Math. Slovaca 28, 1978, No. 4, 407–421

SOME SOLVED AND UNSOLVED PROBLEMS IN COMBINATORIAL NUMBER THEORY

PÁL ERDÖS—ANDRÁS SÁRKÖZY

This paper contains mostly joint problems of the two authors. Also, some problems of the second author are involved (which have arisen mostly starting out from problems of the first author). Finally, a paper of this title would not be complete without some famous unsolved problems (like Problems 1, 6 and 11), which are due mostly to the first author; for further details concerning these problems, see P. Erdös, Problems and results on combinatorial number theory, III, Lectures Notes Math., vol. 626, Springer 1977 ; we shall refer to this paper as [1].

Throughout this paper, $c, c_1, c_2, \ldots$ will denote absolute constants, $c(\alpha, \beta, \ldots)$, $x_0(\alpha, \beta, \ldots)$, $x_1(\alpha, \beta, \ldots), \ldots, k_0(\alpha, \beta, \ldots), \ldots$ constants depending only on the parameters $\alpha, \beta, \ldots$. The counting function of a sequence $\mathcal{A}$ of non-negative integers $a_1 < a_2 < \ldots$ will be denoted by $A(x)$:

$$A(x) = \sum_{\substack{a \leq x \\ a \in \mathcal{A}}} 1.$$

1. Additive problems

Problem 1

Show the existence of an *infinite* sequence $\mathcal{A}$ of positive integers $a_1 < a_2 < \ldots$, such that all the sums $a_i + a_j$ are distinct and

$$\liminf_{n \to +\infty} \frac{A(n)}{n^{1/3}} = +\infty.$$

(This problem is due essentially to Sidon and it is more than 40 years old. For the background of this problem, see [1].)

We remark that for *finite* sequences the situation is different : it is well-known that for $\varepsilon > 0$, $N > N_0(\varepsilon)$, there exists a sequence $\mathcal{A} \subset \{1, 2, \ldots, N\}$ for which all the sums $a_i + a_j$ are distinct and

$$A(N) > (1 - \varepsilon) N^{1/2}$$

holds.

Problem 2

Let $\mathscr{A}=\{a_1, a_2, \dots, a_N\}$ be a sequence of ***positive integers*** for which

$$(1) \qquad (0<)a_1<a_2<\dots<a_N$$

holds and for t and k fixed let $f(N, \mathscr{A}, t)$ and $g(N, \mathscr{A}, t, k)$ denote the number of the solutions of

$$\sum_{i=1}^{N} \varepsilon_i a_i = t \quad \text{(where } \varepsilon_i = 0 \text{ or } 1 \text{ for } i=1,2,\dots,N)$$

and

$$\sum_{i=1}^{N} \varepsilon_i a_i = t, \quad \sum_{i=1}^{N} \varepsilon_i = k \quad \text{(where } \varepsilon_i = 0 \text{ or } 1 \text{ for } i=1,2,\dots,N),$$

respectively.

P. Erdős and L. Moser proved that

$$(2) \qquad f(N, \mathscr{A}, t) < c_1 \frac{2^N}{N^{3/2}} (\log N)^{3/2}$$

(for all $N, \mathscr{A}, t$) and they conjectured that

$$(3) \qquad f(N, \mathscr{A}, t) < c_2 \frac{2^N}{N^{3/2}}.$$

Furthermore, they conjectured that

$$(4) \qquad g(N, \mathscr{A}, t, k) < c_3 \frac{2^N}{N^2}$$

(for all $N, \mathscr{A}, t, k$).

A. Sárközy and E. Szemerédi proved that for $\varepsilon > 0$, $N > N_0(\varepsilon)$, (3) holds with $c_2 = (1+\varepsilon)\frac{8}{\sqrt{\pi}}$, even in the more general case when $a_1, a_2, \dots, a_N$ are real numbers (not necessarily integers) satisfying (1) (Acta Arithm. 11 (1965), pp. 205-208).

J. L. Nicolas improved on the value of the constant c_2 (C. R. Acad. Sci. Paris Sér. A. 282 (1976), pp. 9—12).

G. Halász investigated an n-dimensional generalization of the problems above (Estimates for the concentration function of combinatorial number theory and probability, Periodica Math. Hung. 8 (1978), pp. 197-211; his results involve both estimates (3) and (4) as special cases.

Finally, J. Beck sharpened (3) for sequences satisfying certain restrictions, e.g.

$$a_{i+1} - a_i \neq a_{j+1} - a_j \quad (\text{for } 1 \leqq i < j \leqq N-1)$$

and

$$a_2 - a_1 < a_3 - a_2 < \dots < a_N - a_{N-1},$$

respectively. (His first paper on this subject will appear in Coll. Math. Soc. J. Bolyai.) However, he has not investigated, e.g. the important case when the differences $a_i - a_j$ are distinct.

The first problem is to find the best possible value of the constant c_2 in (3). The second problem is to estimate $f(N, \mathscr{A}, t)$ in the case when the differences $a_i - a_j$ are distinct:

$$a_i - a_j \neq a_u - a_v \quad \text{if} \quad 1 \leqq i < j \leqq N,\ 1 \leqq u < v \leqq N,\ i \neq u,\ j \neq v.$$

Halász's paper suggests further problems. E.g., in the n-dimensional case, his upper estimate contains a constant C = C(n) (corresponding to c_2 and c_3); it would be worth to investigate the behaviour of this constant as $n \to +\infty$.

Problem 3

It can be shown easily that if N is a positive integer, $a_1, a_2, \ldots, a_N$ are real numbers such that $|a_i| \leqq 1$ for $i = 1, 2, \ldots, N$, and we form all the sums $\sum_{i=1}^{N} \varepsilon_i a_i$, where $\varepsilon_i = -1$ or $+1$ (for $i = 1, 2, \ldots, N$); then at least $\binom{N}{N/2}$ of these sums satisfy

$$\left| \sum_{i=1}^{N} \varepsilon_i a_i \right| \leqq 1. \tag{5}$$

(This theorem is best possible in the sense that the number of the sums satisfying (5) need not exceed $\binom{N}{N/2}$, as the sequence $a_1 = a_2 = \ldots = a_N = +1$ shows.)

In fact, this theorem can be proved easily by using *the* following purely combinatorial theorem of G. K a t o n a:

Let us denote the number of the elements of a finite set S by $|S|$. If $N = 2M$, $M = 1, 2, \ldots, |A| = N$, and $B_1, B_2, \ldots, B_x$ are subsets of A such that

$$|B_i \cap B_j| \geqq 2 \quad (\text{for } 1 \leqq i < j \leqq x), \tag{6}$$

then $x \leqq 2^{N-1} - \frac{1}{2}\binom{N}{N/2}$. This theorem is best possible as those subsets B_i show for which $|B_i| \geqq \frac{N}{2} + 1$ holds.

(More exactly, Katona investigated the more general case when (6) is replaced by $|B_i \cap B_j| \geqq k$; the theorem referred to by us is only a special case of this theorem.)

The problem is to investigate the n-dimensional analogue of the result above. In particular, the twodimensional analogue is the following:

Let N be a positive integer, $z_1, z_2, \ldots, z_N$ complex numbers for which $|z_i| \leqq 1$ for $i = 1, 2, \ldots, N$. Let us form all the sums

$$(7) \qquad \sum_{i=1}^{N} \varepsilon_i z_i, \quad \text{where} \quad \varepsilon_i = -1 \text{ or } +1 \text{ for } i = 1, 2, \ldots, N.$$

It can be shown easily (by induction) that at least one of these sums satisfies

$$(8) \qquad \left|\sum_{i=1}^{N} \varepsilon_i z_i\right| = \sqrt{2}.$$

The upper bound $\sqrt{2}$ is best possible as the following construction shows: let

$$(9) \qquad N = 4M-2, \quad z_1 = z_2 = \ldots = z_{2M-1} = 1 \quad \text{and}$$
$$z_{2M} = z_{2M+1} = \ldots = z_{4M-2} = i.$$

The problem is : how many of the sums (7) must satisfy (8) ? We guess that the number of these sums must be greater than $c2^N/N$, and if $N = 4M-2$, then, perhaps, the construction in (9) gives the exact value of the extremum. Our conjecture can be reduced to the following combinatorial problem:

Is it true that if $N = 1, 2, \ldots, |A| = N$, and $B_1, \ldots, B_x, C_1, \ldots, C_y$ are subsets of A such that

$$|B_i \cap B_j| \geqq 2 \quad (\text{for } 1 \leqq i < j \leqq x), \quad |C_i \cap C_j| \geqq 2 \quad (\text{for } 1 \leqq i < j \leqq y)$$
$$\text{and} \quad |B_i \cap C_j| \geqq 1 \quad (\text{for } 1 \leqq i \leqq x,\ 1 \leqq j \leqq y),$$

then

$$x + y \leqq 2^{N-1} - c\frac{2^N}{N}$$

must hold?

We have not been able to prove this conjecture.

Problem 4

An infinite sequence $\mathscr{A}$ of non-negative integers $a_1 < a_2 < \ldots$ is said to be a ***basis of order k*** (where $k = 2, 3, \ldots$) if for $n > n_0$ there exist indices $x_1, x_2, \ldots, x_k$ such that

$$a_{x_1} + a_{x_2} + \ldots + a_{x_k} = n.$$

A simple consideration shows that "almost all" the sequences of non negative integers are bases of finite order. The second author proved . the following sharpening of this fact (Some metric problems in the additive number theory, II, Annales Univ. Sci. Budapest. Eötvös, 20 (1977), pp. 111-129:

For $k = 2, 3, \ldots,$ let Γ_k denote the set of those sequences of non-negative integers which are not bases of order k. Let us map the set Γ_k into the interval [0, 1] in the following way: for each sequence

$$\mathscr{A} = \{a_1, a_2, \ldots\} \in \Gamma_k,$$

let

$$\varphi(\mathscr{A}) = \sum_{i=1}^{+\infty} 2^{-(a_i+1)},$$

and let $\varphi(\Gamma_k)$ denote the set of these points $\varphi(\mathscr{A})$. Then

$$\dim \varphi(\Gamma_2) = \frac{\log 3}{\log 4},$$

furthermore,

$$\lim_{k \to +\infty} \dim \varphi(\Gamma_k) = \frac{1}{2}$$

(where $\dim S$ denotes the Hausdorff-dimension of the set S).

The problem is to determine dim $\varphi(\Gamma_k)$, or at least to estimate $\dim \varphi(\Gamma_k) - \frac{1}{2}$ for $k \geqq 3$. Furthermore, it would be worthwile to decide whether

$$\dim \varphi(\Gamma_k) > \frac{1}{2} \tag{10}$$

for each k or

$$\dim \varphi(\Gamma_k) = \frac{1}{2} \tag{11}$$

for $k > k_0$. In particular, does (10) or (11) hold for $k = 3$?

Problem 5

The second author proved the following theorem:
If $\varepsilon > 0$, $N > N_0(\varepsilon)$, $\mathscr{A} = \{a_1, a_2, \ldots\} \subset \{1, 2, \ldots, N\}$ and

$$A(N) > \varepsilon N \tag{12}$$

then there exist indices x, y and a prime number p for which

$$a_x - a_x = p - 1. \tag{13}$$

(Unpublished yet.) Furthermore, the assumption (12) can be improved slightly.

(We remark that this statement is not true if we write p on the right -hand side, as the sequence $\mathscr{A} = (4, 8, \ldots, 4k, \ldots\}$ shows.)

On the other hand, it can be easily shown that for $N = 1, 2, \ldots$ there exists a sequence $\mathscr{A} \subset \{1, 2, \ldots, N\}$ such that

$$A(N) > c \log N$$

holds and the equation (13) is not solvable.

The problem is to improve on this lower estimate, i.e. to prove the existence of sequences $\mathcal{A} \subset \{1, 2, \ldots, N\}$ such that (13) is not solvable and $A(N)/\log N \to +\infty$ as $N \to +\infty$.

2. Multiplicative problems

A sequence $\mathcal{A}$ of positive integers $a_1 < a_2 < \ldots$ is said to be *primitive* if $a_i \nmid a_j$ for all $i < j$. F. Behrend (J. London Math. Soc. 10 (1935), pp. 42-44) and P. Erdős (J. London Math. Soc. 10 (1935), pp. 126—128) proved that if the sequence $\mathcal{A}$ is primitive, then

$$\sum_{a_i \leqq x} \frac{1}{a_i} < c_1 \frac{\log x}{(\log \log x)^{1/2}} \tag{14}$$

and

$$\sum_i \frac{1}{a_i \log a_i} < c_2 . \tag{15}$$

P. Erdős, A. Sárközy and E. Szemerédi extended these results in various directions. In particular, they determined the infimum of those constants c_1 for which (14) holds. (For further details and references, see Coll. Math. Soc. J. Bolyai 2 (1970), pp. 3549; this paper contains also some further unsolved problems. See also [1].)

The most interesting unsolved problems connected with these results are Problems 6, 7 and 8 below.

Problem 6

Prove that if $\varepsilon > 0$, $X > X_0(\varepsilon)$, $x_1, x_2, \ldots, x_N$ are real numbers (not necessarily integers) such that $1 \leqq x_1 < x_2 < \ldots < x_N \leqq X$ and

$$\sum_{n=1}^{N} \frac{1}{x_n} > \varepsilon \log X, \tag{16}$$

then there exist indices i, j $(1 \leqq i < j \leqq N)$ and a positive integer k such that

$$|kx_i - x_j| < 1 .$$

This conjecture must be true even with $c \dfrac{\log X}{(\log \log X)^{1/2}}$ on the right — hand side of (16). (Obviously, this sharper form of the conjecture would involve Behrend's theorem (14).)

Problem 7

For $\omega > 0$, let A_ω denote the set of those primitive sequences $\mathcal{A}$ for which $\omega < a_1 < a_2 < \ldots$ holds. Determine

$$\lim_{\omega\to+\infty}\left(\sup_{\mathcal{A}\in A_\omega}\sum_i \frac{1}{a_i\log a_i}\right).$$

(It is known that this limit is between 1 and e^{γ}; we guess that it is equal to 1.)

Problem 8

Show that if $a_1<a_2<\dots$ is an infinite sequence of positive integers for which

$$\text{(17)}\qquad \liminf_{x\to+\infty}\left(\sum_{a_i\le x}1\right)\Big/\frac{x}{(\log\log x)^{1/2}}>0,$$

then

$$\text{(18)}\qquad \limsup_{x\to+\infty}\left(\sum_{\substack{a_i\mid a_j\\ a_j\le x}}1\right)\Big/x=+\infty.$$

(It is known that (17) implies that the left – hand side of (18) is >0.)

Problem 9

Starting from a conjecture of G. Halász, the second author proved the following theorem (Studia Sci. Math. Hung. 9 (1974), pp. 161-171):

Let $a_1<a_2<\dots<a_n$ be a sequence of positive integers such that it contains the first k positive integers:

$$\text{(19)}\qquad a_1=1,\ a_2=2,\ \dots\ a_k=k.$$

There exists an absolute constant c such that if $k>k_0$, then there exist at least $n\cdot k^{c/\log\log k}$ distinct products of the form a_ia_j (i, $j=1, 2, \dots, n$).

(We remark that to get "many" distinct products, a condition of the type (19) is necessary; otherwise, e.g., the sequence $a_1=1$, $a_2=2$, $a_3=2^2, \dots, a_n=2^{n-1}$ is a counterexample.)

On the other hand, it can be shown easily that for $\varepsilon>0$, $k>k_0(\varepsilon)$, $n\ge k$, there exists a sequence $a_1<a_2<\dots<a_n$ of positive integers such that it satisfies (19) and the number of the distinct products of the form a_ia_j is less than εnk.

There is a gap between the lower and upper estimates. The upper bound seems to be more precise ; in fact, we guess that for $\varepsilon>0$, $k>k_1(\varepsilon)$ and under the assumption (19), the number of the distinct products must be greater than $nk^{1-\varepsilon}$.

The theorem above suggests the following problem:

Let $a_1<a_2<\dots<a_n$ be a sequence of positive integers for which (19) holds. What can be asserted about the number of the distinct products of form $\prod_{i=1}^{n} a_i^{\varepsilon_i}$ (where $\varepsilon_i=0$ or 1 for $i=1, 2, \dots, n$)? Is it true that for $\omega>0$, $k>k_2(\omega)$, the number of the distinct products must be greater than n^2k^{ω}?

We guess that for $k > k_3$, the number of these products must be greater than $n^2 e^{c_1 k/\log k}$. On the other hand, it can be shown easily that for $k > k_4$, $n \geqq k$, there exists a sequence $a_1, a_2, \dots, a_n$ for which (19) holds and the number of the distinct products is less than $n^2 e^{c_2 k/\log k}$.

3. Irregularities of distribution of sequences relative to arithmetic progressions

The results concerning irregularities of distribution of sequences relative to arithmetic progressions can be divided into two groups.

The papers belonging to the first group investigate short arithmetic progressions ; e.g. Van der Waerden's well-known theorem is a result of this type. The deepest result in this field is due to E. Szemerédi (Acta Arithm., 27 (1975), pp. 199-245) ; see Szemerédi's paper for further details and references. (See also [1].)

In this paper, we shall be interested mostly in long arithmetic progressions. The first results of this type have been proved by K. F. Roth (Acta Arithm., 9 (1964), pp. 257-260). These results have been extended in various directions by Roth, S. L. G. Choi, H. L. Montgomery, M. N. Huxley and A. Sárközy. A typical result of K. F. Roth is the following:

Let N be a positive integer and let $\mathcal{A} \subset \{1, 2, \dots, N\}$. Let us write

$$\eta = \frac{A(N)}{N},$$

$$D_{q,h}(m) = \sum_{\substack{1 \leqq a \leqq m \\ a \equiv h \ (\mathrm{mod}\ q) \\ a \in \mathcal{A}}} 1 - \eta \sum_{\substack{1 \leqq a \leqq m \\ a \equiv h \ (\mathrm{mod}\ q)}} 1$$

and

$$V_q(m) = \sum_{h=0}^{q-1} D^2_{q,h}(m)$$

(where q is a positive integer, h, m are integers). Then for any integer Q,

$$\sum_{q=1}^{Q} q^{-1} \sum_{m=1}^{N} V_q(m) + Q \sum_{q=1}^{Q} V_q(N) > c_1 \eta (1-\eta) Q^2 N.$$

Furthermore, this inequality (with $Q = [N^{1/2}]$) implies the existence of q, m, h such that

$$1 \leqq q \leqq N^{1/2}$$

and

$$|D_{q,h}(m)| = \left| \sum_{\substack{1 \leqq a \leqq m \\ a \equiv h \ (\mathrm{mod}\ q) \\ a \in \mathcal{A}}} 1 - \eta \sum_{\substack{1 \leqq a \leqq m \\ a \equiv h \ (\mathrm{mod}\ q)}} 1 \right| > c_2 \{\eta(1-\eta)\}^{1/2} N^{1/4}.$$

For further details and references, see A. Sárközy, Some remarks concerning irregularities of distribution of sequences of integers in arithmetic progressions, III and IV, Periodica Math. Hung. 9 (1978), pp. 127-144 and Acta Math. Acad. Sci. Hung. 30 (1977), pp. 155-162.

Problem 10

Let N be a positive integer. Let E_N denote the set of the 2^N sequences $\varepsilon = \{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_N\}$, where $\varepsilon_i = -1$ or $+1$ for $i = 1, 2, \ldots, N$. Let

$$F(N) = \min_{\varepsilon \in E_N} \max_{\substack{k,n,q \\ 1 \leqq n \leqq n+(k-1)q \leqq N}} \left| \sum_{i=0}^{k-1} \varepsilon_{n+iq} \right| .$$

Roth's theorem referred to above implies that

$$F(N) \geqq c_3 N^{1/4}. \tag{20}$$

On the other hand, P. Erdős proved that

$$F(N) \leqq c_4 N^{1/2}$$

(Mat. Lapok 17 (1966), pp. 135-155). J. Spencer improved on this estimate by showing that

$$F(N) \leqq c_5 (N \log \log N)^{1/2} (\log N)^{-1/2}$$

(Canad. Math. Bull. 15 (1972), pp. 43-44). Finally, H. L. Montgomery and A. Sárközy proved that

$$F(N) \leqq c_6 N^{1/3} (\log N)^{2/3}. \tag{21}$$

(Unpublished yet.) Montgomery's and Sárközy's construction is the following:

Let p denote the smallest prime number satisfying $p > N^{2/3}(\log N)^{-2/3}$ and for $i = 1, 2, \ldots, N$, let

$$\varepsilon_i = \begin{cases} \left(\frac{i}{p}\right) & \text{for } (i,p)=1 \\ 1 & \text{for } p \mid i \end{cases}$$

(where $\left(\frac{i}{p}\right)$ denotes the Legendre symbol). It can be easily shown that for this sequence ε,

$$\left| \sum_{i=0}^{k-1} \varepsilon_{n+iq} \right| \leqq c_7 N^{1/3} (\log N)^{2/3}$$

for any integers k, n, q satisfying $1 \leqq n \leqq n + (k-1)q \leqq N$.

There is a gap between (20) and (21). It seems that the lower estimate is more precise; in fact, we: guess that for $\varepsilon > 0$, $N > N_0(\varepsilon)$,

$$F(N) \leqq c(\varepsilon) N^{1/4+\varepsilon}$$

and, perhaps, it will not be very difficult to show this.

Problem 11

A collection of problems of this type would not be complete without mentioning an old conjecture of the first author:

If $\varepsilon_1, \varepsilon_2, \ldots$ is an infinite sequence such that $\varepsilon_i = -1$ or $+1$ for $i = 1, 2, \ldots$, then

$$\sup_{\substack{m=1, 2, \ldots \\ n=1, 2, \ldots}} \left| \sum_{i=1}^{n} \varepsilon_{im} \right| = +\infty .$$

This conjecture is about 40 years old, however, no advance has been made yet. The difficulties can be illustrated by the following fact:

There exists an infinite sequence $\varepsilon_1, \varepsilon_2, \ldots$ (where $\varepsilon_i = -1$ or $+1$ for $i = 1, 2, \ldots$) such that

$$\max_{\substack{m, n \\ 1 \leqq m \leqq nm \leqq N}} \left| \sum_{i=1}^{n} \varepsilon_{im} \right| \leqq c \log N .$$

In fact, let us define the sequence $\varepsilon_1, \varepsilon_2, \ldots$ in the following way: for $i = 1, 2, \ldots$, let $i = 3^{\alpha_i} j_i$, where $\alpha_i (\geqq 0)$, $j_i (\geqq 1)$ are integers such that $(3, j_i) = 1$, and let

$$\varepsilon_i = \begin{cases} +1 & \text{if } j_i \equiv 1 \pmod 3, \\ -1 & \text{if } j_i \equiv -1 \pmod 3). \end{cases}$$

It can be easily shown that this sequence satisfies (22).

Comparing (22) with (20) in Problem 10, we see that the situation is different here ; the best possible lower bound for the left-hand side of (22) is much smaller than the one for F(N).

Problem 12

In Roth's referred results, the *moduli* of the quantities D,.,(m) are estimated ; these results are nearly best possible. On the other hand, it is much more difficult to deduce one-sided estimates ; the known results of this type seem to be far from the best possible.

K. F. Roth proved the following theorem (Math. Ann. 169 (1967), pp. 1-25):

Let $\Lambda \geqq 1$ and let k be an integer satisfying $k > (10^2 \Lambda)^4$. Then there exists a number $N_1 = N_1(\Lambda, k)$ such that the following statement is true.

If $N > N_1$ and the set $s_1, s_2, \ldots, s_N$ of real numbers satisfies

$$1 \leqq |s_i| \leqq \Lambda$$

and

(23) $$s_1 + s_2 + \ldots + s_N = 0,$$

then there exist integers n, q satisfying $1 \leqq n < n + (k-1)q \leqq N$ such that

$$\sum_{i=0}^{k-1} s_{n+iq} > \left\{ 10^{-4} \Lambda^{-2} \sum_{i=0}^{k-1} s_{n+iq}^2 \right\}^{1/2}.$$

The weakness of this theorem is that the function $N_1(\Lambda, k)$ is not explicitly given. Improving on Roth's method, S. L. G. Choi showed that the statement of this theorem is valid with $N_1(\Lambda, k) = 2\{2(\Lambda^{-1}s)^{2s+5^s}\}^6$, where $s = 2\Lambda((2k^2)!)^2k^8$ (Math. Ann. 205 (1973), pp. 1-8). This value of $N_1(\Lambda, k)$ is extremely large. In fact, for A fixed and k large, $N > N_1$ implies that $k = O((\log \log \log N / \log \log \log \log N)^{1/2})$. However, the assertion of Roth's theorem should be valid also with $N_1(\Lambda, k) = c(\Lambda)k^c$ (like the case when we estimate the *modufi* of the sums concerned).

A. Sárközy estimated the large positive values of the sums $\sum_{i=0}^{k-1} s_{n+iq}$ in terms of q instead of k:

Let $A \geqq 1$ and let Q, N be any positive integers for which

(24) $$Q \leqq \frac{1}{5}\left(\frac{N}{\Lambda}\right)^{2/5}.$$

If the set $s_1, s_2, \ldots, s_N$ of real numbers satisfies (23) and

$$1 \leqq |s_j| \leqq \Lambda \quad \text{for } j = 1, 2, \ldots, N,$$

then there exist positive integers n, k, q such that $1 \leqq q \leqq Q$, $1 \leqq n < n + (k-1)q \leqq N$ and

$$\sum_{i=0}^{k-1} s_{n+iq} \geqq \frac{1}{40}\left(Q \frac{\sum_{n=1}^{N} s_n^2}{N}\right)^{1/2}.$$

This theorem is best possible (except the value of the constant on the right) for any Q satisfying (24) and it gives a much greater lower bound (in terms of N) for $\max \sum_{i=0}^{k-1} s_{n+iq}$:

$$\max_{n,q,k} \sum_{i=0}^{k-1} s_{n+iq} > c(\Lambda)N^{1/5}.$$

However, the statement of the theorem should be true also for Q in a greater range ; in fact, we guess that (24) can be replaced by $Q < c(\Lambda)N^{1/2}$. This would imply

$$\max_{n,q,k} \sum_{i=0}^{k-1} s_{n+iq} > c(\Lambda)N^{1/4}.$$

(Again, this would correspond to the estimates for the *moduli* except that the dependence on A must be different.)

Problem 13

Let $\mathcal{A} = \{a_1, a_2, \dots\}$ be an infinite sequence of positive integers for which $a_1 < a_2 < \dots$. Let us write

$$E_{q,h}(m) = \sum_{\substack{1 \leq a \leq m \\ a \equiv h \pmod q \\ a \in \mathcal{A}}} 1 - q^{-1}A(m),$$

where q, m are positive integers, h is an integer.

The second author of this paper proved that

$$\limsup_{N \to +\infty} \Bigg(\sum_{\substack{1 \leq a \leq N \\ a \in \mathcal{A},\, a-1 \notin \mathcal{A}}} 1 \Bigg) \Big/ N^{1/2} = +\infty \tag{25}$$

implies that

$$\sup_{q,h,m} |E_{q,h}(m)| = +\infty. \tag{26}$$

Furthermore, let $\log_k x$ denote the k-fold iterated logarithm (i.e. $\log_k x = \log(\log_{k-1} x)$), and for $x > e$ let us define the positive integer $L(x)$ by

$$\log_{L(x)+1} x < 1 \leq \log_{L(x)} x.$$

A. Sárközy constructed also an infinite sequence $\mathcal{A}$ for which

$$\liminf_{N \to +\infty} \Bigg(\sum_{\substack{1 \leq a \leq N \\ a \in \mathcal{A},\, a-1 \notin \mathcal{A}}} 1 \Bigg) \Big/ L(N) > 0 \tag{27}$$

and, on the other hand,

$$|E_{q,h}(m)| \leq 3$$

for any positive integers q, h, m. (See Some remarks concerning irregularities of distribution of sequences of integers in arithmetic progressions, I and II, Coll. Math. Soc. J. Bolyai 13 (1974), 287-303, and Studia Sci. Math. Hung., to appear.)

This construction shows that to obtain (26) it is not sufficient to assume that

$$\sum_{\substack{1 \leq a \leq N \\ a \in \mathcal{A},\, a-1 \notin \mathcal{A}}} 1 \to +\infty. \tag{28}$$

(We remark that to obtain (26) we need an assumption involving rather the function on the left — hand side of (28) than the function ***A(N).)***

There is a considerable gap between (25) and (27) ; the problem is to tighten this gap. In particular, is it true that if only

$$\limsup_{N\to+\infty}\Bigg(\sum_{\substack{1\leq a\leq N\\ a\in\mathcal{A},\, a-1\notin\mathcal{A}}} 1\Bigg)\Big/ N^{\varepsilon}>0$$

is assumed (for any $\varepsilon > 0$), then (26) must hold ?

Problem 14

Using the same notations as in the previous problem, A. Sárközy showed the existence of an infinite sequence $\mathcal{A}$ such that for any $\varepsilon > 0$ and positive integers q, h, m,

(29) $$|E_{q,h}(m)|<ce^{(1+\varepsilon)q}.$$

On the other hand, such a sequence does not exist if we write $\varepsilon q^{1/2}$ on the right — hand side. (See the papers referred to in the previous problem.) The problem is again to tighten the gap. (The lower bound $\varepsilon q^{1/2}$ seems to be more precise ; perhaps, the right — hand side of (29) can be replaced by $q^{1/2+\varepsilon}$.)

4. Distances near integers

Throughout this section, the distance between some points ***P,*** Q in the n-dimensional Euclidean space will be denoted by ϱ ***(P,*** Q). We shall denote the distance from the real number x to the nearest integer by $\|x\|$, i.e. $\|x\| = \min\{x-[x], [x]+1-x\}$. Let 6 be some fixed real number satisfying $0<\delta<1/2$.

For $X(>0)$ and δ fixed, let $P_1, P_2, \ldots, P_k$ be points in the n-dimensional sphere of radius X, such that

$$\|\varrho(P_i, P_j)\|\geqq\delta \quad \text{for} \quad 1\leqq i<j\leqq n$$

(i.e. each of the distances between the given points is further from any integer than δ). Let us denote the maximal number of points with these properties by $F_N(X,$ 6).

Startingfromaproblemof P. Erdős, A. Sárközy proved that for $X>X_0(\delta)$,

(30) $$F_2(X, \delta)<\frac{4\cdot 10^4}{\delta^3}\,\frac{X}{\log\log X},$$

and, on the other hand, for $0<\delta\leqq 1/6\cdot 8^4$ and $X>X_1(\delta)$,

$$F_2(X,\delta)>X^{1/2-\delta^{1/7}}$$

(On distances near integers, I and II, Studia Sci. Math. Hung., to appear).

It can be easily shown that (30) implies that for $X > X_2(\delta, n)$,

$$F_n(X, \delta) < c(\delta, n) \frac{X^{n-1}}{\log \log x}.$$

Furthermore, the authors showed that if $\delta > 0$, $m \geqq 2$ is an integer, $X > X_3(\delta, m)$, k is an integer satisfying

$$k > c(\delta, m) \frac{X}{\log \log x}$$

and $P_1, P_2, \ldots, P_k$ are points in the circle of radius X, then these points contain a "near integer m-tuple", i.e. there exist indices $i_1, i_2, \ldots, i_m$ such that $1 \leqq i_1 < i_2 < \ldots < i_m \leqq k$ and

$$\|\varrho(P_{i_\mu}, P_{i_\nu})\| < \delta \quad \text{for} \quad 1 \leqq \mu < \nu \leqq m.$$

(This result has not been published yet.) However, our proof yields only the existence of "degenerated" m-tuples, i.e. m-tuples such that their vertices are "near" a fixed line.

These results suggest the following problems:

Problem 15

How rapidly does $F_n(X, \delta)$ increase for $n \to +\infty$? Does there exist a positive integer n such that

$$\lim_{N \to +\infty} \frac{F_n(X, \delta)}{X} = +\infty$$

for some $\delta > 0$? (in view of (30), this would imply

$$\left. \lim_{X \to +\infty} \frac{F_n(X, \delta)}{F_2(X, \delta)} = +\infty . \right)$$

Problem 16

Show that there exists a positive number a satisfying the following conditions :

Let $\delta > 0$, $\omega > 0$, $X > X_0(\delta, w)$, and let $P_1, P_2, \ldots, P_k$ be points in the Cartesian plane in the square $0 \leqq x \leqq X$, $0 \leqq y \leqq X$, such that if $0 \leqq u < X - \omega$, $0 \leqq v < X - \omega$, then the square

$$u \leqq x < u + \omega, \quad v \leqq y < x + \omega \tag{31}$$

contains at least one of these points. Then these points contain a "non degenerated near integer triangle", i.e. there ***exist*** indices i_1, i_2, i_3 such that $1 \leqq i_1 < i_2 < i_3 \leqq k$, each angle of the triangle $P_{i_1}P_{i_2}P_{i_3}$ is greater than a, and

$$\|\varrho(P_{i_1}, P_{i_2})\| < \delta, \quad \|\varrho(P_{i_1}, P_{i_3})\| < \delta \quad \text{and} \quad \|\varrho(P_{i_2}, P_{i_3})\| < \delta .$$

(This conjecture should be true even replacing the squares in (31) by the squares

$$u \leqq x < u + \omega\sqrt{X}, \quad v \leqq y < v + \omega\sqrt{x}.)$$

Problem 17

Show that if $\delta > 0$, $\omega > 0$, X $> X_0(\delta, \omega)$ and the points $P_1, P_2, \ldots, P_k$ satisfy the condition in Problem 16, then these points contain an "almost equilateral triangle", i.e. there exist indices j_1, j_2, j_3 such that $1 \leqq j_1 < j_2 < j_3 \leqq k$ and

$$|\varrho(P_{j_1}, P_{j_2}) - \varrho(P_{j_1}, P_{j_3})| < \delta, \quad |\varrho(P_{j_1}, P_{j_2}) - \varrho(P_{j_2}, P_{j_3})| < \delta$$

and

$$|\varrho(P_{j_1}, P_{j_3}) - \varrho(P_{j_2}, P_{j_3})| < \delta .$$

Received February 15, 1977

Magyar *tudományos akadémia*
matematikai kutató intézete
Reáltanoda u. 13-15
1053 Budapest V