An Additive Problem in Different Structures

Vera T. Sòs*

1. INTRODUCTION

Additive number theory deals with the representation of positive integers as sums of terms belonging to a given $A \subseteq N$.

Many of the problems are or can be considered for arbitrary groups, semigroups or for some specified structures, like for set systems. It is not without interest to compare the analogous results and methods, often developed independently, and the difficulties of different nature which arise in the various structures. Here we illustrate this relationship by considering Sidon-type problems on different structures.

Sidon Sets

Let H be a set in which a binary operation is defined, (where this operation is not necessarily commutative or invertible). We will use the additive notation $+$.

Definition A set $S \subseteq H$ is called a Sidon set if the sums $x + y, x, y \in S$ are all different. More precisely

$$x + y \neq u + v \tag{1.1}$$

for any $x, y, u, v \in S$ of which at least three are different.

We mention a few important special cases which will be discussed below.

a) $H = \mathcal{N}$, the set of positive integers or $H = \{1, ..., N\}$ with the usual addition.

*Mathematical Institute of the Hungarian Academy of Sciences. Research was supported by the Hungarian National Foundation for Scientific Research, Grant No. 1812.

b) $H = \mathcal{N}$ or $H = \{1, ..., N\}$ and $+$ is the mod p addition.
c) H is an arbitrary or a commutative group.
d) $H = \{0, 1\}^n$ and the coordinatewise addition $+$ is the usual one:

$$1 + 1 = 2.$$

e) $H = \{0, 1\}^n$ and the coordinatewise addition $+$ is the mod 2 addition; we denote it by $\oplus$:

$$1 \oplus 1 = 0$$

(which corresponds to the symmetric difference of sets).

f) $H = \{0, 1\}^n$ and the coordinatewise addition $+$ is the Boole- addition; we denote it by 1:

$$1 + 1 = 1$$

(which corresponds to the union of sets).

g) $H = \{1, ..., N\}^n$ where the coordinatewise addition is the usual one. (i.e. the common generalisation of a) and d)).

The object of the investigation of Sidon-sets in different structures may be to determine, how "large" a Sidon-set $S \subseteq H$ can be, to construct "large" Sidon-sets, to investigate the structural properties of large Sidon- sets.

Historical Note

In 1932, Sidon [S1] in connection with his work in Fourier-analysis considered power series of type $\sum_1^\infty z^{ai}$ when $(\sum_1^\infty z^{ai})^h$ is of bounded coefficients. This led to the investigation of finite and infinite sequences (a_i) with the property that for h fixed the number of solutions of

$$a_{i1} + \cdots + a_{ih} = n$$

is bounded by K for $n \in \mathcal{N}$. Sidon-sets, as defined above correspond to the case $h = 2$ and $K = 1$.

Another source for Sidon-type questions is coding theory.

In connection with coding theory (superimposed codes) in 1964 Kautz and Singleton [KS] considered e.g. the problem of finding a large number of codewords (of 0-1 vectors) such that the pairwise (resp. h- term) sums are all different. Independently in 1969 Lindström [L1], [L2] investigated this question and its different variations of the problem. Depending on the particular problem, different additions (like Boolean sum or mod 2 sum) are considered.

The following problem of independent interest was asked by Erdös and Moser in 1969 [EM]: how large a family of subsets of an n element set can be, if all the pairwise unions (Boolean sums!) are different. In that setting the problem belongs to extremal set theory and so this also motivates different variations.

Many problems in additive number theory are studied for arbitrary groups. The investigation of Sidon-sets in arbitrary groups is motivated by and applied for the investigation of Cayley-graphs. See [B1], [B2], [BS].

The aim of this short survey is to point out the similarities or dissimilarities of the results or problems in the different structures and call attention to problems which are relevant in one but not in another structure.

Needless to say, most of the known results concern the set of integers. We shall discuss this in §3. In §4 we consider Sidon-type problems for vectors (including set systems). §5 contains the Sidon- type results and problems for groups. In §6 some generalizations, applications and open problems are mentioned.

In §2 we state some facts about Sidon-sets in general.

2. ABOUT SIDON-SETS IN GENERAL

Let $A+A := \{a+b : a, b \in A\}$. Denote by $s(H)$ the maximum size of a Sidon-set $S \subseteq H$. We use $|H|$ to denote the number of elements of H. Below $c_1, \ldots, c_k, \ldots$ will denote positive absolute constants, and in different statements their values are not necessarily the same.

The first question about Sidon-sets is: how large is $s(H)$ for different structures H. In the investigations of different structures it is not so much the methods, but their power and efficiency that vary.

However, there are simple upper and lower bounds which hold for arbitrary H.

Fact 2.1 Let S be a Sidon-set in H. If $S+S \subseteq A$, then

$$s \leq \sqrt{2}|A|^{1/2}. \tag{2.1}$$

Though this upper bound is trivial, in some cases it is the best possible one up to a constant factor, like in the cases (b) and (e) and in some particular cases of (c). In some other cases - though it is the base upper bound known - it still may be far away from the exact value of $s(H)$.

A greedy algorithm provides a simple lower bound.

Fact 2.2 Suppose that fixing any three or any two of x, y, u, v in H,

$$x + y = u + v \tag{2.2}$$

has at most k solutions in H (in the remaining variables). Then

$$s(H) \geq \frac{1}{2k}|H + H|^{1/3}. \tag{2.3}$$

In particular, if H is a commutative group and has no involution, then

$$s(H) \geq |H|^{1/3}. \tag{2.4}$$

Proof We consider only the simplest case, when H is a commutative group and has no involution. Suppose $S \subseteq H$ is a Sidon-set. If $z \in H \backslash S$ and $S \cup \{z\}$ is not a Sidon-set, then z is the solution either of

$$z + a = b + c \tag{2.5}$$

or of

$$2z = b + c \tag{2.6}$$

for some $a, b, c, \in S$. Let $s = |S|$. The equations (2.5) and (2.6) exclude at most $s(s-1)(s-2) + s(s-1)$ elements. Therefore, if $|S| < |H|^{1/3}$, we must have an element z such that $S \cup \{z\}$ is also a Sidon-set.

We do not know any particular structure where this lower bound gives the exact answer or even the order of magnitude of $s(H)$. However, there are structures where not much more is known (e.g. (c)).

Remark 2.3 Up to now, we have only considered the simplest common definition of Sidon-sets. In some structures or in connection with some particular problems it is more appropriate to consider Sidon-sets where we require (1.1) only for $x, y, u, v \in S$ where all four are different, or when $x \neq y, u \neq v$. In the non-commutative case there are further variations of the definition. These modifications may change the situation quite significantly.

3. SIDON-SETS OF INTEGERS

A sequence (a_i) of positive integers (finite or infinite) is called a Sidon-sequence (or B_2 sequence) if the sums $a_i + a_j$ (or what is the same, the $a_i - a_j$ differences) are all different.

Let $s(n)$ denote the largest number k for which there exists a Sidon-sequence $1 \leq a_1 < \cdots < a_k \leq n$.

By the simple arguments in §2,

$$n^{1/3} < s(n) < \sqrt{2n}^{1/2}.$$

The asymptotically best possible result is given in

Theorem 3.1 $[ET], [ECh]$ There exists that for all n, constants $c_1, c_2 > 0$ such that

$$n^{1/2} - c_1 n^{5/6} < s(n) < n^{1/2} + c_2 n^{1/4}. \tag{3.1}$$

For infinitely many n

$$n^{1/2} < s(n). \tag{3.2}$$

The lower bound was proved independently by Erdös and Chowla [Ch] in 1944, using Singer's result.

Theorem 3.2 $[S]$ If q is a power of prime, there exist integers $1 \leq a_1 < \cdots < a_{q+1} \leq q^2 + q + 1$ such that the $a_i - a_j$ $(i \neq j)$ differences are all distinct mod $(q^2 + q + 1)$.

Consequently the $q^2 + q$ differences represent all the nonzero residues mod $(q^2 + q + 1)$. Such sequences are called perfect difference sets. Obviously, a perfect difference set is also a Sidon-set.

The upper bound and the weaker lower bound $(\frac{1}{2}+o(1))n^{1/2}$ was proved earlier, in 1941 in Erdös-Turán [ET] which is the first paper about Sidon-sequences. The construction in [ET] for the lower bound is probably the origin of all the known constructive lower bounds for analogous problems in other structures. This is the following.

Construction 3.3 Let p be a prime and let

$$a_l = 2pl + (l^2) \quad \text{for} \quad l = 1, 2, \ldots, p-1 \tag{3.3}$$

where (x) stands for the least positive residue of $x \pmod p$. From

$$a_1 + a_j = a_r + a_s$$

also

$$i + j = r + s$$

and

$$i^2 + j^2 = r^2 + s^2$$

would follow, which is impossible, if $\{i, j\} \neq \{r, s\}$.

Komlós-Sulyok-Szemerédi [KSSz] proved the following more general and not obvious result:

Theorem 3.4 For an arbitrary n-element set $C \subseteq N$ there exists a Sidon-sequence $S \subseteq C$ such that

$$|S| > c\sqrt{n}.$$

Infinite Sidon-sequences

Let $S = \{1 \leq a_1 \leq ...\}$ be an infinite Sidon-sequence and let $A_s(n)$ denote the number of terms $\leq n$ in S. The question is, how fast can $A_S(n)$ grow. In the infinite case much less is known than in the finite case. Beyond what follows from Theorem 3.1, Erdös proved the

Theorem 3.5 (See [St] There is an absolute constant $c > 0$, such that for every (infinite) Sidon-sequence S

$$A_S(n) < c_1(n/\log n)^{1/2} \tag{3.4}$$

holds infinitely often.

On the other hand Krickerberg, improving a result of Erdös, proved in 1961 the

Theorem 3.6 [K] There is an (infinite) Sidon-sequence S such that

$$A_S(n) > \frac{1}{\sqrt{2}}\sqrt{n} \tag{3.5}$$

holds infinitely often.

It is not known whether or not the factor $\frac{1}{\sqrt{2}}$ is best possible. The greedy algorithm gives the existence of an (infinite) Sidon-sequence for which

$$A_S(n) > n^{1/3} \quad \text{for all } n. \tag{3.6}$$

It was a great achievement when after 30 years, in (1981) Ajtai, Komlós and Szemerédi improved this:

Theorem 3.7 [AKSZ] There is a Sidon-sequence S such that

$$A_S(n) > c(n \log n)^{1/3} \quad \text{for all } n \geq n_o. \tag{3.7}$$

The proof uses the "probabilistic method with deletion" and is based on the following Ramsey-Turán type graph theorem.

Theorem 3.8 [AKSZ] If G is a triangle free graph with n vertices e edges and with average degree $t = 2e/n$, then it has an independent set of size $> c \cdot n/t\log t$.

(Without the requirement that G is triangle free, Turán's theorem yields the lower bound $c\frac{n}{t}$. The above result also led to the best known upper bound of the Ramsey-function: $R(3,k) < ck^2/\log k$).

Though substantial progress has been made with the density question of infinite Sidon-sequences, there is a major gap between the known lower and upper bounds. It is not known whether or not any of (3.4), (3.5) or (3.7) can be improved. As we will see, such an $n^{1/3}$ vs. $n^{1/2}$ (or n^{α} vs. n^{β}) gap is quite typical also in other structures. In most structures the situation is even worse, there is such a gap in the finite case too. (While for integers the finite case is, at least asymptotically, settled.)

4. SIDON-SETS OF VECTORS

Superimposed codes were introduced by Kautz and Singleton in 1964 [KS]. We quote from this:

"A binary superimposed code consists of a set of code words whose digit by digit Boolean sums enjoy a prescribed level at distinguishability. These codes find their main application in the representation of document attributes within an information retrieval system, but might also be used as a basis for channel assignments to relieve congestion in crowded communications bands."

One of the basic definitions is

Definition 4.1 An $m \times n$ $0-1$ matrix A is called a (k,m) superimposed code of length n, if all the Boolean sums composed of k different rows of A are different.

The problem is to determine the minimal possible n of a (k,m) code. Obviously this is dual formulation of a Sidon-type problem. For further reference about superimposed codes see the "Survey of superimposed code theory" by Dyachkov-Rykov [DR] and [DDR]. We already mentioned in the introduction that for 0-1 vectors we may define the coordinate-wise addition several different ways. We will consider the following three additions:

1) the Boolean-sum, when $1+1=1$,

2) the mod 2-sum, when $1 \bigoplus 1 = 0$,

3) the usual addition, when $1+1=2$.

Concerning the problem of Sidon-sets of maximum size in the corresponding three structures, the first one is more difficult than the second and third. In the last two, the problem is solved asymptotically, while in the first there is a significant gap between the best known upper bound and lower bound.

1. Boolean sum (union of sets)

Recall, that $S \subseteq \{0,1\}^n$ is a Sidon-set, if

$$a+b \neq c+d \text{ for } a,b,c,d, \in S \tag{4.1}$$

for every $a,b,c,d, \subseteq \{0,1\}^n$ where at least three of these are different.

Obviously $\{0,1\}^n$ corresponds to the family of all subsets of $\{1,...,n\}$ and the Boole-sum corresponds to the union of two sets.

A family S of subsets of $\{1,...,n\}$ is a *Sidon-family*, if

$$A \cup B \neq C \cup D \tag{4.2}$$

where at least three of A,B,C,D are different. (In [FF1] it is called union-free family.)

Denote by $f(n)$ the maximum number m of vectors in a Sidon-set $S \subseteq \{0,1\}^n$.

Theorem 4.1

$$2^{n-3)/4} \leq f(n) \leq 1 + 2^{(n+1)/2} \tag{4.3}$$

The lower bound was proved in [KS] 1964 and independently in [FF1] in 1984.

The upper bound is trivial, since all the $\binom{f(n)}{2}$ sums must be different. The lower bound in [FF1] is proved by a construction which is a modification of the Erdös-Turán construction. In [KS] parity-check matrices are used. If H^t is the transpose of an $n \times m$ parity check matrix, then the $0 \to 01, \to 10$ substitution leads to a matrix A where the rows form a Sidon-set in $\{0,1\}^{2n}$.

Call a set $S^* \subseteq \{0,1\}^n$ a *weak Sidon-set*, if (4.1) holds for every *four* distinct $a, b, c, d, \in \{0,1\}^n$. Denote by $(F(n)$ the maximum number m of vectors in a weak Sidon-set $C^* \subseteq \{0,1\}^n$. It is clear that

$$f(n) \leq F(n). \tag{4.4}$$

The best known result for $F(n)$ is the following

Theorem 4.2

$$2^{(n-\log\ 3)/3} \leq F(n) \leq (1+o(1))2^{\frac{3}{5}n}. \tag{4.5}$$

The lower bound, as well as a weaker $2^{\frac{3}{3}n}$ upper bound was proved in Frankl-Füredi [FF1], the upper bound in Lindström [L3].

In the lower bound the "probabilistic method with deletion" is used; we take a random set of vectors with appropriate probability and omit one vector from each "bad" four-tuple, which does not satisfy (4.1).

The upper bound is proved by the following idea:

Let $v_1, ..., v_m$ be n-dimensional vectors. Split each v into two vectors, u and w of dimension d and $n-d$, let $v = (u, w)$. For a fixed d dimensional u_i let $w_i^{(1)}, ..., v_i^{(k)}$ be all the $n-d$ dimensional (complementing) 0-1 vectors such that $(u_i, w_i^{(j)}), 1 \leq j \leq k$ belong to our Sidon-set. Consider the $w_i^{(j)} - w_i^{(l)}$ differences, which are $-1, 0, 1$ vectors. The observation , that for a fixed i at each coordinate at least half of the differences are 0, and that all the $w_i^{(j)} - w_i^{(l)}$ differences are distinct, with the choice $d = \lfloor \frac{n}{3} \rfloor$ this gives the upper bound.

Bollobás [B1] considered the analogous question for k-uniform hypergraphs, i.e. for 0-1 vectors of constant weight.

In Frankl-Füredi [FF2] this "uniform" case is *solved* asymptotically. They consider the following three different versions. Let $f_k(n)$ resp. $F_k(n)$ denote the maximum number m of vectors in $\{0,1\}^n$ of constant weight k in a Sidon-set $\subseteq \{0,1\}^n$ resp. in a weak Sidon-set $\subseteq \{0,1\}^n$.

Let $H_k(n)$ denote the maximum number m of vectors in $\{0,1\}^n$ and constant weight k in a set $\mathcal{H} \subseteq \{0,1\}^n$ such that

$$x + y \neq z + w$$

where the addition means the usual one $(1+1=2)$ for any *four* distinct $x, y, z, w \in \mathcal{H}$. (For any two pairs of sets in the family either the unions or the intersections are different.) Obviously,

$$f_k(n) \leq F_k(n) \leq H_k(n).$$

Using the less obvious inequality

$$\frac{k!}{k^k} H_k(n) \leq f_k(n), \qquad ((4.6))$$

it suffices to get bounds for $H_k(n)$.

Theorem 4.3 [FF1] $c_k n^{\lceil 4k/3 \rceil/2} \leq H_k(n) \leq c'_k n^{\lceil 4k/3 \rceil/2}$.

For the lower bound Frankl and Füredi use an ingenious construction based on a theorem they prove about systems (k-tuples) which are solutions of some equations for symmetric polynomials over finite fields.

(4.6) follows from a theorem of Erdös-Kleitman [EK] which states that every k-uniform hypergraph $\mathcal{H}$ contains a k-partite $\mathcal{H}' \subseteq \mathcal{H}$ with $|\mathcal{H}'| \geq k!k^{-k}|\mathcal{H}|$. Other generalizations and applications for information theory (for certain search problems) are considered in [HS].

2. mod 2 sum

Consider $\{0,1\}^n$ with mod 2 addition. This corresponds to the symmetric difference of sets.

Let $S \subseteq \{0,1\}^n$ denote a Sidon-set and $h(n)$ be the maximum number of vectors in a Sidon-set S.

Theorem 4.4 Lindström [L1]

$$\lim_{n \leftarrow \infty} {}^n\sqrt{h(n)} = 2^{1/2}. \tag{4.7}$$

The upper bound $\overline{lim}_n \, {}^n\sqrt{h(n)} = 2^{1/2}$ follows from 2.1 (and is trivial). The lower bound follows from a construction which is a modification of the Erdös-Turán construction [ET1]: Let S be the set of all vectors (u, u^3) with $u \in GF(2^n)$. Obviously S is a set of 0-1 vectors of length $2n, |S| = 2^n$ and S is a Sidon-set in $\{0,1\}^{2n}$.

Remark 4.5 Since $v_1 + v_2 \equiv v_1 + v_3$ mod 2 implies $v_2 = v_3$ and $v_1 + v_1 \equiv v_2 + v_3$ mod 2 cannot hold for $v_2 \neq v_3$, in this structure Sidon-sets and weak Sidon-sets are the same.

3. Addition in $\mathbf{R}^n$

Instead of $\{0,1\}^n$ we consider more generally the set of vectors $\{0,1,\cdots,N-1\}^n$ and the usual addition $+$. Let $S_N(n)$ denote a Sidon-set in $\{0,...,N-1\}^n$ and $s_N(n)$ denote the maximum number of vectors in a Sidon-set $S_N(n)$.

Theorem 4.6 [L1] For fixed dimension n

$$s_N(n) \leq N^{\frac{n}{2}} + 0(N^{n^2/(2n+2)}) \text{ as } N \to \infty. \tag{4.8}$$

For $n = 1$ this is the Erdös-Turán inequality for Sidon sequences of integers.

Remark 4.7 For $N \geq 2$ the trivial upper bound is only

$$s_N(n) < \sqrt{2} N^{\frac{n}{2}}. \tag{4.9}$$

The lower bound

$$s_N(n) > (1 + o(1))N^{\frac{n}{2}} \text{ for } n \leftarrow \infty \tag{4.10}$$

follows from the corresponding lower bound for $n = 1$. I.e., consider the N-ary expansion of the integers $m \in [0, N^n - 1] \; : \; m = \sum_{v=0}^{n-1} x_v(m)N^\nu$. The bijection $\varphi \; : \; [0, N^n - 1] \rightarrow [0, ..., N - 1]^n$, defined by

$$\underline{\varphi}(m) = (x_1(m), ..., x_{n-1}(m))$$

obviously has the property, that

$$\varphi(a) + \varphi(b) = \varphi(c) + \varphi(d)$$

implies $a + b = c + d$.

However the converse implication does not hold, hence (4.8) does not follow from the classical result for $n = 1$.

Remark 4.8 In this case - though Sidon-sets and weak Sidon-set are not the same the maximum sizes are asymptotically the same.

Since now

$$x + y = x + z$$

implies $y = z$, the only difference is that

$$2x = y + z$$

or

$$z - x = x - y$$

is excluded in Sidon-sets, but permitted in weak Sidon-sets. For each x we have at most one such pair (y, z); there are at most $0(N^{\frac{n}{2}})$ differences, which may occur twice. This will not change the order of magnitude of the maximal size.

Remark 4.9 The problem of Sidon-sets resp. weak Sidon-sets is related to anti- Ramsey type problems. If V is the set where a commutative addition $+$ is defined, consider the complete graph $\mathcal{H}$ where $V(\mathcal{H}) = V$ and with the colouring $\varphi : [V]^2 \to V$ where $\varphi(a, b) = a + b$. A Sidon-set $S \subseteq V$ is the vertex set of a TMC (totally multi-coloured) complete subgraph. Weak Sidon-sets $S^* \subseteq V$ correspond to the vertex set of complete subgraphs where only independent edges must have different colours.

If $(V, +)$ is an Abelian group then φ is a "good" edge-colouring; adjacent edges have different colours.

The edge-colourings belonging to the structures $\{0, 1, ..., N-1\}^n$, $+$ and $\{0, 1\}^n$ with the mod 2 addition are obviously "good" edge-colourings, but $\{0, 1\}^n$ with the Boolean sum does not lead to a "good" edge-colouring.

In order to obtain nontrivial information about Sidon-sets in a certain structure we must use some additional information about the coloring given by the addition.

In [ENR1] anti-Ramsey type questions are considered for "good" edge colourings. For the special case of complete graphs (which is related to Sidon-sets) Babai proved the following:

Theorem 4.10 [B1] Let φ be a "good" edge colouring of K_n. Let $r_\varphi(n)$ denote the maximum size of a TMC K_t in that colouring. Put $r(n) = \min_\varphi r_\varphi(n)$. Then

$$(2n)^{1/3} < r(n) < 8(n \log n)^{1/3}.$$

5. SIDON-SETS IN GROUPS It is natural to ask analogous questions for groups (see Babai-Sós [BS]). Beside its own interest, these are motivated also by some applications for Cayley-graphs. Embedding graphs as induced subgraphs in Cayley-graphs was investigated first in Babai [B1], [B2]. Results for Sidon sets in groups provide an improvement of these results ([BS]).

There is a natural way to generalize the notion of Sidon-sequences to Abelian groups. However there are two natural ways in the non-Abelian case.

Definition 5.1 Let G be an *Abelian* group. We call $S \subseteq G$ a Sidon-set if for any $x, y, z, w \in S$ *at least three* of which are different

$$x + y \neq z + w \tag{5.1}$$

(or equivalently, $x - z \neq w - y$.)

Definition 5.2 Let G be a group, $S \subseteq G$ is a Sidon-set of the first kind if for any $x, y, z, w \in S$ at least three of which are different,

$$xy \neq zw. \tag{5.2}$$

Definition 5.3 $S \subseteq G$ is a Sidon set of the second kind if

$$xy^{-1} \neq zw^{-1} \tag{5.3}$$

for any $x, y, z, w \in S$ at least three of which are different.

Remark 5.1 In fact, Sidon-sets of *second kind* are relevant in the problem of embedding graphs as induced subgraphs in Cayley-graphs, Sidon-sets of first kind do not seem to have any relevance to these problems.

Remark 5.2 Let G be an arbitrary group of order n and S be a Sidon-subset of either kind. For $s = |S|$ the upper-bound $s < \sqrt{2n}^{1/2}$ as in general follows trivially from the inequality $\binom{s+1}{2} < n$. On the other hand, the best known lower bound is max $s > cn^{1/3}$ and only for very particular Abelian groups max $s > c\sqrt{n}$ is known.

Observe, that here even the $cn^{1/3}$ lower bound is not quite trivial, the simple greedy algorithm cannot be applied in general. It would be, if we consider weak Sidon-sets, if we exclude only 4-distinct elements which satisfy (5.1), (5.2) or (5.3). However, for a given pair a, b which belongs to our set

$$\begin{gathered} x^2 = ab \\ \text{or } ax = xb \\ \text{resp. } ax^{-1} = xb^{-1} \end{gathered} \tag{5.4}$$

may hold for arbitrary many x, so we cannot apply (2.2).

Theorem 5.3 [BS]. Let G be a (finite or infinite) group and $W \subset G$ a finite subset with $|W| = n$. Then W contains Sidon subsets of both kinds, of size

$$(c + o(1))n^{1/3}$$

where $c = 3^3\sqrt{2}/8 \sim 0,47$. Furthermore, W contains a subset of size $(3/8 + o(1))n^{1/3}$ which is simultaneously a Sidon subset of both kinds.

The proof follows from a proposition that sparse hypergraphs have large independent sets.

For some particular classes of Abelian groups the gap in the exponent between $1/2$ and $1/3$ can be closed, the lower bound being $cn^{1/2}$.

Theorem 5.4 [BS] Let q be a prime power and G the elementary Abelian group of order q^2. Then G has a Sidon-subset of size q.

The proof follows by a slight modification of the original Erdös-Turán construction [ET].

Let $F = F_q$ denote the field of q elements, and G be identified with the additive group of the 2-dimensional space F^2 over F. The set

$$S = \begin{cases} \{(x, x^2), x \in F\} & \text{if } q \text{ is odd,} \\ \{(x, x^3), x \in F\} & \text{if } q \text{ is even,} \end{cases}$$

is a Sidon-set.

If q is odd, this is best possible, since $s(s-1) \leq n-1$ must hold. For elementary Abelian 2-groups, when we have involutions there is a gap of a factor $\sqrt{2}$, since only $\binom{s}{2} \leq n-1$ is the trivial, but best known upper bound.

An improvement of the above idea gives an analogously good construction for some further class of Abelian groups. A more general, but still too specific result is the

Theorem 5.5 [BS] Elementary Abelian groups of order n have Sidon sets of size $n^{1/2+o(1)}$.

For infinite groups we mention only the

Theorem 5.6 [BS] Any infinite subset of a group contains an infinite subset which is a Sidon-subset of both kinds simultaneously.

This can be proved by a simple argument using Ramsey's theorem.

6. SOME GENERALIZATIONS AND OPEN PROBLEMS

1. Extremal problems for systems of solutions of homogeneous linear equations

In extremal graph theory a basic problem is the following: given a family of graphs $\mathcal{L}$ determine $\mathrm{ex}(n;\mathcal{L})$, the maximum number of edges in a graph on n vertices not containing any member of $\mathcal{L}$.

Besides the many results for some specific classes of $\mathcal{L}$, the Erdös-Stone-Simonovits type results ([ES]), [ESS]) have a particular importance and carry some information about the hierarchy. According to these results the order of $\mathrm{ex}(n;\mathcal{L})$ is determined by $\chi_0 = \max_\chi(L); ex(n;\mathcal{L}) = (1+o(1))\frac{1}{2}\frac{\chi_0^{-2}}{\chi_0^{-1}}n^2$; the chromatic number is the relevant parameter in these extremal problems.

For $\chi = 2$ this gives $f(n;\mathcal{L} = o(n^2)$. One of the most difficult open problems in extremal graph theory concerning the order of $f(n;\mathcal{L})$ is to find the hierarchy in the class of bipartite graphs. (When is $f(n;\mathcal{L}_1) < f(n;\mathcal{L}_2)$?).

There is a very extended subject in extremal hypergraph theory or in extremal set theory where analogous problems are considered.

Many classical or recent results and problems in number theory can be considered as extremal problems for integers of the above type. Very often the difficulty of the specific problem depends heavily ont he arithmetic structure of integers. Though various beautiful results are known, we should have a better understanding of the general phenomena. Our point here is to formulate some problems which arose as generalizations of some specific results and may help in the clarification.

Consider homogeneous system of linear equations

$$\sum_{i=1}^{k} a_i^v x_i = 0, \nu = 1, ..., l \tag{6.1}$$

with integer coefficients. Let A denote the matrix in (6.1).

Let $f_A(n)$ resp $f_A^*(n)$ denote the maximum size of a subset $B \subseteq \{1, ..., n\}$ such that the system (*) has no distinct solution (with $x_1, ..., x_k$ all different) resp. non constant $(x_1 = \cdots = x_k = c)$ solution in B.

How does f_A depend on the system A? A few special cases have been investigated intensively. The deepest one is the case of

Arithmetic progressions: The systems of solutions of the system

$$x_{i+1} - 2x_i + x_{i-1} = 0 \text{ for } 1 \leq i \leq k-1 \tag{6.2}$$

are the $k+1$-term arithmetic progressions. The celebrated theorem (Roth [R] for $k = 2, (x + y - 2z = 0)$ and of Szemerédi [Sz] $k > 2$ states that for the system in (6.2)

$$f(n) = o(n) \tag{6.3}$$

or more precisely, for $k = 3$

$$ne^{-c\sqrt{\log n}} < f(n) < n(\log \log n)^{-\alpha}. \tag{6.4}$$

A wea~~l~~Ker upper bound is in Roth [1] and in Szemerédi [SZ], which was improved by Szemerédi and by Hea~~ble~~th Brown. The lower bound is given by a construction in Behrend [Be].

Another special case is the *Sidon-problem* we discussed above, with one equation

$$x + y - z - w = 0 \tag{6.5}$$

and with $f(n) \sim \sqrt{n}$. The so-called sum *free sets*, when the solutions of

$$x + y - z = 0 \tag{6.6}$$

are excluded is a special case when $f(n) > cn$. In fact, it is easy to see that $f(n) = [\frac{n+1}{2}]$.

As stated also in [KSSz], the theorem below gives a general estimation.

Theorem 6.2 [KSSz] $f_A(n) = o(n)$ if and only if the system of solutions of (6.1) is translation invariant, i.e. if and only if

$$\sum_{i=1}^{k} a_i^{(\nu)} = 0 \quad \text{for} \quad \nu = 1, ..., l.$$

It is easy to see, that for one equation of the form

$$a_1x_1 + \cdots + a_kx_k = a_1y_1 + \cdots + a_ky_k$$

$f(n) = 0(n^{\frac{1}{k}})$.

On the other hand, a straightforward modification of Behrend's construction and argument gives, that for one equation of form

$$a_oy = a_1x_1 + \cdots + a_kx_k \text{ with } a_i > 0 \text{ for } 0 \leq i \leq k$$

$f(n) > n^{1-\epsilon}$ for all $\epsilon > 0$, and for $n > n_o(\epsilon)$.

Consider this class of systems, where $f(n) = o(n)$.

Problem 1 When is $f(n) >> n^{1-\epsilon}$ for all $\epsilon > 0$ (like for arithmetic progressions), and when is

$$f(n) << n^{1-\alpha}$$

with some $\alpha > 0$ (like for the Sidon-problem). Does there exist any $\frac{1}{2} < \alpha < 1$, such that for some system $Ax = 0, f(n) = \Omega(n^\alpha)$?

2. Representation function

For a given sequence $A = (a_i)$ let $R_A(n)$ denotes the number of solutions of

$$a_i + a_j = n \text{ with } a_i \leq a_j.$$

$R_A(n)$ is called the representation function of A.

Let S be a Sidon-sequence in [1,n]. Consider the 0-1 sequence given by the representation function R_S.

Problem 2

a) At least how many blocks of 1's are in $(R_S(k))$?

b) How long can blocks of 1's in $(R(k))$ be? Suppose $R_S(k) = R_S(k+1) = \dots = R_S(k + f(k)) = 1$. How large can $f(k)$ be?

c) How large can $f(k)$ be, if $|S| > \sqrt{n}$?

d) How long must blocks of 0's in $R_S(k)$ be?

4. Sidon-sets in families of sets

Let $\mathcal{A} \subseteq \{0,1\}^n$ and $F(m;\mathcal{A}), f(n;\mathcal{A})$ resp. $h(n;\mathcal{A})$ denote the maximum number of vectors in a Sidon-set $S(n;+) \subseteq \mathcal{A}$, in a weak Sidon- set $S(n;+) \subseteq \mathcal{A}$ resp. in Sidon-set $S(n;\oplus) \subseteq \mathcal{A}$.

Put

$$F(n;m) = \min_{|\mathcal{A}|=m} F(n;\mathcal{A})$$

$$f(n;m) = \min_{|\mathcal{A}|=m} f(n;\mathcal{A})$$

and

$$h(n;m) = \min_{|\mathcal{A}|=m} h(n;\mathcal{A}).$$

Determine or estimate $F(N;m), f(n;m), h(n;m)$. It is easy to see that

$$m = \sum_{i=1}^{r} \binom{n}{i}$$

then

$$F(n;m) \le 2^r$$

and

$$f(n;m) \le 2^r$$

Is it true, that

$$F(n; m) > 2^{cr}$$

or

$$f(n; m) > 2^{cr}?$$

5. The structure of $\mathcal{A}+\mathcal{A}$

There is extended literature on the different properties of sum sets $A+A$ and difference sets $A-A$. We mention only the deep results of Freiman [F] on the structure of $A+A$. If $A+A$ is small, close to the minimum possible value (which is attained, when A is an arithmetic progression), then A is contained in a short arithmetic progression.

Let $\mathcal{A} - \{A_1, ..., A_m\}$ be a family of subsets of an n-element set S. Put $\mathcal{A} \cup \mathcal{A} = \{A_i \cup A_j; 1 \le i < j \le m\}$.

How does $|\mathcal{A} \cup \mathcal{A}|$ constrain the structure of $\mathcal{A}$?

Obviously,

$$1 \le |\mathcal{A} \cup \mathcal{A}| \le \binom{|\mathcal{A}|}{2}$$

The upper bound is attained if and only if $\mathcal{A}$ is a Sidon-set. $|\mathcal{A} \cup \mathcal{A}| = 1$ implies $|\mathcal{A}| \le n$ and the unique A extremal family is $A_i = S - \{i\}, 1 \le i \le n$.

Problem 3 Put

$$f(m) := \min_{|\mathcal{A}|=m} |\mathcal{A} \cup \mathcal{A}|.$$

It is especially interesting to estimate $f(m)$ for $m \sim (1+\delta)^n$ or for $m \sim c2^m$ for $\delta > 0$, $c > 0$. Is it true that $f(c2^m) \ge c2^m$?

Erdös observed that

$$|\mathcal{A} \cup \mathcal{A}| \ge \frac{|\mathcal{A}|}{n+1}$$

always holds. Most probably this can be improved for $|\mathcal{A}| > cn$.

For a given $|\mathcal{A}|$, when is $|\mathcal{A} \cup \mathcal{A}|$ "small"?

6. EXTREMAL PROBLEMS AND RAMSEY PROBLEMS

Let $\underline{Ax} = 0$ be a given system. We know that $f_A(n) = o(n)$ if and only if $A1 = 0$. This "density" property obviously implies the Ramsey-property:

$$\text{For any } r \text{ and } n > n_o(A,r) \text{ at any } r \text{ coloring } \varphi : \{1,...,n\} \to \{1,...,r\}$$
$$\text{there is a monochromatic solution of} \underline{Ax} = 0. \tag{6.10}$$

(For further reference see [GRS]). Let $N(A,r)$ denote the smallest integer m such that property (6.10) holds with $n > m$.

$N(A,r)$ was investigated intensively for arithmetic progressions. The best known upper bound is an extremely rapidly growing function. It was a breakthrough in 1988 when Shelah gave a new proof for Van der Waerden's theorem. This also gives an important improvement of the estimate for $N(A,r)$ for k-term arithmetic progressions. To determine $N(A,r)$ in general is just as difficult as to determine $f_A(n)$ in general. But is the hierarchy the same for $(N;r)$ as for $f_A(n)$? There is more experience about the analogous question for graphs where the answer is no, and very probably this is the situation here. We formulate the problem more precisely:

Problem 4 Give two systems $A_1x = 0$ and $A_2x = 0$ such that

$$f_A^1(n)_{A_1} << f_{A_2}(n)$$

and

$$N(A_1,r) >> N(A_2,r).$$

7. THE STRUCTURE OF SOLUTIONS IN LARGE $B \subseteq \{1,...,n\}$

Let $Ax = 0$ be a system of homogeneous linear equations. Suppose $B \subseteq \{1,...,n\}$ and $|B| > f_A(n)$.

Let $\mathcal{H}_B = \{x = (x_1,...,x_k); Ax = 0, \{x_1,...,x_k\} \subset B\}$. Put

$$S_A(n;m) = \min_{|B|=m} |\mathcal{H}_B|.$$

Problem a. Obviously $S_A > m - f_A(n)$.

Determine $s_A(n;m)$.

Problem 6 What is the structure of $\mathcal{H}_B$.

Consider the particular case

$$x + y - z = 0.$$

The sumgraph of $B \subseteq \{1,...,n\}$ is the graph $\mathcal{G}(B;E)$ where $(x,y) \in E \ iff \ x + y \in B$. With the above notation $|E| = |\mathcal{H}_b|$. What can we say on the structure of $\mathcal{G}$? We formulate just one particular conjecture:

Problem 7 Is it true that, if $B > \frac{5}{8}n$, then $\mathcal{G}(B,E)$ contains a triangle? (I.e. a solution of the system

$$\begin{aligned} x + y &= u \\ y + z &= v \\ x + z &= w \end{aligned} \tag{6.11}$$

with $x, y, z, u, v, w, \in B$.)

P. Erdös observed, that if it is true, then it is a sharp result (and gives the exact value of $f_A(n)$ for the system (6.11).

REFERENCES

[AEKSz] M. Ajtai, P. Erdős, J. Komlós, E. Szemerédi: On Turán's theorem for sparse graphs, Combinatorica 1(1981) 313-317

[AKSz] M. Ajtai - J. Komlós - E. Szemerédi: A dense infinite Sidon- sequence

[A] N. Alon: Independent sets in regular graphs and sum-free subsets of finite graphs (to appear)

[AE] N. Alon - P. Erdős: An application of graphtheory to additive numbertheory. Europ. J. Comb. 6 (1985) 201-203

[AK] N. Alon - D. J. Kleitman: Sum-free subsets. Combinatorics 1988, Cambridge Univ. Press (Ed. A. Baker, B. Bollobás, A. Hajnal)

[B1] L. Babai: Embedding graphs in Cayley-graphs. Probl. Comb. Theorie des Graphes, Proc. Conf. Paris-Orsay 1976, J. C. Bermond el al.eds. C.N.R.S. Paris 1978, 13-15.

[B2] L. Babai: Chromatic number and subgraphs of Cayley graphs. Theory and Applcations of Graphs, Y. Alavi, R.R. Lick eds, Springer Lect. Notes in Math. 642., Springer Verlag Berlin, 1978, 10-22.

[B3] L. Babai: An anti-Ramsey theorem. Graphs and Combinatorics 1, (1985) 23-28

[BS] L. Babai - V. T. Sós: Sidon sets in groups and induced subgraphs of Cayley Graphs. Europ. J. Comb. 6 (1985) 101-114

[Ch] S. Chowla: Solution of a problem of Erdős and Turán in additive numbertheory. Proc. Math. Acad. Sci. India 14 (1944) 1-2

[DE] M. Deza, P. Erdős: Extension de quelques theoremes sur les densites de series d'elements de $\mathcal{N}$ a de series de sons ensembles finis de $\mathcal{N}$. Discr. Math. 12 (1975) 295-308

[DRR] A. G. Dyachkov, V. V. Rykov, A. M. Rashed: Superimposed distance codes. Problems of Control and Inform. Theory (to appear)

[DR] A. G. Dyachkov, W. V. Rykov: A survey of superimposed code theory. Problems of Control and Information Th. 12 (1983) 229-242

[DRR] Superimposed distance codes. To appear in "Problems of Control and Information Theory"

[E1] P. Erdős: Some application of Ramsey's theorem to additive numbertheory. Europ. J. Comb. 1 (1980) 43-46

[E2] P. Erdős: Some problems on additive numbertheory, Annals of Discr. Math. 12 (1982) 113-116

[E3] P. Erdős: Some old and new problems on additive and combinatorial number theory. Annals New York Acad. of Sci. (????) 181-186

[EFF] P. Erdős, P. Frankl, Z. Füredi: Families of finite sets in which no set is covered by the union of r others. Israel J. of Math. 51 (1985) 79-89

[EF] P. Erdős - R. Freud: On sums of a Sidon sequence (to appear)

[EM] P. Erdős - L. Moser: Problem 35. Proc. Conf. Comb. Structures and Appl. Calgary, 1969, Gordon and Breach, New York, 1970, 506.

[ENR] P. Erdős, J. Nesetril, V. Rödl: On some problems related to partitions of edges of graphs (Graphs and other combinatorial Topics, Proc. 3rd Czechoslovak Symposium on Graph Theory, Prague, 1982, ed. Fiedler.) 54-63, Teubner Texte in Math., 59, Leipzig 1983.

[ERT] P. Erdős - I. Ruzsa - H. Taylor: Bounds for arrays of dots with distinct slopes (to appear)

[ESS] P. Erdős - A. Sárközy - V. T. Sós: Problems and results on additive properties of general sequences III. Studia Sci. Math. Hung. 22 (1987) 53-63

[F] Z. Füredi: Hypergraphs in which all disjoint pairs have distinct unions. Combinatorica 4 (1984) 161-168

[FGR] P. Frankl, R. L. Graham - V. Rödl: On subsets of abelian groups with no 3-term arithmetic progression. ???

[FF1] P. Frankl - Z. Füredi: A new extremal property of Steiner-triple systems. Discr. Math. 48 (1984) 205-212

[FF2] P. Frankl - Z. Füredi: Union-free hypergraphs and probability theory. Europ. J. Comb. 5 (1984) 127-131

[FF3] P. Frankl - Z. Füredi: Union-free families of sets and equations over fields. J. Numberth. 23 (1986) 210-218

[GRS] R. L. Graham, B. L. Rothschild, J. Spencer: Ramsey Theory. Wiley Intersc. Ser. in Discrete Math. 1980, 1989.

[GS] R. L. Graham - N. J. A. Sloane: Lower bounds for constant weight coads. IEEE Trans. Inform. Theory IT-26 (1980) 37-43

[KS] W. H. Kautz - R. C. Singleton: Nonrandom binary superimposed codes. IEEE Trans. on Information Th. 10 (1964) 363-377

[KSSz] J. Komlós, M. Sulyok, E. Szemerédi: Linear problems in combinatorial numbertheory. Acta Math. Acad. Sci. Hung. 26 (1975) 113-121

[K] F. Krückeberg: B_2-Folgen und verwandten Zahlenfolgen. Z. reine angew. Math. 206 (1961) 53-60

[L1] B. Lindström: Determination of two vectors from the sum. J. Comb. Th. 6 (1969) 402-407

[L2] B. Lindström: An inequality for B_2-sequences. J. Comb. Th. 6 (1969) 211-212

[L3] B. Lindström: On a combinatory detection problem I. ???? 195-

[L4] B. Lindström: On B_2-sequences of Vectors. J. Comb. Th. 4 (1972) 261-265

[N] M. B. Nathonson: Representation functions of sequences in additive number theory. Proc. Amer. Math. Soc. 72 (1978) 16-20

[Ra] R. Rado: Verallgemeinerung eines Satzes von van der Waerden mit anwendungen auf ein Problem zum Zahlentheory. Sonderausg. Sitzungber. Preuss. Akad. Niss(????). Phys. Math. Klasse 17 (1933) 1-10

[Ro] F. Roth: Sur quelques ensembles d'Entiers. C. R. Acad. Sci. Paris 234 (1952) 388-390.

[Ru] I. Z. Ruzsa: A just basis (to appear in the Monatshefte für Math.)

[S1] S. Sidon : Ein Satzüber trigonometrische Polynome und seine Anwendungen in der Theorie der Fourier-reihen, Math. Ann. 106 (1932) 539

[S2] S. Sidon: Über die Fourierkonstanten der Funktionen der Ulane L_p for $p > 1$. Acta Sci. Math. (Szeged) 7 (1935) 175-176

[St] A. Stöhr: Gelöste und ungelöste Fragen über Basen der natürlichen Zahlen, J. Reine Angew. Math. 194 (1955) 40-65, 111-140

[Sz] E. Szemerédi: On sets of integers containing no k elements in arithmetic progression. Acta Arithm. 27 (1975) 199-245