Some Old and New Problems on Additive and Combinatorial Number Theory

P. ERDÖS

Tel Aviv University
Ramat-Aviv
69978 Tel-Aviv, Israel

In this survey paper I will discuss some old and new problems of a somewhat unconventional nature. Proofs will not be given, but I hope they will appear soon in several joint papers of V. T. Sós, Sárközy, and myself. Some of the older results can be found with detailed proofs in the excellent book *Sequences* of Halberstam and Roth.

First of all an old question of Sidon posed to me more than 50 years ago states as follows: Let $A = \{a_1 < a_2 < \cdots\}$ be a basis of order 2; in other words, if $f(n)$ denotes the number of solutions of $n = a_i + a_j$, then for $n > n_0$, $f(n) > 0$. Now Sidon asked: Is there a basis of order 2 for which $f(n)/n^\varepsilon \to 0$ for every $\varepsilon > 0$. In 1953 I proved by probabilistic methods that the answer is affirmative. In fact, I proved that there is a basis of order 2 for which

$$c_1 \log n < f(n) < c_2 \log n \tag{1}$$

holds for every $n > n_0$. A constructive proof of (1) would in my opinion be still of great interest and I offer 100 dollars for the construction of a basis of order 2 for which $f(n)/n^\varepsilon \to 0$. An old and in my opinion very interesting conjecture of Turán and myself states that for every basis of order 2 we have $\limsup f(n) = \infty$. I offered and offer 500 dollars for a proof or disproof of this fascinating conjecture. In fact, I believe that for a basis of order 2 we must have

$$\frac{\limsup f(n)}{\log n} > 0. \tag{2}$$

In view of (1), then (2) if true is essentially best possible. The following slight refinement is perhaps still possible. Can $\limsup f(n)/\log n$ be arbitrarily small or is there an absolute constant so that for every basis of order 2 we have

$$\frac{\limsup f(n)}{\log n} > c? \tag{3}$$

If (3) is true, it would of course be of great interest to determine the smallest possible value of c in (3).

Another old problem of mine states that if A is a basis of order 2, then

$$\frac{\lim f(n)}{\log n} = c \tag{4}$$

cannot hold for $0 < c < \infty$. In other words, (4) would show that for a basis of order 2 the growth of $f(n)$ cannot be too regular. Sárközy and I recently proved the following much weaker result: If A is a basis of order 2, then

$$\frac{|f(n) - \log n|}{\sqrt{\log n}} \to 0 \tag{5}$$

cannot hold. The proof of (5) uses analytic methods and will be published soon.

We also showed that if $g(n) \to \infty$ as slowly as we please and satisfies certain very mild regularity conditions, then there is a basis of order 2 for which

$$\lim \frac{f(n)}{g(n)\log n} = 1 \tag{6}$$

holds. The proof of (6) again uses probability methods.

Let A be a basis of order r. The solution set $S(A, n)$ of n is the set of integers $a \in A$ for which $a_0 + a_1 + a_2 + \cdots + a_{r-1} = n$ is solvable for some $a_i \in A$. M. Nathanson and I showed by probabilistic methods that there is a basis of order r and an absolute constant C_r, so that for every n and m

$$|S_A(n) \cap S_A(m)| \leq C_r. \tag{7}$$

The best value of C_r is not known (even for $r = 2$). Perhaps $C_2 = 2$, but we could only prove $C_2 \leq 4$.

Put $A(n) = \sum_{a_i < n} 1$. Rényi and I proved by probabilistic methods that there is a sequence A for which for all n,

$$A(n) > c_1\sqrt{n} \qquad \text{and} \qquad \sum_{m<n} f^2(m) < c_2 n. \tag{8}$$

Sárközy and I will publish a detailed proof of (8) and various extensions and sharpenings. On the other hand, we conjecture that if A is a basis of order 2, then (8) does not hold, that is, in this case we have

$$\lim \frac{1}{x} \sum_{n=1}^{x} f^2(n) = \infty. \tag{9}$$

At present we do not see how to prove (or disprove) (9). We further conjectured that if A is a "thin" basis of order 2, that is, if $A(n) < c\sqrt{n}$ holds for every n, then

$$\frac{\limsup f(n)}{\log n} = \infty \tag{10}$$

and perhaps even

$$f(n) > n^{\varepsilon} \tag{11}$$

for some $\varepsilon > 0$ and infinitely many n. The condition (11) is perhaps a bit too optimistic, but we have no counterexample at present.

Perhaps (10) holds if A is a basis for which

$$\frac{A(n)}{\sqrt{n \log n}} \to 0. \tag{12}$$

In view of (1) the condition (11) can certainly not be sharpened.

If (9) does not hold, then we certainly must have $A(n) < c\sqrt{n}$. Perhaps if A is a basis of order 2 for which $A(n) < c\sqrt{n}$ holds, then there is a subsequence A_1 of A for which $A_1(n) > c'\sqrt{n}$ for some $c' > 0$ and all n, so that the density of the integers $a_u + a_v$, a_u, $a_v \in A_1$ is 0. This conjecture, which is also perhaps too optimistic, would immediately imply (9). One difficulty in trying to find a counterexample is that we know very few "thin" bases of order 2.

Nathanson, Sárközy, and I considered the following problem: Let $a_1 < a_2 < \cdots$ be any sequence of integers. Consider the sequence of integers $n_1 < n_2 < \cdots$ for which $f(n_k) \neq 1$, that is, $f(n) = 0$ or $f(n) > 1$. We showed that the number $H(X)$ of these integers not exceeding x can be $< cx^{1/2}$, but it must be $> c' \log x$. It is an interesting and perhaps difficult question to try to get better bounds for $H(X)$.

On the other hand, observe that it is easy to construct an infinite sequence $a_1 < a_2 < \cdots$ for which for every n the number of solutions of $n = a_j - a_i$ is 1.

V. T. Sós, Sárközy, and I studied various problems about the number of solutions of $a_i + a_j = n$ and $a_u - a_v = n$. Denote by $S(A)$, $A = \{a_1 < a_2 < \cdots\}$ the possible values of $f(n)$. It is easy to see that if 0 and 1 belong to S (i.e., if there is an n_1 and n_2 with $f(n_1) = 0$, $f(n_2) = 1$), then the other elements of S can be prescribed arbitrarily. The old conjecture with Turán can be restated as follows: If $0 \notin S$, then $|S| = \infty$. Perhaps if $|S| < \infty$ (i.e., if $f(n)$ is bounded), then $f(n) = 1$ for infinitely many integers n. Denote by $g(n)$ the number of solutions of $n = a_j - a_i$. We easily showed that the set of values of $\{g(n)\}$ can be arbitrarily prescribed. We obtained several further results that I hope will be published soon. Here I only state one of our results. Assume that $g(n)$ is 0 or 1, and let T be a sequence of integers that contains for every l a set of l consecutive integers. Then there is a sequence A for which $g(n) = 1$ if $n \in T$ and $g(n) = 0$ if $n \notin T$. If $g(n) \geq 2$ is permitted and the values of $g(n)$ are arbitrarily prescribed, but the set of integers for which $g(n) \geq 2$ contains for every l a set of l consecutive integers, then there is a sequence A for which $g(n)$ takes on the prescribed values. Perhaps there is no simple necessary and sufficient condition that will characterize the set of values $\{g(n)\}$ for which a sequence A exists.

Let A be a sequence of integers for which $f(n) < \infty$ and denote by $v_1 < v_2 < \cdots$ the sequence of integers for which $f(v) > 0$. An old conjecture of mine states that $\limsup v_n/n = \infty$ (or the lower density of the v's is 0). Perhaps even

$$\sum_{v_i < x} \frac{1}{v_i} = o(\log x). \tag{13}$$

In other words, the logarithmic density of the v's is 0. Equation (13) if true is certainly very deep.

Another old problem of Sidon states as follows: Let A be an infinite sequence for which $f(n) = 0$ or 1 (i.e., the sums $a_i + a_j$ are all distinct). Sidon called such a sequence a B_2 sequence. He asked how slowly can a_n tend to infinity for a B_2 sequence. For a long time it was not known if there is a B_2 sequence for which $a_n = o(n^3)$.

A few years ago Ajtai, Komlós, and Szemerédi [1] proved by an ingenious combination of probabilistic and combinatorial methods that there is a B_2 sequence for which

$$a_n < \frac{cn^3}{\log n} \tag{14}$$

holds for all n. The condition (14) is probably very far from being best possible. It seems likely that there is a B_2 sequence for which for every fixed $\varepsilon > 0$

$$a_n < cn^{2+\varepsilon} \tag{15}$$

and perhaps even for sufficiently large c_1 and c_2

$$a_n < c_1 n^2 (\log n)^{c_2}. \tag{16}$$

Rényi and I proved by probabilistic methods that there is a sequence A satisfying (15) for which $f(n) < C_\varepsilon$ (see Halberstam–Roth, chap. 3).

On the other hand, it is easy to see that for every B_2 sequence we must have

$$\frac{\limsup a_n}{n^2 \log n} > 0 \tag{17}$$

(Halberstam–Roth, chap. 2).

Perhaps if (17) does not hold, then $\limsup f(n) = \infty$.

Let A be a B_2 sequence and denote by $u_1 < u_2 < \cdots$ the integers for which $f(u_i) = 1$. It is easy to see that $\sum_{i=1}^{\infty} 1/u_i$ can diverge and in fact

$$\sum_{u_i < x} \frac{1}{u_i} > c \log \log x \tag{18}$$

is possible. Perhaps (18) is best possible, and in fact I have no counterexample to the following (perhaps unlikely) conjecture: Let A be a sequence for which $f(n) < C$ and denote as in (18) by $v_1 < v_2 < \cdots$ the integers for which $f(v_i) > 0$. Is it then true that

$$\sum_{v_i < x} \frac{1}{v_i} < c' \log \log x? \tag{19}$$

I would be very surprised if (19) would be true, but at the moment I have no counterexample. Perhaps the following conjecture holds: There is a sequence A for which $f(n) < C$, for which there do not exist a finite set of B_2 sequences $B^{(1)}, B_1^{(2)}, \ldots, B^{(r)}$, so that every integer n that is the sum of two elements of A is for some i the sum of two elements of B_i, $1 \le i \le r$.

V. T. Sós, Sárközy, and I investigated the following question: Let A be a sequence of integers and denote by $f_1(n)$, $f_2(n)$, and $f_3(n)$, respectively, the number of solutions of

$$a_i + a_j = n,$$

$$a_i + a_j = n, \qquad i < j,$$

$$a_i + a_j = n, \qquad i \le j,$$

respectively.

We proved that if $f_1(n)$ is monotone increasing for some $n > n_0$, then the complement of A is finite. On the other hand, there is an A whose complement is infinite and for which $f_2(n)$ is monotone increasing for $n > n_0$.

We could not decide whether if $f_3(n)$ is monotone increasing, then the complement of A must be finite. However, Sárközy just proved that if the complement of A tends to infinity faster than $\log n$, then $f_3(n)$ cannot be monotone.

Is there a B_2 sequence for which $a_k/k^3 \to 0$ and every integer is of the form $a_i - a_j$?

To finish this paper I state some further unconventional problems in number theory: Let $a_1 < a_2 < \cdots$ be an infinite sequence of integers and assume that no a_t is the sum of consecutive a's, that is,

$$a_t \neq a_i + a_{i+1} + \cdots + a_j \tag{20}$$

for every $i < j < t$.

I conjecture that if **(20)** holds, then the lower density of the a's is 0; in other words, $\lim\sup a_t/t \to \infty$. Perhaps **(20)** implies that the logarithmic density of the a's is 0, that is,

$$\frac{1}{\log x} \sum_{a_i < x} \frac{1}{a_i} \to 0. \tag{21}$$

There is a sequence satisfying **(20)** whose upper density is $\frac{1}{2}$ (I expect that it cannot be $> \frac{1}{2}$). There is a sequence satisfying **(20)** for which

$$\sum_{a_i < x} \frac{1}{a_i} > c \log \log x. \tag{22}$$

Perhaps **(22)** is best possible and **(20)** implies

$$\sum_{a_i < x} \frac{1}{a_i} < c' \log \log x. \tag{23}$$

The condition **(23)** if true is of course very much stronger than **(21)**, and perhaps **(21)** is too optimistic.

Let $1 < a_1 < a_2 < \cdots < a_t \leq n$ be a sequence satisfying **(20)**. Put $\max t = f(n)$. Is it true that $f(n) = (n+1)/2$? Perhaps this is trivial and I overlooked a simple argument.

If instead of **(20)** we assume that no a_t is the distinct sum of smaller a's. Then $\sum_{t=1}^{\infty} (1/a_t) < \infty$ [2]. Perhaps then for infinitely many n, $A(n) < n^{1-\varepsilon}$.

To end this paper I state a problem of Sárközy and myself [3]: A sequence A of integers is said to have property P if no a_i divides the sum of two larger a's. We proved that every infinite sequence of having property P has density 0. We conjecture that if A has property P, then $\sum 1/a_i$ converges and in fact $\sum 1/a_i < c$ for some absolute constant c. Also probably

$$A(X) = \sum_{a_i < x} 1 < x^{1-\varepsilon}$$

for infinitely many x. The choice $a_i = p_i^2$, where p_i is the ith prime $\equiv 3 (\text{mod } 4)$ shows that there is an infinite sequence having property P for which $A(X) > (cx^{1/2}/\log x)$. We have not been able to do better.

Let $1 \leq a_1 < \cdots < a_t \leq x$ be a finite sequence having property P. Is it true that $\max t = [x/3] + 1$? It is very annoying that we have not been able to prove or disprove this simple conjecture.

REFERENCES

1. AJTAI, M., J. KOMLÓS & E. SZEMERÉDI. 1981. A dense infinite Sidon sequence. Eur. J. Comb. **2**: 1–11.
2. ERDÖS, P. 1962. Remarks in number theory III (in Hungarian). Some problems in additive number theory. Mat. Lapok **13**: 28–32; see also LEVINE, E. & J. O'SULLIVAN. 1977. An upper estimate for the reciprocal sum of a sum-free sequence. Acta Arith. **34**: 9–24.
3. ERDÖS, P. & A. SÁRKÖZY. 1979. On the divisibility properties of sequences of integers. Proc. London Math. Soc. **21**(3): 97–101.