Infinite Sidon sequences

Javier Cilleruelo

ICMAT-Universidad Autónoma de Madrid

Erdos Centennial
Budapest, July 1-5, 2013

The origin of the problem

In 1932, Simon Sidon asked to Erdős about the slowest possible growth of an infinite sequence A of positive integers having the property that all the sums

$$a + a', \quad a \le a', \quad a, a' \in A$$

are distinct.

Erdős named them Sidon sequences and they became one of his favorite topics.

The Sidon sequence given by the greedy algorithm

Main problem: Construct (or prove the existence of) an infinite Sidon sequence A with counting function

$$A(x) = |A \cap [1, x]|$$

as large as possible.

Erdős considered the sequence given by the greedy algorithm:

- Starting with $a_1 = 1$, define a_{n+1} as the least positive integer we can add to the set $\{a_1, \dots, a_n\}$ preserving the Sidon property.

$$1, 2, 4, 8, 13, 21, 31, 45, 66, 81, 97, 123, 148, 182, 204, \dots$$

The Sidon sequence given by the greedy algorithm

The forbidden elements for a_{n+1} are those of the form

$$a_i + a_j - a_k, \quad i, j, k \leq n.$$

Since there are at most n^3 of them, certainly we have that

$$a_{n+1} \leq n^3 + 1,$$

which implies

$$A(x) \gg x^{1/3}.$$

An upper bound for the counting function of a Sidon sequence

The number of sums $a + a'$, $a, a' \in A$ with $a < a' \leq x$ is $\binom{A(x)}{2}$, and all of them are distinct and less than $2x$:

$$\binom{A(x)}{2} < 2x \implies A(x) \ll x^{1/2}.$$

Conjecture (Erdős): For each $\epsilon > 0$ there is an infinite Sidon sequence A with

$$A(x) \gg x^{1/2-\epsilon}.$$

- Erdős proved that the conjecture is false for $\epsilon = 0$.

The construction of Ajtai, Komlós and Szemerédi

The greedy Sidon sequence found by Erdős was the densest known during almost 50 years.

Theorem (Ajtai, Komlós and Szemerédi, 1981): There exists an infinite Sidon sequence with counting function

$$A(x) \gg (x \log x)^{1/3}.$$

They wrote: *"The task of constructing a denser sequence has so far resisted all efforts, both constructive and random methods. Here we use a random construction for giving a sequence which is denser than the above trivial one".*

The construction of Ruzsa

Theorem (Ruzsa, 1998): There exists an infinite Sidon sequence with counting function

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

The starting point of Ruzsa was the observation that the sequence of the prime numbers is a multiplicative Sidon sequence, or equivalently the sequence $(\log p)_{p \in P}$ is a Sidon sequence of real numbers.

The construction of Ruzsa

Ruzsa introduced a random parameter $\alpha \in [1, 2]$ and constructed a sequence $A_\alpha = (a_p)_{p \in P}$, where each integer a_p is built using the binary digits of $\alpha \log p$.

Then he proved that for almost all $\alpha \in [1, 2]$ it is possible to extract a dense Sidon sequence from A_α.

The constructions of Ruzsa and Ajtai, Komlós and Szemerédi are probabilistic. They are not explicit.

An explicit construction

It was a open problem to construct an explicit Sidon sequence with counting function $A(x) \gg x^c$ for some $c > 1/3$.

We construct an explicit Sidon sequence as dense as Ruzsa's sequence:

Theorem (C., 2012): There exists an infinite Sidon sequence A, **which can be explicitly constructed**, with counting function

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

Generalized basis

Given a sequence of positive integers

$$\overline{q} := 4q_1, \ldots, 4q_j, \ldots \qquad \text{(the base)},$$

any positive integer a can be written, in only a way, in the form

$$a = x_1 + x_2(4q_1) + x_3(4q_1)(4q_2) + \cdots + x_j(4q_1)\cdots(4q_{j-1}) + \cdots$$

where the digits x_j satisfy

$$0 \leq x_j < 4q_j.$$

We represent the integer a in the form

$$a := \ldots x_j \ldots x_1.$$

Summing integers as vectors

If all the digits of a, a' satisfy

$$q_j < x_j, x_j' < 2q_j,$$

$$\begin{array}{rcr} a & = & x_{k_1} \ldots\ldots\ldots x_1 \\ a' & = & x'_{k_2} \ldots x'_1 \end{array}$$

we have

$$a + a' = {\color{blue}(}x_{k_1} + 0{\color{blue})} \ldots {\color{blue}(}x_{k_2+1} + 0{\color{blue})}{\color{red}(}x_{k_2} + x'_{k_2}{\color{red})} \ldots {\color{red}(}x_1 + x'_1{\color{red})}.$$

Furthermore, the digits of $a + a'$ determine the lengths k_1, k_2 of a, a', $a \leq a'$.

The construction: the base and the set of indexes

1) We consider a fix generalized base

$$\overline{q} := 4q_1, \ldots, 4q_j, \ldots,$$

where the q_j are primes satisfying

$$2^{2j-1} < q_j \leq 2^{2j}.$$

2) We use the set of the primes P as the indices

$$A = (a_p)_{p \in P}$$

and represent the elements a_p in the base $\overline{q}$ as:

$$a_p = \ldots x_j(p) \ldots x_1(p).$$

The construction: the growth

3) Fix c, $0 < c < 1/2$ and make a partition of the set of the primes (the set of indices):

$$P = \bigcup_k P_k, \qquad P_k = \{p : \ 2^{c(k-1)^2} < p \le 2^{ck^2}\}.$$

Proposition: Assume that the elements a_p with $p \in P_k$ have exactly k digits in the base $\overline{q}$,

$$a_p = x_k(p) \dots x_1(p).$$

Then we have

$$A_{\overline{q},c}(x) = x^{c+o(1)}.$$

The construction: the digits

4) For $p \in P_k$ we define the digits of

$$a_p := x_k(p) \dots x_1(p)$$

as follows: the digit $x_j(p)$ is given by the solution of

$$g_j^{x_j(p)} \equiv p \pmod{q_j}, \qquad q_j < x_j(p) < 2q_j,$$

where g_j is a given generator of $\mathbb{F}_{q_j}^*$.

(The digit $x_j(p)$ is the discrete logarithm of p modulo q_j and it is unique modulo $q_j - 1$.)

- We will prove that if there is a repeated sum

$$a_{p_1} + a_{p_2} = a_{p'_1} + a_{p'_2}$$

  then the primes involved, p_1, p_2, p'_1, p'_2, must satisfy some relations.

- We will prove that these relations cannot hold if

$$c \leq \text{ some value } c_0.$$

Lemma 1: If $a_{p_1} + a_{p_2} = a_{p'_1} + a_{p'_2}$ then there exist $k_2 \leq k_1$ such that
$$p_1, p'_1 \in P_{k_1}, \quad p_2, p'_2 \in P_{k_2}.$$

$$\begin{aligned}
a_{p_1} &= x_{k_1}(p_1) \cdots x_{k_2}(p_1) \cdots x_1(p_1) \\
a_{p_2} &= x_{k_2}(p_2) \cdots x_1(p_2) \\
a_{p'_1} &= x_{k_1}(p'_1) \cdots x_{k_2}(p'_1) \cdots x_1(p'_1) \\
a_{p'_2} &= x_{k_2}(p'_2) \cdots x_1(p'_2)
\end{aligned}$$

Lemma 2: If $a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}$, $p_1, p_1' \in P_{k_1}$, $p_2, p_2' \in P_{k_2}$ then

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

$$\begin{aligned}
a_{p_1} &= x_{k_1}(p_1) \cdots x_{k_2}(p_1) \cdots x_1(p_1) \\
a_{p_2} &= x_{k_2}(p_2) \cdots x_1(p_2) \\
a_{p_1'} &= x_{k_1}(p_1') \cdots x_{k_2}(p_1') \cdots x_1(p_1') \\
a_{p_2'} &= x_{k_2}(p_2') \cdots x_1(p_2')
\end{aligned}$$

For $1 \leq j \leq k_2$ we have

$$\begin{aligned}
x_j(p_1) + x_j(p_2) &= x_j(p_1') + x_j(p_2') \\
g_j^{x_j(p_1)+x_j(p_2)} &\equiv g_j^{x_j(p_1')+x_j(p_2')} \pmod{q_j} \\
p_1 p_2 &\equiv p_1' p_2' \pmod{q_j}
\end{aligned}$$

Lemma 3: If $a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}$, $p_1, p_1' \in P_{k_1}$, $p_2, p_2' \in P_{k_2}$ then

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

$$\begin{aligned}
a_{p_1} &= x_{k_1}(p_1) \cdots x_{k_2}(p_1) \cdots x_1(p_1) \\
a_{p_2} &= x_{k_2}(p_2) \cdots x_1(p_2) \\
a_{p_1'} &= x_{k_1}(p_1') \cdots x_{k_2}(p_1') \cdots x_1(p_1') \\
a_{p_2'} &= x_{k_2}(p_2') \cdots x_1(p_2')
\end{aligned}$$

For $k_2 + 1 \leq j \leq k_1$ we have

$$\begin{aligned}
x_j(p_1) &= x_j(p_1') \\
g_j^{x_j(p_1)} &\equiv g_j^{x_j(p_1')} \pmod{q_j} \\
p_1 &\equiv p_1' \pmod{q_j}
\end{aligned}$$

If $a_{p_1} + a_{p_2} = a_{p'_1} + a_{p'_2}$

i) $p_1, p'_1 \in P_{k_1} = \{p : \ 2^{c(k_1-1)^2} < p \le 2^{ck_1^2}\}$ (Lemma 1)
$p_2, p'_2 \in P_{k_2} = \{p : \ 2^{c(k_2-1)^2} < p \le 2^{ck_2^2}\}$

ii) $p_1 p_2 \equiv p'_1 p'_2 \pmod{q_1 \cdots q_{k_2}}$ (Lemma 2)

iii) $p_1 \equiv p'_1 \pmod{q_{k_2+1} \cdots q_{k_1}}$ (Lemma 3)

iv) $2^{2j-1} < q_j \le 2^{2j}$ (by construction)

$$2^{ck_1^2 + ck_2^2} \underset{\substack{\uparrow \\ i)}}{\ge} |p_1 p_2 - p'_1 p'_2| \underset{\substack{\uparrow \\ ii)}}{\ge} q_1 \cdots q_{k_2} \underset{\substack{\uparrow \\ iv)}}{>} 2^{1+3+\cdots+(2k_2-1)} = 2^{k_2^2}$$

$$\Longrightarrow \quad k_2^2 < \frac{c}{1-c} k_1^2.$$

If $a_{p_1} + a_{p_2} = a_{p'_1} + a_{p'_2}$

i) $p_1, p'_1 \in P_{k_1} = \{p : \ 2^{c(k_1-1)^2} < p \leq 2^{ck_1^2}\}$ (Lemma 1)
$p_2, p'_2 \in P_{k_2} = \{p : \ 2^{c(k_2-1)^2} < p \leq 2^{ck_2^2}\}$

ii) $p_1 p_2 \equiv p'_1 p'_2 \pmod{q_1 \cdots q_{k_2}}$ (Lemma 2)

iii) $p_1 \equiv p'_1 \pmod{q_{k_2+1} \cdots q_{k_1}}$ (Lemma 3)

iv) $2^{2j-1} < q_j \leq 2^{2j}$ (by construction)

$$\begin{array}{ccccccc} & i) & & iii) & & iv) & \\ & \downarrow & & \downarrow & & \downarrow & \\ 2^{ck_1^2} & \geq & |p_1 - p'_1| & \geq & q_{k_2+1} \cdots q_{k_1} & > & 2^{(2k_2+1)+\cdots(2k_1-1)} = 2^{k_1^2-k_2^2} \end{array}$$

$$\implies \quad (1-c)k_1^2 < k_2^2.$$

An explicit infinite Sidon sequence

$$(1-c)k_1^2 < k_2^2 < \frac{c}{1-c}k_1^2 \quad \Longrightarrow \quad 1-c < \frac{c}{1-c}$$
$$\Longrightarrow \quad c > \frac{3-\sqrt{5}}{2} = 0.381966..$$

Corollary (C., 2012): The sequence $A_{\overline{q},c}$ is a Sidon sequence for $c = \frac{3-\sqrt{5}}{2} = 0.3819..$ with counting function

$$A_{\overline{q},c}(x) = x^{\frac{3-\sqrt{5}}{2}+o(1)}.$$

An explicit infinite Sidon sequence

If $c > \frac{3-\sqrt{5}}{2}$ then $A_{\overline{q},c}$ is not a Sidon sequence. Some repeated sums $a_{p_1} + a_{p_2} = a_{p_1'} + a_{p_2'}$ may appear.

We remove the bad a_{p_1} involved in these sums to obtain a Sidon sequence.

If $c \leq \sqrt{2} - 1$, the removed elements are not too many.

Theorem (C., 2012): For $c = \sqrt{2} - 1$ the sequence $A_{\overline{q},c}$ contains a Sidon sequence A, **which can be explicitly constructed**, with

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

If $a_{p_1} + a_{p_2} = a_{p'_1} + a_{p'_2}$

$$p_1(p_2 - p'_2) = \frac{p_1 p_2 - p'_1 p'_2}{Q_1} \cdot Q_1 + \frac{(p'_1 - p_1)p'_2}{Q_2} \cdot Q_2,$$

where $Q_1 = q_1 \cdots q_{k_2}, \quad Q_2 = q_{k_2+1} \cdots q_{k_1}.$

$$p_1(p_2 - p'_2) \in \left\{ s_1 \cdot Q_1 + s_2 \cdot Q_2 : \ |s_1| < \frac{2^{c(k_1^2+k_2^2)}}{Q_1}, \ |s_2| \leq \frac{2^{c(k_1^2+k_2^2)}}{Q_2} \right\}$$

for some $k_2, \quad k_2^2 < \frac{c}{1-c} k_1^2.$

We remove from each P_{k_1} all the primes $p_1 \in P_{k_1}$ dividing some integer of these sets.

It can be checked easily that the number of bad p_1 we have to remove is $o\left(|P_{k_1}|\right)$ for $c = \sqrt{2} - 1$.

B_h sequences

They are those sequences A such that all the sums of h elements of A are distinct. The greedy algorithm for B_h sequences gives one with

$$A(x) \gg x^{1/(2h-1)}.$$

Our approach also extends to B_h sequences:

Teorema (C., 2012) For each $h \geq 3$, **there exists** a B_h sequence A with

$$A(x) \gg x^{\sqrt{(h-1)^2+1}-(h-1)+o(1)}.$$

The cases $h = 3$ and $h = 4$ had been proved previously (C. and R, Tesoro, 2012) using a variant of Ruzsa's method, but that proof does not generalize to $h > 4$.

B_h sequences

For $h \geq 3$, our construction is not explicit. The problem is that we are not able to estimate the number of bad p_1 we have to remove from each P_{k_1} for a given basis $\overline{q}$.

We overcome this difficulty considering the probabilistic space of all basis

$$\overline{q} = h^2 q_1, \ldots, h^2 q_j, \ldots \quad \text{with} \quad 2^{2j-1} < q_j \leq 2^{2j}$$

and proving that for almost all basis $\overline{q}$ the number of bad p_1 in each P_{k_1} is $o(|P_{k_1}|)$.

The finite Sidon set that motivated our construction

Let q be a prime and g a generator of $\mathbb{F}_q^*$ and let $\log_g p$ be the discrete logarithm of p modulo q, which is unique modulo $q-1$.

Theorem (C., 2012): The set

$$\mathcal{A} = \{\log_g p : \ p \text{ prime} \ , \ p \le \sqrt{q}\}$$

is a Sidon set in $\mathbb{Z}_{q-1}$ of size $\pi(\sqrt{q}) \sim \frac{\sqrt{q}}{\log \sqrt{q}}$.

$$\begin{aligned}
\log_g p_1 + \log_g p_2 &\equiv \log_g p_1' + \log_g p_2' \pmod{q-1} \\
p_1 p_2 &\equiv p_1' p_2' \pmod{q} \\
p_1 p_2 &= p_1' p_2'
\end{aligned}$$