COLLOQUAI MATHEMATICA SOCIETATIS JÁNOS BOLYAI.
2. NUMBER THEORY, DEBRECEN (HUNGARY) 1968.

On divisibility properties of sequences of integers

by

P. Erdős, A. Sárközi and E. Szemerédi
Budapest, Hungary

In this paper we discuss the results which we obtained on sequences of integers in the last few years and also state some of the problems which we could not settle. First we review the older work on this subject, most of which can be found in the excellent book of Halberstam and Roth [13].

Let $A = \{a_1 < a_2 < \cdots\}$ be a sequence of integers. Put $A(x) = \sum_{a_i < x} 1$ The density of A (if it exists) is defined as

$$\lim_{x = \infty} \frac{A(x)}{x} .$$

The logarithmic density is defined as

$$\lim_{x = \infty} \frac{1}{\log x} \sum_{a_i < x} \frac{1}{a_i} .$$

It is obvious that if the density exists then the logarithmic density exists too, but the converse is not true.

Throughout this paper $c, c_1, \dots, C$ will denote positive absolute constants not necessarily the same at each occurrence. $\log_k x$ will denote the k-fold iterated logarithm. $\nu(n)$ denotes the number of distinct prime factors of n and $\nu_u(n)$ denotes the number of distinct prime factors of n not exceeding u.

A sequence of integers A is called primitive if no term divides any other. More than thirty years ago Chowla, Davenport and Erdős raised the question if every primitive sequence has density 0. This guess seemed reasonable at the time (it certainly holds for the integers having exactly k prime factors). It certainly was a great surprise to the senior author of this paper (the junior authors were then not yet alive) when Besicovitch [17] constructed a primitive sequence of positive upper density, he also constructed a sequence A so that the set of integers which are multiples of some $a \in A$ do not have a density. Behrend [16] and Erdős [18] proved that every primitive sequence has lower density 0. In fact Behrend proved that for every primitive sequence

$$(1) \qquad \sum_{a_i < x} \frac{1}{a_i} < c_1 \log x / \sqrt{\log \log x}$$

and Erdős proved that (for a sharpening of (2) see Alexander [1])

$$(2) \qquad \sum_{i=1}^{+\infty} \frac{1}{a_i \log a_i} < C .$$

Pillai proved that (1) is best possible i.e. there is a c_2 so that for every x there is a primitive sequence $a_1 < \dots < a_k < x$ satisfying

$$(3) \qquad \sum_{a_i < x} \frac{1}{a_i} > c_2 \log x / \sqrt{\log \log x} .$$

We can now ask the following. Let A be a primitive sequence. It is easy to see that $\max_A \sum_{a_i \leq x} 1 = \left[\frac{x+1}{2} \right]$. The following question is much more difficult. What is the maximum of $\sum_{a_i<x} \frac{1}{a_i}$

and for which sequence is this maximum assumed? Perhaps this question has no reasonable answer but we proved [3] sharpening a previous result of Anderson [2] that

(4)
$$\max_A \sum_{a_i<x} \frac{1}{a_i} = (1+o(1)) \frac{\log x}{(2\pi \log\log x)^{1/2}} .$$

Asymptotically the maximising sequence is the one which has $[\log\log x]$ prime factors, multiple factors counted multiply. By our method we can prove the following theorem: Let A be the union of k primitive sequences, then the value of $\max_A \sum \frac{1}{a_i}$ is asymptotic to the case when A consists of the integers the number of prime factors of which is between

$$[\log\log x - \frac{k}{2}] \text{ and } [\log\log x + \frac{k}{2}].$$

Through (3) and (4) are best possible for fixed x we proved that if A is an infinite primitive sequence then [4]

(5)
$$\lim_{x=\infty} \sum_{a_i<x} \frac{1}{a_i} \left(\frac{\log x}{(\log\log x)^{1/2}} \right)^{-1} = 0 .$$

It is easy to see that (5) is best possible.

The following problem seems difficult: Let $b_1 < \cdots$ be an infinite sequence of integers. What is the necessary and sufficient condition that there should exist a primitive sequence $a_1 < \cdots$ satisfying $a_n < cb_n$ for every n ?

From (2) and (5) we obtain that we must have

(6)
$$\sum_{i=1}^{\infty} \frac{1}{b_i \log b_i} < \infty \qquad \sum_{b_i<x} \frac{1}{b_i} = o\left(\frac{\log x}{(\log\log x)^{1/2}}\right) .$$

We know that (6) is not sufficient - it is not clear if a simple necessary and sufficient condition exists.

In [4] we state the following result. Let A be a primitive sequence and $x_1 < x_2 < \dots$ any sequence satisfying

(7) $$\log\log x_{\nu+1} > (1+c)\log\log x_\nu .$$

Put

$$\varepsilon_\nu = \sum_{a_i < x_\nu} \frac{1}{a_i} \frac{(\log\log x_\nu)^{1/2}}{\log x_\nu} .$$

Then $\sum \varepsilon_\nu < \infty$.

We thought that (7) can be weakened, but in the mean time we showed that (7) is best possible. In other words: if $\log\log x_{\nu+1} / \log\log x_\nu \to 1$ there always is a primitive sequence for which $\sum \varepsilon_\nu = \infty$.

In [4] we further proved the following theorems: Let $g(x)$ be an increasing function for which

$$\sum_n g(2^{2^n})/2^n < \infty .$$

Then for every primitive A

$$\liminf \sum_{a_i < x} \frac{1}{a_i} \Big/ g(x) = 0 .$$

On the other hand if $g_1(x) = \dfrac{\log x}{\log\log x \; h(x)}$, $h(x)$ increasing is such that

$$\sum_n g_1(2^{2^n})/2^n$$

converges then there is a primitive sequence for which

$$\lim_{x=\infty} \sum_{a_i<x} \frac{1}{a_i} / g(x) = \infty .$$

The method of Behrend easily gives that for every primitive sequence $(y = tx)$

$$\sum_{x<a_i<y} \frac{1}{a_i} < c \log t / (\log\log t)^{1/2}$$

This is all we know about primitive sequences.

Let A be any sequence of integers. Denote by $B(A)$ the set of all integers which have at least one divisor in A. Davenport and Erdős [19] proved that $B(A)$ always has a logarithmic density and that this density equals the lower density of $B(A)$. From this fact they deduced that if A has positive upper logarithmic density then there is an a_i in A so that the set of integers t for which $a_i t \in A$ also has positive upper logarithmic density. Then they deduced that if A has positive upper logarithmic density it must contain an infinite divisibility chain, i.e. a subsequence a_{i_j}, satisfying $a_{i_j} / a_{i_{j+1}}$.

We proved [5] the following sharpening of this result: If A has positive upper logarithmic density then A contains a divisibility chain a_{i_j} satisfying

(8) $$\sum_{a_{i_j}<y} 1 > c(\log\log y)^{1/2}$$

for infinitely many y.

We also show that (8) is best possible.

If A satisfies

(9) $$\limsup_{x=\infty} \frac{1}{\log\log x} \sum_{a_i>x} \frac{1}{a_i \log a_i} = c_1 > 0$$

then there is a divisibility chain a_{i_j} so that

(10) $$\limsup_{y\to\infty} \sum_{a_{i_j}<y} 1/\log\log y \geq c_2 , \quad c_2 = c_2(c_1) .$$

In these theorems $\limsup$ can not be replaced by $\lim$ -we in fact construct for every $g(n)\to\infty$ a sequence of density 1 so that for every divisibility chain we have for infinitely many n

$$\sum_{a_{i_j}<n} 1 = o(g(n)) .$$

Also in (10) $\log\log y$ can not be replaced by any function tending to infinity faster than $\log\log y$.

It is possible though that in (10) $c_2 = c_1$ (it is easy to see that $e^{-\gamma}c_1 \leq c_2 \leq c_1$, where γ is Eulers constant). would probably follow if we could prove the following conjecture: To every $\varepsilon > 0$ there is a k so that if $k < a_1 < \ldots$ is any primitive sequence then

(11) $$\sum_{i=1}^{+\infty} \frac{1}{a_i \log a_i} < 1 + \varepsilon .$$

We conjectured [5] that perhaps the following strengthening f the Davenport-Erdős theorem holds: Let A be a sequence of upper logarithmic density α, then there is an $a_i \in A$ so that the upper logarithmic density of the t's satisfying $a_i t \in A$ is $\geq \alpha$. Recently we observed that this conjecture is completely wrong-headed and fails even with $\varepsilon\alpha$ instead of α. To see this it suffices to let n be sufficiently large and consider the integers m for which

$$(1-\eta)\log\log n < \nu_n(m) < (1+\eta)\log\log n .$$

The density of these integers is by the result of Turán [15] as close to 1 as we please (if n is sufficiently large), but the density of the t's for which $a_i t \in A$ is as small as we please (if $n > n_0$).

By the methods of [6] we can show that if the logarithmic density of A is α then there is c so that for infinitely many $a_i \in A$ the logarithmic density of the t's satisfying $a_i t \in A$ is greater than $\frac{c}{\log a_i}$ It seems possible that this result can be slightly strengthened, perhaps to every c there is an a_i so that the logarithmic density of the t satisfying $a_i t \in A$ is $> \frac{c}{\log a_i}$. If true this conjecture is close to being best possible, since it is false with $\frac{c}{(\log a_i)^{1-\varepsilon}}$. This can be seen by a slight modification of the previous example. Let $\eta > 0$ be fixed, $n > n_0(\eta, \varepsilon)$ is sufficiently large. Our sequence A consists of the integers n satisfying for every $n_0 < N \leqq n$

$$\log\log N - (\log\log N)^{1/2+\eta} < \nu_N(n) < \log\log N + (\log\log N)^{1/2+\eta}.$$

It follows from [7] that the density of A is $> 1-\varepsilon$ and it is easy to see that the density of the t's for which $a_i t \in A$ is less than

$$\exp((\log\log a_i)^{1/2+c\eta})/\log a_i .$$

All the moment we can not decide about (11).

Using Kleitmans combinatioral results [8] we proved the following old conjecture:

Let A be an infinite sequence so that for infinitely many n_k

$$\sum_{a_i < n_k} \frac{1}{a_i} > c_2 \log n_k/(\log\log n_k)^{1/2}$$

then both equations

$$(a_i, a_j) = a_r \qquad [a_i, a_j] = a_s$$

have infinitely many solutions. (If we write henceforth $(a_i, a_j) = a_r$ or $[a_i, a_j] = a_s$ we always assume $a_i \nmid a_j$, $a_j \nmid a_i$.)

Further if for infinitely many n_k

$$\sum_{a_i < n_k} \frac{1}{a_i} > c_2 \log n_k/(\log\log n_k)^{1/4}$$

then the system of equations

$$(a_i, a_j) = a_r, \quad [a_i, a_j] = a_s$$

has infinitely many solutions. The exponents $\frac{1}{2}$ and $\frac{1}{4}$ in the denominators can not be diminished and in fact we constructed a sequence A satisfying

$$\sum_{a_i < x} 1 > cx/(\log\log x)^{1/2} \tag{12}$$

such that

$$[a_i, a_j] = a_r$$

has no solutions.

At first we thought that the same holds for the equation $(a_i, a_j) = a_r$ but later we proved [6] that if $(a_i, a_j) = a_r$ is not solvable then

$$\sum_{i=1}^{\infty} \frac{1}{a_i \log a_i} < c \tag{13}$$

(12) shows that the (13) can not hold for the equation $[a_i, a_j] = a_r$. Further we proved [9] that if $(a_i, a_j) = a_r$ is not solvable then

$$\sum_{a_i < x} \frac{1}{a_i} = o\left(\frac{\log x}{(\log\log x)^{1/2}}\right).$$

One would guess that the condition that $(a_i, a_j) = a_r$ is not solvable, is much weaker than the condition that A is primitive nevertheless these theorems seem to show that the sequences for which $(a_i, a_j) = a_r$ is not solvable seem to behave very much like the primitive sequences. . In fact we can not decide the following question: Let $b_1 < b_2 < \cdots$ be a sequence for which $(b_i, b_j) = b_r$ is not solvable. Does there then exist a primitive sequence $a_1 < \cdots$ satisfying

$$a_k < cb_k, \qquad k = 1, 2, \ldots ?$$

Finally we want to state one of our recent results which in some sense is definitive [10] . Let A have positive upper logarithmic density. Then there is an infinite subsequence $a_{i_1} < a_{i_2} < \cdots$ so that the least common multiple and greatest common divisor of any set of a_{i_j}'s is again in A . Further every two least common multiples are distinct. This in particular implies that no a_{i_j} divides any other. Our proof does not use the results of Kleitman. Our principal lemma is the following result of independent interest: Let A have positive upper logarithmic density. Then there is an $a_i \in A$ so that the sequence of a_j's satisfying

$$[a_i, a_j] \in A, \qquad a_i \nmid a_j$$

also has positive upper logarithmic density.

Before we leave this subject we would like to call attention to the following problems of a diophantine nature. Let $a_1 < a_2 < \cdots$ be a sequence of real numbers so that for every integer i, j and k

(14) $$|a_j - k a_i| \geqq 1 .$$

If the a's are integers then (14) means that $a_1 < a_2 < \cdots$ is a primitive sequence.

Is it now true that (14) implies

$$\sum_{i=1}^{\infty} \frac{1}{a_i \log a_i} < \infty$$

and

$$\sum_{a_i < x} \frac{1}{a_i} < c \log x / (\log \log x)^{1/2} \ ?$$

In fact we can not even prove that (14) implies

$$\liminf \frac{1}{x} \sum_{a_i < x} 1 = 0 .$$

Very recently Schmidt [14] asked the following question: Is there a set S of infinite measure on the line so that $x = ny$, $x \in S$, $y \in S$ integer is not solvable? Szemerédi proved (unpublished) that such a set exists.*

* We recently heard that the same result was obtained independently and simultaneously by Haight.

Let now S be a measurable set in $(0,\infty)$. Denote by $m(S,x)$ the measure of the intersection of S with the interval $(0,x)$. It is easy to see that if there is a sequence $x_n \to \infty$ satisfying $m(S,x_n) > cx_n$ then there is a sequence $y_n \in S$, $n = 1,2,\ldots$ so that for every n y_{n+1}/y_n is an integer. In view of this result and Szemerédi's example the following question remains open: Determine a function $f(x)$ tending to infinity as slowly as possible so that if $m(S,x_n) > f(x_n)$ for a sequence $x_n \to \infty$ then there is a $y_1 \in S$, $y_2 \in S$, y_2/y_1 integral. Clearly many similar questions can be formulated, but we leave these for the reader.

We conclude our report by stating some results of more analytic character. Denote

$$f(x) = \sum_{\substack{a_i/a_j \\ a_j < x}} 1$$

We proved [11] that for every sequence A of positive logarithmic density we have for infinitely many x

$$f(x) > x \exp\left(c_1 (\log_2 x)^{1/2} \log_3 x\right).$$

This result is best possible. There is in fact a sequence of positive density so that for every x

$$f(x) < x \exp\left(c_2 (\log_2 x)^{1/2} \log_3 x\right).$$

This theorem does not imply $f(x)/x \longrightarrow \infty$. Here we proved the following theorem of surprising accuracy: Put $\liminf A(x)/x = \alpha$. Assume $\frac{1}{k+1} < \alpha \leq \frac{1}{k}$. Then there is a $c_1 = c_1(\alpha)$ so that for every sufficiently large x

$$f(x) > x \exp\left(c_1 (\log_{k+1} x)^{1/2} \log_{k+2} x\right). \tag{15}$$

It was to us very surprising when we found that this theorem is nearly best possible. Let $\frac{1}{k+1} < \alpha < \frac{1}{k}$. Then there is a sequence A of density α and a constant $c_2 = c_2(\alpha)$ satisfying

(16) $$\liminf_{x=\infty} f(x)\left(x \exp\left(c_2(\log_{k+1} x)^{1/2} \log_{k+2} x\right)\right)^{-1} = 0 .$$

$c_2 = c_2(\alpha)$ tends to 0 if $\alpha \to 1/k+1$ and $k > 1$. Very likely this holds for $k = 1$, too.

Let finally $\alpha = \frac{1}{k}$ and $g(x)$ any function tending to infinity as x tends to infinity. Then there exsists a sequence of density $\frac{1}{k}$ for which

$$\liminf_{x=\infty} f(x) \left(x \exp\left(g(x)(\log_{k+1} x)^{1/2} \log_{k+2} x\right)\right)^{-1} = 0 .$$

Denote by $L(\alpha)$ the upper limit of the values of c_1 for which (15) holds. Clearly for every $c_2 > L(\alpha)$ (16) holds. It would be of interest to determine $L(\alpha)$ explicitely and to decide what happens for $c = L(\alpha)$. Very likely $L(\alpha)$ tends to infinity as α tends to $\frac{1}{k}$ but we can only prove this if $k = 1$.

Denote by $\ell(x)$ the smallest integer k for which $1 \le \log_k x < e$. It is very likely that the methods of [12] enable one to prove the following results: Let A be a sequence of satisfying for all large x

$$A(x) > (1+\varepsilon)\frac{x}{\ell(x)}$$

then $f(x)/x \to \infty$. On the other hand there exists a sequence A satisfying for all large x

$$A(x) > (1-\varepsilon)\frac{x}{\ell(x)}$$

and nevertheless $\liminf_{x=\infty} f(x)/x = 0$. We have not carried out the details of the proofs of these results and are not absolutely sure that they are correct. We have no idea what happens to $f(x)/x$ if $A(x) = (1+o(1))\frac{x}{\ell(x)}$.

Before we conclude this paper we would like to state a few unpublished problems and results. By the method of [4] we can prove that if the sequence A is such that for infinitely many x

$$\sum_{a_i < x} \frac{1}{a_i} > \varepsilon \log x / (\log \log x)^{1/2}$$

then

$$\limsup_{x = \infty} f(x)/A(x) = \infty \tag{17}$$

By the methods of [1] we can also prove that $\sum_i 1/a_i \log a_i = \infty$ also implies (17). We can further show by the methods of [3] that if ($k \geq 1$ integer)

$$\sum_{a_i < x} \frac{1}{a_i} > (k + \varepsilon) \log x / (2\pi \log \log x)^{1/2}$$

then

$$f(x) > c_\varepsilon x (\log \log x)^{k - \frac{1}{2}} \tag{18}$$

and if for an infinite sequence A, for infinitely many x

$$\sum_{a_i < x} \frac{1}{a_i} \geq (k + o(1)) \log x / (2\pi \log \log x)^{1/2}$$

then

(19)

(19) $$\limsup_{x=\infty} f(x)/A(x)(\log\log x)^{k-1} = \infty .$$

It is not difficult to see that both (18) and (19) are best, possible.

It would be interesting to prove the following conjecture: Assume that

(19) $$A(x) > \varepsilon x/(\log\log x)^{1/2}$$

for every x. Then

(20) $$\limsup_{x=\infty} f(x)/x = \infty .$$

We can only show that (19) implies

$$\limsup_{x=\infty} f(x)/x > 0 ,$$

but we think it likely that a considerably weaker condition than (19) will imply (20).

REFERENCES

[1] R. ALEXANDER, Density and multiplicative structure of sets of integers, Acta Arith. 12 (1967), 321-332.

[2] I. ANDERSON, On primitive sequences, Journal London Math. Soc. 42 (1967), 137-148.

[3] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On an extremal problem concerning primitive sequences, ibid. 42. (1967), 484-488.

[4] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On a theorem of Behrend, Journal Australian Math. Soc. 7. (1967), 9-16.

[5] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On divisibility properties of sequences of integers, Studia Sci. Math. Hung. 1 (1966), 431-425.

[6] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, Doklady Acad. Sci. USSR 176 (1967), 541-544.

[7] P. ERDŐS, On the distribution function of additive functions. Annals of Math. 97 (1946), 1-20. see p. 4.

[8] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On the solvability of the equations $[a_i, a_j] = a_r$ and $(a'_i, a'_j) = a'_r$ in sequences of positive density, Journal Math. Anal. and Applications, 15 (1956), 60-64.

[9] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, Über Folgen Ganzen Zahlen, Abhandlungen aus Zahlentheorie und Analysis (Zur Erinnerung an E. Landau), VEB Deutscher Verlag der Wissenschaften, Berlin 1968, 79-86.

[10] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On the solvability of certain equations in sequences of positive upper logarithmic density, Journal London Math. Soc. 43 (1968), 71-78.

[11] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On the divisibility properties of sequences of integers I, Acta Arith. 11 (1966), 411-418.

[12] P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI, On the divisibility properties of sequences of integers II, ibid 14 (1968), 1-12.

[13] H. HALBERSTAM and K.F. ROTH, Sequences, Clarendon Press, Oxford 1966, see Chapter V.

[14] Schmidts paper will appear in Studia Sci. Math. Hung.

[15] P. TURÁN, On a theorem of Hardy and Ramanujan, Journal London Math. Soc. 9 (1934), 274-276.

[16] F. BEHREND, On sequences of numbers not divisible one by another. Journal London Math. Soc. 10/1935/. 42-44.

[17] A.S. BESICOVITCH, On the density of certain sequences, Math. Ann. 110 (1934), 336-341.

[18] P. ERDŐS, Note on sequences of integers no one of which is divisible by any other, J. London Math. Soc. 10. (1935), 126-128.

[19] H. DAVENPORT and P. ERDŐS, On sequences of positive integers, Acta Arith. 2. (1936), 147-151.