ON SOME EXTREMAL PROPERTIES OF SEQUENCES OF INTEGERS

By

P. ERDŐS, A. SÁRKÖZI and E. SZEMERÉDI

Mathematical Institute of the Hungarian Academy of Sciences, Budapest

(Received 15 March, 1969)

Let $\{A\} = a_1 < a_2 < \dots$ be a sequence of positive integers. Put $A(n) = \sum_{a_i \le n} 1$. Denote by $f_k(n)$ the smallest integer so that every sequence A satisfying $A(n) = f_k(n)$ contains a subsequence of k terms which are pairwise relatively prime. It is easy to see that $f_2(n) = \left[\frac{n}{2}\right] + 1$ and it seems likely that

$$f_k(n) = 1 + \psi_{k-1}(n) \tag{1}$$

where $\psi_{k-1}(n)$ denotes the number of integers not exceeding n which are multiples of at least one of the first $k-1$ primes $2, 3, \dots, p_{k-1}$. Clearly (1) if true is best possible. (1) is easy to show for $k = 3$, but we have not been able to prove it in general. On the other hand we prove in a sharper and more general form several conjectures stated in [1]. First we introduce some notations. $A_{(m,u)}$ denotes the integers $a_i \in A$, $a_i \equiv u (\operatorname{mod} m)$ ($A_{(m,u)}(n)$ denotes the number of terms of the sequence $A_{(m,u)}$). $A_{(2,1)}$ respectively $A_{(2,2)}$, we will denote by A_1 respectively A_2. $\varphi(n)$ denotes Euler's φ function.

$$\varphi(A, k) = \sum_{\substack{a_i \le n \\ (a_i, k) = 1}} 1.$$

$\Phi(A)$ denotes the number of pairs $(a_i, a_j) = 1$, $a_i < a_j \le n$. Put

$$F(n) = \min_{A} \max_{a_j \in A} \varphi(A, a_j)$$

where the minimum is to be taken over all sequences A satisfying $A(n) \ge \left[\frac{n}{2}\right] + 1$.

For simplicity we will henceforth assume that n is even, all our results could easily be extended for odd n. $c_1, c_2 \dots$ denote suitable positive absolute constants.

Let $A(n) = \left[\frac{n}{2}\right] + 1$. P. ERDŐS proved that for $n > n_0$ $\Phi(A) > c_1 n / \log \log$

and in fact the minimum of $\Phi(A)$ is assumed if A consists of the even numbers and u_r where

$$u_r = 3\cdot 5, \ldots, p_r, \; 3\ldots p_r \le n < 3\ldots p_r \cdot p_{r+1}.$$

He also conjectured that

$$\lim_{n=\infty} F(n) = \infty. \tag{2}$$

We now prove (2). In fact we prove the following sharper

THEOREM 1.

$$F(n) > c_2 n/\log\log n.$$

We first prove two other theorems which will easily imply Theorem 1.

THEOREM 2. *Let A satisfy*

$$A_1(n) = s, \qquad 1 \le s < c_3 n, \tag{3}$$

$$A(n) > \frac{n}{2}. \tag{4}$$

Then for $n > n_0$

$$\max_{a_i \in A} \varphi(A, a_i) > c_4 n/\log\log\frac{n}{s} \tag{5}$$

and

$$\Phi(A) > c_5 sn/\log\log\frac{n}{s}. \tag{6}$$

We need the following known

LEMMA 1. *The number of integers $1 \le k \le n$ satisfying $\varphi(k)/k < 1/t$ is less than* ($\exp z = e^z$) $n \exp(-\exp c_6 t)$, *uniformly in $t > 1$.*

Choose

$$t = \frac{1}{c_6}\log\log\frac{2n}{s}. \tag{7}$$

We obtain from Lemma 1 that the number of integers $1 \le k \le n$ which satisfy $\varphi(k)/k < 1/t$ (where t is defined by (7)) is less that $s/2$. Thus the number of integers $a_i \in A_1$ for which $\varphi(a_i)/a_i > 1/t$ is by (3) greater than $s/2$. Denote now by $b_1 < \ldots < b_r \le n$, $r > s/2$ the integers in A_1 satisfying $\varphi(b_i)/b_i > 1/t$. Clearly the number of integers $2u \le n$ satisfying $(2u, b_i) = 1$ is greater than $\frac{n}{2}\,\frac{1}{2}\,\varphi(b_i)/b_i > n/4t$. Thus (in Σ' a_j runs through the numbers of A_2)

$$\sum_{(b_i, a_j)=1}' 1 > A_2(n) - \frac{n}{2} + \frac{n}{4t}, \tag{8}$$

or by (8) and (7) for sufficiently small c_3 and c_4 we obtain by a simple computation for sufficiently large n

$$(9) \quad \varphi(A, b_i) > A_2(n) - \frac{n}{2} + \frac{n}{4t} = A(n) - A_1(n) - \frac{n}{2} + \frac{n}{4t} > \frac{n}{4t} - s > c_4 n/\log\log\frac{n}{s},$$

which proves (5).

To prove (6) observe that (9) holds for every $1 \leq i \leq r$ and $r > \frac{s}{2}$, hence from (9)

$$\Phi(A) > \frac{c_4}{2}\, sn/\log\log\frac{n}{s},$$

which proves (6) and hence the proof of Theorem 2 is complete.

THEOREM 3. *To every c_9 there is a $c_{10} = c_{10}(c_9)\left(c_{10}\right.$ is bounded in terms of $\left.\frac{1}{c_9}\right)$ so that if $A_1(n) = s > c_9 n$ and $A(n) > \frac{n}{2}$ then for $n > n_0$*

$$\Phi(A) > c_{10} n^2.$$

For $s < c_3 n$ Theorem 3 would follow from Theorem 2, but for the large values of s we need a separate proof.

Denote by P_r the product of the primes not exceeding r. We first prove

LEMMA 2. To every $\varepsilon > 0$ and $\delta > 0$ there is an $r = r(\varepsilon, \delta)$ so that if $n > m_0(\varepsilon, \delta, r)$ then for all but $\varepsilon\frac{n}{P_r}$ integers k satisfying

$$1 \leq k \leq n, \qquad k \equiv u (\mathrm{mod}\; P_r)$$

we have

$$\alpha(k) = \prod_{\substack{p \mid k \\ p > r}} \left(1 - \frac{1}{p}\right) > 1 - \delta.$$

The Lemma is very easy to prove and we only outline it. We evidently have (in $\prod'$ $k \equiv u (\mathrm{mod}\; P_r)$, $1 \leq k \leq n$)

$$(10) \quad \prod{}' \alpha(k) > \prod_{r<p<n} \left(1 - \frac{1}{p}\right)^{\frac{n}{pP_r}+1} > \prod_{p<n}\left(1 - \frac{1}{p}\right)\left(\prod_{r<p<n}\left(1 - \frac{1}{p}\right)^{1/p}\right)^{n/P_r} > (1 - \eta_r)^{n/P_r},$$

where η_r can be chosen as small as we wish if r is sufficiently large. (10) implies Lemma 2 by a simple argument.

Now we prove Theorem 3. We evidently have

$$(11) \quad \sum_{i=1}^{\frac{1}{2}P_r} \left(A_{(P_r,\, 2i-1)}(n) + A_{(P_r,\, 2i)}(n) + A_{(P_r,\, 2i+1)}(n)\right) = A_2(n) + 2A_1(n) = A(n) + s > \frac{n}{2} + s.$$

Hence by (11) there is an i_0 for which

$$A_{(P_r, 2i_0-1)}(n)+A_{(P_r, 2i_0)}(n)+A_{(P_r, 2i_0+1)}(n)>\frac{n+2s}{P_r}. \tag{12}$$

Clearly for every u $A_{(P_r, u)}(n)<\frac{n}{P_r}+1$. Thus we obtain from (12) that there are two integers u_1 and u_2, u_1 odd, $|u_1-u_2|=1$ or 2 satisfying

$$A_{(P_r, u_l)}(n)\geq\frac{1}{2}\left(\frac{2s}{P_r}-1\right)\quad (l=1, 2). \tag{13}$$

Denote now by $a_1^*<\ldots<a_t^*$ the sequence of integers for which

$$k\in A_{(P_r, u_1)} \text{ and } \prod_{\substack{p|k\\ p>r}}\left(1-\frac{1}{p}\right)>1-c_9/10. \tag{14}$$

From Lemma 2 and (13) we have for $r>r_0$, $\varepsilon=\frac{1}{3}c_9$ $(s>c_9n)$

$$t>A_{(P_r, u_1)}(n)-\frac{\varepsilon n}{P_r}>\frac{2}{3}\frac{s}{P_r}-\frac{\varepsilon n}{P_r}>c_9n/3P_r. \tag{15}$$

Now we estimate from below the number of solutions of

$$(a_i^*, a_j)=1,\qquad a_j\in A_{(P_r, u_2)}. \tag{16}$$

Assume $p|(a_i^*, b)$, $b\equiv u_2 (\bmod P_r)$. As $|u_1-u_2|\leq 2$ and u_1 is odd, we have $p>r$. Denote by $B_i(P_r, u_2)$ the number of integers $b\leq n$, $b\equiv u_2 (\bmod P_r)$ for which $(b, a_i^*)=1$. We have by a simple argument

$$\left|B_i(P_r, u_2)-\frac{n}{P_r}\prod_{\substack{p|a_i^*\\ p>r}}\left(1-\frac{1}{p}\right)\right|<2^{V(a_i^*)}<2^{2\log n/\log\log n} \tag{17}$$

since it is well known (and follows from the prime number theorem or a more elementary theorem) that for $m<n$ $V(m)<2\log n/\log\log n$.

Thus from (14) and (17) for sufficiently large n

$$B_i(P_r, u_2)>(1-c_9/7)n/P_r. \tag{18}$$

From (18) and (13) we obtain that the number of solutions of (16) is greater than $(s>c_9n)$

$$\frac{1}{2}\left(\frac{2s}{P_r}-1\right)-c_9n/4P_r>c_9n/2P_r. \tag{19}$$

From (15) and (19) wo evidently have

$$\Phi(A)>c_9^2n^2/6P_r^2$$

where r is bounded in terms of $1/c_9$ which proves Theorem 3.

It is now easy to prove Theorem 1. Let A be any sequence satisfying $A(n) \geq \frac{n}{2}+1$. We distinguish two cases. Assume first $A_1(n) < c_3 n$. In this case (5) and the definition of $F(n)$ implies Theorem 1. Assume next $A(n) \geq c_3 n$. Then from Theorem 3 we have

$$\max_{a_j} \varphi(A, a_j) \geq \Phi(A)/n > c_{10}(c_3)n$$

which completes the proof of Theorem 1.

We outline the following sharpening of Theorem 1.

THEOREM 4. *Let $n > n_0$. The only class of sequences A^* for which $F(n)$ is assumed is defined as follows: $A^* = A_1^* \cup A_2^*$,* where *A_1^* consists of all odd multiples not exceeding n of u_r $(3 \dots p_r \leq n < 3 \dots p_r p_{r+1},\ u_r = 3 \dots p_r)$ and A_2^* consists of the set of even numbers* (not exceeding n) *from which $A_1^*(n)-1$ even numbers relatively prime to u_r have been omitted.*

Theorem 4 clearly implies that

$$F(n) = \varphi_n^{(2)}(u_r) - \left[\frac{n}{u_r}\right] + \left[\frac{n}{2u_r}\right] + 1 \tag{20}$$

where $\varphi_n^{(2)}(u_r)$ denotes the number of even integers not exceeding n which are relatively prime to u_r.

We only outline the proof of Theorem 4. Let A $\left(A(n) \geq \frac{n}{2}+1\right)$ be any sequence which contains an odd number u which is not a multiple of u_r. A simple argument shows (see [1]) that

$$\varphi_n^{(2)}(u) > \varphi_n^{(2)}(u_r) + c_{11} n/(\log n)^2.$$

Thus if

$$\varphi(A, u) \leq \varphi_n^{(2)}(u_r) - \left[\frac{n}{u_r}\right] + \left[\frac{n}{2u_r}\right] + 1$$

we must have $A_1(n) = s > c_{11} \frac{n}{(\log n)^2}$. But then from (5) and Theorem 3 we have

$$\max_{a_i \in A} \varphi(A, a_i) > c_{12} n/\log\log\log n > \varphi_n^{(2)}(u_r)$$

which proves Theorem 4. Theorem 4 implies by a well known theorem of Mertens that (C is Euler's constant)

$$F(n) = (1+o(1))e^{-C} n/\log\log n.$$

References

[1] P. ERDŐS, Remarks on number theory, V. Extremal problems in number theory II. (in Hungarian), *Mat. Lapok*, **17** (1966), 135–155.

[2] P. ERDŐS, Some remarks about additive and multiplicative functions, *Bull. Amer. Math. Soc.*, **52** (1946), 527–537.