ck_1^2-3}}{Q_2}, p_2' \in \mathcal{P}_{k_2}\}.$$

We now define the set of primes. Let

$$\mathcal{B}_{k_1} := \{p_1 \in \mathcal{P}_{k_1} : p_1|s, s \in S_{k_2,k_1}, s \neq 0 \text{ with } k_2^2 < \frac{c}{1-c}k_1^2\}.$$

Then for

$$\mathcal{P}^* = \bigcup_{k_1}(\mathcal{P}_{k_1} \setminus \mathcal{B}_{k_1}),$$

the set $A^*_{\bar{q},c} = (a_p)_{p\in\mathcal{P}^*}$ is clearly a Sidon set. To prove that $A^*_{\bar{q},c}(x) = x^{c+o(1)}$ as $x \to \infty$, it suffices to show that $|\mathcal{B}_{k_1}| \leq (\frac{1}{2} + o(1))|\mathcal{P}_{k_1}|$ as $k_1 \to \infty$. We will prove this holds for $c = \sqrt{2} - 1$. Note that an integer $s \neq 0$ of S_{k_2,k_1} cannot be divisible by two primes $p, p' \in \mathcal{P}_{k_1}$ otherwise we get that

$$2^{2c(k_1-1)^2-6} < pp' \leq |s| \leq 2(2^{c(k_1^2+k_2^2)-6}) < 2^{\frac{c}{1-c}k_1^2-5},$$

which does not hold for large enough k_1 since $2c > \frac{c}{1-c}$ for $c < \frac{1}{2}$. Using $Q_1Q_2 = q_1 \dots q_{k_1} > 2^{k_1^2}$ and $|\mathcal{P}_{k_2}| \leq \pi(2^{ck_2^2}) \leq 2\frac{2^{ck_2^2}}{\log(2^{ck_2^2})}$ [7] we have

$$|\mathcal{B}_{k_1}| \leq \sum_{k_2<\sqrt{\frac{c}{1-c}}k_1} |S_{k_2,k_1}| \leq \sum_{k_2<\sqrt{\frac{c}{1-c}}k_1} (2\frac{2^{c(k_1^2+k_2^2)-6}}{Q_1})(2\frac{2^{ck_1^2-3}}{Q_2})|\mathcal{P}_{k_2}|$$

$$\leq \frac{2^{-6}}{c\log(2)}2^{(2c-1)k_1^2} \sum_{k_2<\sqrt{\frac{c}{1-c}}k_1} \frac{2^{2ck_2^2}}{k_2^2} \leq \frac{2^{-6}}{c\log(2)}\frac{(1+o(1))(1-c)}{c}\frac{2^{(\frac{2c}{1-c}-1)k_1^2}}{k_1^2}. \tag{3.2.6}$$

To estimate the sum over k_2, note that since $([x]-1)^2 - x^2 \leq ([x]-1)^2 - [x]^2 = 1 - 2[x] < 3 - 2x$, we find

$$\begin{aligned}\sum_{k_2<x} \frac{2^{2ck_2^2}}{k_2^2} &< \frac{2^{2c[x]^2}}{[x]^2} + \sum_{k_2\leq[x]-1} \frac{2^{2c[x]^2}}{[x]^2} < \frac{2^{2c[x]^2}}{[x]^2} + \sum_{k_2\leq[x]-1} 2^{2c([x]-1)^2} \\ &< \frac{2^{2cx^2}}{x^2}\left(\frac{x^2}{[x]^2} + x^3 2^{2c([x]-1)^2-x^2)}\right) < \frac{2^{2cx^2}}{x^2}\left(\frac{x^2}{[x]^2} + x^3 2^{2c(3-2x)}\right) \\ &\leq (1+o(1))\frac{2^{2cx^2}}{x^2} \text{ as } x \to \infty\end{aligned} \tag{3.2.7}$$

Now, using the fact that for $\frac{2c}{1-c} - 1 = c$ and $\frac{1-c}{c} = \sqrt{2}$ for $c = \sqrt{2} - 1$ and the estimate

$$|\mathcal{P}_{k_1}| = \pi(2^{ck_1^2-3}) - \pi(2^{c(k_1-1)^2-3}) = \frac{2^{ck_1^2-3}}{ck_1^2\log(2)}(1+o(1)),$$

we have

$$|\mathcal{B}_{k_1}| \leq \frac{2^{-6}(1+o(1))2^{ck_1^2}\sqrt{2}}{c\log(2)k_1^2} \leq (\frac{1}{2}+o(1))|\mathcal{P}_{k_1}|.$$

□

4 Alternative proof

In this last section we will prove some of the theorems from the previous section using a different approach. We will use irreducible polynomials from $\mathbb{F}_2[X]$ instead of prime numbers. This idea also comes from Cilleruelo. We will again construct a sequence $A_{\bar{q},c}$.

4.1 Polynomial construction

We again consider a base, this time $\bar{q} = (2^{2j+1})_{j=1}^{\infty}$, and we use an infinite sequence of irreducible polynomials $(q_j)_{j=0}^{\infty}$ from $\mathbb{F}_2[X]$ with $\deg(q_j) = 2j-1$ in $\mathbb{F}_2[X]$. For each j we will choose a generator g_j of $(\mathbb{F}_2[X]/q_j(X))^*$.
Recall that each element of the finite field $\mathbb{F}_2[X]/q_j(X)$ is uniquely represented by a polynomial in $\mathbb{F}_2[X]$ of degree $< 2j-1$. Since $(\mathbb{F}_2[X]/(q_j))^*$ is of cyclic order $2^{2j-1}-1$, there is a $g_j \in \mathbb{F}_2[X]$ of degree $< 2j-1$ such that $(\mathbb{F}_2[X]/(q_j))^*$
$= \langle g_j \pmod{q_j} \rangle$

We fix c such that $0 < c < \frac{1}{2}$ and for each $k \geq 3$ we now consider,

$$\mathcal{P} = \bigcup_{k\geq 3} \mathcal{P}_k, \text{ where } \mathcal{P}_k = \{p \in \mathbb{F}_2[X] : p \text{ irreducible}, c(k-1)^2-3 < \deg(p) \leq ck^2-3\}$$

For each $p \in \mathcal{P}$ we define a_p as follows. Choose k such that $p \in \mathcal{P}_k$ and then define

$$a_p = (x_k(p) \dots x_1(p))_{\bar{q}} \tag{4.1.1}$$

where $x_j(p)$ is the unique solution of the polynomial congruence

$$g_j^{x_j(p)} \equiv p \pmod{q_j},\ 2^{2j-1}+1 \leq x_j(p) \leq 2^{2j}-1. \tag{4.1.2}$$

Again we define $x_j(p) = 0$ if $j > k$. Let $A_{\bar{q},c} := (a_p)_{p\in\mathcal{P}}$

4.2 Properties of $A_{\bar{q},c}$ with the polynomial construction

We will prove the same properties as before, only now for our new sequence $A_{\bar{q},c}$. Most properties will be proven similarly to their earlier counterparts.

Proposition 4.0.1. *The terms a_p of $A_{\bar{q},c}$ are pairwise distinct.*

Proof. Suppose $a_p = a_{p'}$ for some $p, p' \in \mathcal{P}$ with $p \neq p'$. Then all digits of a_p and $a_{p'}$ must be equal, i.e., $x_j(p) = x_j(p')$ for all $j \geq 1$. By construction we have that

$p, p' \in \mathcal{P}_k$ such that k is the largest j such that the digit $x_j \neq 0$. We also know that

$$p \equiv g_j^{x_j(p)} \equiv g_j^{x_j(p')} \equiv p' \pmod{q_j} \text{ for } j \leq k.$$

Hence $p \equiv p' \pmod{q_1 ... q_k}$. Suppose that $p \neq p'$. Using the fact that $p, p' \in \mathcal{P}_k$ and $\deg(q_j) \geq 2^{2j-1}$ we find that

$$ck^2 \geq \deg(p - p') \geq \deg(q_1 \dots q_k) \geq 1 + \dots + (2k-1) = k^2 \tag{4.2.1}$$

which is impossible because $0 < c < \frac{1}{2}$. Hence, the elements a_p of $A_{\bar{q},c}$ are pairwise distinct. □

For the following proposition we will use an analogue of the prime number theorem for irreducible polynomials.

Theorem 4.1. *(Prime Number Theorem Analogue)*
Let N_j denote the number of monic irreducible polynomials of degree $j \in \mathbb{F}_q[X]$. Then we have:

$$N_j \sim \frac{q^j}{j} \text{ as } j \to \infty. \tag{4.2.2}$$

Proof. The subfields of $\mathbb{F}_{q^n}$ are $\mathbb{F}_{q^m}$ where $m|n$. Let $a \in \mathbb{F}_{q^n} \setminus \bigcup_{m|n, m<n} \mathbb{F}_{q^m}$. Now a has degree n over $\mathbb{F}_q$. The minimum polynomial I_a of a over $\mathbb{F}_2$ is the monic polynomial over $\mathbb{F}_2$ of minimal degree such that $p(a) = 0$. The polynomial I_a is irreducible of degree n.
Now let $p \in \mathbb{F}_q[X]$ be a monic irreducible polynomial, such that $\deg(p) = n$. Then p has a zero $a \in \mathbb{F}_{q^n} \setminus \bigcup_{m|n, m<n} \mathbb{F}_{q^m}$. p has in fact precisely n zeroes in $\mathbb{F}_q^n$, i.e., $a, a^q, \dots a^{q^{n-1}}$. Indeed, p cannot have more than n zeroes in $\mathbb{F}_q^n$, and from the fact that $u \mapsto u^{q^i}$, where $i \in \{0, \dots, n-1\}$ are $\mathbb{F}_q$-invariant automorphisms of $\mathbb{F}_q^n$, it follows that $p(a) = p(a^q) = \dots = p(a^{q^{n-1}}) = 0$. Thus, there is a bijection between monic irreducible polynomials in $\mathbb{F}_q[X]$ of degree n and subsets $\{a, a^q, \dots, a^{q^{n-1}}\}$ of $\mathbb{F}_{q^n} \setminus \bigcup_{m|n, m<n} \mathbb{F}_{q^m}$.
Let $N_n(q)$ be the number of monic irreducible polynomials in $\mathbb{F}_q[X]$ of degree n. Then we have

$$N_n(q) \geq \frac{1}{n}(q^n - \sum_{m|n, m<n} q^m) \geq \frac{1}{n}(q^n - nq^{\frac{n}{2}}).$$

Clearly we have also $N_n(q) \leq \frac{q^n}{n}$. The squeeze theorem gives us that $N_n \sim \frac{q^n}{n}$ as $n \to \infty$.

□

Proposition 4.1.1. *We have $A_{\bar{q},c}(x) = x^{c+o(1)}$ as $x \to \infty$.*

Proof. Let $x \in \mathbb{R}$ and consider the integer k such that

$$2^{(k+1)^2} - 1 < x \leq 2^{(k+2)^2} - 1.$$

Hence $x = 2^{k^2(1+O(\frac{1}{k}))}$ and thus $2^{k^2} = x^{1+o(1)}$. If $\deg(p) \leq ck^2 - 3$ then $p \in \mathcal{P}_l$ for some $l \leq k$ so

$$a_p = x_1(p) + \sum_{j \leq l} x_j(p) 2^{3+\cdots+2j-1} = x_1(p) + \sum_{j \leq l} x_j(p) 2^{j^2-1} \leq 2^{k^2} \leq x.$$

Now using Theorem 4.1 we find

$$A_{\bar{q},c}(x) \geq N_{[ck^2-3]} = \frac{2^{[ck^2-3]}}{[ck^2]}(1 + o(1)) = 2^{[ck^2](1+o(1))} = x^{c+o(1)}.$$

Note that if $c(k+1)^2 - 3 \leq \deg(p) < c(k+2)^2 - 3$, then $p \in \mathcal{P}_{k+2}$ and we obtain $a_p \geq 2^{3+\cdots+2k+3} \geq 2^{(k+2)^2-1} \geq x$. This gives an upper bound

$$\begin{aligned} A_{\bar{q},c}(x) &\leq \sum_{j \leq c(k+1)^2-3} N_j \leq \sum_{j \leq c(k+1)^2-3} \frac{2^j}{j} \\ &\leq 2^{c(k+1)^2} = 2^{ck^2(1+o(1))} = x^{c+o(1)}. \end{aligned} \tag{4.2.3}$$

□

Proposition 4.1.2. *Suppose that there exist $a_{p_1}, a_{p_2}, a_{p_1'}, a_{p_2'} \in A_{\bar{q},c}$ such that $a_{p_1} > a_{p_1'} \geq a_{p_2'} > a_{p_2}$ and*

$$a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}.$$

Then we have:

1. *There exist k_1, k_2 with $k_2 \leq k_1$ such that $p_1, p_1' \in \mathcal{P}_{k_1}$ and $p_2, p_2' \in \mathcal{P}_{k_2}$,*
2. *$p_1 p_2 \equiv p_1' p_2' \pmod{q_1 \dots q_{k_2}}$,*
3. *$p_1 \equiv p_1' \pmod{q_{k_{2+1}} \dots q_{k_1}}$ if $k_2 < k_1$,*
4. *$(1-c)k_1^2 < k_2^2 < \frac{c}{1-c} k_1^2$.*

Proof. We prove the four claims

1. Since $0 \leq x_j(p_1) + x_j(p_2) < 2^{2j+1}$, $0 \leq x_j(p_1') + x_j(p_2') < 2^{2j+1}$ for all j, and $a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}$, we infer that the digits of the sums are equal. We find

$$x_j(p_1) + x_j(p_2) = x_j(p_1') + x_j(p_2') \text{ for all } j. \tag{4.2.4}$$

   By construction, $p_1 \in \mathcal{P}_{k_1}$, where k_1 is the largest j such that

$$x_j(p_1) + x_j(p_2) \geq 2^{2j-1} + 1.$$

   This is because taking $j > k_1$ gives $x_j(p_1) + x_j(p_2) = 0$ and $j \leq k_1$ gives $x_j(p_1) + x_j(p_2) \geq x_j(p_1) \geq 2^{2j-1} + 1$.
   Further, we have $p_2 \in \mathcal{P}_{k_2}$ and k_2 is the largest j such that

$$x_j(p_1) + x_j(p_2) \geq 2^{2j} + 2.$$

   This is because taking $k_2 < j \leq k_1$ gives $x_j(p_1) + x_j(p_2) = x_j(p) \leq 2^{2j} - 1$ and $j < k_2$ gives that both $x_j(p_1), x_j(p_2) \geq 2^{2j-1} + 1$ hence $x_j(p_1) + x_j(p_2) \geq 2^{2j} + 2$.
   This proves claim 1.

2. For the second claim we observe that (4.2.4) implies the following congruence for all j:

$$g_j^{x_j(p_1)+x_j(p_2)} \equiv g_j^{x_j(p_1')+x_j(p_2')} \pmod{q_j}.$$

   If $p \in \mathcal{P}_k$, then $g_j^{x_j(p)} \equiv p \pmod{q_j}$ for $j \leq k$ and $g_j^{x_j(p)} \equiv 1 \pmod{q_j}$ when $j > k$. Thus, for all $j < k_2$ we find $p_1 p_2 \equiv p_1' p_2' \pmod{q_j}$, this leads to

$$p_1 p_2 \equiv p_1' p_2' \pmod{q_1 \dots q_{k_2}}$$

3. For claim 3 we note that if $k_2 < k_1$, for $k_2 + 1 \leq j \leq k_1$ we have that $p_1 \equiv p_1' \pmod{q_j}$. Hence,

$$p_1 \equiv p_1' \pmod{q_{k_2+1} \dots q_{k_1}}.$$

4. Combining claims 2 and 3 we get that

$$\begin{gathered} c(k_1^2 + k_2^2) \geq \deg(p_1 p_2 - p_1' p_2') \geq \deg(q_1 \dots q_{k_2}) > 2^{1+\dots+2k_2-1} = k_2^2, \\ \text{implying } k_2^2 < \frac{c}{1-c} k_1^2. \end{gathered} \tag{4.2.5}$$

This implies that $k_2 < k_1$, so we can apply claim 3 and get,

$$ck_1^2 \geq \deg(p_1 - p_1') \geq \deg(q_{k_2+1} \dots q_{k_1}) > (2k_2+1) + \dots + (2k_1-1) = k_1^2 - k_2^2,$$
$$\text{which implies } k_2^2 > (1-c)k_1^2. \tag{4.2.6}$$

□

Alternative proof of Theorem 3.2. Proposition 4.1.2 implies that all repeated sums that appear are of the form:

$$a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'} \text{ where } (p_1, p_2) \neq (p_1', p_2'), \tag{4.2.7}$$

where $p_1, p_1' \in \mathcal{P}_{k_1}, p_2, p_2' \in \mathcal{P}_{k_2}$ and $k_2^2 < \frac{c}{1-c}k_1^2$. Let $Q_1 := q_1 \dots q_{k_2+1}$ and $Q_2 := q_{k_2+1} \dots q_{k_1}$. By substituting this we get

$$p_1(p_2 - p_2') = p_1p_2 - p_1'p_2' + (p_1' - p_1)p_2' = \frac{p_1p_2 - p_1'p_2'}{Q_1}Q_1 + \frac{p_1' - p_1}{Q_2}Q_2p_2'.$$

By Proposition 4.1.2 we also have that if there is a repeated sum, then $s_1 = \frac{p_1p_2 - p_1'p_2'}{Q_1}$ and $s_2 = \frac{p_1' - p_1}{Q_2}$ are non-zero polynomials such that

$$\deg(s_1) \leq c(k_1^2 + k_2^2) - \deg(Q_1) - 6, \deg(s_2) \leq ck_1^2 - \deg(Q_2) - 3.$$

Hence we know that if $p_1 \in \mathcal{P}_{k_1}$ appears in a repeated sum it must divide a polynomial $s \neq 0$ of the set

$$S_{k_2,k_1} = \left\{ s = s_1Q_1 + s_2Q_2p_2' : \begin{array}{c} 0 \leq \deg(s_1) \leq c(k_1^2 + k_2^2) - \deg(Q_1) - 6 \\ 0 \leq \deg(s_2) \leq ck_1^2 - \deg(Q_2) - 3 \\ p_2' \in \mathcal{P}_{k_2} \end{array} \right\}.$$

Let B_{k_1} the set of $p_1 \in \mathcal{P}_{k_1}$ with the property that there are k_2 with $0 \leq k_2^2 < \frac{c}{1-c}k_1^2$ and $s \in S_{k_2,k_1}$ with $s \neq 0$ such that p_1 divides s. Then for

$$\mathcal{P}^* = \bigcup_{k_1} (\mathcal{P}_{k_1} \setminus \mathcal{B}_{k_1}),$$

the set $A^*_{\bar{q},c} = (a_p)_{p \in \mathcal{P}^*}$ is a Sidon set. To prove that $A^*_{\bar{q},c}(x) = x^{c+o(1)}$ as $x \to \infty$, it suffices to show that $|\mathcal{B}_{k_1}| \leq (\frac{1}{2^{\frac{9}{2}}} + o(1))|\mathcal{P}_{k_1}|$. We will prove this holds for

$c = \sqrt{2} - 1$. Note that a polynomial $s \neq 0$ of S_{k_2,k_1} cannot be divisible by two polynomials $p, p' \in \mathcal{P}_{k_1}$ otherwise we have

$$2c(k_1-1)^2 - 6 < \deg(pp') \leq \deg(s) \leq c(k_1^2 + k_2^2) < \frac{c}{1-c}k_1^2 - 6,$$

which does not hold for large enough k_1 since $2c > \frac{c}{1-c}$ for $c < \frac{1}{2}$. Using $\deg(Q_1Q_2) = \deg(q_1 \dots q_{k_1}) = k_1^2$ and

$$\begin{aligned} |P_{k_2}| &= \sum_{c(k_2-1)^2-3<j\leq ck_2^2-3} N_j \leq \sum_{c(k_2-1)^2-3<j\leq ck_2^2-3} \frac{2^j}{j} \\ &\leq \frac{1}{c(k_2-1)^2-3} \sum_{c(k_2-1)^2-3<j\leq ck_2^2-3} 2^j \leq \frac{1}{c(k_2-1)^2-3} 2^{[ck_2^2-2]} \\ &\leq \frac{1}{4} \cdot \frac{ck_2^2}{c(k_2-1)^2-3} \cdot \frac{2^{ck_2^2}}{ck_2^2}. \end{aligned} \tag{4.2.8}$$

For $k_2 \geq 5$ this is a decreasing function and < 1. So $|\mathcal{P}_{k_2}| < \frac{2^{ck_2^2}}{ck_2^2}$ for $k_2 \geq 5$. For $0 \leq k_2 \leq 4$ we have

$$|\mathcal{P}_{k_2}| = \sum_{c(k_2-1)^2-3<j\leq ck_2^2-3} N_j \leq \sum_{1<j\leq \lceil ck_2^2-3 \rceil} \frac{2^j}{j}$$

this gives $|\mathcal{P}_1| = 0 \leq \frac{2^c}{c}$, $|\mathcal{P}_2| \leq 1 \leq \frac{2^{c\cdot 2^2}}{c\cdot 2^2}$, $|\mathcal{P}_3| \leq 2 \leq \frac{2^{c\cdot 3^2}}{c\cdot 3^2}$, $|\mathcal{P}_4| \leq \frac{32}{3} \leq \frac{2^{c\cdot 4^2}}{c\cdot 4^2}$. Hence $|\mathcal{P}_{k_2}| \leq \frac{2^{ck_2^2}}{ck_2^2}$ also for $k_2 \leq 4$. This gives

$$\begin{aligned} |\mathcal{B}_{k_1}| \leq \sum_{k_2<\sqrt{\frac{c}{1-c}}k_1} |S_{k_2,k_1}| &\leq \sum_{k_2<\sqrt{\frac{c}{1-c}}k_1} (2^{c(k_1^2+k_2^2)-\deg(Q_1)-6})(2^{ck_1^2-\deg(Q_2)-3})|\mathcal{P}_{k_2}| \\ &\leq \frac{2^{-9}}{c} 2^{(2c-1)k_1^2} \sum_{k_2<\sqrt{\frac{c}{1-c}}k_1} \frac{2^{2ck_2^2}}{k_2^2}. \end{aligned} \tag{4.2.9}$$

Again using

$$\sum_{k_2<x} \frac{2^{2ck_2^2}}{k_2^2} < (1+o(1))\frac{2^{2cx^2}}{x^2} \text{ as } x \to \infty$$

we find

$$|B_{k_1}| \leq \frac{2^{-9}(1+o(1))(1-c)}{c^2} \frac{2^{(\frac{2c}{1-c}-1)k_1^2}}{k_1^2}$$

Now, using $\frac{2c}{1-c} - 1 = c$ and $\frac{1-c}{c} = \sqrt{2}$ for $\sqrt{2} - 1$ and the estimate

$$|\mathcal{P}_{k_1}| \geq N_{[ck_1^2]-3} = \frac{2^{[ck_1^2]-3}}{[ck_1^2]-3}(1+o(1)) = \frac{2^{ck_1^2-4}}{ck_1^2}(1+o(1)) = 2^{-4}(1+o(1))\frac{2^{ck_1^2}}{ck_1^2}.$$

We find,

$$|\mathcal{B}_{k_1}| \leq \frac{2^{-9}(1+o(1))2^{ck_1^2}\sqrt{2}}{ck_1^2} \leq (\frac{1}{2^{\frac{9}{2}}} + o(1))|\mathcal{P}_{k_1}|.$$

□

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi. A dense infinite Sidon sequence. *European Journal of Combinatorics*, 2(1):1–11, 1981.

[2] J. Cilleruelo. Infinite Sidon sequences. *Advances in Mathematics*, pages 474–486, 2014.

[3] H. Davenport. *Multiplicative Number Theory: Second Edition*. Springer, second edition, 1980.

[4] P. Erdős. Solved and unsolved problems in combinatorics and combinatorial number theory. *Congressus Numerantium*, 32:49–62, 1981.

[5] P. Erdős and P. Turán. On a problem of Sidon in additive number theory, and on some related problems. *Journal of The London Mathematical Society-second Series*, pages 212–215, 1941.

[6] H. Halberstam and K. F. Roth. *Sequences*. Oxford Clarendon Press, first edition, 1966.

[7] G. H. Hardy and E. M. Wright. *An Introduction to the Theory of Numbers*. Oxford University Press, fourth edition, 1975.

[8] J.Singer. A theorem in finite projective geometry and some applications to number theory. *TAMS*, 43:377–385, 1938.

[9] I. Z. Ruzsa. An infinite Sidon sequence. *Journal of Number Theory*, 68:63–71, 1998.